INTRODUCTION TO TIME-SERIES ANALYSIS

G. Rowlands
J.C. Sprott

PLP 1079

December 1990

Plasma Studies

University of Wisconsin

These PLP Reports are informal and preliminary and as such may contain errors not yet eliminated. They are for private circulation only and are not to be further transmitted without consent of the authors and major professor.

Introduction to Time-Series Analysis

G. Rowlands and J. C. Sprott

In most experimental situations it is usually only possible to measure a limited number of variables for a limited interval of time. Furthermore, these variables may not be directly related to the variables that control the dynamics (the time evolution of the system). For example in the experimental study of fluid flows it is usually only the velocity of the fluid at a few particular points in space which is measured as a function of time. However, to obtain a complete understanding of the fluid flow, it is necessary to know the velocity at all points in the fluid.

An obvious question is what information can one glean from restricted information such as a single time series? This of course is not a new question, and the problem has received much attention in the past.

A method which has been found, in the right circumstance, to be very powerful is where the time series is expressed in terms of its Fourier modes. If the number of modes is found to be small then the dynamics of the system can be expressed in terms of the dynamics of these modes. Mathematically this means that the time series for a quantity T(t) may be expressed in the form

$$T(t) = \sum_{m=1}^{M} A_m e^{i\omega_m t}$$
(1)

with M a small integer. The power spectrum $P(\omega) = |\int T(t)e^{i\omega t} \, dt|^2$ is then just a series of delta functions at $\omega = \omega_m$ (m taking all integer values between 1 and M) of strength $|A_m|^2$. In fact the way this method of representation is found to be appropriate is to use T(t) to evaluate the power spectrum and see if it is of the form of a finite number of peaks. In practice the presence of external noise broadens the delta-functions, but, if the noise is not too strong, the peaks in $P(\omega)$ can still be recognized and hence A_m and ω_m calculated. The above expression for T(t) then gives a smoothed or noise-free representation of the original data.

An important question remains of how best to represent these smoothed data? One way is to use the phase plane where T(t) for example is plotted against dT/dt. This latter quantity is readily obtained from Eq.(1) but not from the original noisy data. For M=1 (a single mode) such a plot would reveal a simple closed loop, while for $M=m_0$ the (m_0+1) -dimensional phase-space portrait whose ordinates are T(t), dT/dt, ... dT^{MO}/dt^{MO} , would also be a simple m_0 -dimensional loop or torus (depending on whether the frequencies are harmonically related or not). Incidentally, this type of analysis illustrates how a single time-varying quantity T(t) can give multi-dimensional information. However, in a significant number of experimental

situations which fall under the umbrella of turbulence, the power spectrum $P(\omega)$ does not have very much structure. That is, it falls off monotonically with ω , and Fourier peaks are poorly resolved or not present. In such a case an expression in the form of Eq. (1) is not appropriate, although a Fourier transform integral rather than a sum does convey some information.

An alternative is to use the method of singular value decomposition. 1,2 This method, originally used in statistics to study linear problems, has more recently been applied, under a number of disguises, to inherently nonlinear physical problems. The method replaces the expansion of T(t) in Fourier modes by an expansion in terms of another complete set of functions. Importantly this set is obtained from a numerical analysis of the data and is not imposed, as in the Fourier series representation, from outside. Thus instead of Eq. (1) we write

$$T(t) = \sum_{m=1}^{\infty} \psi_m(t)$$
 (2)

where the ψ_m 's form a complete orthogonal set of functions. By analogy with $P(\omega)$ we define $P_m = (\int T(t)\psi_m(t)dt)^2$, which is just $P_m = \int \psi_m^2 dt$, and this latter form we can interpret as the probability of the system being in the state ψ_m . Thus again by analogy with Fourier modes one can anticipate there are situations where T(t) can be adequately represented by just a few of the ψ_m 's. Then just as a few peaks in the power spectrum indicates the possibility of an expansion in a small number of Fourier modes, the existence of a few dominant P_m 's indicates the possibility of a useful expansion in a small number of the ψ 's.

