Estimating fractal dimension

James Theiler

Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts 02173-9108

Received September 13, 1989; accepted November 21, 1989

Fractals arise from a variety of sources and have been observed in nature and on computer screens. One of the exceptional characteristics of fractals is that they can be described by a noninteger dimension. The geometry of fractals and the mathematics of fractal dimension have provided useful tools for a variety of scientific disciplines, among which is chaos. Chaotic dynamical systems exhibit trajectories in their phase space that converge to a strange attractor. The fractal dimension of this attractor counts the effective number of degrees of freedom in the dynamical system and thus quantifies its complexity. In recent years, numerical methods have been developed for estimating the dimension directly from the observed behavior of the physical system. The purpose of this paper is to survey briefly the kinds of fractals that appear in scientific research, to discuss the application of fractals to nonlinear dynamical systems, and finally to review more comprehensively the state of the art in numerical methods for estimating the fractal dimension of a strange attractor.

Confusion is a word we have invented for an order which is not understood.

—Henry Miller, "Interlude,"
Tropic of Capricorn

Numerical coincidence is a common path to intellectual perdition in our quest for meaning. We delight in catalogs of disparate items united by the same number, and often feel in our gut that some unity must underlie it all.

—Stephen Jay Gould, "The Rule of Five,"
The Flamingo's Smile

INTRODUCTION

Fractals are crinkly objects that defy conventional measures, such as length and area, and are most often characterized by their fractional dimension. This paper will arbitrarily categorize fractals into two main types: solid objects and strange attractors.

Of the first type, coastlines and clouds are typical examples. Further examples include electrochemical deposition,[1,2] viscous fingering,[3] and dielectric breakdown[4] as well as porous rocks[5] and spleenwort ferns.[6] These are physical objects that exist in ordinary physical space. Percolation clusters,[7] diffusion-limited aggregates,[8] and iterated function systems[9] provide mathematical models of fractals that correspond to these physical objects. An overview of many of these topics can be found in Ref. 10, and a recent bibliography has been compiled in Ref. 11.

Strange attractors, by contrast, are conceptual objects that exist in the state space of chaotic dynamical systems. The emphasis of this paper will be on strange-attractor dimensions, although some of what is discussed can also be applied to solid fractal objects.

Self-Similarity and Self-Affinity

Fractal objects, as Kadanoff notes, "contain structures nested within one another like Chinese boxes or Russian dolls."[12] This self-similar structure is perhaps the main reason for the striking beauty of so many fractals. Self-similarity also implies a scale-invariant property. There are crinkles upon crinkles and therefore no natural or preferred crinkle size. A set is strictly self-similar if it can be expressed as a union of sets, each of which is a reduced copy of (is geometrically similar to) the full set. See Fig. 1(a). However most fractal-looking objects in nature do not display quite this precise form. In a coastline, for instance, there is an irregular nesting of gulfs, bays, harbors, coves, and inlets that are observed over a broad range of spatial scales. See Fig. 1(b). A magnified view of one part of the coastline will not precisely reproduce the full picture, but it will have the same qualitative appearance. A coastline displays the kind of fractal behavior that is called statistical self-similarity.

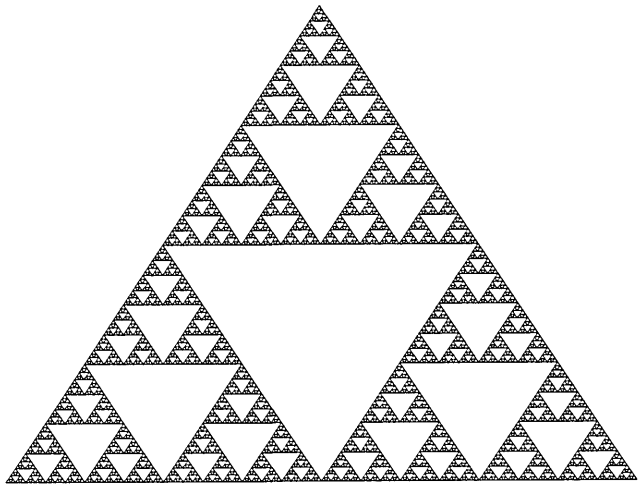
A fractal is self-affine if it can be decomposed into subsets that can be linearly mapped into the full figure. If this linear map involves only rotation, translation, and (isotropic) dilation, then the figure is self-similar. For a self-affine map, the contraction in one direction may differ from the contraction in another direction. The class of self-affine fractals therefore includes the class of self-similar fractals. (A visually impressive collection of self-affine fractals, and a method of generating them that goes under the name "the chaos game," can be found in Refs. 6 and 9.) The distinction between self-similar and self-affine fractals is not always made in practice. The concept is most useful in cases for which there are preferred global directions: in fractal surfaces,[5] for example, or for fractal profiles [these are continuous nowhere-differentiable functions $f(x)$ for which the graph $(x, f(x))$ is a set of fractal dimension[13,14]].

Quantifying Fractals

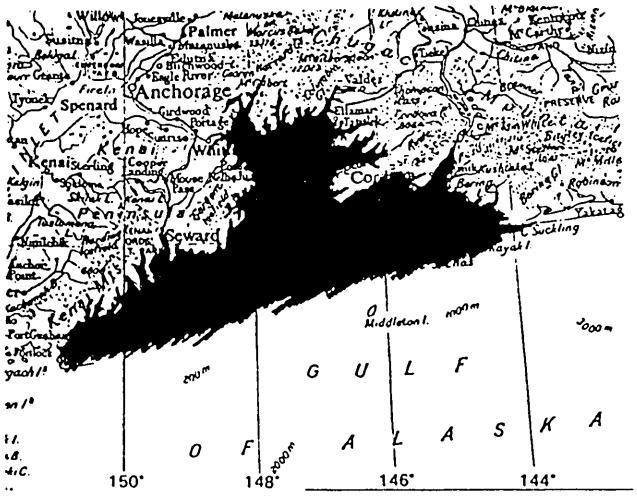
Stanley[10] has outlined the program of the practicing scientist who wants to study fractals:

> If you are an experimentalist, you try to measure the fractal dimension of things in nature. If you are a theorist, you try to calculate the fractal dimension of models chosen to describe experimental situations; if there is no agreement then you try another model.

0740-3232/90/061055-19$02.00 © 1990 Optical Society of America



(a)



(b)

Fig. 1. Self-similar fractals. (a) The Sierpinski gasket. Here, the inner triangles are small copies of the full figure. (b) Prince William Sound. The sound has been darkened to highlight the fractal appearance of the coastline.

This pithy advice applies to strange attractors as well as to solid fractal objects: Fractal dimension provides the benchmark against which theories are compared with experiments. In the case of strange attractors, however, there are further reasons for wanting to know the dimension.

CHAOS AND STRANGE ATTRACTORS

Why Quantify Chaos?: A Manifesto

Chaos is the irregular behavior of simple equations, and irregular behavior is ubiquitous in nature. A primary motivation for studying chaos is this: Given an observation of irregular behavior, is there a simple explanation that can account for it? And if so, how simple? There is a growing consensus that a useful understanding of the physical world will require more than finally uncovering the fundamental laws of physics. Simple systems, which obey simple laws, can nonetheless exhibit exotic and unexpected behavior. Nature is filled with surprises that turn out to be direct consequences of Newton's laws.

One usually measures the complexity of a physical system by the number of degrees of freedom that the system possesses. However, it is useful to distinguish nominal degrees of freedom from effective (or active) degrees of freedom. Although there may be many nominal degrees of freedom available, the physics of the system may organize the motion into only a few effective degrees of freedom. This collective behavior, which can be observed in the laminar flow of a fluid or in the spiral arms of a galaxy, is often termed self-organization. (Much of the interest in the principle of self-organization has been stimulated by the writings of Haken,[15] an early and enthusiastic advocate.) Self-organization is interesting because the laws of thermodynamics seem to forbid it: A self-organizing system decreases its own entropy. (For real physical systems, this apparent decrease in entropy is achieved by dissipation to an external bath. In this case the total entropy—of the system and of the bath—does increase, and the second law of thermodynamics is satisfied.)

Self-organization arises in dissipative dynamical systems whose posttransient behavior involves fewer degrees of freedom than are nominally available. The system is attracted to a lower-dimensional phase space, and the dimension of this reduced phase space represents the number of active degrees of freedom in the self-organized system. A system that is nominally complex may in fact relax to a state of chaotic but low-dimensional motion. Distinguishing behavior that is irregular but low dimensional from behavior that is irregular because it is essentially stochastic (many effective degrees of freedom) is the motivation for quantifying chaos. Estimating dimension from a time series is one way to detect and quantify the self-organizational properties of natural and artificial complex systems.

Strange Attractors

Strange attractors arise from nonlinear dynamical systems. Physically, a dynamical system is anything that moves. (And if it does not move, then it is a dynamical system at a fixed point.) Mathematically, a dynamical system is defined by a state space $\mathbf{R}^M$ (also called phase space) that describes the instantaneous states available to the system and an evolution operator ϕ that tells how the state of the system changes in time. (One usually thinks of ϕ as the physics of the system.) An element $\mathbf{X} \in \mathbf{R}^M$ of the state space specifies the current state of the system, and M characterizes the number of degrees of freedom in the system. For a particle in space, $\mathbf{X}$ might represent the three coordinates of the particle's position, and the three coordinates of its momentum, for a total of $M = 6$ degrees of freedom. The evolution operator is a family of functions $\phi_t:\mathbf{R}^M \to \mathbf{R}^M$ that map the current state of the system into its future state at a time t units later. The operator ϕ_t satisfies $\phi_0(\mathbf{X}) = \mathbf{X}$ and $\phi_{t+s}(\mathbf{X}) = \phi_t[\phi_s(\mathbf{X})]$. The dynamical system is nonlinear if, in general, $\phi_t(c_1\mathbf{X}_1 + c_2\mathbf{X}_2) \neq c_1\phi_t(\mathbf{X}_1) + c_2\phi_t(\mathbf{X}_2)$.

The function ϕ_t can be defined either as a discrete map or in terms of a set of ordinary differential equations, although partial differential equations and differential delay equations have also been studied (in these last two cases, M is infinite).

Dissipative Dynamics

Although conservative dynamical systems can also exhibit chaos,[16] only dissipative dynamical systems have strange

attractors. A system is dissipative if the volume of a fiducial chunk of phase space tends to zero as $t \to \infty$. In other words, if $\mathcal{B}$ is a bounded subset of $\mathbf{R}^M$ with (Lebesque) volume $\mu_L(\mathcal{B})$, then

$$\lim_{t\to\infty} \mu_L[\phi_t(\mathcal{B})] = 0. \tag{1}$$

Usually the dissipation is exponential (at least on the average for a long time): $\mu_L[\phi_t(\mathcal{B})] \sim \mu_L(\mathcal{B})e^{-\Lambda t}$, with the rate of dissipation Λ.

Informally the attractor $\mathcal{A}$ of a dynamical system is the subset of phase space toward which the system evolves. An initial condition $\mathbf{X}_0$ that is sufficiently near the attractor will evolve in time so that $\phi_t(\mathbf{X}_0)$ comes arbitrarily close to the set $\mathcal{A}$ as $t \to \infty$. If the set $\mathcal{A}$ is a fractal, then the attractor is said to be strange. A more formal treatment of strange attractors in dynamical systems can be found in Ref. 17.

Natural Invariant Measure

Equation (1) implies that the phase-space volume of the attractor $\mu_L(\mathcal{A})$ is zero. To quantify the dynamics on the attractor requires first the introduction of a new measure μ that is concentrated on the attractor. A measure μ is defined on the set $\mathcal{A}$ if the subsets $\mathcal{B}$ of the set $\mathcal{A}$ can be associated with real values $\mu(\mathcal{B})$ that represent how much of $\mathcal{A}$ is contained in $\mathcal{B}$. The measure that is defined on a set reflects the varying density over the set and can intuitively be regarded as a mass. See Fig. 2. The reader is referred to Ref. 20 for a more rigorous discussion of measure in the context of fractal sets.

A useful measure μ should be invariant under time evolution: The proper way to define this is to write

$$\mu(\mathcal{B}) = \mu[\phi_{-t}(\mathcal{B})], \tag{2}$$

where $\phi_{-t}(\mathcal{B}) \equiv \{\mathbf{X} \in \mathbf{R}^M : \phi_t(\mathbf{X}) \in \mathcal{B}\}$. In general, Eq. (2) is not enough to define uniquely the measure for a dynamical system; for instance, a measure that is concentrated on an unstable fixed point satisfies Eq. (2) but has little to do with the generic posttransient motion of the system.

The physically relevant measure for a dynamical attractor counts how often and for how long a typical trajectory visits various parts of the set. A constructive definition is given by

$$\mu(\mathcal{B}) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \mathbf{I}_{\mathcal{B}}[\phi_t(\mathbf{X}_0)]\,dt, \tag{3}$$

where $\mathbf{X}_0$ is a typical initial condition and $\mathbf{I}_{\mathcal{B}}(\mathbf{X})$ is the indicator function for $\mathcal{B}$: It is unity if $\mathbf{X} \in \mathcal{B}$ and is zero otherwise. The natural invariant measure is given by Eq. (3) for almost all $\mathbf{X}_0$. An alternative definition that avoids the notion of typical trajectories by adding infinitesimal noise to the dynamics is mentioned in Ref. 21.

