Hausdorff dimension from the minimal spanning tree

Vicent J. Martínez
Departament de Matemàtica Aplicada i Astronomia, Universitat de València, Burjassot, 46100 València, Spain

R. Domínguez-Tenreiro
Departamento de Física Teórica, Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain

L. J. Roy
Departamento de Física Aplicada, Universidad de Zaragoza, Corona de Aragón 35, 50009 Zaragoza, Spain
(Received 24 January 1992; revised manuscript received 1 July 1992)

A technique to estimate the Hausdorff dimension of strange attractors, based on the minimal spanning tree of the point distribution, is extensively tested in this work. This method takes into account in some sense the infimum requirement appearing in the definition of the Hausdorff dimension. It provides accurate estimates even for a low number of data points and it is especially suited to high-dimensional systems.

PACS number(s): 05.45.+b, 02.50+s

A peculiar feature of dynamical systems which exhibit chaotic behavior is the appearance of strange attractors for some values of their control parameters. These strange attractors are characterized by a noninteger dimension, the Hausdorff dimension, which distinguishes between regular motion and deterministic chaos.

The Hausdorff dimension is mathematically defined as follows [1]: Let us consider a set A with a measure μ embedded in a Euclidean d-dimensional space. Let $R(\epsilon)$ be the family of all countable coverings of A formed by domains of diameters $\delta_i \leq \epsilon$. For $\beta > 0$ we form the partition sums:

$$H^{\beta}(A) = \lim_{\epsilon \rightarrow 0} \inf{}_{R(\epsilon)} \sum_{i \in I} \delta_i^{\beta}, \qquad (1)$$

where the infimum is taken on the set of all the possible countable coverings of A with diameters $\delta_i \leq \epsilon$, $R(\epsilon)$. It can be shown that for any A there always exists a critical value D_H such that

$$H^{\beta}(A) = 0 \;\text{ if } \beta > D_H, \;\; H^{\beta}(A) = \infty \;\text{ if } \beta < D_H. \qquad (2)$$

This critical exponent is known as the Hausdorff dimension of the set A.

When one intends to calculate the Hausdorff dimension of strange attractors appearing in dynamical systems (either maps or autonomous ordinary differential equations), in most cases what we have at our disposal is a time series formed by a discrete set of N points which sample the attractor. The Hausdorff dimension of any countable set is trivially zero; nevertheless, we are interested in the estimation of the dimensionality of the attractor itself, i.e., the support of the point distribution. The mathematical definition of the Hausdorff dimension given in Eqs. (1) and (2) cannot be applied to our problem and, in consequence, we do not aim at a calculation of the Hausdorff dimension in the sense of this definition, but only at an estimate of the Hausdorff dimension of the strange attractor from its finite point realization. This estimate has to remain as close as possible to the original definition.

In the literature about this subject, apart from the calculations based on local Lyapunov exponents [2–4], the following approaches are found.

(a) Traditionally, D_H has been calculated through the box-counting algorithm, which leads to the so-called attractor capacity [5] D_C. However, this method is very memory consuming and it is not useful when d is greater than 2 [6,7].

(b) Another possibility is to randomly choose N_R points among the N which sample the attractor and to take as domains balls centered on them. A covering is formed by fixing a natural number n and taking the ball radii $\epsilon_i(N_R, n)$ in such a way that each ball contains n points. There are two possibilities: (i) either the n points are taken out of the N points which sample the attractor [4,8,9] or (ii) the n points are taken out of the N_R randomly chosen among the N [10]. In this way we assign to each ball a weight or probability:

$$p = \frac{n}{N} \qquad (3)$$

in the first case, and

$$p = \frac{n}{N_R} \qquad (4)$$

in the second case. The radii $\epsilon_i(N_R, n)$ are the distances of point i to its n-nearest neighbor. The partition sums which allow us to determine the Hausdorff dimension are

$$W^{\beta}(A, p) = \frac{1}{M} \sum_{i=1}^{M} \epsilon_i(p)^{\beta} \qquad (5)$$

with $M = N_R$ in case (i) and $M = N$ in case (ii).

For finite point distributions, the limit appearing in Eq. (1) (equivalent in this case to making $N \to \infty$) cannot be taken, and we have to resort to some approximating methods. The sums in Eq. (5) present a scaling behavior when the probability p is varied within a certain range of its values, termed the scaling region. The probability p can be varied either by changing n in Eq. (3) or by changing N_R in Eq. (4). In both cases we have, for p within its scaling region,

$$W^{\beta}(A,p) = K(\beta)p^{\beta/g(\beta)}. \tag{6}$$

The fixed point of the $g(\beta)$ function, defined by this equation, gives us an estimate of the Hausdorff dimension $D_{(0)} = g(D_{(0)})$ [4,8,10]. Strictly speaking, the method gives an upper limit [11] to D_H when $N \to \infty$. Notice that even if the methods (i) and (ii) to construct the probabilities are equivalent from a mathematical point of view, when dealing with finite point sets and due to finite sample effects, they could lead to results which are not exactly the same.