Singular value decomposition is a method for the generation of an orthogonal set of functions. The essentials of the method are as follows. The data are assumed to be known at a finite number, N, of equally spaced intervals of time, Δt . Writing $T_n = T(t=n\Delta t)$ we assume T_n is known for $n=1,\ldots N$. From these data we construct a set of N vectors (actually it is N - M vectors, but in practice N >> M, and so we will simply refer to N vectors) \underline{V}_ℓ of dimension M defined such that

$$\underline{\mathbf{y}}_{\ell} = \{\mathbf{T}_{\ell}, \ \mathbf{T}_{\ell+1}, \ \dots \ \mathbf{T}_{M+\ell-1}\}$$
 (3)

We also construct the auto-correlation function defined by

$$C(n) = \sum_{\ell=1}^{N} T_{\ell} T_{\ell+n}$$
(4)

The choice of Δt and M is important, but a discussion is left to later.

Using the values of C(n) one constructs the symmetric M X M correlation Matrix $\underline{\underline{M}}$ with elements $\underline{M}_{\ell p} = C(|\ell-p|)$. This matrix has M eigenvalues which we denote by λ_m and corresponding eigenfunctions $\underline{\alpha}_m$. Using these, one defines the functions $\psi_m(t)$ such that

$$\psi_{\mathbf{m}}(\mathsf{t}=\mathsf{n}\Delta\mathsf{t}) = \underline{\alpha}_{\mathbf{m}} \cdot \underline{\mathbf{v}}_{\mathbf{n}} \tag{5}$$

These functions are orthogonal and normalized such that

$$\begin{array}{ll} 1 & N \\ - & \Sigma \psi_{\rm m}^2(\ell \triangle t) = \lambda_{\rm m} \\ N & \ell = 1 \end{array} \tag{6}$$

and so $P_m = \lambda_m$, and hence λ_m is a measure of the importance of the mode ψ_m in the expansion of T(t). Thus by analogy with the usual procedure of expanding in a finite number of Fourier modes, we now expand in a finite number, d, say, of the ψ 's and select those ψ 's which correspond to the d largest values of λ_m . Thus as an approximation to the original data T(t) we write

$$T_{D}(t) = \sum_{m=1}^{d} \psi_{m}(t)$$
(7)

Alternatively, λ_{m} may be considered as a measure of the time the function T(t) lies in a direction parallel to ψ_{m} , in which case the choice of ψ_{m} as defined above corresponds to maximizing this time.

Besides giving the best set of orthogonal functions, the above method, as does the Fourier series expansion, involves some smoothing of the original data T(t). This is of course desirable since most experimental data have a significant noisy component. A purely random time series leads to a correlation matrix \underline{M} which is diagonal ($C(n) = C_0 \delta_{n,0}$) and whose eigenvalues are all equal to C_0 . Thus for any other data, the existence of eigenvalues, λ_m , which are greater than C_0 reveal the presence of structure in the data. By limiting the summation in Eq. (7) to such eigenvalues one automatically removes a substantial amount of the noise, so that $T_D(t)$ is a smoothed version of T(t). Furthermore, since the correlation function as defined by Eq. (4) is essentially a time average of the data, some of the noise will be averaged away. The singular value decomposition method is identical to the Karhussen-Loéve expansion and as such was originally suggested by Lumley to study turbulence. 3 , 4

A very powerful way of identifying any underlying structure in the original data T(t) is to examine the d-dimensional phase space constructed using the functions $\psi_1, \psi_2, \ldots, \psi_d$. Such plots reveal the topological structure of the solution. Importantly this is the same structure as one would obtain by considering a phase space constructed using the vectors \underline{V} as defined above, that is, the original data. However, topological details in this latter phase-space plot could be masked due to the presence of noise in the \underline{V} 's. The fact that the topological structure is conserved in going from a \underline{V} -based phase space to a ψ -based phase space is simply due to the fact that, by construction, the ψ 's are just linear combinations of the \underline{V} 's. An additional feature of singular value decomposition is that a two-dimensional plot of ψ_2 versus ψ_1 amounts to a rotation of the attractor in such a way that it is viewed from its broadest side, and thus its

structure is most readily apparent.

The Fourier mode method is useful when the number of modes needed for a description of the problem is small. The system can then be described in terms of a small number of weakly coupled oscillators, a picture which usually corresponds to one's physical intuition. One would of course like an analogous situation to exist when the system is described in terms of a small number of the ψ 's. However, this is not so, and at present it seems that a better understanding can best be achieved by studying some of the results of dynamical systems theory.

