RAPID COMMUNICATIONS

PHYSICAL REVIEW A VOLUME 29, NUMBER 3 MARCH 1984

Kolmogorov entropy from a time series

Yves Termonia
Polymer Department, Weizmann Institute of Science, Rehovot 76100, Israel
(Received 1 September 1983)

A practical algorithm for estimating the Kolmogorov (K) entropy directly from a time series is presented. In all the cases studied, we verify the conjecture that K is the sum of the positive Lyapunov exponents.

The Kolmogorov, or K entropy,[1] is an important measure for the study of nonlinear dynamical systems which are characterized by changes from regular to irregular behavior upon variation of parameters. Examples of such changes are the transition to irregular internal dynamics in polyatomic molecules at high levels of excitation[2] and the appearance of chaos or turbulence in highly nonlinear dissipative systems.[3-7] Qualitatively, the K entropy permits to distinguish between regular and irregular systems: K is zero in regular systems, positive and finite for irregular (chaotic) systems. Quantitatively, K provides a measure of the rate of uncoupling of correlations in phase space.[8] It turns out, however, that it is very difficult to measure K from the time series of points, representing a system's trajectory at regular time intervals. An estimate of K for systems of two degrees of freedom has been obtained by Benettin.[9] For higher degrees of freedom, an algorithm for calculating a quantity $K_2 \leq K$ has been proposed recently.[10]

The K entropy is defined as follows. Consider a dynamical system whose trajectory $\vec{x}(t)$ is confined to a well-defined subspace (called "attractor") of some arbitrary d-dimensional phase space. Let that attractor be partitioned into small boxes of volume ϵ^d and labeled $1,2,3,\ldots M(\epsilon)$. Let $\{\vec{x}(i+1),\vec{x}(i+2),\ldots,\vec{x}(i+l)\}$ denote a time series of l successive points representing the state of the trajectory at discrete time intervals Δt, i.e., $\vec{x}(i)=\vec{x}(t_0+i\Delta t)$. Assume that the first point falls into box $i(1)$, the second into box $i(2)$, etc., so that the series of points leads to the sequence of boxes $I_l=\{i(1),i(2),\ldots i(l)\}$. The total number of distinct sequences I_l which might occur is $M(\epsilon)^l$. Let $P(I_l)$ denote the probability of a given I_l, where

$$\sum_{I_l} P(I_l)=1 \ . \tag{1}$$

The K entropy is then[11]

$$K=-\lim_{\epsilon\to 0}\lim_{l\to\infty}\lim_{\Delta t\to 0}\left(\frac{1}{\Delta t}\right)\sum_{I_l}P(I_l)\ln\left(\frac{P(I_l)}{P(I_{l-1})}\right) \ , \tag{2}$$

where $I_{l-1}=\{i(1),i(2),\ldots i(l-1)\}$. When $M(\epsilon)$ and l are large, as required by Eq. (2), the calculation of K requires the examination of a very large number [which can reach $M(\epsilon)^l$] of distinct sequences I. For simple low-dimensional systems, the need for large $M(\epsilon)$ in Eq. (2) can be circumvented by constructing "generating" two-element partitions.[12] In such a case, the number of distinct I's is considerably reduced and convergent values for K in the range $5\leq l\leq 15$ have been obtained, using a reasonable amount of data, for the logistic[12] and the Lorenz[13] maps. In experimental situations, however, typically high-dimensional systems exist and procedures for the construction of generating two-element partitions are computationally intractable.[12] In the present Rapid Communication, we propose a practical algorithm for the calculation of K through Eq. (2) in the large $M(\epsilon)$ limit. The efficiency of that algorithm has been tested for various chaotic systems[3-7] and, in all the cases, our results verify the conjecture[14-16] that K is the sum of the positive Lyapunov exponents.

Our algorithm is based on the following considerations:

(i) The $P(I)$'s in Eq. (2) have been originally defined for a given partitioning of the attractor's volume into boxes, which is extremely difficult to realize for the high-dimensional systems encountered in experiment. For that reason, we assume that those $P(I)$'s can be estimated through a division of the attractor's volume into (possibly) overlapping balls of radius ϵ centered around each of the data points. In that representation, a point $\vec{x}(j)$ is said to "fall in the same box" as a point $\vec{x}(i)$ if it lies inside a ball of radius ϵ centered around $\vec{x}(i)$, i.e., if $|\vec{x}(j)-\vec{x}(i)|\leq\epsilon$. A similar technique has been used in Ref. 10 for the estimation of K_2 and also in previous calculations[17] of the attractor's dimensions.

