

Local exponential divergence plot and optimal embedding of a chaotic time series

Jianbo Gao and Zhemin Zheng

Laboratory for Nonlinear Mechanics of Continuous Media, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100080, China

Received 2 October 1992; revised manuscript received 26 July 1993; accepted for publication 4 August 1993
Communicated by A.R. Bishop

We propose here a local exponential divergence plot which is capable of providing a new means of characterizing chaotic time series. The suggested plot defines a time dependent exponent λ and a “plus” exponent λ_+ which serves as a criterion for estimating simultaneously the minimal acceptable embedding dimension, the proper delay time and the largest Lyapunov exponent.

In recent years much progress has been made in understanding and characterizing chaotic dynamical systems. Much of the progress has been brought about by the important discovery that an appropriate state space can be reconstructed from a scalar time series [1]. Calculation of the correlation dimension and of the K_2 entropy by the Grassberger–Procaccia algorithm [2], and estimation of the Lyapunov exponents [3–6] have become standard procedures for analyzing chaotic signals. However, one may not gain much understanding by routine calculations under certain circumstances, since there are difficulties in interpreting correlation integral results [7], which are intimately related to the problem of how to distinguish chaos from stochastic processes. Therefore, it would be very helpful if a geometric method could be devised to view the dynamics, especially the local exponential divergence dominated behavior of a time series, so that a glance at this divergence plot would provide some insight into the dynamic system.

We report here a kind of local exponential divergence plot which enables one to view the dynamics on a chaotic attractor. The simple plot provides a criterion for the selection of the minimal acceptable embedding dimension and an optimal delay time. When the unstable motion on the chaotic attractor only is extracted, a proper estimation of the largest positive Lyapunov exponent can also be obtained.

Assume we have a time series x_1, x_2, \dots , with sam-

pling time δt and construct vectors of the form $X_i = (x_i, x_{i+L}, \dots, x_{i+(m-1)L})$, with m the embedding dimension and L delay time. Hence a dynamics $F: X_i \rightarrow X_{i+1}$ is defined, which is assumed to be representative of the original system. The distance between X_i and X_j , denoted by $\text{dis}(X_i, X_j)$, is mapped to $\text{dis}(X_{i+k}, X_{j+k})$ after k iterations of F . The local exponential divergence plot is defined by plotting $\ln[\text{dis}(X_{i+k}, X_{j+k})/\text{dis}(X_i, X_j)]$ versus $\ln[\text{dis}(X_i, X_j)]$ when $\text{dis}(X_i, X_j)$ is smaller than a prescribed small distance r^* . If we assume that most of these sufficiently small distances $\text{dis}(X_i, X_j)$ can be regarded as distances between orbits, then if the motion is truly chaotic, points with $\text{dis}(X_{i+k}, X_{j+k}) > \text{dis}(X_i, X_j)$ will dominate and lie above the zero level line in the plot.

Figure 1 shows divergence plots with different m and L for the Rössler attractor ($a=0.15$, $b=0.20$, $c=10.0$, $\delta t = \pi/25$, the dynamics is reconstructed from the x component of the flow). We will show below that the difference between these plots gives a hint to optimal embedding, and $m=3$, $L=8$ correspond to optimal parameter values. The zero level line is added to fig. 1b for a clear view of the divergence dominated behavior.

A problem of significant practical importance is to determine the minimum acceptable embedding dimension m_e . A basic idea is that in the passage from dimension m to $m+1$ one can differentiate between