There is now much evidence from the theory of dynamical systems that apparently complicated time behavior can result from the solution of relatively simple differential or difference equations. For example, the logistic equation 5

$$x_{n+1} = \lambda x_n (1-x_n) \tag{8}$$

is known to have chaotic (pseudo-random) solutions for certain values of the parameter λ . The Lorenz equations 5 , 6

$$dx/dt = \sigma(y-x)$$

$$dy/dt = rx - y - xz$$

$$dz/dt = xy - bz$$
(9)

where σ , r and b are constants, also have chaotic solutions. Furthermore, these latter solutions are associated with the existence of a strange attractor in the x, y, z phase space.

Though neither of these equations has been shown to be a good model for a real physical system in the sense that the variables have a direct physical interpretation, the fact that such simple equations show complicated time behavior leads one to hope that complicated time behavior, as found in the real world for example in turbulence, may be modeled by simple equations. For this reason it seems reasonable to expect that equations for the time evolution of the $\psi_{\rm m}$'s, as defined above, may be simple. Given that the $\psi_{\rm m}$'s retain the topological features of the original data T(t), these simple equations, if they exist, would also retain these features. Thus such equations would model many of the important aspects of T(t).

From a physical point of view one can interpret the ψ_m 's for $1 \leq m \leq d$ as coherent structures, and then the expansion of T(t) as expressed by Eq. (7) is an expansion in terms of these structures. This is to be contrasted with the expansion as given by Eq. (1) which is in terms of Fourier modes. The coherent modes could themselves be Fourier modes, but then nothing has been achieved, but more generally they could be complicated linear combinations of such modes. In this way of thinking, the x, y and z of the Lorenz equations would be identified as relating to these distinct coherent structures, not to the amplitude of simple sinusoidal modes as assumed in the original deviation.

Given the nature of their evaluation it is convenient to assume the equations of motion for the ψ 's are of the form

$$\overline{\psi}_{m}(t+\Delta t) = F_{m}(\overline{\psi}_{n}(t))$$
 (10)

where the bar over the ψ 's is used to denote a value satisfying these model equations and F_m is a nonlinear function of all the $\overline{\psi}_n$'s from n=1 to d. Guided by the argument that simple nonlinear functions are sufficient to produce chaotic behavior we assume a form

$$F_{m}(\vec{\psi}_{n}) = a_{mo} + \sum_{q=1}^{d} a_{mq} \vec{\psi}_{q} + \sum_{q \neq \ell} b_{mq\ell} \vec{\psi}_{q} \vec{\psi}_{\ell} + \sum_{q \neq \ell} c_{mq\ell} p \vec{\psi}_{q} \vec{\psi}_{\ell} \vec{\psi}_{p}$$
(11)

that is a cubic polynomial characterized by the constants a, b and c. These constants are determined by obtaining, in the least squares sense, the best fit of the solutions of the equations of motion (10) to the known form for the ψ_m 's. That is, the quantities J_m , where

$$J_{m} = - \sum_{\substack{N \text{ s=1}}} \{\psi_{m}((s+1)\Delta t) - F_{m}(\psi_{n}(s\Delta t))\}^{2}$$
(12)

are minimized with respect to the coefficients a, b and c. Computer software that carries out the procedure described above as well as many other tests for chaotic time series is available.⁷

Since the ψ_m 's are normalized to the eigenvalues λ_m , it is natural to impose the condition given by Eq. (6) on the solution of the equations of motion, namely Eq. (10). This is readily achieved by introducing Lagrange multipliers η and minimizing

$$I_{m} = J_{m} + \eta_{m} \begin{vmatrix} 1 & N \\ - & \Sigma \psi_{m}^{2}((s+1)\Delta t) - \lambda_{m} \end{vmatrix}$$

$$N = 1$$
(13)

In this particular case, this procedure is equivalent to minimizing the J_m 's and multiplying all the coefficients by a common factor to ensure the normalization. Such a normalization procedure is useful in that it removes damping introduced by numerical procedures (such as the finite data-sample rate) and not present in the original data. In principle other constraints, such as symmetry requirements, may be imposed on the assumed form of the functions F_m .