Sensitivity to Initial Conditions

The hallmark of a chaotic system is its sensitivity to initial conditions. This sensitivity is usually quantified in terms of the Lyapunov exponents and the Kolmogorov entropy. The Lyapunov exponents measure the rate of exponential divergence of nearby trajectories, and the Kolmogorov entropy measures the rate of information flow in the dynamical system.

Long-term predictions of chaotic systems are virtually impossible, even if the physics (that is, ϕ_t) is known completely, because errors in measurement of the initial state propagate exponentially fast.

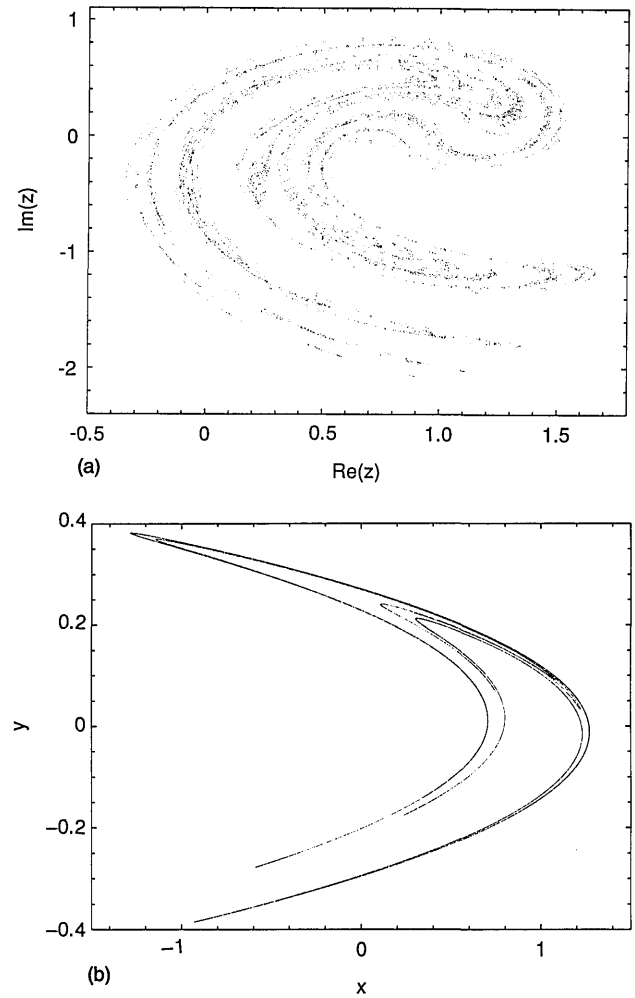


Fig. 2. Two strange attractors. (a) The Ikeda map. The complex map, $z_{n+1} = a + Rz_n \exp\{i[\phi - p/(1 + |z_n|^2)]\}$, which derives from a model of the plane-wave interactivity field in an optical ring laser,[18] is iterated many times, and the points $[\mathrm{Re}(z_n), \mathrm{Im}(z_n)]$ are plotted for $n \geq 1000$. Here, $a = 1.0$, $R = 0.9$, $\phi = 0.4$, $p = 6$. The fractal dimension of this attractor is approximately 1.7. (b) The Hénon map.[19] This map, $x_{n+1} = 1.0 - ax_n^2 + y_n$; $y_{n+1} = bx_n$, with $a = 1.4$ and $b = 0.3$, gives an attractor with fractal dimension of approximately 1.3. Note how much thinner the attractor of lower dimension appears.

Consider two nearly initial conditions, $\mathbf{X}_0$ and $\mathbf{X}_0 + \epsilon$, and evolve both forward in time. A Taylor-series expansion gives

$$\phi_t(\mathbf{X}_0 + \epsilon) = \phi_t(\mathbf{X}_0) + \mathbf{J}(t)\,\epsilon + O(|\epsilon|^2), \tag{4}$$

where $\mathbf{J}(t)$ is the Jacobian matrix given by the linearization of ϕ_t about the point $\mathbf{X}_0$:

$$\mathbf{J}(t) = \frac{\partial \phi_t(\mathbf{X}_0)}{\partial \mathbf{X}_0} = \frac{\partial \mathbf{X}(t)}{\partial \mathbf{X}(0)}. \tag{5}$$

The ij element of this matrix is

$$J_{ij}(t) = \frac{\partial X_i(t)}{\partial X_j(0)}, \tag{6}$$

where $X_i(t)$ is the ith component of the state vector $\mathbf{X}$ at time t. Thus an initial small separation ϵ is magnified to $\mathbf{J}(t)\epsilon$. The determinant of the $\mathbf{J}(t)$ describes the overall contraction of phase-space volume (dissipation), and the eigenvalues describe the divergence of nearby trajectories. The Lyapunov exponents quantify the average rate of expansion of these eigenvalues:

$$\lambda_n = \lim_{t\to\infty} \frac{1}{t} \log |n\text{th eigenvalue of } \mathbf{J}(t)|. \tag{7}$$

Conventionally the Lyapunov exponents are indexed in descending order: $\lambda_1 \geq \lambda_2 \geq \ldots$. The largest Lyapunov exponent is the clearly most important since, if the difference between a pair of initial conditions has a nonzero component in the eigendirection associated with the largest eigenvalue, then that rate of divergence will dominate. If the largest Lyapunov exponent is positive, then chaotic motion is ensured. (Otherwise, trajectories will collapse to a fixed point, a limit cycle, or a limit torus.) The sum of all the exponents $\lambda_1 + \ldots + \lambda_M = -\Lambda$ is negative for a dissipative dynamical system and defines the rate Λ of phase-space contractions.

If the state $\mathbf{X}(t)$ of the system is known at time t to an accuracy ϵ, then the future can be predicted by $\mathbf{X}(t + \Delta t) = \phi_{\Delta t}[X(t)]$ but to an accuracy that is usually worse than ϵ. Thus to observe (or measure) the state of the system again at time $t + \Delta t$ to the original accuracy ϵ is to learn information that was previously unavailable. The system is in this sense creating information. (Some authors prefer to say that the system is losing information since the initial accuracy is lost; this choice is largely a matter of taste, since the numerical results are the same.) The average rate of information gain (or loss) is quantified by the Kolmogorov entropy, which is given by the sum of the positive Lyapunov exponents.[21]

The numerical estimation of Lyapunov exponents and Kolmogorov entropy from a time series is discussed in Refs. 22–24. A method for estimating Kolmogorov entropy that is related to the correlation dimension is developed in Refs. 25 and 26. The notion of a generalized entropy K_q is discussed in Refs. 27 and 28.

Delay-Time Embedding

An experimentalist, confronted with a physical system, measures at regular and discrete intervals of time the value of some state variable (voltage, say) and records the time series: $x(t_0), x(t_1), x(t_2), \ldots$, with $x(t_i) \in \mathbf{R}$ and $t_i = t_0 + i\Delta t$. [This section discusses a stroboscopic time discretization of a continuous flow. The method of Poincaré section (see Ref. 17 or any textbook) is another way to time discretize a flow, but it is a method that is rarely available to the experimentalist.] The measurement $x(t)$ represents a projection

$$\pi: \mathbf{R}^M \to \mathbf{R} \tag{8}$$

from the full state vector $X(t) \in \mathbf{R}^M$. A time series is manifestly one dimensional and, as such, provides an incomplete description of the system in its time evolution. On the other hand, many properties of the system can be inferred from the time series.

Packard *et al.*[29] devised a delay scheme to reconstruct the state space by embedding the time series into a higher-dimensional space. From time-delayed values of the scalar time series, vectors $\hat{\mathbf{X}} \in \mathbf{R}^m$ are created:

$$\hat{\mathbf{X}}(t) = [x(t), x(t-\tau), \ldots, x(t-(m-1)\tau)]^T, \tag{9}$$

where the delay time τ and the embedding dimension m are parameters of the embedding procedure. (The superscript T denotes the transpose.) Here $\hat{\mathbf{X}}(t)$ represents a more comprehensive description of the state of the system at time t than does $x(t)$ and can be thought of as a map of

$$\pi^{(m)}: \mathbf{R}^M \to \mathbf{R}^m \tag{10}$$

from the full state $\mathbf{X}(t)$ to the reconstructed state $\hat{\mathbf{X}}(t)$.

Takens[30] and Mañé[31] have provided often-cited proofs that this procedure does (almost always) reconstruct the original state space of a dynamical system, as long as the embedding dimension $m > 2D + 1$, where D is the fractal dimension of the attractor. That is, $\pi^{(m)}$, restricted to $\mathcal{A}$, is a smooth one-to-one map from the original attractor to the reconstructed attractor. It should be noted, however, that as long as $m > D$, the reconstructed set will almost always have the same dimension as the attractor.[21]

The delay-time embedding procedure is a powerful tool. If one believes that the brain (say) is a deterministic system, then it may be possible to study the brain by looking at the electrical output of a single neuron. This example is an ambitious one, but the point is that the delay-time embedding makes it possible for one to analyze the self-organizing behavior of a complex dynamical system without knowing the full state at any given time.

Although almost any delay time τ and embedding dimension $m > D$ will in principle work (with unlimited, infinitely precise data), it is nontrivial to choose the embedding parameters in an optimal way. For instance, if the product $(m - 1)\tau$ is too large, then the components $x(t)$ and $x[t + (m - 1)\tau]$ of the reconstructed vector $\hat{\mathbf{X}}$ will be effectively decorrelated, and this result will be an inflated estimate of dimension. By the same token, if $(m - 1)\tau$ is too small, then the components $x(t), \ldots, x[t + (m - 1)\tau]$ will all be nearly equal, and the reconstructed attractor will look like one long diagonal line. It is also inefficient to take τ too small, even if m is large, since $x(t) \approx x(t + \tau)$ means that successive components in the reconstructed vector become effectively redundant.

In general, one wants τ to be not too much less than, and $(m - 1)\tau$ not too much greater than, some characteristic decorrelation time. The (linear) autocorrelation time is one such characteristic, although Fraser and Swinney[32] have introduced a more sophisticated rule based on the mutual information time. A topological rule, based on maintaining nearest-neighborhood relations as embedding dimension is increased, has recently been suggested.[33]

Typically, having chosen τ, one performs the dimension analysis for increasing values of m and looks for a plateau in the plot of D versus m. Some authors choose to increase m and decrease τ in such a way as to preserve $(m - 1)\tau$ as a constant.[34,35]

As a preprocessing step, linear transforms of the time-delayed variables have been suggested. One class of such transforms is based on a principal-value decomposition of the autocorrelation function.[34,36,37] (Engineers will recognize this decomposition as a Karhunen–Loève expansion.) However, it has been argued that the information available in the (linear) autocorrelation function is not necessarily relevant to optimal processing of a time series that arise from

nonlinear systems.[38,39] A transformation called relevance weighting was suggested by Farmer and Sidorowich.[40] In their scheme,

$$\hat{\mathbf{X}}(t) = \{x(t), e^{-h\tau}x(t-\tau), \ldots, \exp[-h(m-1)\tau]x[t-(m-1)\tau]\}^T. \quad (11)$$

The idea behind this embedding procedure is that the state $\mathbf{X}(t)$ depends most significantly on measurements that are taken most recently in time.

DEFINITIONS OF DIMENSION: FORMAL AND INFORMAL

A completely rigorous definition of dimension was given in 1919 by Hausdorff,[41] but it is a definition that does not lend itself to numerical estimation. Much interest in the past decade has focused on numerical estimates of dimension, and it is natural to consider more operational definitions of dimension, i.e., those that can be more readily translated into algorithms for estimating dimension from a finite sample of points.

This section will describe various definitions of fractal dimension, with some discussion of how they are related to one another and to Hausdorff's original definition and how they can be extended to generalized dimension. An algorithm based on box counting will also be presented as a way to illustrate some of the notions. The next section will then take a more thorough look at the various algorithms that have been suggested for the practical estimation of fractal dimension.

In what follows, three different ways of thinking about dimension will be pursued. First, dimension will be defined intuitively as a scaling of bulk with size. Then, a more formal definition, which involves coarse-grained volumes, will be given. Finally, the notion of counting degrees of freedom will be related to the information dimension. Each approach yields equivalent definitions for dimension and for generalized dimension, but each attempts to provide a different intuition for understanding what the dimension means.

Local Scaling Comparison of Bulk with Size

A geometrically intuitive notion of dimension is as an exponent that expresses the scaling of an object's bulk with its size:

$$\text{bulk} \sim \text{size}^{\text{dimension}}. \quad (12)$$

Here bulk may correspond to a volume, a mass, or even a measure of information content, and size is a linear distance. For example, the area (bulk) of a plane figure scales quadratically with its diameter (size), and so it is two dimensional. The definition of dimension is usually cast as an equation of the form

$$\text{dimension} = \lim_{\text{size}\to 0} \frac{\log \text{bulk}}{\log \text{size}}, \quad (13)$$

where the limit of small size is taken to ensure invariance over smooth coordinate changes. This small-size limit also implies that dimension is a local quantity and that any global definition of fractal dimension will require some kind of averaging.