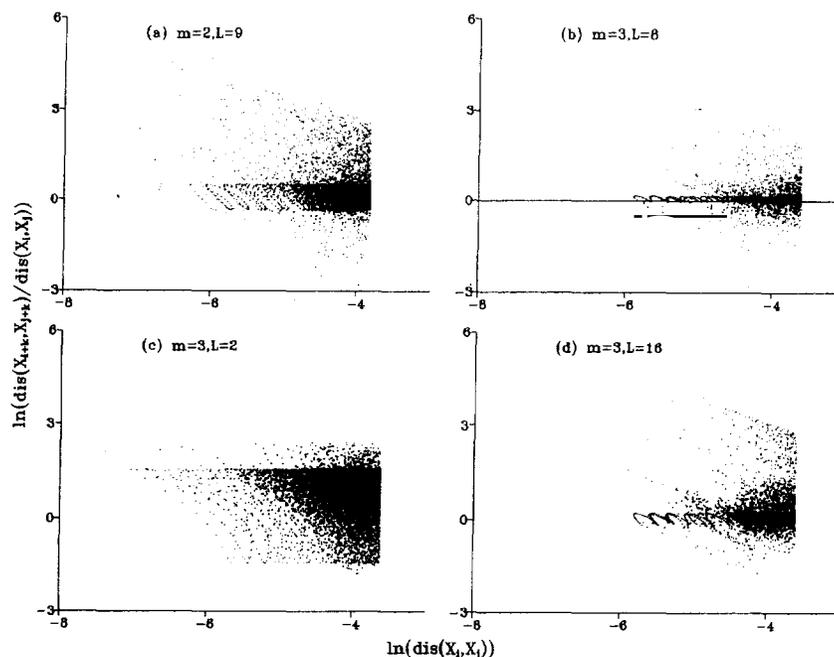


Fig. 1. Local exponential divergence plots for the Rössler attractor, $k=9$.

points on the orbit $X(n)$ that are “true” neighbors and points which are “false” neighbors – points which appear to be neighbors because the orbit is being viewed in too small an embedding space. Based on this basic idea, several methods are now available [8–11], which differ in implementations either by way of graphic display or by defining some appropriate statistical quantity. When the embedding dimension is increased from $m_e - 1$ to m_e , the structure of the graphic representation and/or the value of the statistical quantity will undergo a radical change while further increasing m causes little change. Our divergence plot implements this basic idea dynamically. When the embedding space is too small, the ill-defined dynamics F and the false neighbors will generate many points of $\ln[\text{dis}(X_{i+k}, X_{j+k})/\text{dis}(X_i, X_j)]$ with excessively large positive values in the divergence plot. This is clearly shown by the difference between figs. 1a and 1b.

Figure 1 also points out how to select the proper delay time L . Dynamically, when L is either too small or too large the dynamics F will not be very well defined, in the sense that excessively large values of $\ln[\text{dis}(X_{i+k}, X_{j+k})/\text{dis}(X_i, X_j)]$ frequently appear.

The key to selecting a proper delay time L is that the orbital motion should be as uniform as possible, and distortion be small. This can be achieved by requiring that the number of points $\ln[\text{dis}(X_{i+k}, X_{j+k})/\text{dis}(X_i, X_j)]$ with excessively large positive values in the divergence plot be as small as possible and the structure of the divergence plot be as compact as possible. This is the reason that the structure of fig. 1b is preferred to fig. 1c or 1d. Actually, fig. 1b is representative of that constructed from the original system and does not change much when m is further increased.

For a quantitative description, we define the time dependent divergence exponent λ by

$$\lambda(k, m, L) = \langle \ln[\text{dis}(X_{i+k}, X_{j+k})/\text{dis}(X_i, X_j)] \rangle, \quad (1)$$

with $\text{dis}(X_i, X_j) \leq r^*$. The evolution time corresponding to k is $k\delta t$ and the angular brackets denote the ensemble average of all possible pairs (X_i, X_j) . Points with positive values of $\ln[\text{dis}(X_{i+k}, X_{j+k})/\text{dis}(X_i, X_j)]$ are especially important since we are more concerned with points $\ln[\text{dis}(X_{i+k}, X_{j+k})/$

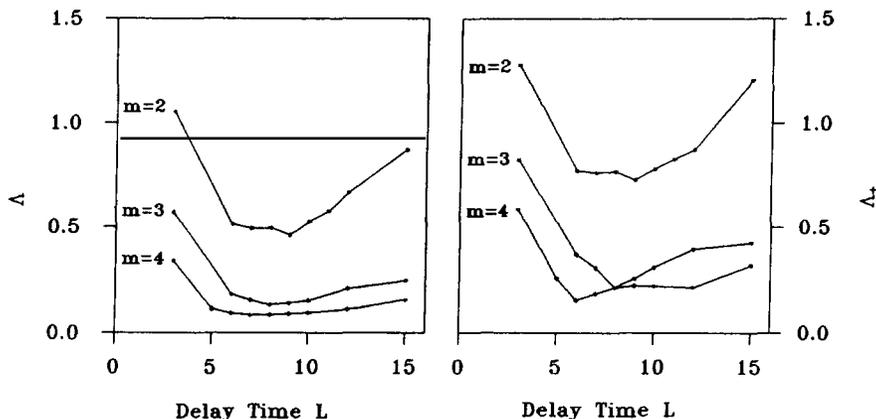


Fig. 2. $\Lambda(L)$ and $\Lambda_+(L)$ curves with different m for the Rössler attractor, $k=9$.