The original data T(t) are given for a finite time interval only (N is finite). However, once the equations of motion for the ψ_m 's are known they can be solved for all time. Then using Eq. (7) we have an estimate $T_D(t)$, for T(t) for all time. Thus in principle we have an extrapolation scheme. However, this scheme should be used with caution. If the underlying system is chaotic then due to sensitivity to initial conditions we only expect $T_D(t)$ to be a good representation of T(t) for a short period of time after $t = N \Delta t$. On the other hand since the $\overline{\psi}$'s and the model Eqs. (10) capture the topology of the

original data, these model equations should provide good topological information. Thus for example a major obstacle in the calculation of fractal information is that there are usually not sufficient data, that is, N is too small. However, one can now use the model equations to generate as many data points as required and use them to calculate the various measures of the attractor's dimension. This whole procedure raises an interesting mathematical point about the errors in the various measures as obtained directly and as obtained using the model equations.

It is much simpler from a computational point of view to calculate Lyapunov exponents⁹ from equations than directly from data. The above procedure of first obtaining equations thus leads to a significant simplification.

There is another way in which the above methods can be used to extrapolate experimental data. Time series are usually known for the same experiment but where some adjustable parameter or parameters can be changed. For simplicity, consider the case where the data T(t) depend on a single parameter μ and the T(t)'s are known for a number of different values of μ .

The methods outlined above can be applied to the data for each value of μ . This will result in values of $\bar{\psi}$, λ , a, b and c which also depend on μ . There are good reasons to believe that the dependence of the a's, b's and c's on μ is simple. In this case by using the experimental data, namely T(t) for a range of values of μ , and a least squares fitting procedure these constants can be expressed as simple polynomial functions of μ .

An extrapolation scheme based on extrapolation of these polynomials in μ can then be used to study the changes in the topology of the solution T(t), for example, the onset of chaos or the disappearance of chaos and the emergence of periodic behavior. This extrapolation scheme is insensitive to initial conditions and does not have the associated restrictions of the direct extrapolation procedure outlined above.

The reasons for believing that the dependence of the coefficients such as a on μ is simple comes from a study of equations such as (8) and (9). As is well known, the solution to Eq. (8) can go through a complicated period-doubling sequence leading to chaos simply by changing the parameter λ . But λ corresponds to μ , and the coefficients a and b are just linearly proportional to μ . The same behavior is found for the Lorenz equations. The solutions of these equations show a whole range of behavior as the quantity r is allowed to vary. In the Landau theory of phase transformations the phase transformation arises because a parameter, namely, the temperature as measured relative to the critical temperature, goes linearly through zero. Thus a simple linear variation in a parameter in the equations is sufficient to cause a dramatic change (catastrophe) in the state of the system.

In the above method there are two quantities, namely M the order of the correlation matrix and d the number of significant eigenfunctions retained. These are to be considered as parameters of the method which can be adjusted to obtain the best fit between the real system under investigation through the data T(t) and the solution of the model Eqs. (10). Since we envisage applying the method to situations where the auto-correlation function shows little structure, we hope the complicated time variation can be attributed to the presence of a strange attractor. Then the parameters M and d are chosen to represent best the topological features of the attractor.

To illustrate some of the above techniques we have applied the methods to a few selected model situations. In the first case we consider the time series generated by iterating the logistic Eq. (8). Then with $T(t=n\Delta) = T_n$ we simply identify T_n with x_n .

An almost trivial procedure is to consider the case where the correlation matrix is of order one. The eigenfunction must then be x_n . If the data are then fitted to an equation of the form of Eq. (10), one recovers the logistic equation with coefficients accurate to within the The result of this procedure with $\lambda = 4$ is plotted round-off errors. in Fig. 1(a). Importantly this applies to all types of solution of Eg. not just those which are chaotic. For the case of a 2 X 2 correlation matrix there are two eigenfunctions. A phase-space plot as shown in Fig. 1(b) reveals a single closed loop which is just a distortion of the loop in a phase space generated by x_n and x_{n-1} , for this case the two eigenfunctions are just different linear combinations of x_n and x_{n-1} . For this case, $x = \psi_1 + \psi_2$. One can then proceed to get two coupled equations to describe the time evolution of the ψ 's. These are not readily recognizable as related to the logistic equation, but they can be factored and reduced to this equation. factorization is to be expected since the phase-space trajectory is a The same considerations apply if one takes higher-order The simple loop structure in the phase space correlation matrices. still remains if one considers higher-order correlation matrices (M>2) but restricts the number of eigenfunctions used in the model equations (10) to be two.