The obvious relevant measure of bulk for a subset $\mathcal{B}$ of a dynamical attractor is its natural invariant measure $\mu(\mathcal{B})$, although other notions of bulk are also possible and will be discussed. A good quantity for the size of a set is its radius or its diameter, the latter of which is defined by

$$\delta(\mathcal{B}) \equiv \sup\{\|\mathbf{X}-\mathbf{Y}\|:\mathbf{X},\mathbf{Y}\in\mathcal{B}\}, \quad (14)$$

where sup is the supremum, or maximum, and $\|\mathbf{X}-\mathbf{Y}\|$ is the distance between $\mathbf{X}$ and $\mathbf{Y}$. How this distance is calculated depends on the norm of the embedding space. If X_i is the ith component of the vector $\mathbf{X}\in\mathbf{R}^m$, then the L_s norm gives distance according to

$$\|\mathbf{X}-\mathbf{Y}\| \equiv \left(\sum_{i=1}^{m} |X_i - Y_i|^s\right)^{1/s}, \quad (15)$$

where $|\cdot|$ is the absolute value. The most useful of these norms are L_2, the Euclidean norm, which gives distances that are rotation invariant; L_1, the taxicab norm, which is easy to compute; and L_∞, the maximum norm, which is also easy to compute. It is not difficult to show that fractal dimension is invariant to choice of norm.

Pointwise Dimension

The pointwise dimension is a local measure of the dimension of the fractal set at a point on the attractor. Let $\mathcal{B}_X(r)$ denote the ball of radius r centered at the point $\mathbf{X}$. (Whether this ball is a hypersphere or a hypercube will depend on the norm L_s.) Define the pointwise mass function $B_X(r)$ as the measure

$$B_X(r) = \mu[\mathcal{B}_X(r)]. \quad (16)$$

The scaling of the mass function at $\mathbf{X}$ with the radius r defines the pointwise dimension at $\mathbf{X}$ (Ref. 42):

$$D_p(\mathbf{X}) = \lim_{r\to 0} \frac{\log B_X(r)}{\log r}. \quad (17)$$

The pointwise dimension is a local quantity, but one can define the average pointwise dimension to be a global quantity:

$$D_p = \int_{\mathcal{A}} D_p(\mathbf{X})\,d\mu(\mathbf{X}). \quad (18)$$

Coarse-Grained Volume and Hausdorff Dimension

Although Hausdorff's rigorous definition of dimension does not immediately lead to numerical estimates, it does provide a framework from which dimension-estimation algorithms can be derived.

The Hausdorff dimension is purely a description of the geometry of the fractal set and makes no reference to the *a priori* measure μ that may be defined on the attractor. In fact, the Hausdorff definition begins by defining its own measure Γ, which corresponds to uniform density over the fractal set.

Suppose that $\mathcal{A}$ is the fractal whose dimension one wishes to calculate. Let $\mathcal{C}(r, \mathcal{A}) = \{\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_k\}$ be a finite covering of $\mathcal{A}$ into sets whose diameters are less than r. That is, $\mathcal{A} \subset \bigcup_{i=1}^{k}\mathcal{B}_i$, and the diameter of each set satisfies $\delta_i \equiv \delta(\mathcal{B}_i) < r$. Then the function

$$\Gamma(\mathcal{A}, D, r) = \inf_{\mathcal{C}(r,\mathcal{A})} \sum_i \delta_i{}^D, \tag{19}$$

where inf is the infimum (or minimum) over all coverings satisfying $\delta_i < r$, defines a kind of coarse-grained measure for the set $\mathcal{A}$. For example, if $D = 1$, then $\Gamma(\mathcal{A}, D, r)$ gives the length of the set $\mathcal{A}$ as measured with a ruler of length r, and, as $r \to 0$, Γ approaches the actual length of $\mathcal{A}$. For most values of D, the $r \to 0$ limit leads to a degenerate measure: either $\Gamma \to 0$ or $\Gamma \to \infty$. That is not surprising: A figure with finite length will have zero area, and a finite area is covered by a curve of infinite length. Something that is similar to a coastline will have infinite length and zero area. One can think of $\Gamma(\mathcal{A}, D)$ as the D-dimensional volume of the set $\mathcal{A}$. In fact, since $\Gamma(\mathcal{A}, D, r)$ is a function that decreases monotonically with D, there is a unique transition point D_H that defines the Hausdorff dimension:

$$\Gamma(\mathcal{A}, D) = \limsup_{r\to 0} \Gamma(\mathcal{A}, D, r) = \begin{cases} \infty & \text{for } D < D_H \\ 0 & \text{for } D > D_H \end{cases}, \tag{20}$$

so that $D_H = \inf\{D:\Gamma(\mathcal{A}, D) = 0\}$ defines the Hausdorff dimension.

For instance, the fractal coastline may have a one-dimensional volume (length) of infinity and a two-dimensional volume (area) of zero, but there is a D between 1 and 2 at which the D volume crosses over from ∞ to 0, and that value of D is the Hausdorff dimension of the coastline.

The coarse-grained measure defined in Eq. (19) typically exhibits the scaling $\Gamma(\mathcal{A}, D, r) \sim r^{D-D_H}$, which gives another way to estimate dimension. For instance, taking $D = 1$, one can estimate $\Gamma(\mathcal{A}, D, r)$ for a coastline, say, by measuring its length with rulers of ever-decreasing length r. The measured length will be ever increasing at a rate $L(r) \sim r^{1-D_H}$, which gives D_H.

Having defined the Hausdorff dimension D_H, we give the Hausdorff measure on the set $\mathcal{A}$ by $\Gamma(\mathcal{A}, D_H)$. The Hausdorff measure of a subset $\mathcal{B}$ of $\mathcal{A}$ is given by $\Gamma(\mathcal{B}, D_H) = \lim \sup_{r\to 0} \inf_{\mathcal{C}}(r, \mathcal{B}) \sum_i \delta_i{}^l$, and this measure is one way to describe the bulk of the set $\mathcal{B}$. This equation suggests that $\Gamma(\mathcal{B}, D_H) \sim \delta(\mathcal{B})^{D_H}$, which echoes the form of Eq. (12) and more prominently displays the role of the Hausdorff dimension as a local scaling exponent.

Definition of Box-Counting Dimension

The difficulty with implementing the Hausdorff dimension numerically is the infimum over all coverings that is taken in Eq. (19). If this requirement is relaxed and instead one chooses a covering that is simply a fixed-size grid, one obtains an upper bound on the Hausdorff dimension that has been variously referred to as the capacity, the box-counting dimension, and the fractal dimension. The last term, however, has come to be used in a generic sense for any dimension that may be nonintegral. For most fractal sets of interest, the capacity and the Hausdorff dimension are equal.[42]

With the grid size r, Eq. (19) becomes

$$\Gamma(\mathcal{A}, D, r) = \sum_i \delta_i{}^D = \sum_i r^D = n(r) r^D, \tag{21}$$

where $n(r)$ is the number of nonempty grid boxes. The box-counting dimension is the value of D on the transition between $\Gamma \to 0$ and $\Gamma \to \infty$. Here $\Gamma(\mathcal{A}, D_H, r) \sim 1$ implies that

$$n(r) \sim r^{-D_H} \tag{22}$$

or, more formally, that

$$D_H = \lim_{r\to 0} \frac{\log[1/n(r)]}{\log r}. \tag{23}$$

Here the local notion of bulk is replaced with a global one: $1/n(r)$ is the average bulk of each nonempty box, since each box contains, on average, $1/n(r)$ of the whole fractal.

Generalized Dimension

In computing the box-counting dimension, one either counts or does not count a box according to whether there are some points or no points in the box. No provision is made for weighting the box count according to how many points are inside a box. In other words, the geometrical structure of the fractal set is analyzed but the underlying measure is ignored.

The generalized dimension that was introduced in Ref. 43, and independently in Ref. 44, does take into account the number of points in the box. Let $\mathcal{B}_i$ denote the ith box, and let $P_i = \mu(\mathcal{B}_i)/\mu(\mathcal{A})$ be the normalized measure of this box. Equivalently, it is the probability for a randomly chosen point on the attractor to be in $\mathcal{B}_i$, and it is usually estimated by counting the number of points that are in the ith box and dividing by the total number of points.

The generalized dimension is defined by

$$D_q = \frac{1}{q-1} \lim_{r\to 0} \frac{\log \sum_i P_i{}^q}{\log r}. \tag{24}$$

Writing the sum of $P_i{}^q$ as a weighted average $\sum_i P_i{}^q = \sum_i P_i(P_i{}^{q-1}) = \langle P_i{}^{q-1}\rangle$, one can associate bulk with the generalized average probability per box $(\langle P_i{}^{q-1}\rangle)^{1/(q-1)}$ and identify D_q as a scaling of bulk with size. For $q = 2$ the generalized average is the ordinary arithmetic average, and for $q = 3$ it is a root mean square. It is not hard to show that the limit $q \to 1$ leads to a geometric average. Finally, it is noted that $q = 0$ corresponds to the plain box-counting dimension defined above.

For a uniform fractal, with all P_i equal, one obtains a generalized dimension D_q that does not vary with q. For a nonuniform fractal, however, the variation of D_q with q quantifies the nonuniformity. For instance,

$$D_\infty = \lim_{r\to 0} \frac{\log\left(\max_i P_i\right)}{\log r}, \tag{25}$$

$$D_{-\infty} = \lim_{r\to 0} \frac{\log\left(\min_i P_i\right)}{\log r}. \tag{26}$$

It is clear from Eq. (24) that D_q decreases with increasing q. From this fact and the above equations, it is clear that the maximum dimension $D_{-\infty}$ is associated with the least-dense points on the fractal and the minimum dimension D_∞ corresponds to the most-dense points. This should not be surprising: The densest set possible is a point, which has dimension zero.

The notion of generalized dimension first arose out of a need to understand why various algorithms gave different

answers for dimension. A further motivation came from the need to characterize more fully fractals with nonuniform measure. These sets are sometimes called multifractals and are characterized by an *a priori* measure μ that differs from the Hausdorff measure Γ. An extensive review of multifractals in a variety of physical contexts (this includes solid fractal objects as well as strange attractors) can be found in Ref. 45. The point is that, rather than measure just one dimension, one can compute the full spectrum of generalized dimensions from $D_{-\infty}$ to D_{∞}.

The formalism of coarse-grained measure introduced in Hausdorff's definition of dimension can be generalized. Instead of Eq. (19), one can write

$$\Gamma_q(\mathcal{A}, D, r) = \sum_i \mu_i{}^q \delta_i{}^{(1-q)D}, \tag{27}$$

where $\mu_i = \mu(\mathcal{B}_i)$ is the *a priori* measure of the ith element of the covering. Properly the sum in the above equation is preceeded by either an inf or a sup over all coverings, according to whether q is less than or greater than one.

As before, the $r \to 0$ limit gives $\Gamma_q(\mathcal{A}, D) \to 0$ for $(1-q)D < (1-q)D_q$ and $\Gamma_q(\mathcal{A}, D) \to \infty$ for $(1-q)D > (1-q)D_q$. And the transition between $\Gamma \to 0$ and $\Gamma \to \infty$ defines the generalized dimension:

$$D_q = \frac{1}{1-q} \inf\{(1-q)D : \lim_{r\to 0} \Gamma_q(\mathcal{A}, D, r) = 0\}. \tag{28}$$

The advantage of Eqs. (27) and (28) is that they provide a definition of generalized dimension without requiring fixed-size boxes.

Spectrum of Scaling Indices: $f(\alpha)$

Halsey *et al.*[46] introduced a change of variables that provides a new interpretation of generalized dimension. Let $\tau = (q-1)D_q$ and take a Legendre transformation from the variables (q, τ) into a new set of variables (α, f):

$$\alpha = \frac{\partial \tau}{\partial q} \qquad f = \alpha q - \tau \tag{29}$$

and

$$q = \frac{\partial f}{\partial \alpha} \qquad \tau = \alpha q - f. \tag{30}$$

In case either $\tau(q)$ or $f(\alpha)$ is not differentiable, a more robust formulation is given by

$$f(\alpha) = \min_q\{q\alpha - \tau(q)\}, \tag{31}$$

$$\tau(q) = \min_\alpha\{q\alpha - f(\alpha)\}. \tag{32}$$

The authors of Ref. 46 interpret $f(\alpha)$ as a "spectrum of scaling indices." The variable α is the scaling index: $\mu_i = \delta_i{}^{\alpha_i}$ defines the local scaling at the ith element of the cover. Then, the number $n(\alpha, r)$ of cover elements with scaling index between α and $\alpha + \Delta\alpha$ scales as $n(\alpha, r) \sim r^{-f(\alpha)}\Delta\alpha$.

To see that this interpretation leads to Eqs. (29)–(32), consider the fixed-box-size sum $\sum_i P_i{}^q$. The number of terms in this sum for which $P_i = r^\alpha$ is given by $n(\alpha, r)$. Thus

$$\sum_i P_i{}^q = \int n(\alpha, r) r^{q\alpha} d\alpha \tag{33}$$

$$\sim \int r^{-f(\alpha)} r^{q\alpha} d\alpha \sim r^\theta, \tag{34}$$

where $\theta = \min_\alpha\{q\alpha - f(\alpha)\}$ since the integral will be dominated by the smallest exponent of r when $r \to 0$. Comparing this to the definition of generalized dimension in Eq. (24), which gives $\sum_i P_i{}^q \sim r^{(q-1)D_q}$, one obtains $(q-1)D_q = \min_\alpha\{q\alpha - f(\alpha)\}$, which is the formula given in Eq. (32). From this formula, Eqs. (29)–(31) follow. See Fig. 3.