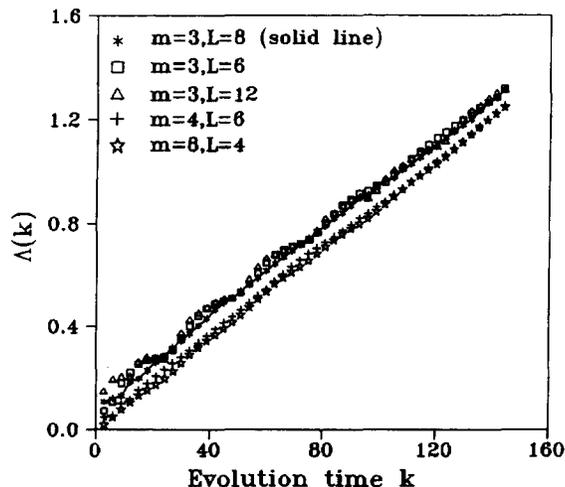


Fig. 3. $\Lambda(k)$ curves with different m and L for the Rössler attractor.

$\text{dis}(X_i, X_j)$ with excessively large positive values. Thus we also define the “plus” exponent Λ_+ by

$$\Lambda_+(k, m, L) = \langle \ln[\text{dis}(X_{i+k}, X_{j+k})/\text{dis}(X_i, X_j)]_+ \rangle, \quad (2)$$

where $+$ simply denotes that points with positive values of $\ln[\text{dis}(X_{i+k}, X_{j+k})/\text{dis}(X_i, X_j)]$ only are averaged. In the following discussion, when only one variable is considered, we will simply write $\Lambda(k)$, $\Lambda(L)$ and $\Lambda_+(L)$ for convenience.

Now the problem of properly reconstructing the

state space can be stated as follows. It is required that F be a continuous mapping preserving neighborhood relations. The minimal acceptable embedding dimension m_e is determined by requiring that the structure of the divergence plot no longer changes radically, or equivalently, that Λ and Λ_+ do not decrease significantly by further increasing m . When m is thus selected, for a series of L , the minima of $\Lambda_+(L)$ and $\Lambda(L)$ determine an optimal delay time.

The physical significance of the quantity Λ is obvious. When the evolution time $k\delta t$ is very small, $\Lambda/k\delta t$ is the mean value of the so-called largest local Lyapunov exponent [12,13]. After several iterations, the separation vector between X_i and X_j will align with the eigendirection for the largest positive Lyapunov exponent, and $\Lambda/k\delta t$ is equivalent to the standard estimation of this exponent. Hence when the proper reconstruction of the state space has been achieved, we would require that the $\Lambda(k)$ curve for $k\delta t$ not very small be a straight line which passes through the origin when extrapolated. Otherwise, different values of the largest Lyapunov exponent will be obtained with different selections of k . This kind of situation has been observed by Zeng et al. [6], however. Therefore, in practice at least we would require that $\Lambda(k)$ depend linearly on k for $k \in (k_{\min}, k_{\max})$, and the interval $(k_{\max} - k_{\min})\delta t$ be not too small. The largest positive Lyapunov exponent can be objectively estimated by $[\Lambda(k_1) - \Lambda(k_2)] / (k_1 - k_2)\delta t$, with $k_1, k_2 \in (k_{\min}, k_{\max})$.

Let us continue to discuss the Rössler system. Fig-

Table 1

Simultaneous minimal embedding dimension m and optical delay time L , and the largest Lyapunov exponent λ . The total number of points N is given in the table for different model systems.

System	Optimal embedding parameters		λ	
	others	present results ($N=2000$)	others	present results
Henon map [15] ($a=1.4, b=0.3$)	$m=2$ ($N=5000$) [9]	$m=2, L=1$	0.418 [5]	0.421 ± 0.003 ($N=2000$)
Rössler [16] ($\delta t = \pi/25, a=0.15,$ $b=0.20, c=10.0$)	$m=3, L=7$ ($N=10000$) [10]	$m=3, L=8$	0.09 [5]	0.067 ± 0.006 ($N=2000$)
Lorenz [17] ($\delta t=0.03, \sigma=10,$ $b=\frac{8}{3}, r=45.92$)		$m=3, L=3$	1.497 [5]	1.48 ± 0.03 ($N=3000$)
Mackey–Glass [18] ($\delta t=1.5, a=0.2,$ $b=0.1, c=10, \Gamma=30$)	$m=4, L\delta t=0.3\Gamma$ ($N=9000$) [10]	$m=4, L\delta t=0.3\Gamma$	0.0071 [4]	0.0072 ± 0.0006 ($N=4000$)