A slightly less trivial example is the Hénon map⁵, 10

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

 $y_{n+1} = bx_n$ (14)

which is a two-dimensional generalization of the logistic equation. The data values obtained by iterating Eqs. (14) with a = 1.4 and b = 0.3 are shown in Fig. 2(a). This case requires a 2 X 2 correlation matrix and produces values of ψ_1 and ψ_2 whose phase-space plot as shown in Fig. 2(b) is topologically equivalent to the original attractor. In this case, $x = \psi_1 + \psi_2$. The equations produced by a least squares fit using Eqs. (11) and (12) have a solution that fits the original data to within round-off errors and is indistinguishable from Fig. 2(b).

The Hénon map provides an opportunity to illustrate how the analysis method discriminates against noise in the data. A time series of 2300 values of x was generated from successive iterates of Eq. (14), to which was added normally distributed deviates with zero mean and standard deviation of 0.1. The resulting input data as shown in Fig. 3(a) is a noisy version of Fig. 2(a). Two eigenfunctions were used, and the resulting model equations were solved to generate a new time series given by $x = \overline{\psi}_1 + \overline{\psi}_2$ whose values are plotted in Fig. 3(b). The noise is totally removed, and the map is indistinguishable from the This reduction of noise is a result of forcing the data to Hénon map. fit a relatively simple functional form involving only two equations whose solution must therefore be an attractor with at most dimension two.

The next case considered is where the time series was generated by solving the Lorenz equations (9) and identifying T(t) with x(t). The neglect of the information contained in the solutions y(t) and z(t) mirrors the experimental situation where only a limited amount of information is available. The traditional values of $\sigma = 10$, r = 28 and b = 8/3 have been chosen so as to lead to a strange attractor. time series with $\Delta t = 0.05$ has been analyzed, and it is found that for a range of M values, three, or sometimes four, of the eigenvalues are significantly larger than the others. Thus we assume that a reasonable approximation of the results can be allowed by restricting the analysis to just three eigenfunctions and values, that is, d = 3. In Fig. 4(a) we plot the phase space defined by the functions x_n , x_{n-1} and x_{n-2} . This reveals the topological nature of the strange attractor. A phase space constructed from $\psi_1(t)$, $\psi_2(t)$ and $\psi_3(t)$ is shown in Fig. 4(b). For this case $x(t) = 0.813 \ \psi_1(t) - 1.338 \ \psi_2(t) - 1.507 \ \psi_3(t)$. The topological features are the same, and the general shape is seen to be a distortion of Fig. 4(a). This phase plot is insensitive to the value of M, and in fact a value of M as low as three is sufficient to capture the general features.

Using these three ψ 's, a model set of dynamical equations of the form of Eq. (10) was derived. These equations were then solved to obtain the time variation of the model ψ 's, and these were used to construct a new phase-space plot. This is shown in Fig. 4(c) and is seen to be in good agreement with the phase space portrait using the ψ 's as shown in Fig. 4(b).

A similar set of phase-space portraits for the Rossler equations 11 with $\Delta t = 0.2$ is shown in Fig. 5. This set of equations has the form

$$dx/dt = -(y+z)$$

$$dy/dt = x + \alpha y$$

$$dz/dt = \beta + z(x-y)$$
(15)

with $\alpha = \beta = 1/5$ and $\gamma = 5.7$ so as to generate a strange attractor. As with the Lorenz attractor, three eigenfunctions suffice, and $x(t) = 1.33 \ \psi_1(t) - 0.035 \ \psi_2(t) + 1.567 \ \psi_3(t)$. Comparison of the phase

portraits in Fig. 5 shows that the model equations capture the essential features of the strange attractor.

The results for the Lorenz and Rossler equations have been obtained using the value of \mathbf{x}_n at only 1000 distinct points. The phase-space portraits for the model equations are shown for times longer than a thousand time intervals, illustrating the stability of the equations.