This scaling $n(\alpha, r) \sim r^{-f(\alpha)}$ suggests interpreting $f(\alpha)$ as the Hausdorff dimension of the points with scaling index α [compare with formula (22)]. Indeed, the scaling index α is associated with the pointwise dimension D_p, and the set

$$S_\alpha = \{X \in \mathcal{A} : D_p(X) = \alpha\} \tag{35}$$

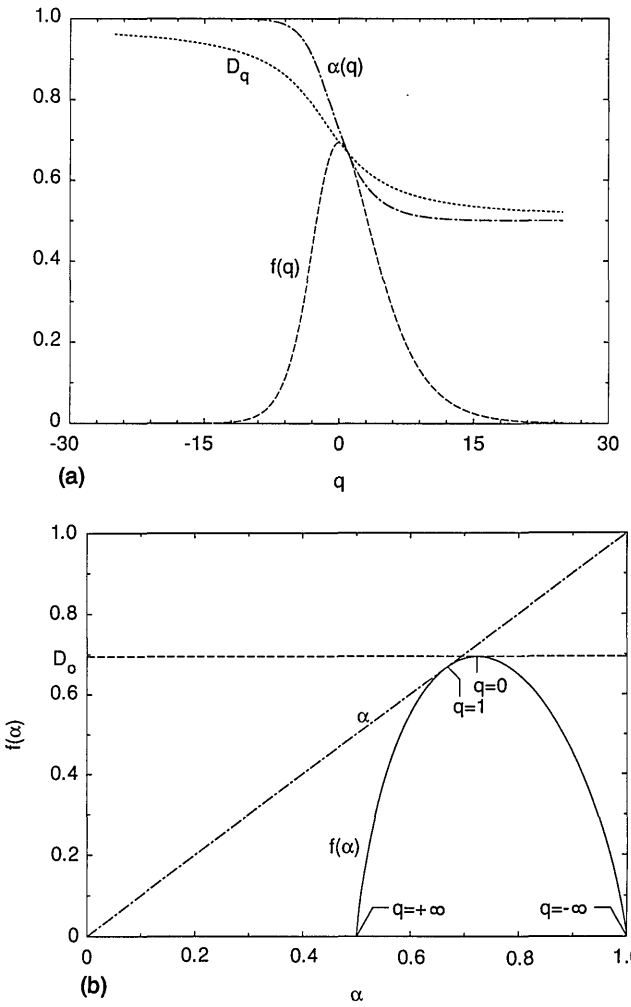


Fig. 3. Generalized dimension. (a) D_q as a function of q for a typical multifractal. Also shown are both f and α as a function of q for the same multifractal. (b) f as a function of α. The curve is always convex upward, and the peak of the curve occurs at $q = 0$. At this point f is equal to the fractal dimension D_0. Also, the $f(\alpha)$ curve is tangent to the curve $f = \alpha$, and the point of tangency occurs at $q = 1$. In general, the left-hand branch corresponds to $q > 0$ and the right-hand branch to $q < 0$.

is the set of all points in $\mathcal{A}$ for which the pointwise dimension is α. The Hausdorff dimension of the set S_α is given by $f(\alpha)$.

The reader who is hoping for a dramatic picture of this fractal set S_α will probably be disappointed. As Sakar[47] and others have pointed out, S_α is not necessarily a closed set and may even be dense in the original fractal $\mathcal{A}$. Thus, although the set S_α may have a lower Hausdorff dimension than $\mathcal{A}$, it is possible for the box-counting dimension to be the same. And a picture of S_α would look just like the picture of the original set $\mathcal{A}$.

The $f(\alpha)$ formalism provides a tool for testing the notion of universality, which states that a wide variety of dynamical systems should behave in a similar way and should leave the same characteristic signatures. Indeed, several researchers[48–50] have found physical systems whose $f(\alpha)$ curves precisely matched the $f(\alpha)$ associated with a theoretical model of a circle map undergoing a transition from quasi-periodicity to chaos.

Discussions of $f(\alpha)$ from the viewpoint of a thermodynamic formalism can be found in Refs. 51–54. Here the authors attempt to work backward by constructing a dynamical system that exhibits the desired $f(\alpha)$ structure.

Information Dimension

As an alternative to the scaling of mass with size, one can also think of the dimension of a set in terms of how many real numbers are needed to specify a point on that set. For instance, the position of a point on a line can be labeled by a single real number, the position on a plane by two Cartesian coordinates, and the position in ordinary (three-dimensional) space by three coordinates. Here, dimension is something that counts the number of degrees of freedom. For sets more complicated than lines, surfaces, and volumes, however, this informal definition of dimension needs to be extended.

One way to extend this definition is to determine not how many real numbers but how many bits of information are needed to specify a point to a given accuracy. On a line segment of unit length, k bits are needed to specify the position of a point to within $r = 2^{-k}$. For a unit square, $2k$ bits are needed to achieve the same accuracy (k bits for each of the two coordinates specified). And similarly, $3k$ bits are needed for a three-dimensional cube. In general, $S(r) = -d \log_2(r)$ bits of information are needed to specify the position of a unit d-dimensional hypercube to an accuracy of r. This example leads to a natural definition for the information dimension of a set; it is given by the small r limit of $-S(r)/\log_2(r)$, where $S(r)$ is the information (in bits) needed to specify a point on the set to an accuracy r. If $S(r)$ is the entropy, then $2^{S(r)}$ is the total number of available states, and $2^{-S(r)}$ can then be interpreted as the average bulk of each state. This interpretation permits one to express the information dimension as a scaling of bulk with size.

Consider partitioning the fractal into boxes $\mathcal{B}_i$ of size r. To specify the position of a point to an accuracy r requires that one specify in which box the point is. The average information needed to specify one box is given by Shannon's formula:

$$S(r) = -\sum_i P_i \log_2 P_i, \tag{36}$$

where P_i is the probability measure of the ith box: $P_i = \mu(\mathcal{B}_i)/\mu(\mathcal{A})$. This relation leads directly to an expression for the information dimension of the attractor:

$$D_I = \lim_{r\to 0} \frac{-S(r)}{\log_2 r} \tag{37}$$

$$= \lim_{r\to 0} \frac{\sum_i P_i \log_2 P_i}{\log_2 r}. \tag{38}$$

Renyi[55] has defined a generalized information measure:

$$S_q(r) = \frac{1}{q-1} \log \sum_i P_i^q, \tag{39}$$

which reduces to Shannon's formula in the limit $q \to 1$. The generalized information dimension associated with the Renyi entropy is just the generalized dimension that has been defined above by another approach. Thus

$$D_q = \lim_{r\to 0} \frac{-S_q(r)}{\log r} = \frac{1}{q-1} \lim_{r\to 0} \frac{\log \sum_i P_i^q}{\log r}, \tag{40}$$

which is the same as Eq. (24).

ALGORITHMS FOR ESTIMATING DIMENSION

It is emphasized that numerical techniques can only estimate the dimension of a fractal. Practical estimation of dimension begins with a finite description of the fractal object. This description may be a digitized photograph with finite resolution, an aggregation with a finite number of aggregrates, or a finite sample of points from the trajectory of a dynamical system. In any case, what is sought is the dimension not of the finite description but of the underlying set.

In the previous section an algorithm based on box counting was introduced. However, the box-counting algorithm has a number of practical limitations,[56] particularly at a high embedding dimension, and so a variety of other algorithms have also been developed.

The most popular way to compute dimension is to use the correlation algorithm, which estimates dimension based on the statistics of pairwise distances. The box-counting algorithm and the correlation algorithm are both in the class of fixed-size algorithms because they are based on the scaling of mass with size for fixed-sized balls (or grids). An alternative approach uses fixed-mass balls, usually by looking at the statistics of distances to kth nearest neighbors. Both fixed-size and fixed-mass algorithms can be applied to estimation of generalized dimension D_q, although fixed-size algorithms do not work well for $q < 1$.[57]

Also discussed are methods that directly involve the dynamical properties of the strange attractor. The Kaplan–Yorke conjecture, for example, relates dimension to the Lyapunov exponents. Recently interest has focused on trying to determine the unstable periodic orbits (these compose the "Cheshire set"[58]) of the attractor. The most direct use of the dynamics is to make predictions of the future of the time

series. Successful predictions provide a reliable indication that the dynamics is deterministic and low dimensional.

Finally the notion of intrinsic dimension is introduced. This is an integer dimension that provides an upper bound on the fractal dimension of the attractor by looking for the lowest-dimensional manifold that can (at least locally) confine the data.

A number of these algorithms can be used to compute generalized dimension, but from the point of view of practical estimation it bears remarking that this application is not always useful. A generalized dimension is useful for quantifying the nonuniformity of the fractal or, in general, for characterizing its multifractal properties. And this use is important if one wants to compare an exact and predictive theory with an experimental result. On the other hand, the goal of dimension estimation is often more qualitative in nature. One wants to know only whether the number of degrees of freedom is large or reasonably small. To answer the question "Is it chaos or is it noise?" a robust estimate of dimension is more important than a precise estimate. In these cases the subtle distinction between information dimension at $q = 1$ and correlation dimension at $q = 2$, say, may not be so important as the more basic issues that arise from experimental noise, finite samples, or even computational efficiency.

Average Pointwise Mass Algorithms

Recall the definition of the pointwise mass function in Eq. (16): $B_X(r) = \mu[\mathcal{B}_X(r)]$, where $\mathcal{B}_X(r)$ is the ball of radius r about the point $\mathbf{X}$.

The scaling of $B_X(r)$ with r is what defines the pointwise dimension at $\mathbf{X}$. If this scaling is the same for all $\mathbf{X}$, then the fractal is uniform, and it follows that D_q is a constant independent of q and has the value of the pointwise dimension. For most fractals, however, pointwise dimension is not a global quantity, and some averaging is necessary for the quantity to be made global.

From the point of view of practical computation, it is possible to compute the pointwise dimension at a sample of points on the attractor and then to average these values to obtain an estimate of the dimension of the attractor. This average is one that in principle gives the information dimension D_1. This approach has been advocated, for example, in Ref. 59.

However, it is probably more efficient to estimate dimension by averaging the mass function before the limit in Eq. (17) is taken. It is in this way that the statistics of the interpoint distances can be most effectively exploited.

Correlation Dimension

The most natural such averaging strategy was introduced by Grassberger and Procaccia[60,61] and independently by Takens.[62] Here, a direct arithmetic average of the pointwise mass function gives what Grassberger and Procaccia call the correlation integral:

$$C(r) = \langle B_X(r) \rangle. \tag{41}$$

From this, the correlation dimension ν is defined:

$$\nu = \lim_{r \to 0} \frac{\log C(r)}{\log r}. \tag{42}$$

Here, $B_{X_j}(r)$ can be approximated from a finite data set of size N by

$$B_{X_j}(r) \approx \frac{\#\{X_i : i \neq j \text{ and } \|X_i - X_j\| \leq r\}}{N-1} = \frac{1}{N-1} \sum_{\substack{i=1 \\ i \neq j}}^{N} \Theta(r - \|X_i - X_j\|), \tag{43}$$

where Θ is the Heaviside step function: $\Theta(x)$ is zero for $x < 0$ and one for $x \geq 0$. The importance of excluding $i = j$ has been overlooked by some authors, although Grassberger has stressed this point.[63] In fact, the case is made in Ref. 64 for excluding all values of i for which $|i - j| < W$ with $W > 1$. It is now straightforward to approximate $C(r)$ with a finite data set:

$$C(N, r) = \frac{1}{N} \sum_{j=1}^{N} B_{X_j}(r) = \frac{1}{N(N-1)} \sum_{i \neq j} \Theta(r - \|X_i - X_j\|). \tag{44}$$

In words,

$$C(N, r) = \frac{\#\text{ of distances less than } r}{\#\text{ of distances altogether}}. \tag{45}$$

Thus the correlation algorithm provides an estimate of dimension based purely on the statistics of pairwise distances. Not only is this a particularly elegant formulation but it has the substantial advantage that the function $C(r)$ is approximated even for r as small as the minimum interpoint distance. For N points, $C(N, r)$ has a dynamic range of $O(N^2)$. Logarithmically speaking, this range is twice that available to $\hat{n}(N, r)$ in the box-counting method. It is also twice the range available in an estimate of the pointwise mass function $B_X(r)$ for a single point $\mathbf{X}$. This greater range is the one advantage that the correlation integral has over the average pointwise dimension.

Generalized Dimension from Averaged Pointwise Mass

A more general average than the direct arithmetic average used above is given by

$$G_q(r) = [\langle B_X(r)^{q-1} \rangle]^{1/(q-1)}. \tag{46}$$

The scaling of this average with r gives the generalized dimension: $G_q(r) \sim r^{D_q}$ or

$$D_q = \lim_{r \to 0} \frac{\log G_q(r)}{\log r}. \tag{47}$$

Note that $q = 2$ gives the direct arithmetic average that defines the correlation dimension and that the $q = 1$ average (which is associated with the information dimension),

$$G_1(r) = \lim_{q \to 1} G_q(r) = \exp\langle \log B_X(r) \rangle, \tag{48}$$

is a geometric average of the pointwise mass function.