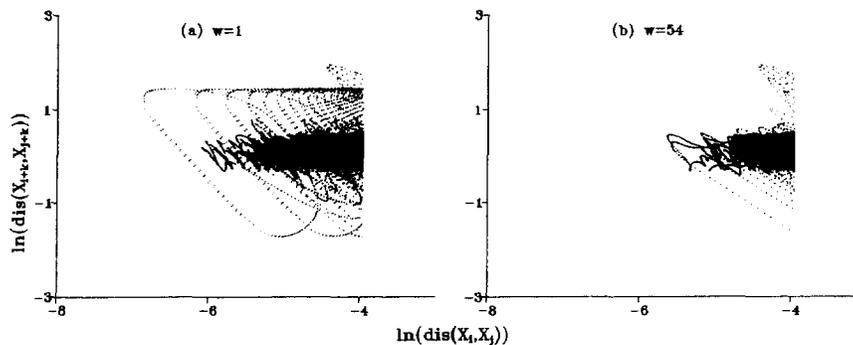


Fig. 4. Original ($w=1$) and modified ($w=54$) local exponential divergence plots for the Lorenz system constructed from the x component of the flow, $m=3, L=27, k=30$.

ure 2 shows the $A(L)$ and $A_+(L)$ curves, and we see that the combination of $m=3$ and $L=8$ is an optimal choice. Note that improper embedding (underestimated m or improper L) always results in overestimated positive exponent. Figure 3 gives the $A(k)$ curves for different m and L . We see that when $m=3$, the $A(k)$ curve for $L=8$ shows a clear linear dependence, while curves for the smaller value of $L=6$ or the larger value of $L=12$ (approximately the optimal value suggested by ref. [14]) are less satisfactory. Figure 3 also shows two curves for $m=4, L=6$ and $m=8, L=4$, with improved linearity. Note that the degree of the linearity is “saturated” when m is increased to 4, while the minimum acceptable value of m is 3. The estimated value of the largest Lyapunov

exponent from the slope of the linear $A(k)$ curve of $m=3$ and $L=8$ is 0.067. These results, together with results for other known model chaotic systems, are summarized in table 1.

We note that, with regard to the proper reconstruction of the state space, the results of refs. [9,10] can be easily obtained by our method with a very small data set. Our method has the additional advantage that the approach is simpler, more natural and easier to understand and implement, and capable of providing more information.

Finally we examine the assumption that most sufficiently small distances $\text{dis}(X_i, X_j)$ can be regarded as distances between orbits. A problem related to this is that whether and how the divergence plot for a

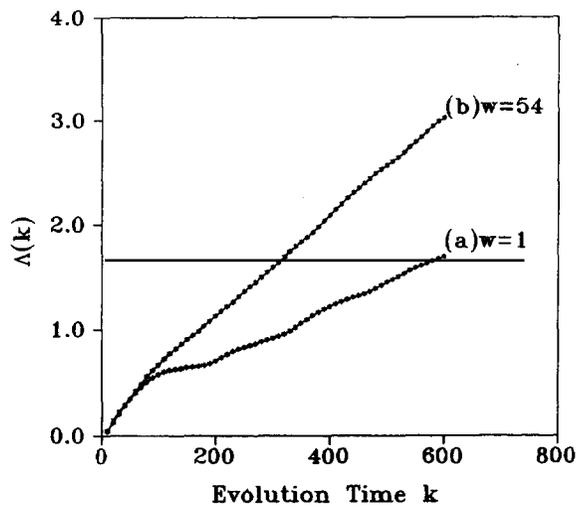


Fig. 5. $A(k)$ curves corresponding to fig. 4.

continuous system changes with the sampling time δt . Let us discuss the Lorenz system with a much smaller δt of 0.003 than the one taken in table 1. Figure 4a shows the divergence plot for $m=3$, $L=27$. We see that the plot consists of two parts, a dotted curve-like part and a heavy black part. Do they have the same origin? Figure 5a shows the $A(k)$ curve, which is by no means linear. Hence we should conclude that each part has its distinct origin, and that the above assumption does not hold in this case.

The answer is rather simple, however. There are certain small distances like $\text{dis}(X_i, X_{i+w})$, with w very small, and their k th iterations can also be very small. These points obviously correspond to the orbital motion, and cannot be regarded as small distances between orbits. Points in the divergence plot corresponding to these points merely reflect changes of the phase velocity along the orbit, and such points will increase if δt is decreased. Let us call this part of the motion tangential motion. The tangential motion contributes a dimension nearly one, corresponding to a Lyapunov exponent equal to zero [19]. Hence this motion should be excluded when calculating the fractal dimension and estimating the largest Lyapunov exponent. A possible way of doing this is to add an additional condition to expression (1), namely

$$j-i \geq w. \quad (3)$$

Expression (1) corresponds to $w=1$. Theiler [20] has proposed a similar improvement to the calculation of the correlation dimension, and suggests w to be selected as the auto-correlation time. In our case, theoretically, it does not matter if a very big w is selected. However, we have tested numerically that putting w equal to the embedding window $(m-1)L$ already does the job.

