AUTOREGRESSIVE METHODS FOR CHAOS ON BINARY SEQUENCES FOR THE LORENZ ATTRACTOR ☆

P. SINGH and D.D. JOSEPH

Department of Aerospace Engineering and Mechanics, University of Minnesota, 107 Akerman Hall, Minneapolis, MN 55455, USA

A binary sequence is defined for the Lorenz attractor. This binary sequence contains some information about the original system. To extract this information we have used autoregressive methods from the theory of signal processing. The binary sequences and the associated methods could also be used to estimate the system characteristics when one does not have access to all the variables involved in the underlying process; this is usually the case in an experimental study. We introduce an autocorrelation function for binary sequences, a one-step predictor and associated power spectra and a macroscopic approximation of the largest Lyapunov exponent.

1. Binary sequence for the Lorenz equations

The Lorenz equations

$$\frac{dx}{dt} = \sigma(y-x) , \quad \frac{dy}{dt} = rx - y - xz , \quad \frac{dz}{dt} = xy - bz ,$$

$$(1.1)$$

with $\sigma = 10.0$, $b = \frac{8}{3}$, $r = 28.0$, were integrated numerically using the NAG library, subroutine D02BBF with different tolerance levels in the range 10^{-4} to 10^{-10}.

We looked at the projection of trajectories generated by (1.1) onto the xz-plane. The z-axis divides the xz-plane into two half-planes, called left and right half-planes. The projected trajectory makes closed loops in the left or right plane, or switches between half-planes. We define a binary sequence by assigning the number one to a loop in the right half-plane and the number minus one to a loop in the left half-plane. This binary sequence still contains some information about the original system with some information being lost in the process of defining the

binary sequence. We want to extract the information contained in the binary sequence. Symbolic sequences have been used before to characterize dynamical systems [1–3]. We do not know any previous works on autocorrelations, predictors and their associated power spectrum for binary sequences in general, or applied to the Lorenz attractor. Our formulas (4.9) and (4.10) for the macroscopic Lyapunov exponent are new, though some referees have suggested they are related to metric entropy. Shimada [2] has used binary sequences for the Lorenz system to discuss the system in the frame of the statistical mechanics of the Ising model. He does not work directly on the measures of chaos on binary sequences discussed in this paper. There are many problems of chaos in which the observable variables can be precisely if not broadly defined by binary sequences. This might be the case in problems defined by a double well potential or in experiments [4] in which the underlying dynamics are not understood. For such problems, and others, the approach taken in this paper should be useful.

2. Autocorrelation

Now we are going to use the methods of estimation theory to characterize chaos on a binary se-

quence. A convenient reference for these methods is, for example, ref. [5]. An estimate of the autocorrelation function on an ergodic binary sequence can be obtained as follows,

$$r(n) = \frac{1}{N} \sum_{k=1}^{N} u(k+n)u(k) ,$$

$$n = 1, 2, ..., \quad N \gg n , \tag{2.1}$$

where

$$\{u(i)\}, \quad i = 1, 2, ..., N , \tag{2.2}$$

is a binary sequence, $u(i)$ has values 1 and -1.

The value $r(1)$ represents the correlation between immediate neighbors (1,2), (2,3), (3,4), etc. The value $r(2)$ gives the correlation between separated pairs (1,3), (2,4), etc. A chaotic response is one for which $r(1) \neq 0$ and $r(n) \to 0$ for large n, predictability only in the short run.

3. One-step prediction method

The problem of prediction from chaotic sequences has been discussed in a general context by Farmer and Sidorowich [6]. They show that in many situations nonlinear predictors are much better than linear ones. They do not consider binary sequences explicitly. In the appendix to this paper we show that if the autocorrelation for the binary sequence decays rapidly and is nonnegative, then the predictive power of a polynomial predictor is not significantly better than the linear one. Our binary sequence can be treated like a sequence of real numbers. Hence the one-step linear prediction method can be used to measure the deterministic part. This gives a second measure of chaos on a binary sequence. We assume that the sequence is ergodic, so we may use time averages as an approximation to the corresponding ensemble averages. For our sequence we found

$$E\{u(n)\} = \frac{1}{N} \sum_{k=1}^{N} u(k) \approx 0 . \tag{3.3}$$

This tells us that left and right (± 1) have the same probability of occurring. Also

$$E\{u^2(n)\} = \frac{1}{N} \sum_{k=1}^{N} u^2(k) = 1 , \tag{3.4}$$

where E stands for the expected value or ensemble average.