(c) Another possible choice is to form a covering by fixing the ball radii to a value ϵ, and then to assign to each ball a probability $p_i(\epsilon)$ as in Eq. (3), with $n_i(\epsilon)$ now depending on the particular ball considered. The partition sum is now

$$Z(A,\epsilon) = \frac{1}{N_R}\sum_{i=1}^{N_R} p_i(\epsilon)^{-1}. \tag{7}$$

If it scales as

$$Z(A,\epsilon) \propto \epsilon^{-D^{(0)}}, \tag{8}$$

then we get another estimate of D_H, $D^{(0)}$ [12,13].

We would like to point out that in the methods outlined so far, the infimum on the coverings of the attractor, demanded by the definition of the Hausdorff dimension, has not been considered. We now present another method where this infimum is in some sense taken into account: the minimal-spanning-tree (MST) technique.

The MST of a point distribution is a graph-theoretical construct which was introduced by Kruskal [14] and Prim [14]. Stated in terms of graph theory, given a set of N points, a spanning tree is a network of $N-1$ edges, each of them linking two points in the distribution, such that it provides a path between any pair of points in the set and contains no closed loops or circuits. The minimal-spanning-tree is the tree which satisfies the condition that the sum of the lengths of its edges is minimum. For a given distribution of N points, only one MST can be constructed. A spanning tree defines a set of balls whose diameters are the edge size and which are centered on the middle point of these edges. These balls provide a covering for the attractor.

Given N points which sample an attractor and N_R points randomly chosen among them, their MST is formed by the N_R points and $m = N_R - 1$ segments of lengths $\{l_i\}_{i=1}^{m}$. When the N points have compact support on $\mathbb{R}^d$, with $d \geq 2$, it has been shown [15] that the MST provides us with a method to calculate the dimension d. We now conjeture that when, instead, the N points have support in a strange attractor, the MST provides us with a technique to estimate its Hausdorff dimension. The Hausdorff dimension of the attractor A can be determined [16] by forming the sums

$$S^{\beta}(A,m) = \frac{1}{m}\sum_{i=1}^{m} l_i(m)^{\beta}. \tag{9}$$

In these sums, the infimum condition appearing in Eq. (1) has not been made explicit, but due to the fact that the MST is the spanning tree with minimal length, the infimum requirement has somehow been considered, because $l_i(m)$ are the branch lengths of the MST. In fact, the sums in Eq. (9) give us the infimum taken over all possible coverings of the attractor formed by spanning trees as explained above. These sums present a scaling behavior as m is varied:

$$S^{\beta}(A,m) = K(\beta)m^{-\beta/h(\beta)}, \tag{10}$$

which defines the function $h(\beta)$. The fixed point of this function $h(D(0)) = D(0)$ is a good estimate of the Hausdorff dimension. Due to the fact that in the sums of Eq. (9) we have considered the infimum not over all possible coverings of the attractor but only over those defined by spanning trees, the same arguments of Ref. [11] can be used to show that, as far as $m \to \infty$ makes sense, the following inequality holds $D_H \leq D(0)$. Note, however, that when m does not go to infinity, there is no guarantee that the estimate is an upper bound to D_H as defined by Eqs. (1) and (2) above.

Equations (9) and (10) are a particularization of the general definition of the Hausdorff dimension when the covering provided by the MST of the N points is used. The condition of minimum on the different possible coverings is now taken into account by demanding that the total length of the spanning tree is minimum. This remains as close as possible to realizing the definition of the Hausdorff dimension within present computer capacities.

We have calculated the MST of point sets using the algorithm described by Prim in the Whitney [17] version. An arbitrary point in the set is first chosen, and its nearest neighbor is found. These two points and the edge which links them form the first subtree T_1. For each isolated point (point not yet in the subtree), the algorithm calculates its distance to the subtree T_1, that is, the distance to its nearest neighbor within the subtree. These distances and the identity of these neighbors are stored. The Mth subtree T_M is built up by adding to the $(M-1)$th subtree T_{M-1} the isolated point whose distance to T_{M-1} is minimum, together with the corresponding link. Then, the distances from isolated points to T_M are calculated, and so on. The MST is T_{N_R-1}. The total computer CPU time required goes as $N_R(N_R-1)/2$ in a scalar machine and as N_R in a vectorial machine, and it is also proportional to the embedding dimension of the point set d.