(ii) The ensemble average in Eq. (2) requires the labor of constructing a frequency histogram for the occurrence of all the possible sequences I_l of length l. In practice, it is not necessary to go to the large l limit in Eq. (2), that limit being already reached when l is of the order of the correlation length between successive box numbers. Assuming that the ergodic limit can be attained for those short sequences, we then propose to replace the ensemble average in Eq. (2) by a time average over a very long series of length $N>>l$.

Our algorithm thus proceeds as follows. Consider a very long time series $\vec{X}=\{\vec{x}(i)\}_{i=1}^{N}$ of length $N>>l$. Since the ensemble average in Eq. (2) has now been replaced by a time average, K is to be evaluated through a random sampling of $\ln[P(I_l)/PI_{l-1})]$ over $\vec{X}$. We thus start by picking a sequence $\{\vec{x}(i+1),\vec{x}(i+2),\ldots\vec{x}(i+l)\}$ of l points at random out of $\vec{X}$. $P(I_l)$ for that sequence is estimated from the frequency of occurrence in $\vec{X}$ of sequences $\{\vec{x}(j+1),\vec{x}(j+2),\ldots\vec{x}(j+l)\}$, such that $|\vec{x}(j+\alpha)-\vec{x}(i+\alpha)|\leq\epsilon$ for $\alpha=1,2,\ldots l$. For the estimation of $P(I_{l-1})$, α runs from 1 to $l-1$ only. The process is then repeated for another sequence picked at random until the (time) average value of $\ln[P(I_l)/P(I_{l-1})]$ converges. Typically, a few hundred repetitions suffice.

Before presenting our results, we turn to give a rough estimate of the amount of data required by our algorithm for calculating K in a high-dimensional system. For the sake of simplicity and without any loss of generality, our argument will be presented for a partitioning of the system's attractor into $M(\epsilon)$ boxes and not, as actually done, into overlapping balls centered around data points. Consider a sequence

TABLE. I. Kolmogorov entropy for various chaotic systems. Values listed in the second column are those obtained from the Lyapunov exponents. K is the sum of the positive Lyapunov exponents (Refs. 14–16). l^* has been determined visually from the value of l at which a plateau in our K vs l curves (see Fig. 1, for example) starts to appear. The error bars have been calculated from the standard deviation of the K values for different ϵ in the plateau region, i.e., at several $l \geq l^*$.

System	K (Lyapunov exponents)	K [Our algorithm based on Eq. (2)]
Hénon map ($a=1.4$, $b=0.3$)	0.417 ±0.006 (Ref. 21)	0.414 ±0.010 ($N=3.10^4$, $l^*=5$)
Kaplan-Yorke map ($\alpha=0.2$)	0.69315 (Ref. 22)	0.689 ±0.007 ($N=3.10^4$, $l^*=3$)
Lorenz equations		
($R=28$, $\sigma=10$, $b=\frac{8}{3}$)	0.907 ±0.006 (Ref. 23)	0.901 ±0.007 ($N=6.10^4$, $l^*=9$)
($R=40$, $\sigma=16$, $b=4$)	1.373 ±0.005 (Ref. 15)	1.36 ±0.05 ($N=7.10^4$, $l^*=5$)
Rabinovitch-Fabrikant equations		
($\gamma=0.87$), $\alpha=1.1$)	. . .	0.303 ±0.006 ($N=6.10^4$, $l^*=10$)
Mackey-Glass equations		
($a=0.2$, $b=0.1$, $\tau=23$)	0.0095 ±0.0005 (Ref. 14)	0.0092 ±0.0010 ($N=4.10^4$, $l^*=5$)
($a=0.2$, $b=0.1$, $\tau=30$)	0.0090 ±0.0005 (Ref. 14)	0.0086 ±0.0009 ($N=4.10^4$, $l^*=5$)