Suppose that M past values $u(n-1)$, $u(n-2)$, ..., $u(n-M)$ are known; the problem is to predict $u(n)$. We note that the predicted value will be a real number between -1 and 1. Let us denote it by $\hat{u}(n)$. We assume that our predictor is linear, that is, the predicted value is given by

$$\hat{u}(n) = \sum_{k=1}^{M} a(k)u(n-k) , \tag{3.5}$$

where $a(k)$, $k = 1, 2, ..., M$, are unknown constants and M is the order of the predictor. The prediction error is defined as

$$f_M(n) = u(n) - \hat{u}(n) . \tag{3.6}$$

Let P_M denote the expected value of the mean squared prediction error,

$$P_M(a) = E\{f_M^2(n)\} , \tag{3.7}$$

and $\hat{P}_M$ be its minimum value,

$$\hat{P}_M = \min(E\{f_M^2(n)\}) . \tag{3.8}$$

The right-hand side of (3.7) is quadratic in the $a(k)$'s; for it to be minimum it is enough that

$$\frac{\partial}{\partial a(k)} E\{f_M^2(n)\} = 0 , \quad k = 1, ..., M . \tag{3.9}$$

Using time averages to approximate ensemble averages, the above equation reduces to

$$\frac{\partial}{\partial a(k)} \left(\frac{1}{N-M} \sum_{I=M}^{N} f_M^2(I) \right) = 0 ,$$

$$k = 1, ..., M . \tag{3.10}$$

This gives M equations for M unknowns, $a(1)$, ..., $a(M)$. Once the $a(k)$'s are known, we can compute the minimum $\hat{P}_M$ of (3.8).

A third measure of chaos, the autoregressive power spectral density, is given by

$$S(\theta) = \hat{P}_M \left| 1 + \sum_{k=1}^{M} a(k) \exp(-ik\theta) \right|^{-2} , \tag{3.11}$$

where θ is the normalized frequency variable.

There are several methods available for determining the order of the predictor. For the binary sequences we have considered in this paper, the slope
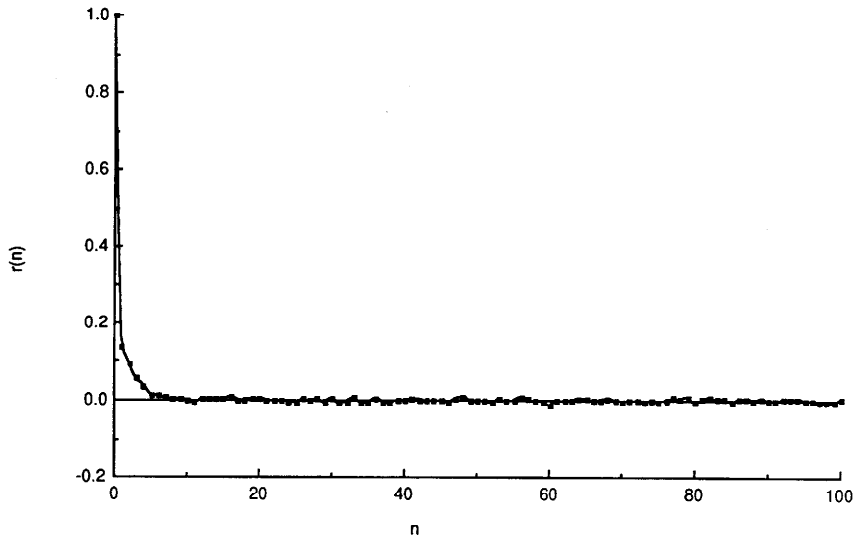
Fig. 1. The autocorrelation sequence for the Lorenz attractor. $N = 76\,000$.