However, the coefficients in the model equations and hence the solution of these equations depend sensitively on the order of the correlation matrix \underline{M} . For small values of \underline{M} , though the value of $\underline{T}(t)$ (that is x) generated using Eq. (7) with d = 3 is in good agreement (over the time where x(t) is given) with the original data, the associated model equations do not reconstruct the strange attractor. Usually after a short interval of time the solutions tend to become infinite or attract to a fixed point or limit cycle. There appears to be a optimal choice of M for the reconstruction of the attractor. This value has been found empirically to be associated with (a) the maximum difference between λ_3 and the higher eigenvalues and (b) that the elements C(n) used in the correlation matrices span the region where the major variation of C occurs. For the results presented in the case of the Lorenz equation, a value of 9 has been found to be appropriate, while for the Rossler equation, because of the longer correlation time, it was found necessary to make M = 16.

This whole procedure has been carried out for the Lorenz equations for a range of r values between 25 and 90, and in particular the coefficients a, b and c appearing in Eq. (11) were evaluated as a function of r. The variation with r of the coefficients of the largest nine terms is shown in Fig. 6(a) from which it is seen that this variation is reasonably smooth. From the symmetry of the Lorenz equations, the terms involving even powers of ψ (a_{mo} and b_{mql}) are negligibly small. Using the least squares method, the coefficients are readily fitted to simple polynomials in r. Such a cubic fit is shown in Fig. 6(b).

One now has a set of dynamical equations of the form given by Eqs. (10) and (11) where the coefficients a, b and c are known in the form of simple polynomials in the parameter r. It is on this set of equations that one can base an interpolation or extrapolation procedure. By taking r values other than the ones measured, and solving the model equations, the behavior of the system can be predicted. This can be in the form of the relevant phase plot or by using Eq. (7) to form x(t).

The phase portrait for r=57 obtained directly from the values of the ψ 's is shown in Fig. 7(a), while the form predicted using the above procedure is shown in Fig. 7(b). The agreement is quite good. Extrapolation outside the range of measured values should be applied with caution, however.

References

- 1. D. S. Broomhead and G. P. King, Physica D 20, 217 (1986).
- 2. R. Vautard and M. Ghil, Physica D 35, 395 (1989).
- 3. J. L. Lumley, Stochastic Tools in Turbulence, Academic Press: New York (1970).
- 4. N. Aubry, P. Holmes, J. L. Lumley and E. Stone, J. Fluid Mechanics 192, 115 (1988).
- 5. G. Rowlands, Non-Linear Phenomena in Science and Engineering, Ellis Horwood: London (1990).
- 6. E. N. Lorenz, J. Atmos. Sci. 20, 130 (1963).
- 7. J. C. Sprott and G. Rowlands, *Chaos Data Analyzer*, Physics Academic Software, Box 8202, North Carolina State University, Raleigh, NC 27695 (in press).
- 8. L. A. Smith, Phys. Lett. A <u>133</u>, 283 (1988).
- 9. A. M. Lyapunov, Ann. Math. Studies <u>17</u> (Princeton U.P., Princeton, NJ, 1949).
- 10. M. Hénon, Commun. Math. Phys. <u>50</u>, 69 (1976).
- 11. O. E. Rossler, Phys. Lett. A 57, 397 (1976).

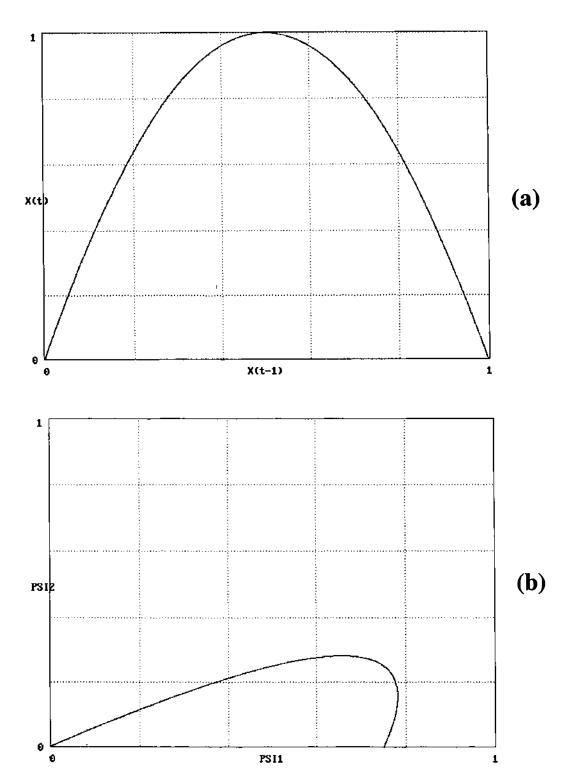
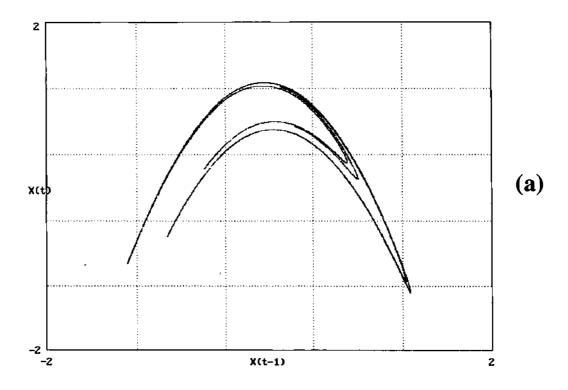