The suitability of the MST technique to estimate the Hausdorff dimension has been tested on four well-known two-dimensional maps: the Baker [18] transformation with $a = 0.3$; the Lozi [19] map with $a = 1.7, b = 0.5$; the Hénon [20] map with $a = 1.4, b = 0.3$; and the Kaplan and Yorke map [21] with $\alpha = 0.2$; and on one three-dimensional system of ordinary differential equations: the Lorenz [22] system with parameters $s = 10$, $b = 8/3$, and $r = 28$.

In all cases, time series with $N = 2 \times 10^4$ points were generated (in the case of the Lorenz attractor, a delay time of $\tau = 0.25$ was employed). We used as the scaling region the interval $100 \leq N_R \leq 7500$, where nine different logarithmically equidistant values of N_R were fixed. Then, minimal-spanning trees with N_R points randomly chosen among the N were calculated for the nine different values of N_R, and the sums in Eq. (9) were formed. The $h(\beta)$ function was determined from Eq. (10) by performing linear regressions in the log-log plane at fixed β values.

Due to finite sample effects, the value of the $S^\beta(A,m)$ sums depends to a small degree on the particular choice of the N_R points. In order to alleviate this difficulty, ten different random selections were performed for each N_R value, and then we have calculated $h(\beta)$ by means of a linear-regression fit

$$\langle \log_{10} S^\beta(A,m) \rangle = \log_{10} K(\beta) - \frac{\beta}{h(\beta)} \log_{10} m, \qquad (11)$$

where the angular brackets stands for an average performed on the different random selections. We gave the same weight to each of the nine N_R values because the changes in the partition sums when the result of the N_R points selection varies are similar for any N_R.

The resulting $D(0)$ values we obtained for the five attractors are given in column 2 of Table I. Column 3 gives the errors obtained when the scaling region is varied and we take seven points within either $294 \leq N_R \leq 7500$ or $100 \leq N_R \leq 2548$. In order to test the stability of our results against the change in the choice of the N_R random points, the results of the ten above-mentioned selection processes were taken in groups of five. Then, the averages of $\log_{10} S^\beta(A,m)$ were calculated independently over each group of five selections, and the results obtained for $D(0)$ were compared. In column 4 of Table I we give the maximum differences of these results with respect to the $D(0)$ values shown in column 2. Finally, in column 5, we present different D_H estimates obtained by other methods.

Comparison of columns 2 and 5 in Table I indicates that the MST technique gives very good estimates of the Haussdorff dimension, even for a relatively low number of points in the time series. This is particularly remarkable for the Hénon attractor: the method produces a $D(0)$ value which agrees with those found through local Lyapunov exponents.

Some comments are necessary concerning the Lorenz attractor. Because of the paucity of estimates of its Hausdorff dimension (the box-counting technique does not converge in this case), we repeated the calculation with $N = 10^5$ points in the time series, taking as the scaling region the range $400 \leq N_R \leq 40\,000$, where 20 logarithmically equidistant values of N_R were fixed. We repeated the calculation procedure explained above, and obtained the results given in Table I. We are led to the conclusion that the method already gives a good accuracy for $N = 2 \times 10^4$ points. Moreover, the Hausdorff dimension of the Lorenz attractor was evaluated by means of th n-nearest-neighbor technique[23] [method (b) above], and, as shown in Table I, taking $N = 10^5$ for the time series the results of this method agree within the errors with the values reported from the MST. Note, however, that both $D(0)$ and $D_{(0)}$ are different estimates of the Hausdorff dimension D_H, so it is not necessary that they coincide. Also, the overall scaling properties of Eqs. (6) and (10) are not the same; in fact, for the Lorenz and other attractors, they produce $g(\beta)$ and $h(\beta)$ functions which are different [23].

To conclude, an estimate of the Hausdorff dimension for point distributions can be calculated using the MST algorithm, for any value of the phase-space dimension d

TABLE I. The Hausdorff dimension of different attractors calculated by means of the MST. Other estimates are given for comparison.