of $\hat{P}_M$ versus M curve becomes approximately zero for some value of $M = \hat{M}$ and remains approximately zero for M larger than $\hat{M}$. This tells us that there is no increase in the performance of the predictor for M larger than $\hat{M}$. So the predictor order is taken to be $\hat{M}$. These discrete measures were computed for the Lorenz attractor with the $a(k)$'s determined by the Levenson–Durbin recursion (see ref. [5]), which is a fast algorithm for solving the system (3.9). It is based on exploiting the Toeplitz nature of the matrix.
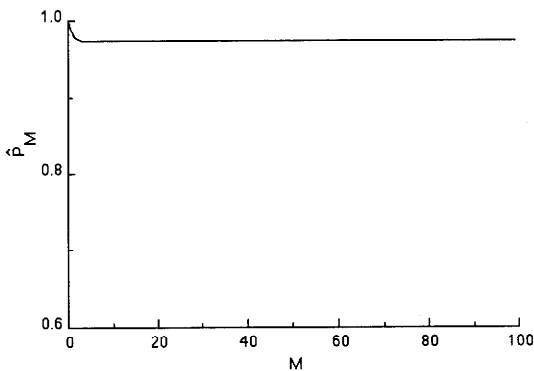
Our binary sequence had 76 000 points. The discrete measures of chaos were computed. These are shown in figs. 1, 2 and 3. The predictor order was found to be five. The tolerance level in the numerical scheme had absolutely no effect on the nature of the autocorrelation sequence, even though the sequences generated were quite different for different tolerance levels. For large n, $r(n)$ approached zero uniformly with the increase in the length of sequence, N.

The decay in the autocorrelation values is very rapid. For large n, autocorrelation values decrease monotonically with the length of the sequence. The minimum mean squared prediction error decreases

Fig. 2. The minimum mean square prediction error as a function of the predictor order for the Lorenz attractor. For $M = 5$, the slope of this curve becomes approximately zero and it remains close to zero for $M > 5$. So the order of the predictor is taken to be 5. The finite slope for large M is due to the finite length of the sequence.
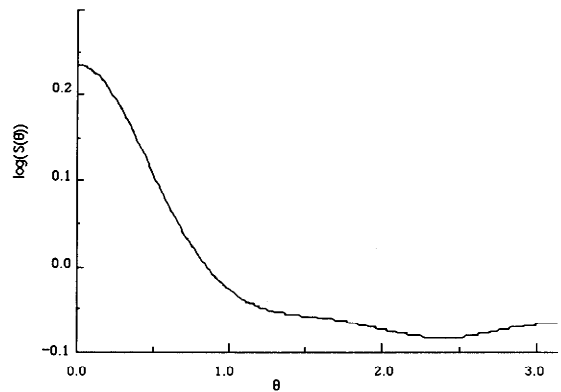

Fig. 3. The power spectral density for the Lorenz attractor.

rapidly with M, for M smaller than $\hat{M}$, and remains almost constant for M larger than $\hat{M}$. $\hat{P}_M$ equal to zero corresponds to a totally predictable sequence and $\hat{P}_M$ equals one when the sequence is completely unpredictable. For example, the binary sequence generated by tossing a coin is a completely unpredictable sequence. Totally predictable sequences are deterministic, so $\hat{P}_M$ gives some measure of randomness in the sequence. For the Lorenz attractor $\hat{P}_{\hat{M}}$ is very close to one, which tells us that the sequence is predominantly random with a very small deterministic part.

4. Lyapunov exponent

In this section the binary sequence is used to approximate the value of the Lyapunov exponent by incorporating some of the information about the original Lorenz process projected onto the xz-plane. The projection of the Lorenz attractor onto the xz-plane is shown in fig. 4. The trajectories projected onto the xz-plane remain in the area shown in fig. 4. Let AD be the line segment passing through the two stationary points for the Lorenz attractor and AB and CD be as shown in fig. 4. Then it is clear that all trajectories in the right half-plane intersect the line segment CD and all trajectories in the left half-plane intersect the line segment AB. Consider two trajectories, selected randomly, starting at the line segment CD. Since trajectories are selected randomly the distance between them is random, with mean value equal to one half of $|CD|$, where $|CD|$

Fig. 4. The projected trajectories of the Lorenz attractor remain inside the closed region shown. Two stationary points are marked by crosses.

is the length of the line segment CD. Similarly, the trajectories starting at the line segment AB are on average $|AB|/2$ apart. For the Lorenz attractor $|AB| = |CD|$. Two trajectories selected randomly stay in the same half-plane or both move over to the other half-plane or one remains in the same half-plane and the other moves over to the other half-plane. When trajectories stay in the same half-plane or both of them move over to the other half-plane the distance between them, on average, does not change. But if one of the trajectories moves over to the other half-plane the distance between them, on average, after one period is equal to $|AD| - (|AB| + |CD|)/2 = |AB| - |CD|$.