FIG. 1



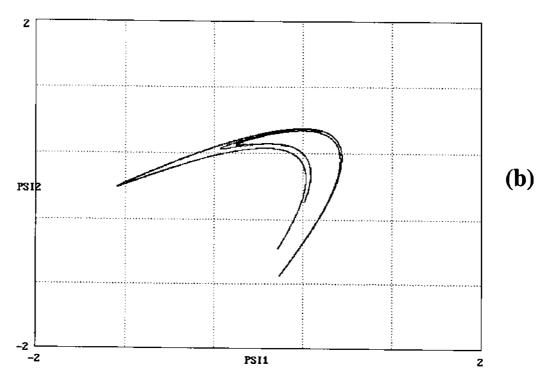


FIG. 2

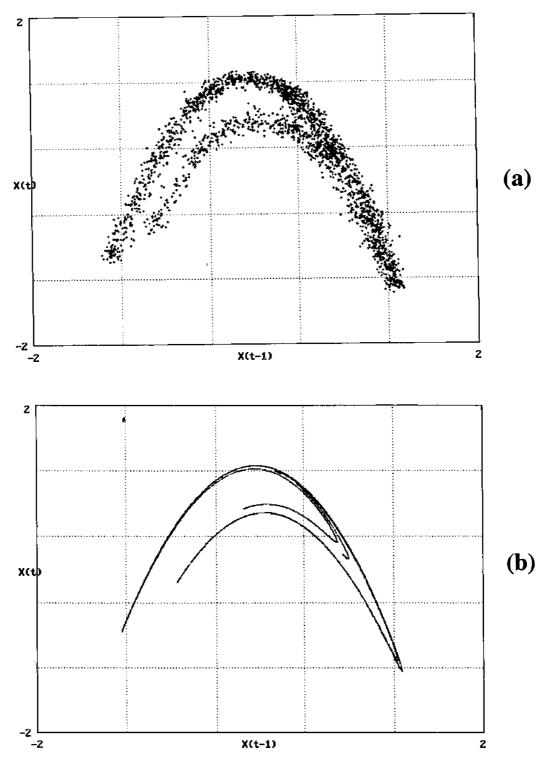
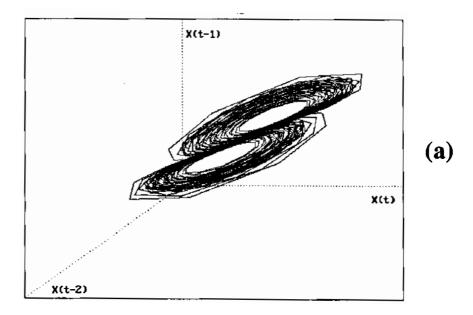
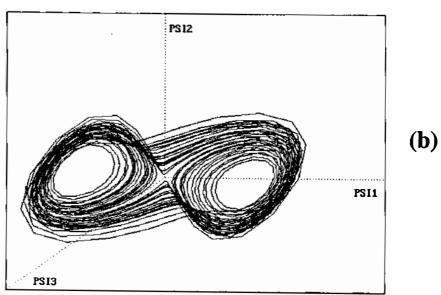