Figure 4b is a modified divergence plot. It can be seen easily that the curve-like part of fig. 4a, especially the part corresponding to very small distances where statistics is poor, is largely suppressed. Figure 5b is the modified $A(k)$ curve. It is now nearly linear. Actually the slope of the modified $A(k)$ curve gives the correct Lyapunov exponent. Thus we know that the heavy black part in the divergence plot originates from the unstable motion, and a linear $A(k)$ curve is a property of the unstable motion and characteristic of chaotic motions.

Let us summarize. Two kinds of motion, tangential motion and unstable motion, can be discerned from the structure of the divergence plot. The former is irrelevant in the calculation of the fractal dimension and estimation of the largest positive Lyapunov exponent. Hence, the former should be removed as much as possible in these calculations. Since the damaging effect of the tangential motion is enhanced when δt is decreased, too small a sampling time is not recommended. Also δt is suggested not to exceed the optimal delay time.

A note on the Lorenz system needs be made. This system is very complicated in that, even when δt is not small (for example, $\delta t=0.009$), the tangential motion still occupies a large fraction in the divergence plot, and the $A(k)$ curve is not linear by expression (1), if condition (3) is not imposed. This was probably the reason that Wolf et al. [5] used a very large δt to estimate the largest Lyapunov exponent for this attractor.

Another important note should also be made. The optimal values of m and $L\delta t$ do not change with the sampling time δt . Though condition (3) is suggested when calculating the $A(k)$ curve to estimate the Lyapunov exponent, it is not needed when reconstructing a state space, since the tangential motion is an integral part of the motion on the attractor.

This work is partly supported by the National Ba-

sic Research Project "Nonlinear Science".

References

- [1] N.H. Packard, J.P. Crutchfield, J.D. Farmer and R.S. Shaw, *Phys. Rev. Lett.* 45 (1980) 712; F. Takens, in: *Lecture notes in mathematics*, Vol. 989. *Dynamical systems and turbulence*, eds. D.A. Rand and L.S. Young (Springer, Berlin, 1981) p. 366.
- [2] P. Grassberger and I. Procaccia, *Physica D* 9 (1983) 189; *Phys. Rev. A* 28 (1983) 2591.
- [3] J.-P. Eckmann and D. Ruelle, *Rev. Mod. Phys.* 57 (1985) 617.
- [4] M. Sano and Y. Sawada, *Phys. Rev. Lett.* 55 (1985) 1082.
- [5] A. Wolf, J.B. Swift, H.L. Swinney and J.A. Vastano, *Physica D* 16 (1985) 285.
- [6] X. Zeng, R. Eykholt and R.A. Pieke, *Phys. Rev. Lett.* 66 (1991) 3229.
- [7] A. Provenzale, L.A. Smith, R. Vio and G. Murante, *Physica D* 58 (1992) 31.
- [8] A. Cenys and K. Pyragus, *Phys. Lett. A* 129 (1988) 227.
- [9] Z. Aleksic, *Physica D* 52 (1991) 362.
- [10] W. Liebert, K. Pawelzik and H.G. Schuster, *Europhys. Lett.* 14 (1991) 521.
- [11] M.B. Kennel, R. Brown and H.D.I. Abarbanel, *Phys. Rev. A* 45 (1992) 3403.
- [12] J.M. Nese, *Physica D* 35 (1989) 237.
- [13] H.D.I. Abarbanel, R. Brown and M.B. Kennel, *J. Nonlinear Sci.* 1 (1991) 175; 2 (1992) 343.
- [14] A.M. Fraser and H.L. Swinney, *Phys. Rev. A* 33 (1986) 1134.
- [15] M. Hénon, *Commun. Math. Phys.* 50 (1979) 69.
- [16] O.E. Rössler, *Phys. Lett. A* 71 (1979) 155.
- [17] E.N. Lorenz, *J. Atmos. Sci.* 20 (1963) 130.
- [18] M.C. Mackey and L. Glass, *Science* 197 (1977) 287.
- [19] H. Haken, *Phys. Lett. A* 94 (1983) 71.
- [20] J. Theiler, *Phys. Rev. A* 34 (1986) 2427; *J. Opt. Soc. Am. A* 7 (1990) 1055.