One can define the Lyapunov exponent by the expression

$$\lambda = \frac{1}{N} \sum_{i=1}^{N} \log_2 \frac{d(n)}{d_0(n-1)}, \qquad (4.1)$$

where $d_0(n-1)$ is the infinitesimal distance between the trajectories at the beginning and $d(n)$ is the infinitesimal distance between them after one time period. We modify the above expression for the Lyapunov exponent to obtain an expression for the macroscopic Lyapunov exponent,

$$\lambda_m = \frac{1}{N} \sum_{i=1}^{N} \log_2 \frac{\bar{d}(n)}{\bar{d}_0(n-1)}, \qquad (4.2)$$

where $\bar{d}_0(n)$ is the average distance between the trajectories starting at the line segments AB or CD and $\bar{d}(n)$ is the average distance between them after one time period.

Our aim is to estimate the Lyapunov exponent using the binary sequence obtained in section 1. For obtaining the Lyapunov exponent we need to select two statistically independent sequences from the given binary sequence, since for the given binary sequence $r(n) = 0$ for $n \geqslant \hat{M}$, where $\hat{M} = 5$ for the Lorenz attractor. If we consider two sub-sequences, obtained from the given sequence such that they are greater than or equal to $\hat{M}$ distance apart then these sub-sequences are statistically independent. Consider the product $u(i)u(i+k)$, $k \geqslant \hat{M}$. If this product is one then sub-sequences starting from index i and $i+k$, in the main sequence, are in the same half-plane; otherwise they are in different half-planes. To find the sum in eq. (4.2) we need statistically independent sequences which start in the same half-plane. If
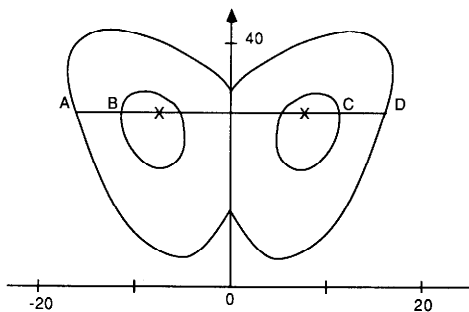
the product $u(i+1)u(i+k+1)$ is one then the sequences remain in the same half-plane and if the product is minus one then the sequences end up in different half-planes after one time period. Let

$$\alpha = \log_2 \frac{|AD| - |AB|}{|AB|},$$

then

$$\log_2[\bar{d}(n)/\bar{d}_0(n-1)] = \alpha,$$

if sequences after one time period are in different half-planes, and

$$\log_2[\bar{d}(n)/\bar{d}_0(n-1)] = 0,$$

if sequences after one time period are in the same half-plane. Note that the above function can also be written as

$$\log_2 \frac{\bar{d}(n)}{\bar{d}_0(n-1)} = \tfrac{1}{2}\alpha[1 - u(i+1)u(i+k+1)],$$

(4.3)

provided $u(i)u(i+k) = 1$. Let $S_1 = \{i: u(i)u(i+k) = 1\}$ and $S_2 = \{i: u(i)u(i+1) = -1\}$ with i numbered in an increasing order in each set and let N_1 and N_2 be the sizes of sets S_1 and S_2 respectively. Then, from (4.2),

$$\lambda_m = \frac{1}{N_1} \sum_{i=1}^{N_1} \log_2 \frac{\bar{d}(n)}{\bar{d}_0(n-1)}$$

$$= \frac{\alpha}{2N_1} \sum_{i \in S_1} [1 - u(i+1)u(i+k+1)].$$

(4.4)