FIG. 3





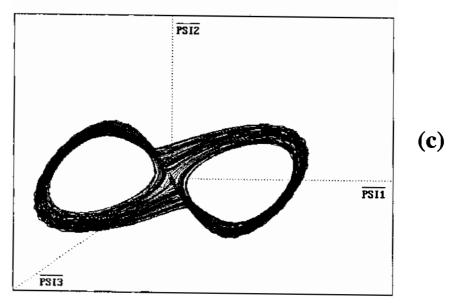
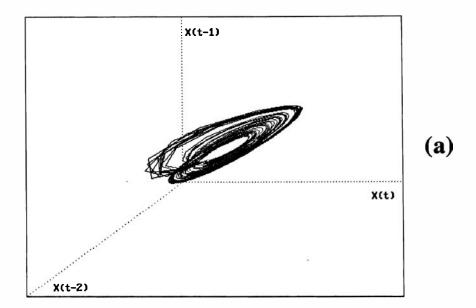
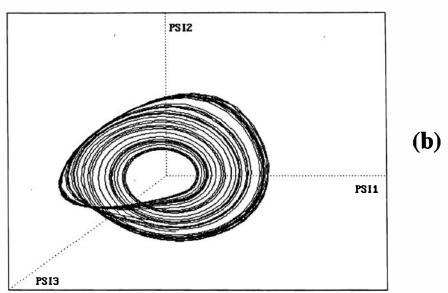


FIG. 4





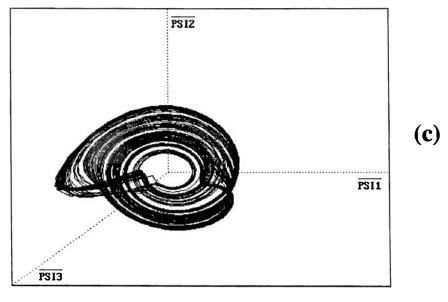
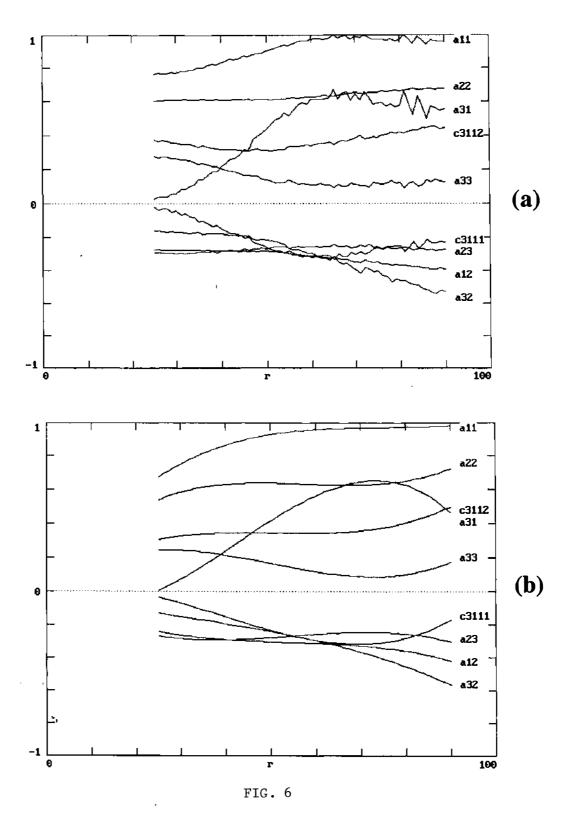
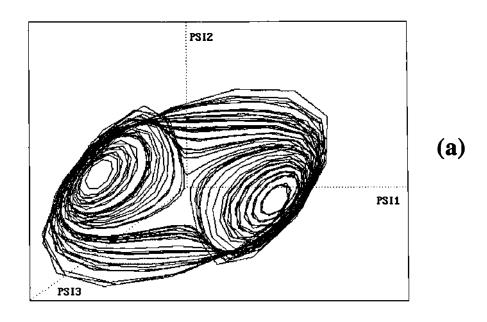


FIG. 5





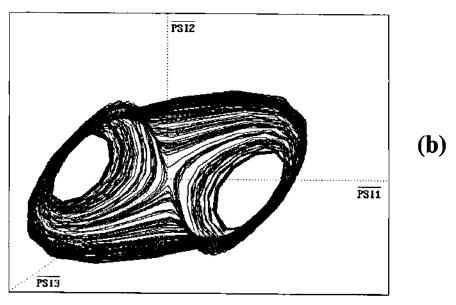


FIG. 7