$\{\vec{x}(i+1), \vec{x}(i+2), \ldots \vec{x}(i+l)\}$ of points picked at random out of a long time series $\vec{X}$. That sequence of points leads to a sequence of boxes $I_l = \{i(1), i(2), \ldots i(l)\}$ and we focus on the number N of data points in $\vec{X}$ that are necessary in order to obtain a reliable estimate for $P(I_l)$. Assuming that the points in $\vec{X}$ are uniformly distributed on the attractor, the probability of finding a point in $i(1)$ is $P[i(1)] = 1/M(\epsilon)$. At small resolution [$M(\epsilon)$ large], box numbers are strongly correlated in time and, for example, our results for the Mackey-Glass[7] map at $\tau = 30$ indicate that $P[i(1), i(2), \ldots i(l)]$ decreases with l roughly as $P[i(1)]/l$. Therefore, the number of points that must be examined to make it likely to find a single series of points falling in the given I_l is $N = lM(\epsilon)$. For a reliable estimation of the $P(I_l)$'s, at least ten occurrences of a given I_l are needed. The number N of data points that must be taken is therefore $N = 10lM(\epsilon)$. In order to quantify these ideas, let us consider our example of the Mackey-Glass map with $\tau = 30$, which is a typically high-dimensional system ($d = 10$, see below). The fractal dimension of that system is around 3.6. Suppose ϵ is of order a tenth the maximum variation in $\vec{x}(i)$. The number of boxes necessary to cover the attractor is then $M(\epsilon) = 10^{3.6} \simeq 4 \times 10^3$. The amount of data required for that system at $l = 5$ is therefore $N = 2 \times 10^5$. Very often, however, the points on an attractor do not have uniform distribution, i.e., different sequences I of box numbers have widely different probabilities. In this case, only the most probable I's are of importance in Eq. (2) and this considerably reduces the number of data points that must be gathered.[18] Indeed, our study of the Mackey-Glass equations at $\tau = 30$ shows that $N = 4 \times 10^4$ points (see Table I) already suffice for a reliable estimation of K. Such amounts of data are readily obtainable from experiment.[19] A last remark on the application of our algorithm to experimental situations should be made. Indeed, typical experimental data are in the form of a time series of a single variable. To apply our method, one must construct from the time series a set of points of dimension d greater than the dimension of the attractor. This is relatively easy to do using derivatives or time delays.[20]

Our algorithm has been tested for various chaotic systems[3–7] which have been widely studied in the literature. The dependence of K on the length l of the sequences being examined is depicted in Fig. 1 for the Hénon[3] and the Kaplan-Yorke[4] maps. Inspection of the figure shows that

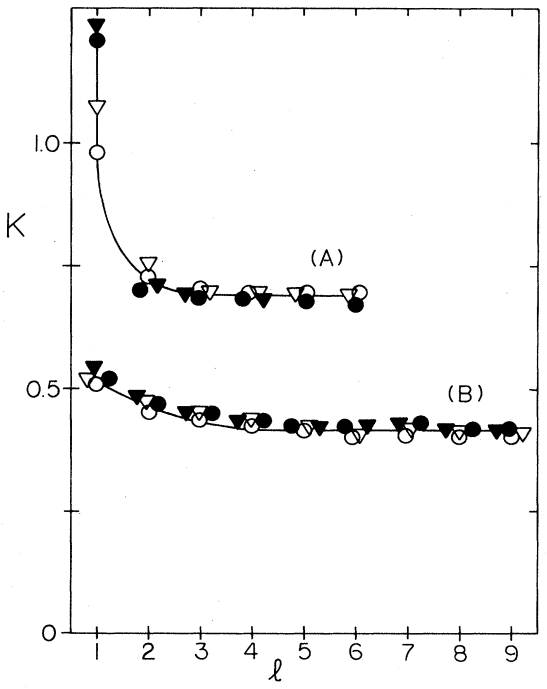

FIG. 1. Dependence of the Kolmogorov entropy K on the sequences length l for the Kaplan-Yorke (A) and the Hénon (B) maps. The values were obtained using our algorithm based on Eq. (2). Values of the parameters are as follows: $\Delta t = 1$ and $\epsilon^2 = 10^{-1}$(○), 5×10^{-2}(▽), 10^{-2}(●), and 5×10^{-3}(▼). The curves are drawn to guide the eye.

the large-l limit in Eq. (2) is rapidly attained (results for the other maps studied below are similar, see also Table I). A typical calculation of K took, depending on the system studied, between 30- and 60-min central-processing-unit time on an IBM 370/165.

Our results for K at the large-l limit are presented in Table I. Integration of the differential equations for the Lorenz,[5] Rabinovitch-Fabrikant,[6] and Mackey-Glass[7] systems were performed using a fourth-order Runge-Kutta procedure. The time step δt for the numerical integration and the time interval Δt chosen between successive measurements were as follows: $\delta t=0.05$ with $\Delta t=0.5$ for the Lorenz map and $\Delta t=0.75$ for the Rabinovitch-Fabrikant equations; $\delta t=\tau/100$ and $\Delta t=\tau$ for the Mackey-Glass equations. For the latter, the d-dimensional vectors in the time series were chosen as $\vec{x}(t)=[x(t),x(t+\tau), x(t+2\tau),\ldots x(t+d\tau)]$ with $d=10$. Inspection of Table I shows that our K values are in excellent agreement with those obtained from the conjecture[14-16] that K is the sum of the positive Lyapunov exponents. This is in contrast with the method proposed in Ref. 10 which led, for example, to $K_2=0.325$ for the Hénon map. As already observed in Ref. 14, the K entropy for the Mackey-Glass equations, in contrast to the fractal dimensionality,[14,24] is rather insensitive to the delay time τ. The constancy of K, as parameters are varied, however does not seem to be a general property of a chaotic system. This is exemplified by the Lorenz equations, for which two widely different values for K are obtained (see Table I), depending on the system's parameters.