From a finite set of points, $G_q(N, r)$ can be approximated by[65]

$$G_q(N,r) \approx \left\{ \frac{1}{N} \sum_{i=1}^{N} \left[\frac{1}{N-1} \sum_{\substack{j=1 \\ j \neq i}}^{N} \Theta(r - \|\mathbf{X}_i - \mathbf{X}_j\|) \right]^{q-1} \right\}^{1/(q-1)}. \quad (49)$$

This formula is easiest to evaluate when $q = 2$, in which case it reduces to the simple form given in Eq. (44).

Algorithms based on this equation for the general case $q \neq 2$ have been pursued in Refs. 35, 65, and 66. This approach does not work well for $q \leq 1$ since the term that is raised to the power of $q - 1$ can be zero for r much larger than the smallest interpoint distance. The formula ought to work well for large positive values of q, however.

q-Point Correlation

Grassberger[44] suggested a q-point correlation integral defined by counting q-tuples of points that have the property that every pair of points in the q-tuple is separated by a distance of less than r:

$$C_q(N,r) = \frac{1}{N^q} \#\{(\mathbf{X}_{i_1}, \ldots, \mathbf{X}_{i_q}): \|\mathbf{X}_n - \mathbf{X}_m\| \leq r \text{ for all } n, m \in \{i_1, \ldots, i_q\}\}, \quad (50)$$

which has the scaling behavior $C_q(r) \sim r^{(q-1)D_q}$. Thus

$$D_q = \frac{1}{q-1} \lim_{r \to 0} \lim_{N \to \infty} \frac{\log C_q(N,r)}{\log r}. \quad (51)$$

This equation can in principle be applied to all integers $q \geq 2$, although its implementation is awkward for $q \geq 3$, since the number of q-tuples N^q grows so rapidly with N.

k-Nearest-Neighbor (Fixed-Mass) Algorithms

In contrast to the average pointwise mass functions described above, in which the variation of mass inside fixed-size balls is considered, the nearest-neighbor algorithms consider the scaling of sizes in fixed-mass balls.

An early implementation of this notion in a chaotic-dynamics context was suggested in Ref. 67. (Guckenheimer and Buzyna[68] used a method like this one on experimental data.) Here one computes $\langle r_k \rangle$, which is the average distance to the kth nearest neighbor, as a function of k. Let $R(\mathbf{X}_i, k)$ denote the distance between point $\mathbf{X}_i$ and its kth nearest neighbor. [By convention, the 0th nearest neighbor is the point itself, so $R(\mathbf{X}, 0) = 0$ for all $\mathbf{X}$.] The average is computed:

$$\langle r_k \rangle = \frac{1}{N} \sum_{i=1}^{N} R(\mathbf{X}_i, k). \quad (52)$$

The scaling $\langle r_k \rangle \sim k^{1/D}$ defines the dimension. (Actually, what was computed in Ref. 67 was $\langle r_k^2 \rangle$, since for Euclidean distances this equation is computationally more efficient. And the scaling was taken to be $\langle r_k^2 \rangle \sim k^{2/D}$.) The authors conjectured that they had estimated the Hausdorff dimension, although this statement was corrected by Badii and Politi[69] and by Grassberger.[70]

Badii and Politi[69] consider moments of the average distance to the kth nearest neighbor and recommend keeping k fixed and computing a dimension function from the scaling of average moments of r_k with the total number of points N:

$$\langle r_k^{\gamma} \rangle \sim (k/N)^{\gamma/D(\gamma)}. \quad (53)$$

This dimension function $D(\gamma)$ is related to the generalized dimension by the implicit formulas

$$\gamma = (q-1)D_q, \qquad D(\gamma) = D_q. \quad (54)$$

Grassberger[70] derived Eqs. (54) independently and further provided a small k correction to the scaling of $\langle r_k^{\gamma} \rangle$ with k:

$$\langle r_k^{\gamma} \rangle \sim \left[\frac{\Gamma(k + \gamma/D)}{\Gamma(k) k^{\gamma/D}} \right] (k/N)^{\gamma/D}. \quad (55)$$

Further pursuit of the dimension function can be found in Ref. 71. Further corrections and discussions of efficient implementation are provided in Ref. 72; see also Ref. 73. Applications of nearest-neighbor methods for estimating intrinsic dimension can be traced back to Ref. 74 (see the cross-disciplinary review in Ref. 75), but these do not give a fractal dimension and were not applied to chaotic systems.

Algorithms That Use Dynamical Information

Recurrence Time

Let $T_{X_0}(r)$ be the r-recurrence time of an initial condition $\mathbf{X}_0$. This is the minimum amount of time necessary for the trajectory of $\mathbf{X}_0$ to come back to within r of $\mathbf{X}_0$. The inverse recurrence time provides an estimate of the pointwise mass function, and one expects

$$\frac{1}{T_{X_0}(r)} \sim B_{X_0}(r). \quad (56)$$

Thus the scaling of the average recurrence time with r can be related to the generalized dimension[44,48]:

$$[\langle T_X(r)^{1-q} \rangle]^{1/(1-q)} \sim [\langle B_X(r)^{q-1} \rangle]^{1/(1-q)} \sim r^{D_q}. \quad (57)$$

Lyapunov Dimension

The Kaplan–Yorke conjecture[76,77] relates the fractal dimension of a strange attractor to the Lyapunov exponents. The conjectured formula defines what has come to be called the Lyapunov dimension:

$$d_L = j + \frac{\sum_{i=1}^{j} \lambda_i}{|\lambda_{j+1}|}, \quad (58)$$

where $\{\lambda_1, \ldots, \lambda_m\}$ are the Lyapunov exponents in decreasing order and j is given by

$$j = \sup\left\{ k: \sum_{i=1}^{k} \lambda_i > 0 \right\}.$$

The conjecture is that the Lyapunov dimension corresponds to the information dimension D_q with $q = 1$.

Farmer[78] has used this method to achieve dimension estimates of as many as 20. Matsumoto[79] studies a model of optical turbulence that leads to an attractor with an estimated Lyapunov dimension of approximately 45. Such numbers are far too large to be estimated by conventional techniques. Badii and Politi[80] used Lyapunov scaling to obtain extremely precise measurements of generalized dimension for constant Jacobian maps. In each of these cases, however, the equations of motion themselves were required.

When the equations of motion are available, computation of Lyapunov exponents, and therefore of Lyapunov dimension, is much easier.

Unstable Periodic Orbits

It has been suggested[81] that the natural invariant measure μ (and from this, the fractal dimension) of a chaotic attractor can be systematically approximated by sets of unstable periodic orbits. This tool is a powerful analytical one and shows promise computationally, as well. See Refs. 82–85 for further discussion.

Prediction Methods

In discussing algorithms for predicting the future of time series, Farmer and Sidorowich[40] suggested several ways in which these prediction algorithms might be used to make more accurate and reliable dimension estimates. One promising possibility is to use bootstrapping: Here a short time series of length N is extrapolated far into the future, and the longer time series is used in a conventional dimension algorithm. This method is appropriate when the data set is small and computation time is not the limiting factor. A second possibility is to use the scaling of prediction error with N, the number of points from which the prediction is made. This scaling can be related to the dimension, and it has the desirable property that if it is observed, that is, if the prediction error is noticably reduced with increasing N, then the system can be reliably considered low dimensional.

Intrinsic Dimension

Unlike the algorithms discussed above, which seek to compute dimension as accurately as possible from the given data, the algorithms in this section seek only to bound the fractal dimension with an integer. The estimate is coarser but, it is hoped, more reliable than the previous estimates.

As Somorjai's review[75] points out, the notion of local intrinsic dimension was first discussed by Fukunaga and Olsen.[86] The approach was first applied to dynamical systems by Froehling *et al.*[87] Here the authors attempt to find D dimensional hyperplanes that confine the data in a local region.

A variety of similar approaches have been considered. Broomhead *et al.*[88] use principal-value decomposition to make the fit to the D dimensional hyperplane. An approach more closely related to the correlation dimension is described in Ref. 89. Another approach based on topological considerations is described in Ref. 33. An algorithm that averages the local intrinsic dimension is discussed in Ref. 90. Although intrinsic dimension is necessarily an integer quantity, the averaging in this algorithm gives a noninteger result that the authors conjecture to be related to the fractal dimension. Brock *et al.* have developed a statistic based on the delay-time embedding procedure and the correlation integral for distinguishing m-dimensional deterministic time series from uncorrelated noise.[91] An analog electronic estimation of intrinsic dimension is described in Ref. 92. An anecdotal comparison of an intrinsic dimension estimator and the correlation dimension estimator can be found in Ref. 93.

The prediction algorithms of Farmer and Sidorowich[40,94] can also be used to estimate intrinsic dimension. Basically, the smallest embedding dimension for which good predictions can be made is a reliable upper bound on the number of degrees of freedom in the time series.

IMPLEMENTATION OF BOX-COUNTING ALGORITHMS

Some of the first numerical estimates of fractal dimension were obtained with the box-counting method.[95] In this section some of the practical issues that arise in the implementation of box-counting algorithms are discussed. The next section will discuss some of the same issues for the more popular correlation dimension.

To compute the box-counting dimension, break up the embedding space $\mathbf{R}^m$ into a grid of boxes of size r. Count the number of boxes $n(r)$ inside which at least one point from the attractor lies. The scaling $n(r) \sim r^{-D_0}$ defines the box-counting dimension, which is formally defined by the limit in Eq. (23). There are some obstacles, however, in applying Eq. (23) directly to an experimentally observed fractal.

Finite Resolution

For a solid fractal object, only a finite resolution is available, so the limit $r \to 0$ cannot be taken. A natural and direct approximation is just to apply Eq. (23) directly but with the smallest r available. That is,

$$D_0 \approx \frac{\log[1/n(r)]}{\log r}. \tag{59}$$

The problem with this approximation is that it converges with logarithmic slowness in r. For instance, $n(r) = n_0 r^{-D_0}$ gives

$$\frac{\log[(1/n_0)r^{D_0}]}{\log r} = \frac{D_0 \log r + \log(1/n_0)}{\log r} = D_0 + \overbrace{\frac{\log(1/n_0)}{\log r}}^{\text{slow to vanish as } r \to 0} \approx D_0. \tag{60}$$

Instead, $\log n(r)$ versus $\log r$ is usually plotted. The negative slope of this curve will give D_0 for small r:

$$D_0 \approx \frac{-\Delta[\log n(r)]}{\Delta(\log r)}. \tag{61}$$

The criterion for choosing the best slope is not immediately obvious. It is clear that the greater the range (the larger the lever arm) the better, but this range is limited on one end by finite resolution and on the other end by the fractal's own finite size.

Finite N

For a dynamical attractor, although the points may be known to a high precision, only a finite sample of N points is available. The natural approximation in this case is to estimate $n(r)$ with $\hat{n}(N, r)$, the number of boxes inside which at least one of the sample points lies. Here, although $\hat{n}(N, r)$ clearly underestimates $n(r)$, it is expected to be a good approximation for large N since for fixed r:

$$n(r) = \lim_{N \to \infty} \hat{n}(N, r). \tag{62}$$

Returning to the formal definition in Eq. (23), we find that

$$D_0 = \lim_{r\to 0}\lim_{N\to\infty} \frac{\log[1/\hat{n}(N,r)]}{\log r}. \tag{63}$$

The order of the limits is crucial. For finite N, $\hat{n}(N, r)$ is bounded by N (there can be no more nonempty boxes than there are things to put in the boxes), and $r \to 0$ gives a dimension of zero. In a formal sense, this result is not surprising: A finite set of points does have dimension zero. But what is wanted is the dimension of the underlying attractor, not the dimension of the finite sample of points.

The finite sample is a serious limitation because, particularly for multifractal attractors that are nonuniformly populated with points, the limit in Eq. (62) converges slowly. Grassberger[96] has fitted the form

$$n(r) - \hat{n}(N,r) \sim r^{-\alpha}N^{-\beta} \tag{64}$$

and used the fit to accelerate the convergence. A similar approach appears in Ref. 97. Caswell and Yorke[98] point out that one way to increase the speed of convergence is to use spheres instead of boxes.

Application to Generalized Dimension

The box-counting algorithm is one of a class of algorithms based on fixed size and as such is not so well suited as the so-called fixed-mass algorithms for estimation of generalized dimension of D_q when $q < 1$. Since the box-counting dimension corresponds to $q = 0$, the box-counting algorithm is a poor way to estimate the box-counting dimension. The modification of the box-counting algorithm for generalized dimension will achieve better performance if $q > 1$.

Indeed, the current interest in multifractals and their $f(\alpha)$ characterization has led to a revival of the box-counting algorithm for estimation of generalized dimension. For example, Chhabra and Jensen[99] use box counting to compute $f(q)$ and $\alpha(q)$ directly for experimental data. Their approach avoids the problem of having to evaluate the derivative in Eqs. (29) numerically by analytically differentiating Eq. (24) for the generalized dimension.

Computational Efficiency

Depending on the details of the computer program that implements the box-counting algorithm, either one must keep track of a large number of empty boxes, which can be a severe memory burden (in the small r limit almost all the boxes will be empty), or else one must maintain a list of nonempty boxes. The latter case can be computationally expensive since, for each point, the program has to check whether the point belongs in any of the nonempty boxes on the list. An efficient compromise for low-dimensional embedding spaces is discussed in Ref. 96.