If N is large we may translate the origin of summation without changing $r(k)$, equivalent to renumbering. Hence

$$r(k) = \frac{1}{N} \sum_{i=1}^{N} u(i+1)u(i+k+1)$$

$$= \frac{1}{N}\left(\sum_{i \in S_1} u(i+1)u(i+k+1)\right.$$

$$\left. + \sum_{i \in S_2} u(i+1)u(i+k+1)\right),$$

implies that

$$\sum_{i \in S_2} u(i+1)u(i+k+1)$$

$$= Nr(k) - \sum_{i \in S_1} u(i+1)u(i+k+1).$$

(4.5)

Since for $k > \hat{M}$, $r(k) = 0$, we have

$$0 = r(k) = \frac{1}{N} \sum_{i=1}^{N} u(i)u(i+k)$$

$$= \frac{1}{N}\left(\sum_{i \in S_1} u(i)u(i+k) + \sum_{i \in S_2} u(i)u(i+k)\right)$$

$$= \frac{1}{N}(N_1 - N_2),$$

hence $N_1 = N_2$. But $N = N_1 + N_2$, so we get

$$N = \tfrac{1}{2}N_1 = \tfrac{1}{2}N_2.$$

(4.6)

Now we relate the Lyapunov exponent to the autocorrelation function of the given sequence. Consider

$$\sum_{i=1}^{N} \sum_{m=1}^{N} u(i+m)u(i+m+1)u(i)u(i+1)$$

$$= \sum_{i=1}^{N}\left(u(i)u(i+1)\sum_{m=1}^{N} u(i+m)u(i+m+1)\right)$$

$$= \sum_{i=1}^{N} u(i)u(i+1)Nr(1) = N^2 r^2(1).$$

(4.7)

By changing the order of summation in (4.7) we get

$$N^2 r^2(1) =$$

$$\sum_{m=1}^{N}\left(\sum_{i=1}^{N} u(i+m)u(i+m+1)u(i)u(i+1)\right).$$

The breakup described in (4.3) reduces the above to

$$N^2 r^2(1) = \sum_{m=1}^{N}\left(\sum_{i \in S_1} u(i+1)u(i+m+1)\right.$$

$$\left. - \sum_{i \in S_2} u(i+1)u(i+m+1)\right).$$

Using (4.5) and the above equation we have

$N^2 r^2(1)$

$$= \sum_{m=1}^{N} \left(2 \sum_{i \in S_1} u(i+1)u(i+m+1) - Nr(m) \right) .$$

After rearranging the above equation, we have

$$\sum_{m=1}^{N} \left(2 \sum_{i \in S_1} u(i+1)u(i+m+1) \right)$$

$$= N^2 r^2(1) - \sum_{m=1}^{N} Nr(m) . \tag{4.8}$$

For our sequence $r(m) = 0$ for $m \geqslant \hat{M}$. Moreover, the sum inside the square brackets in the above equation is a constant for $m \geqslant \hat{M}$. Let

$$\frac{1}{N_1} \sum_{i \in S_1} u(i+1)u(i+m+1) = T(m) ,$$

where $T(m) = T = \text{const}$, for $m \geqslant \hat{M}$. By using the above two results in (4.8), we have

$$2\{(N-\hat{M})N_1 T + N_1[T(1) + T(2) + ... + T(\hat{M}-1)]\}$$

$$= N^2 r^2(1) - N[r(1) + r(2) + ... + r(\hat{M}-1)] .$$

Using the fact that $\hat{M}$, $T(1)$, $T(2)$, ..., $T(\hat{M}-1)$, $r(1), r(2), ..., r(\hat{M}-1)$ are finite if we take the limit $N \to \infty$ we have after using (4.6)

$$T = r^2(1) = \frac{1}{N_1} \sum_{i \in S_1} u(i+1)u(i+m+1) .$$

By putting this in (4.4) we get the desired result

$$\lambda_m = \tfrac{1}{2}\alpha[1 - r^2(1)] . \tag{4.9}$$

Even though we have assumed N to be infinite this result is approximately correct for large N.