To conclude, the numerical results presented above show that the Kolmogorov entropy can be computed with great ease using our algorithm based on Eq. (2). A comparison of that algorithm with previous ones seems appropriate here. First, comparing our method with that proposed in Ref. 10 for the calculation of K_2, we note that the amount of data required by each method is about the same. However, in addition to providing, in principle, the exact value of K, our method is also more convenient than that of Ref. 10. Indeed, the calculation of K_2 requires the additional labor of a scaling analysis for the estimation of the correlation dimension ν, which is often problematic due to the presence of systematic errors.[17] Second, in contrast to the method based on Lyaponov exponents, the application of our algorithm to any experimental signal does not present any difficulty. Thus the present method should be also useful in laser-induced selective chemistry[2] where the K entropy is of great importance in studies of the rate of intramolecular energy relaxation. Work in that direction will be published elsewhere.[25]

This work has been supported by a Sir Charles Clore Postdoctoral Fellowship. I wish to thank Professor Z. Alexandrowicz for his suggestions and comments. Thanks are also due to Dr. H. Meirovitch for valuable discussions.

[1]A. N. Kolmogorov, Dokl. Akad. Nauk. SSSR 124, 774 (1959).
[2]For reviews and references, see P. Brumer, Adv. Chem. Phys. 47, 201 (1981).
[3]M. Hénon, Commun. Math. Phys. 50, 69 (1976).
[4]J. L. Kaplan and J. A. Yorke, in *Functional Differential Equations and Approximations of Fixed Points,* edited by H. O. Peitgen and H. O. Walther (Springer, Berlin, 1979), p. 228.
[5]E. N. Lorenz, J. Atmos. Sci. 20, 130 (1963).
[6]M. I. Rabinovitch and A. L. Fabrikant, Zh. Exp. Theor. Fiz. 77, 617 (1979) [Sov. Phys. JETP 50, 311 (1979)].
[7]M. C. Mackey and L. Glass, Science 197, 287 (1977).
[8]G. M. Zaslavskii and B. V. Chirikov, Usp. Fiz. Nauk. 105, 3 (1971) [Sov. Phys. Usp. 14, 549 (1972)].
[9]G. Benettin, L. Galgani, and J-M. Strelcyn, Phys. Rev. A 14, 2338 (1976).
[10]P. Grassberger and I. Procaccia, Phys. Rev. A 28, 2591 (1983).
[11]J. D. Farmer, Z. Naturforsch. 37A, 1304 (1982).
[12]J. Crutchfield and N. Packard, Int. J. Theor. Phys. 21, 433 (1982).
[13]I. Shimada, Progr. Theor. Phys. 62, 61 (1979).
[14]J. D. Farmer, Physica D 4, 366 (1982).
[15]I. Shimada and T. Nagashima, Prog. Theor. Phys. 61, 1605 (1979).
[16]R. Bowen and D. Ruelle, Inventiones Math. 29, 181 (1975).
[17]P. Grassberger and I. Procaccia, Phys. Rev. Lett. 50, 346 (1983); Physica D (to be published); J. D. Farmer, E. Ott, and J. A. Yorke, *ibid.* (to be published).
[18]This is not to say, however, that our method gets around the limitation facing any dimension-entropy algorithm, namely, that the number of boxes needed to cover an attractor grows exponentially with the dimension of the attractor.
[19]Up to $N=3\times10^4$ data have been gathered in Couette-Taylor flow experiments; see A. Brandstäter, J. Swift, H. L. Swinney, A. Wolf, J. D. Farmer, E. Jen, and P. J. Crutchfield, Phys. Rev. Lett. 51, 1442 (1983). The dripping faucet could produce $\sim10^4$ good points; see R. Shaw, University of California at Santa Cruz report (unpublished).
[20]R. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, Phys. Rev. Lett. 45, 712 (1980); F. Takens, in *Dynamical Systems and Turbulence,* Lecture Notes in Mathematics, Vol. 898, edited by D. A. Rand and L. S. Young (Springer, Berlin, 1981), p. 366.
[21]S. D. Feit, Commun. Math. Phys. 61, 249 (1978); C. Simo, J. Stat. Phys. 21, 465 (1979).
[22]D. A. Russel, J. D. Hanson, and E. Ott, Phys. Rev. Lett. 45, 1175 (1980).
[23]Y. Aizawa, Progr. Theor. Phys. 68, 64 (1982).
[24]Y. Termonia and Z. Alexandrowicz, Phys. Rev. Lett. 51, 1265 (1983).
[25]Y. Termonia, J. Chem. Phys. (to be published).