One can always compute $n(N, r)$ in $O(N \log N)$ time by first discretizing the input points to the nearest integer multiple of r (this effectively converts the point's position to the position of the appropriate box), then sorting the discretized list (which is the same as sorting the list of boxes—a lexicographic or dictionary ordering is useful here), and finally counting the number of unique elements in the sorted list (the number of boxes).

IMPLEMENTATION OF THE CORRELATION ALGORITHM

One reason for the popularity of the correlation algorithm is that it is easy to implement. One computes the correlation integral $C(N, r)$ merely by counting distances. Another and more fundamental advantage is that the correlation algorithm treats all distances equally, so one can probe to distance scales that are limited only by the smallest interpoint distance (which will be of the order of $N^{-2/D}$). Most other algorithms are limited by the average distance to the nearest neighbor (which is of the order of $N^{-1/D}$). That is, the correlation integral $C(N, r)$ has a dynamic range of $O(N^2)$; this compares, for instance, with the $O(N)$ range of the function $\hat{n}(N, r)$ that appears in the box-counting algorithm. Recall Eq. (45), which defines $C(N, r)$ as the fraction of distances less than r. Then, the correlation dimension is given by

$$\nu = \lim_{r\to 0}\lim_{N\to\infty} \frac{\log C(N,r)}{\log r}. \tag{65}$$

However, there are a variety of practical issues and potential pitfalls that come with making an estimate from finite data.

Extraction of the Slope

Extracting dimension directly from the correlation integral according to Eq. (65) is extremely inefficient, since the convergence to ν as $r \to 0$ is logarithmically slow. See relation (60) for the equivalent situation in the box-counting algorithm.

Taking a slope solves this problem. Either one can take the local slope:

$$\hat{\nu}(r) = \frac{\mathrm{d}\log C(N,r)}{\mathrm{d}\log r} = \frac{\mathrm{d}C(N,r)/\mathrm{d}r}{C(N,r)/r}, \tag{66}$$

or one can fit a chord through two points on the curve:

$$\hat{\nu}(r) = \frac{\Delta \log C(N,r)}{\Delta \log r} = \frac{\log C(N,r_2) - \log C(N,r_1)}{\log r_2 - \log r_1}. \tag{67}$$

However, to implement this strategy requires the choice of two length scales. The larger scale is limited by the size of the attractor, and the smaller scale is limited by the smallest interpoint distance. The expression usually converges in the limit as r_1 and r_2 both go to zero, and convergence is guaranteed as long as the ratio r_1/r_2 also goes to zero.

The direct difference in Eq. (67) fails to take full advantage of the information available in $C(N, r)$ for values of r between r_1 and r_2. It is natural to attempt to fit a slope by some kind of least-squares method, but this approach is problematic. An unweighted, least-squares method is particularly poor, since the estimate of $C(r)$ by $C(N, r)$ is usually much better for large r than for small r. Weighted fits can compensate for this effect, but there is still a problem because successive values of $C(N, r)$ are not independent. Since $C(N, r + \Delta r)$ is equal to $C(N, r)$ plus the fraction of the distances between r and $r + \Delta r$, it is a mistake to assume that

$C(N, r + \Delta r)$ is independent of $C(N, r)$. An estimate of dimension will depend on whether $C(N, r)$ is sampled at logarithmic intervals $r = (0.001, 0.01, 0.1,$ etc.) or at uniform intervals $r = (0.001, 0.002, 0.003,$ etc.). Furthermore, the error estimate that the least-squares fit naturally provides will have little to do with the actual error in the dimension estimate.

Takens[100] has shown how this intermediate distance information can be incorporated in a consistent way. Using the theory of best estimators, Takens derived an estimator for ν that makes optimal use of the information in the interpoint distances r_{ij}. In fact, the method uses all distances less than an upper bound r_0, so that a lower cutoff distance need not be specified:

$$\hat{\nu}(r_0) = \frac{-1}{\langle \log(r_{ij}/r_0) \rangle}, \tag{68}$$

where the angle brackets here refer to an average over all distances r_{ij} that are less than r_0. In terms of the correlation integral, this can be written[101] as

$$\hat{\nu}(r_0) = \frac{C(r_0)}{\int_0^{r_0} [C(r)/r] dr} \tag{69}$$

$$= \frac{\int_0^{r_0} [dC(r)/dr] dr}{\int_0^{r_0} [C(r)/r] dr}. \tag{70}$$

The slightly more unwieldy form that is given in Eq. (70) is meant to be compared with Eq. (66) for the local slope.

The Takens estimator not only provides an efficient means to squeeze an optimal estimate out of the correlation integral but also provides an estimate of the statistical error, namely, $\sigma_\nu = \nu/\sqrt{N^*}$, where N^* is the number of distances less than r_0. However this error estimate makes the assumption that all those N^* distances are statistically independent. It is argued in Ref. 102 that this assumption is valid only if $N^* \leq N$ because there cannot be N^2 statistically independent quantities derived from only N independent points. This argument implies, for instance, that even for small dimensions one cannot expect to get a precision of better than 1% with fewer than 10,000 points. (Accuracy is an entirely separate issue. The variety of sources of systematic error is vast. Some of these errors will be discussed in later sections.)

In Ref. 103 the Takens estimator is compared with the local slope for a typical strange attractor. Ellner[104] extends the Takens estimator to a generalized dimension D_q, though in a way that works only for integer $q > 2$ and efficiently only for $q = 2$. He also introduces a so-called local estimator based on the same principle. In Ref. 105 it is pointed out that even though the Takens estimator uses all distances less than r_0, and therefore has in effect an arbitrarily long lever arm, it is still sensitive to oscillations in the correlation integral.

Computation

Although there are $O(N^2)$ interpoint distances to be considered in a time series of N points, most of these distances are large: of the same order as the size of the attractor. Since it is the short distances that are most important, and since there are so few of them, it would be advantageous to organize the points on the attractor so that only the short distances are computed. The algorithm in Ref. 106 provides a way to compute the $O(N)$ shortest distances with only $O(N \log N)$ effort by placing the points into boxes of size r_0. An alternative approach is to use multidimensional trees,[107] such as were introduced by Farmer and Sidorowich[40] for nonlinear prediction. Such an organization of points would also be useful for the k nearest-neighbor algorithms that were discussed above. Hunt and Sullivan[108] have also argued for the importance of efficient data organization.

Franaszek[109] has suggested an algorithm that computes the correlation integral for a range of embedding dimensions simultaneously. A massively parallel analog optical computation of the correlation integral was performed by Lee and Moon.[110]

Comparison with Filtered Noise

It is a common and well-advised tactic to compare the correlation dimension computed for a given time series against a set of test data whose properties are known in advance. For example, if the given time series has a significantly larger $C(N, r)$ at small r than does white noise, then one can rule out the null hypothesis that the original time series is white noise. However, one cannot rule out the possibility that the original time series is simply autocorrelated or colored noise. A more stringent test is to create a time series with the same Fourier spectra as the original time series (for instance, one could take a Fourier transform, randomize the phases, and then invert the transform). If the $C(N, r)$ obtained from this time series is significantly different from that of the original time series, then a stronger statement can be made: Not only is the original time series not white noise, it could not have been produced by any linear filter of white noise. A simple version of this more stringent test was applied by Grassberger[111] in rejecting earlier claims of low-dimensional chaos in a climatic time series.

Osborne and Provenzale[112] recently made the striking observation that stochastic systems with $1/f^\alpha$ power-law spectra can exhibit a finite correlation dimension when $\alpha > 1$. This observation is not an artifact of finite N approximation but is a fundamental property of power-law correlation. This observation makes it all the more imperative that stochastic analogs of the original time series be created and tested.

Sources of Error

There are two fundamental sources of error in the estimation of dimension from a time series: statistical imprecision and systematic bias. The first arises directly from the finite sampling of the data (and as such is reasonably tractable); the second comes from a wide variety of sources. Indeed, it is this wide variety that has led to so much difficulty in the field of verifying results as reliable.

Computing fractal dimension is a tricky business. In the words of Brandstater and Swinney,[113]

> It is not difficult to develop an algorithm that will yield numbers that can be called dimension, but it

> is far more difficult to be confident that those numbers truly represent the dynamics of the system.

Here is an incomplete survey of the kinds of problems that may arise, usually in the computation of correlation dimension. In some cases, remedies are suggested. Figure 4 shows what the ideal correlation integral curves look like and how they are distorted by the various sources of error.

Finite Sampling and Statistical Error

Ultimately, what limits most estimates of dimension is that there are a limited number of points sampling the attractor. Many of the systematic effects discussed below can be eliminated in the $N \to \infty$ limit.

First, a finite sample of N points limits the range over which interpoint distances can scale. The correlation integral itself varies from $2/N^2$ to 1, so that any fit to a slope has only this range. In particular, any fit over a range R of distances (R is the ratio of largest to smallest distance scales) requires that $N^2/2 \geq R^D$, so that at least $N = \sqrt{2R^D}$ points are required. This scaling, however, is an absolute lower bound. The minimum number of points necessary to get a reasonable estimate of dimension for a D-dimensional fractal is probably given by a formula of the form

$$N_{\min} = N_0 \Theta^D, \tag{71}$$

although it is difficult to provide generically good values for N_0 and Θ. Experience indicates that Θ should be of the order of 10, but experience also indicates the need for more experience. See Refs. 114–116 for other points of view.

The statistical error in an estimate of dimension typically scales as $O(1/\sqrt{N})$, although the coefficient of that scaling can be surprisingly small; in some special cases, $O(1/N)$ scaling is observed.[102] Reference 102 also discusses the trade-off between statistical and systematic error and its implications in the context of Eq. (71).

Noise

Noise, the ultimate corrupter of measurements, is usually the first concern of the experimentalist. In the case of dimension estimation, however, the effect of low-amplitude noise is often not so significant as other effects.

Although the particular trajectory of an initial condition is extremely sensitive to noise, the global structure of the strange attractor is robust. Noise will perturb the trajectory away from the strange attractor, but the dynamics will always pull it back toward the attractor. Fractal scaling (of bulk with size) will break down only at length scales equal to the noise amplitude. Thus, at relatively high signal-to-noise ratios, there is still a good range over which fractal scales may be observed.

Because the noise often possesses a much higher characteristic frequency than the deterministic attractor, it is tempting to low-pass filter the signal to reduce the effects of noise. As a rule this is not recommended. As was pointed out in Refs. 117 and 118, linear filtering can artifically raise the measured dimension of the time series. On the other hand, nonlinear filtering techniques have been successfully employed by Kostelich and Yorke[119] to reduce the noise in the time-series characterization of a strange attractor. Similar methods have been proposed in Ref. 40.

Discretization

Discretized time series are of the form $x_i = k_i\epsilon$, where k_i is an integer and ϵ is the discretization level. Such discretization is a natural artifact of digital measuring devices. (In fact, many algorithms work much faster with integer input than real input, so that it may be computationally wise to convert. Usually this involves multiplication by some large factor, followed by rounding to the nearest integer. The multiplicative factor obviously does not affect the slope of a log–log plot, but the rounding to an integer is equivalent to a discretization.) It follows that distances between pairs of discretized points will themselves be discrete multiples of ϵ. This effect is most prominent at small r—indeed, pairs of points with $r = 0$ occur with finite probability, and a log $C(r)$ versus log r plot must deal with the $r = 0$ points. A model of points on an m-dimensional lattice, with lattice points separated by ϵ, leads to a scaling of $C(r) \sim (r + \epsilon/2)^m$ to leading order in ϵ.[101] This suggests that an appropriate plot for a general discretized time series is log $C(r)$ versus $\log(r + \epsilon/2)$.

An alternative approach[120] is deliberately to add noise of amplitude ϵ (this is called dithering) to the original time series.

Edges and Finite Size

The finite size of a compact fractal object limits the range over which $C(r)$ scales as r^ν. For $r > \delta_0$, where δ_0 is the diameter of the attractor, the correlation integral saturates at $C(r) = 1$. This finite size effect is not necessarily a problem in dimension calculations. As long as the effect is confined to length scales that are larger than some $r_0 \approx \delta_0$, then accurate estimates of dimension can still be obtained from the slope of a log $C(r)$ versus log r plot in the $r \leq r_0$ regime.

The real problem stems from the edges that all finite-sized objects in $\mathbf{R}^m$ have. The neighborhoods around points near the edge have different scaling from neighborhoods farther into the interior.