Finally, we relate this value of the macroscopic Lyapunov exponent to that in the continuous case,

$$\lambda_{mc} = \frac{1}{t_N - t_0} \sum_{i=1}^{N} \log_2 \frac{d(t_n)}{d_0(t_{n-1})} .$$

If we assume the t_n to be equally T time units apart, then the above equation becomes

$$\lambda_{mc} = \frac{1}{TN} \sum_{i=1}^{N} \log_2 \frac{d(n)}{d_0(n-1)} = \frac{\lambda_m}{T} . \tag{4.10}$$

For the binary sequence T is equal to the mean time period. We obtained the following results for the Lorenz attractor: $r(1) = 0.13853$, $\alpha = 2.5$, $T = 0.7519$

s, $\lambda_m = 1.226$ bits/period, $\lambda_{mc} = 1.630$ bits/s. The largest Lyapunov exponent computed directly for the Lorenz attractor is $\lambda_c = 1.30$ bits/s (fig. VI.18 in ref. [7]).

Appendix. A note on the binary nonlinear polynomial predictor

For a binary sequence the most general form of the Mth order nonlinear polynomial predictor is

$$\hat{u}(n) = \sum_{(k_1 = 0, ..., k_M = 0)}^{1} a(k_1, k_2, ..., k_m)$$

$$\times u(n-1)^{k_1} u(n-2)^{k_2} ... u(n-M)^{k_3} , \tag{1}$$

where we have used the fact that $u(k)^n = 1$ if n is an even integer and $u(k)^n = u(k)$ if n is an odd integer. So the degree of the binary nonlinear predictor is less than or equal to its order.

The contribution of the nonlinear terms in prediction depends on the magnitude of the higher order correlations. For a binary sequence in which 1 and -1 are equally likely (i.e. $P(1) = P(-1) = 0.5$) the higher order correlations with odd number of terms are zero,

$$E[x(n)x(n-1)...x(n-m)]$$

$$= E[x(n)x(n-1)...x(n-m+1)|x(n-m) = 1]$$

$$\times P(x(n-m) = 1)$$

$$- E[x(n)x(n-1)...x(n-m+1)|x(n-m) = -1]$$

$$\times P(x(n-m) = -1)$$

$$= \tfrac{1}{2}\{E[x(n)x(n-1)...$$

$$\times x(n-m+1)|x(n-m) = 1] - E[x(n)x(n-1)...$$

$$\times x(n-m+1)|x(n-m) = -1]\}$$

$$= 0 ,$$

where m is an even integer and $P(\ )$ stands for the probability. In the last step, the symmetry of the problem is used to equate the two conditional expectations. If least squares are used to find the coefficients of the predictor the contribution of an even degree term in (1) towards prediction depends on the correlation of the above form. So we can drop all terms of even degree in (1). Only the terms of odd degree remain. For a binary sequence with a nonnegative autocorrelation sequence one can further show that the following bound holds on the higher order correlations with an even number of terms, m odd,

$$E[x(n)x(n-1)...x(n-m)]$$

$$\leqslant E[x(n)1x(n-m)]=r(m) \ .$$

It follows that the higher order correlations for binary sequences will decay rapidly whenever the autocorrelation decays rapidly. This is the case for the Lorenz attractor. In such a situation the largest nonlinear contribution will come from the fourth order correlation which is smaller than $r(3)$. But $r(3)$ is already small, so the improvement obtained by including the nonlinear terms will not be significant.

References

[1] J.P. Crutchfield and N.H. Packard, Physica D 7 (1983) 201.
[2] I. Shimada, Prog. Theor. Phys. 62 (1979) 1.
[3] C. Sparrow, The Lorenz equations: bifurcation, chaos, and strange attractor (Springer, Berlin, 1982).
[4] P. Singh, K. Chen and D.D. Joseph, Strange Couette flows, strange Taylor cells and a strange attractor in Taylor–Couette flow of oil and water (1988), unpublished.
[5] R.A. Roberts and C.T. Mullis, Digital signal processing (Addison-Wesley, Reading, 1987).
[6] J.D. Farmer and J.J. Sidorowich, Exploiting chaos to predict the future and reduce noise (1988), unpublished.
[7] P. Bergé, Y. Pomeau and C. Vidal, Order within chaos (Wiley, New York, 1984).