System	$D(0)$	$\Delta_1 D(0)$	$\Delta_2 D(0)$	Other estimates
Baker	1.596	0.029	0.003	$1 + \log_{10} 2/\lvert \log_{10} 0.3\rvert = 1.575\,716\,642$[a]
Hénon	1.270	0.022	0.003	1.272 ± 0.006[b] 1.2755 ± 0.0005[a] 1.28 ± 0.01[c] 1.261 ± 0.003[d]
Lozi	1.431	0.001	0.004	
Kaplan and Yorke	1.396	0.017	0.005	1.4316 ± 0.0016[d]
Lorenz	2.165	0.002	0.009	2.148 ± 0.002[e]
Lorenz ($N = 10^5$)	2.155	0.002	0.005	

[a]See Ref. [3].
[b]See Ref. [2].
[c]See Ref. [7].
[d]See Ref. [24].
[e]See Ref. [23].

and for any kind of attractor. It is much more efficient than the classical box-counting algorithm, and produces accurate estimates of D_H even with only a low number of points in the time series. The computational efficiency of the MST technique is similar to the efficiency of the method based in the n-nearest-neighbor distances, when binary trees are used for the searching algorithm. In fact, one run of this former method with $N = 20\,000$ points performed in order to store the distances from each point to its n-nearest neighbor, for nine values of n in the scaling region $2^4 \leq n \leq 2^8$, takes 5.8 times more CPU time than one run of the MST method, with also nine points in the scaling interval $100 \leq N_R \leq 7500$. Nevertheless, as it is recommendable to average for several choices of the N_R random points [see Eq. (11)], the efficiency is rather equivalent. It is also possible to generalize Eq. (10) in order to extract information about the Rényi dimensions from the MST method [23,25].

We want to thanks Bernard Jones and Rien van de Weygaert for interesting discussions. Computer facilities provided by the IFIC and the CICYT are acknowledged. This work was partially supported by the Dirección General de Investigación Científica y Técnica (Project No. PB86-0292-C04-03-04), Spain.

[1] F. Hausdorff, Math. Ann. **79**, 157 (1919).
[2] P. Grassberger and I. Procaccia, Physica D **13**, 34 (1984).
[3] R. Badii and A. Politi, Phys. Rev. A **35**, 1288 (1987).
[4] P. Grassberger, R. Badii, and A. Politi, J. Stat. Phys. **51**, 135 (1988).
[5] B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, 1982).
[6] H. Greenside, A. Wolf, J. Swift, and T. Pignataro, Phys. Rev. A **25**, 3453 (1982).
[7] P. Grassberger, Phys. Lett. **97A**, 224 (1983).
[8] P. Grassberger, Phys. Lett. **107A**, 101 (1985).
[9] Y. Termonia and Z. Alexandrowicz, Phys. Rev. Lett. **51**, 1265 (1983).
[10] R. Badii and A. Politi, Phys. Rev. Lett. **52**, 1661 (1984); J. Stat. Phys. **40**, 725 (1985).
[11] T.C. Hasley, M.H. Jensen, L.P. Kadanoff, I. Procaccia, and B.I. Shraiman, Phys. Rev. A **33**, 1141 (1986).
[12] R. Benzi, G. Paladin, G. Parisi, and A. Vulpiani, J. Phys. A **17**, 3521 (1984).
[13] S. Sato, M. Sano, and Y. Sawada, Prog. Theor. Phys. **77**, 1 (1987).
[14] J.B. Kruskal, Proc. Am. Math. Soc. **7**, 48 (1956), R.C. Prim, Bell Syst. Tech. J. **3**, 1389 (1957).
[15] J.M. Steele, Ann. Prob. **16**, 1767 (1988).
[16] V.J. Martínez and B.J.T. Jones, Mon. Not. R. Astron. Soc. **242**, 517, (1990).
[17] V.K.M. Whitney, Commun. ACM **15**, 273 (1972).
[18] J.D. Farmer, E. Ott, and J. A. Yorke, Physica D **7**, 153 (1983).
[19] R. Lozi, J. Phys. (Paris) Colloq. **39**, C5-9 (1978).
[20] M. Hénon, Commun. Math. Phys. **50**, 69 (1976).
[21] J.C. Kaplan and J.A. Yorke, in *Functional Differential Equations and Approximations of Fixed Points*, edited by H.-O. Peitgen and H.-O. Walther, Lecture Notes in Mathematics Vol. 730 (Springer, Berlin, 1979).
[22] E.N. Lorenz, J. Atmos. Sci. **20**, 130 (1963).
[23] R. Domínguez-Tenreiro, L. J. Roy, and V.J. Martínez, Prog. Theor. Phys. **87**, 1107 (1992).
[24] D.A. Russell, J.D. Hanson, and E. Ott, Phys. Rev. Lett. **45**, 1175 (1980).
[25] R. van de Weygaert, B.J.T. Jones, and V.J. Martínez, Phys. Lett. **169A**, 145 (1992).