Any model of edge effect will depend on the shape of the fractal, but a tractible and reasonably generic model is that the fractal is a uniform hypercube of length 1 and dimension m. The correlation integral can be derived exactly in that case: $C(N, r) = (2r - r^2)^m$. The local slope at r of the log–log curve is given by

$$\nu(r) = \frac{dC/dr}{C(r)/r} = \frac{m(2-2r)}{2-r} \approx m\left(1 - \frac{r}{2}\right), \tag{72}$$

so that the relative error is $|\nu(r) - m|/m \approx r/2$. Since $C(N, r)$ varies from $2/N^2$ to 1, a natural choice of r for the local slope is that value for which $C(N, r) = 1/N$. This gives $r \approx N^{(-1/m)}/2$, and the relative error ρ in the dimension estimate becomes $\rho = N^{(-1/m)}/4$. Thus, for a given accuracy to be achieved, ρ requires at least $N = (4\rho)^{-m}$ points. For example, for 5% error the minimum number of points needed is

$$N_{\min} = 5^m. \tag{73}$$

If one were to use a value of r for which $C(N, r)$ were of the order of $1/N^2$ and if one employed the Takens best estimator instead of the local slope, the minimum number of points needed for 5% error would be $N_{\min} \equiv (1.25)^m$. It is possible to reduce this value even further if one's model for the strange attractor is, instead of an m-dimensional hypercube,

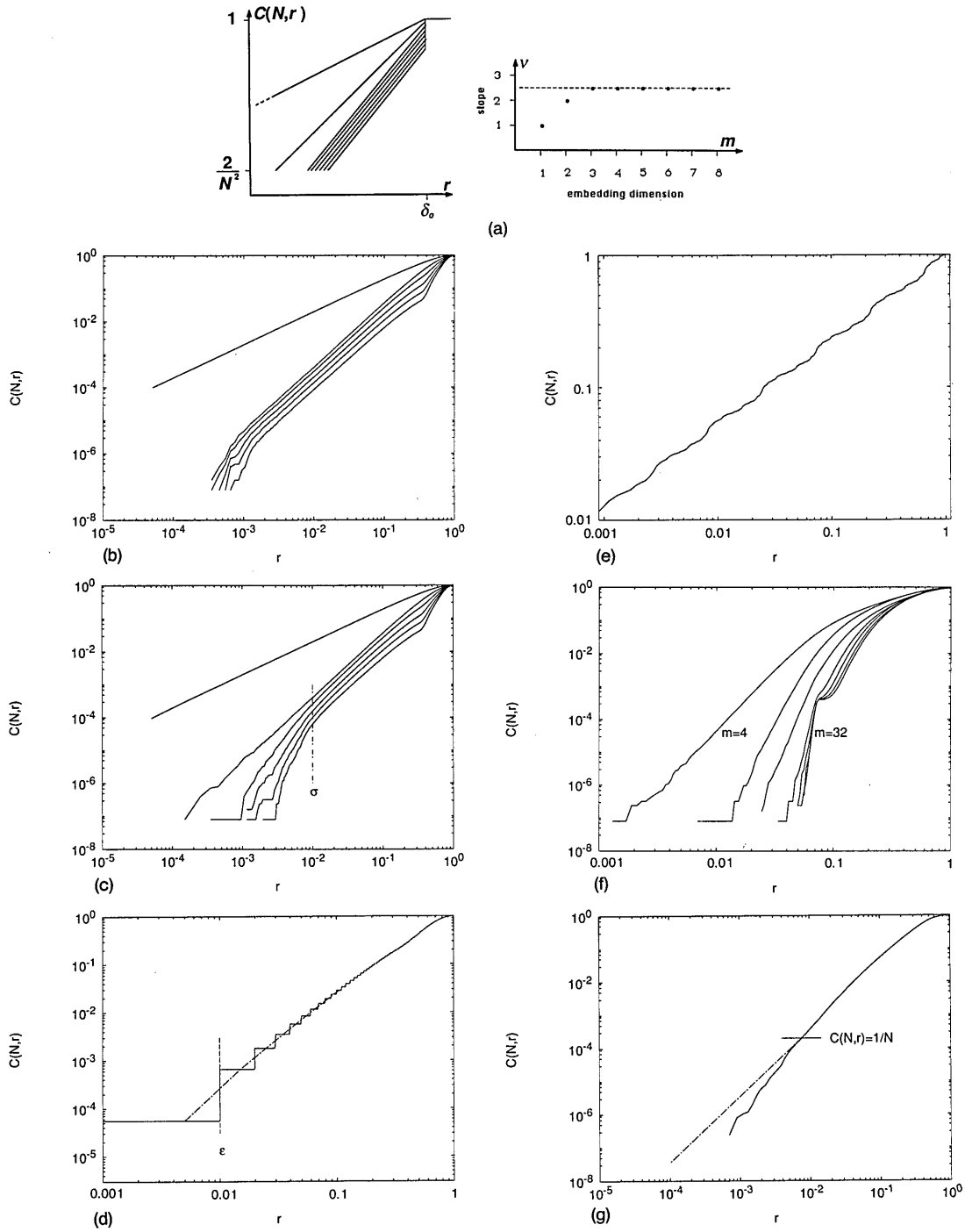


Fig. 4. Sources of error in the correlation integral. (a) Ideally, the correlation integral $C(N, r)$ scales as r^m for $m < \nu$, and as r^ν for $m > \nu$ over a range from $C(N, r) = 2/N^2$ to saturation at $C(N, r) = 1$. Here the dimension is somewhere between 2 and 3. This idealization, however, is only approximated by correlation integrals computed from actual samples of time series data. (b) An actual correlation integral for a two-dimensional chaotic attractor is shown here, with embedding dimensions $m = 1$ through $m = 6$. The finite sample size leads to poor statistics at small r, and the finite size of the attractor (the edge effect) limits the scaling at large r. Nonetheless, the slopes are more or less constant over a range of $C(N, r)$ of the order of N^2. (c) The effect of noise. With σ the amplitude of the noise, one sees that, for $r \ll \sigma$, a slope that approaches the embedding dimension m is observed. For $r \gg \sigma$, the effect of the noise is unimportant. (d) The effect of discretization is to introduce a stair step into the correlation integral (solid curve). The steps are all of equal width, but the log–log plot magnifies those at small r. The effect is minimized if one plots $\log C(N, r)$ versus $\log r$ for $r = (k + \frac{1}{2}\epsilon)$, where k is an integer and ϵ is the level of discretization (dashed–dotted curve). (e) Lacunarity leads to an intrinsic oscillation in the correlation integral that inhibits accurate determinations of slope. The example here is the correlation integral of the middle-thirds Cantor set. (f) Autocorrelation in the time-series data can lead to an anomalous shoulder in the correlation integral. This effect is most highly pronounced for high-dimensional attractors. In this case the input time series was autocorrelated Gaussian noise, and the correlation integral was computed for various (large) embedding dimensions of $m = 4$ to $m = 32$. Equation (75) corrects for this effect. (g) If the time-series data arise from a nonchaotic attractor, then the scaling of $C(N, r)$ as r^ν begins to break down for $C(N, r) < 1/N$. The dotted–dashed curve here has a slope of 2, corresponding to the two-dimensional quasi-periodic input data.

an m torus. Indeed, for some fractals, such as the middle-thirds Cantor set, there is no edge effect at all.

On the other hand, the edge effect can be even stronger than this, if the edge is singularly sharp. An example of this effect is the edge of the attractor that arises from the logistic map $x_{n+1} = 4x_n(1 - x_n)$. The square-root singularity of this edge makes numerical computation of dimension even slower than was predicted by the uniform-measure model.[61]

Smith[121] has used the uniform-measure model to argue that an estimate within 5% of the correct value requires a minimum of $N_{\min} = 42^m$ data points. This value is quite a bit different from that given in Eq. (73), which points to the sensitivity of these kinds of estimate to the assumptions in the model. It is not at all clear that the edge effect is a good model for making *a priori* estimates of errors that arise from dimension-estimation methods.

Cawley and Licht[103] have considered an algorithm that computes correlation dimension with a truncated data set that attempts to sidestep the edge effect by deliberately avoiding points that are far from the centroid of the data points.

Lacunarity

Dimension is not the only way to gauge the fractal-like appearance of a set. Mandelbrot[122] has pointed to lacunarity as another measure: He comments that for two fractals having the same dimension, the one having the greater lacunarity is more textured and appears more fractallike. Further discussion of lacunarity as a means of characterizing fractals can be found in Refs. 123–125.

From the point of view of dimension computation, lacunarity has the effect of introducing an intrinsic oscillation into the correlation integral.[126,127] If the lever arm over which the slope is estimated is long enough to encompass several periods of these oscillations, then the effect of the oscillation will be minimized; but if attempts to compute dimension are based on the local slope of the correlation integral, lacunarity can prevent convergence of the dimension estimator. For further discussion of lacunarity as an impediment to good dimension calculations, see Refs. 105 and 128.

On the other hand, it has been argued in Refs. 63 and 129 that the amplitudes of these oscillations generically damp out in the limit $r \to 0$.

Nonchaotic Attractors

The correlation algorithm does not efficiently compute dimension when the attractor is nonchaotic. Such attractors (limit cycles and limit tori) typically have integer dimension and a Fourier spectrum that is not broadband.

The problem with using the correlation algorithm is that the trajectories of nonchaotic dynamical systems are too regular, and the points that sample the attractor are not even approximately independent. The observed effect is that the correlation integral $C(N, r)$ scales as r^ν from $C(N, r) = O(1/N)$ to $C(N, r) = O(1)$ for nonchaotic attractors, whereas the scaling is complete [from $O(1/N^2)$ to $O(1)$] for chaotic attractors. For further discussion of this effect, see Ref. 101.

Autocorrelation

Autocorrelation is inevitable in time-series data. For continuous signals $x(t)$, there is always some time τ over which $x(t)$ and $x(t + \tau)$ are strongly correlated. If this autocorrelation time τ is long compared with the sampling time, then an anomalous shoulder can appear in the correlation integral and can lead to inaccurate and possibly spurious estimates of dimension.[64]

One approach is to increase the sampling time,[130] although this may introduce other problems: It can further limit the available data and can also affect the delay-time embedding strategy. However, there is a more effective solution. Rewrite the definition of the correlation integral:

$$C(W, N, r) = \frac{2}{(N + 1 - W)(N - W)} \times \sum_{n=W}^{N-1} \sum_{i=0}^{N-1-n} \Theta(r - \|\mathbf{X}_i - \mathbf{X}_{i+n}\|). \quad (74)$$

Note that $W = 1$ is just the standard algorithm. This algorithm computes distances between all pairs of points except for those that are closer together in time than W sampling units. Eliminating this small selection of offending pairs eliminates the anomalous shoulder without sacrificing the statistics of $O(N^2)$ distance calculations.

The definition, in words, of correlation integral in Eq. (45) is now adjusted to

$$C(W, N, r) = \frac{\#\text{ of distances less than } r \text{ except for those from pairs of points closer together in time than } W}{\#\text{ of distances altogether}}. \quad (75)$$

REVIEWS

The reader is referred to Gleick's best-seller *Chaos*[131] as the easiest and undoubtedly the most enjoyable book from which to start learning about chaos, fractals, and the dynamic personalities of scientists in the field. Mandelbrot's book[122] thoroughly discusses the topic of fractals *qua* fractals. This "indispensible, encyclopedic, exasperating" (Gleick's words) volume contains many nice pictures and technical and historical information. Other books with good pictures include Refs. 6, 132, and 133; these books also provide significant technical discussion of fractals.

A recent tutorial on nonlinear dynamics aimed at the engineer can be found in Ref. 134. This includes a discussion of dimension and Lyapunov exponents and has recently been expanded into a book.[135] Also see Ref. 136 for a similarly motivated paper aimed at optical scientists.

Reviews of chaos in general abound: A small sample includes Refs. 21 and 137–140. A recommended text with many references is Ref. 141. Reprint books provide more advanced references that are taken directly from the research literature.[16,142,143] Reviews of dimension and its calculation can be found in Refs. 59, 101, and 144; a particularly rich (and not outdated) source is the workshop proceedings edited by Mayer-Kress.[145] See Ref. 146 for an early review of experimental observations of chaos.

ACKNOWLEDGMENTS

I am grateful to Bette Korber, Michael Holz, and Courosh Mehanian for useful comments on this manuscript. I also

thank Joyce Michaels for help with the bibliography and Al Gschwendtner for his generosity and encouragement. This study was supported by the U.S. Air Force and the Defense Advanced Research Projects Agency.

REFERENCES

1. L. M. Sander, "Fractal growth," Sci. Am. **256,** 94 (1987).
2. L. M. Sander, "Fractal growth processes," Nature (London) **322,** 789 (1986).
3. J. Nittman, G. Daccord, and H. E. Stanley, "Fractal growth of viscous fingers: quantative characterization of a fluid instability phenomenon," Nature (London) **314,** 141 (1985).
4. L. Niemeyer, L. Pietronero, and H. J. Wiesmann, "Fractal dimension of dielectric breakdown," Phys. Rev. Lett. **52,** 1033 (1984).
5. P.-Z. Wong, "The statistical physics of sedimentary rock," Phys. Today 41(12), 24 (1988).
6. M. F. Barnsley, *Fractals Everywhere* (Academic, Boston, Mass., 1988).
7. D. Stauffer, *Introduction to Percolation Theory* (Taylor & Francis, London, 1985).
8. T. A. Witten and L. M. Sander, "Diffusion limited aggregation, a kinetic critical phenomenon," Phys. Rev. Lett. **47,** 1400 (1981).
9. M. F. Barnsley and S. Demko, "Iterated function systems and the global construction of fractals," Proc. R. Soc. London Ser. A **399,** 243 (1985).
10. H. E. Stanley and N. Ostrosky, *On Growth and Form: Fractal and Non-Fractal Patterns in Physics* (Nijhoff, Boston, Mass., 1986).
11. A. J. Hurd, "Resource letter FR-1: fractals," Am. J. Phys. **56,** 969 (1988).
12. L. Kadanoff, "Where is the physics of fractals," Phys. Today **39**(2), 6 (1986).
13. P. H. Carter, R. Cawley, and R. D. Mauldin, "Mathematics of dimension measurement for graphs of functions," in *Fractal Aspects of Materials*, B. B. Mandelbrot and D. E. Passoja, eds. (Materials Research Society, Pittsburgh, Pa., 1985).
14. B. Dubuc, J. F. Quiniou, C. Roques-Carmes, C. Tricot, and S. W. Zucker, "Evaluating the fractal dimension of profiles," Phys. Rev. A **39,** 1500 (1989).
15. H. Haken, *Information and Self-Organization: A Macroscopic Approach to Complex Systems*, Vol. 40 of Springer Series in Synergetics (Springer-Verlag, Berlin, 1988).
16. R. S. MacKay and J. D. Meiss, *Hamiltonian Dynamical Systems* (Hilger, Philadelphia, 1987).
17. J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Vol. 42 of Springer Series in Applied Mathematical Sciences (Springer-Verlag, New York, 1983).
18. S. M. Hammel, C. K. R. T. Jones, and J. V. Moloney, "Global dynamical behavior of the optical field in a ring cavity," J. Opt. Soc. Am. B **2,** 552 (1985).
19. M. Hénon, "A two-dimensional mapping with a strange attractor," Commun. Math. Phys. **50,** 69 (1976).
20. K. J. Falconer, *The Geometry of Fractal Sets*, Vol. 85 of Cambridge Tracts in Mathematics (Cambridge U. Press, Cambridge, 1985).
21. J.-P. Eckmann and D. Ruelle, "Ergodic theory of chaos and strange attractors," Rev. Mod. Phys. **57,** 617 (1985).
22. A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, "Determining Lyapunov exponents from a time series," Physica **16D,** 285 (1985).
23. J.-P. Eckmann, S. O. Kamphorst, D. Ruelle, and S. Ciliberto, "Liapunov exponents from a time series," Phys. Rev. A **34,** 4971 (1986).
24. R. Stoop and P. F. Meier, "Evaluation of Lyapunov exponents and scaling functions from time series," J. Opt. Soc. Am. B **5,** 1037 (1988).
25. P. Grassberger and I. Procaccia, "Estimation of the Kolmogorov entropy from a chaotic signal," Phys. Rev. A **28,** 2591 (1983).
26. A. Cohen and I. Procaccia, "Computing the Kolmogorov entropy from time signals of dissipative and conservative dynamical systems," Phys. Rev. A **31,** 1872 (1985).
27. J. P. Eckmann and I. Procaccia, "Fluctuations of dynamical scaling indices in nonlinear systems," Phys. Rev. A **34,** 659 (1986).
28. P. Szépfalusy and T. Tél, "Dynamical fractal properties of one-dimensional maps," Phys. Rev. A **35,** 477 (1987).
29. N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, "Geometry from a time series," Phys. Rev. Lett. **45,** 712 (1980).
30. F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence, Warwick, 1980*, D. A. Rand and L.-S. Young, eds., Vol. 898 of Springer Lecture Notes in Mathematics (Springer-Verlag, Berlin, 1981), p. 366.
31. R. Mañé, "On the dimension of the compact invariant sets of certain non-linear maps," in *Dynamical Systems and Turbulence, Warwick, 1980*, D. A. Rand and L.-S. Young, eds., Vol. 898 of Springer Lecture Notes in Mathematics (Springer-Verlag, Berlin, 1981), p. 320.
32. A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," Phys. Rev. A **33,** 1134 (1986).
33. W. Liebert, K. Pawelzik, and H. G. Schuster, Institut für Theoretische Physik, Universität Frankfurt, Frankfurt, Federal Republic of Germany, "Optimal embeddings of chaotic attractors from topological considerations," preprint (1989).
34. D. S. Broomhead and G. P. King, "Extracting qualitative dynamics from experimental data," Physica **20D,** 217 (1986).
35. S. Sato, M. Sano, and Y. Sawada, "Practical methods of measuring the generalized dimension and the largest Lyapunov exponent in high dimensional chaotic systems," Prog. Theor. Phys. **77,** 1 (1987).
36. A. I. Mees, P. E. Rapp, and L. S. Jennings, "Singular-value decomposition and embedding dimension," Phys. Rev. A **36,** 340 (1987).
37. A. M. Albano, J. Muench, C. Schwartz, A. I. Mees, and P. E. Rapp, "Singular-value decomposition and the Grassberger–Procaccia algorithm," Phys. Rev. A **38,** 3017 (1988).
38. A. M. Fraser, "Reconstructing attractors from scalar time series: a comparison of singular system and redundancy criteria," Physica **34D,** 391 (1989).
39. A. M. Fraser, "Information and entropy in strange attractors," Ph.D. dissertation (University of Texas at Austin, Austin, Tex., 1988).
40. J. D. Farmer and J. J. Sidorowich, "Exploiting chaos to predict the future and reduce noise," in *Evolution, Learning and Cognition*, Y. C. Lee, ed. (World Scientific, Singapore, 1988), p. 227.
41. F. Hausdorff, "Dimension und äusseres Mass," Math. Annalen **79,** 157 (1919).
42. J. D. Farmer, E. Ott, and J. A. Yorke, "The dimension of chaotic attractors," Physica **7D,** 153 (1983).
43. H. G. E. Hentschel and I. Procaccia, "The infinite number of generalized dimensions of fractals and strange attractors," Physica **8D,** 435 (1983).
44. P. Grassberger, "Generalized dimensions of strange attractors," Phys. Lett. A **97,** 227 (1983).
45. G. Paladin and A. Vulpiani, "Anomalous scaling laws in multifractal objects," Phys. Rep. **156,** 147 (1987).
46. T. C. Halsey, M. H. Jensen, L. P. Kadanoff, I. Procaccia, and B. I. Shraiman, "Fractal measures and their singularities: the characterization of strange sets," Phys. Rev. A **33,** 1141 (1986).
47. S. K. Sakar, "Multifractal description of singular measures in dynamical systems," Phys. Rev. A **36,** 4104 (1987).
48. M. H. Jensen, L. P. Kadanoff, A. Libchaber, I. Procaccia, and J. Stavans, "Global universality at the onset of chaos: results of a forced Rayleigh–Bernard experiment," Phys. Rev. Lett. **55,** 2798 (1985).
49. E. G. Gwinn and R. M. Westervelt, "Scaling structure of attractors at the transition from quasiperiodicity to chaos in electronic transport in Ge," Phys. Rev. Lett. **59,** 157 (1987).
50. J. A. Glazier, G. Gunaratne, and A. Libchaber, "$f(\alpha)$ curves: experimental results," Phys. Rev. A **37,** 523 (1988).
51. M. J. Feigenbaum, M. H. Jensen, and I. Procaccia, "Time ordering and the thermodynamics of strange sets: theory and experimental tests," Phys. Rev. Lett. **57,** 1503 (1986).
52. D. Katzen and I. Procaccia, "Phase transitions in the thermo-

dynamic formalism of multifractals," Phys. Rev. Lett. **58,** 1169 (1987).
53. M. H. Jensen, L. P. Kadanoff, and I. Procaccia, "Scaling structure and thermodynamics of strange sets," Phys. Rev. A **36,** 1409 (1987).
54. T. Bohr and M. H. Jensen, "Order parameter, symmetry breaking, and phase transitions in the description of multifractal sets," Phys. Rev. A **36,** 4904 (1987).
55. A. Renyi, *Probability Theory* (North-Holland, Amsterdam, 1970).
56. H. S. Greenside, A. Wolf, J. Swift, and T. Pignataro, "Impracticality of a box-counting algorithm for calculating the dimensionality of strange attractors," Phys. Rev. A **25,** 3453 (1982).
57. P. Grassberger, R. Badii, and A. Politi, "Scaling laws for invariant measures on hyperbolic and non-hyperbolic attractors," J. Stat. Phys. **51,** 135 (1988).
58. P. Cvitanović, G. H. Gunarante, and I. Procaccia, "Topological and metric properties of Hénon-type strange attractors," preprint (University of Chicago, Chicago, Ill., 1988).
59. G. Mayer-Kress, "Application of dimension algorithms to experimental chaos," in *Directions in Chaos,* Hao Bai-lin, ed. (World Scientific, Singapore, 1987), p. 122.
60. P. Grassberger and I. Procaccia, "Characterization of strange attractors," Phys. Rev. Lett. **50,** 346 (1983).
61. P. Grassberger and I. Procaccia, "Measuring the strangeness of strange attractors," Physica **9D,** 189 (1983).
62. F. Takens, "Invariants related to dimension and entropy," in *Atas do 13°* (Colóqkio Brasiliero do Matemática, Rio de Janeiro, 1983).
63. P. Grassberger, "Finite sample corrections to entropy and dimension estimates," Phys. Lett. A **128,** 369 (1988).
64. J. Theiler, "Spurious dimension from correlation algorithms applied to limited time series data," Phys. Rev. A **34,** 2427 (1986).
65. K. Pawelzik and H. G. Schuster, "Generalized dimensions and entropies from a measured time series," Phys. Rev. A **35,** 481 (1987).
66. H. Atmanspacher, H. Scheingraber, and W. Voges, "Global scaling properties of a chaotic attractor reconstructed from experimental data," Phys. Rev. A **37,** 1314 (1988).
67. Y. Termonia and Z. Alexandrowicz, "Fractal dimension of strange attractors from radius versus," Phys. Rev. Lett. **51,** 1265 (1983).
68. J. Guckenheimer and G. Buzyna, "Dimension measurements for geostrophic turbulence," Phys. Rev. Lett. **51,** 1483 (1983).
69. R. Badii and A. Politi, "Statistical description of chaotic attractors: the dimension function," J. Stat. Phys. **40,** 725 (1985).
70. P. Grassberger, "Generalizations of the Hausdorff dimension of fractal measures," Phys. Lett. A **107,** 101 (1985).
71. G. Broggi, "Evaluation of dimensions and entropies of chaotic systems," J. Opt. Soc. Am. B **5,** 1020 (1988).
72. W. van de Water and P. Schram, "Generalized dimensions from near-neighbor information," Phys. Rev. A **37,** 3118 (1988).
73. R. Badii and G. Broggi, "Measurement of the dimension spectrum $f(\alpha)$: fixed-mass approach," Phys. Lett. A **131,** 339 (1988).
74. K. W. Pettis, T. A. Bailey, A. K. Jain, and R. C. Dubes, "An intrinsic dimensionality estimator from near neighbor information," IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-1,** 25 (1979).
75. R. L. Somorjai, "Methods for estimating the intrinsic dimensionality of high-dimensional point sets," in *Dimensions and Entropies in Chaotic Systems—Quantification of Complex Behavior,* G. Mayer-Kress, ed., Vol. 32 of Springer Series in Synergetics (Springer-Verlag, Berlin, 1986), p. 137.
76. J. L. Kaplan and J. A. Yorke, in *Functional Differential Equations and Approximations of Fixed Points,* H. O. Peitgen and H. O. Walther, eds., Vol. 730 of Springer Lecture Notes in Mathematics (Springer-Verlag, Berlin, 1979), p. 204.
77. P. Fredrickson, J. L. Kaplan, E. D. Yorke, and J. A. Yorke, "The Liapunov dimension of strange attractors," J. Diff. Eq. **49,** 185 (1983).
78. J. D. Farmer, "Chaotic attractors of an infinite dimensional dynamical system," Physica **4D,** 366 (1982).
79. K. Ikeda and K. Matsumoto, "Study of a high-dimensional chaotic attractor," J. Stat. Phys. **44,** 955 (1986).
80. R. Badii and A. Politi, "Renyi dimensions from local expansion rates," Phys. Rev. A **35,** 1288 (1987).
81. D. Auerbach, P. Cvitanović, J.-P. Eckmann, G. Gunarante, and I. Procaccia, "Exploring chaotic motion through periodic orbits," Phys. Rev. Lett. **58,** 2387 (1987).
82. G. Gunarante and I. Procaccia, "Organization of chaos," Phys. Rev. Lett. **59,** 1377 (1987).
83. D. Auerbach, B. O'Shaughnessy, and I. Procaccia, "Scaling structure of strange attractors," Phys. Rev. A **37,** 2234 (1988).
84. C. Grebogi, E. Ott, and J. A. Yorke, "Unstable periodic orbits and the dimension of chaotic attractors," Phys. Rev. A **36,** 3522 (1987).
85. C. Grebogi, E. Ott, and J. A. Yorke, "Unstable periodic orbits and the dimensions of multifractal chaotic attractors," Phys. Rev. A **37,** 1711 (1988).
86. K. Fukunaga and D. R. Olsen, "An algorithm for finding intrinsic dimensionality of data," IEEE Trans. Comput. **C-20,** 176 (1971).
87. H. Froehling, J. P. Crutchfield, D. Farmer, N. H. Packard, and R. Shaw, "On determining the dimension of chaotic flows," Physica **3D,** 605 (1981).
88. D. S. Broomhead, R. Jones, and G. P. King, "Topological dimension and local coordinates from time series data," J. Phys. A **20,** L563 (1987).
89. A. Čenys and K. Pyragas, "Estimation of the number of degrees of freedom from chaotic time series," Phys. Lett. A **129,** 227 (1988).
90. A. Passamante, T. Hediger, and M. Gollub, "Fractal dimension and local intrinsic dimension," Phys. Rev. A **39,** 3640 (1989).
91. W. A. Brock, W. D. Dechert, and J. A. Scheinkman, "A test for independence based on the correlation dimension," preprint SSRI 8702 (University of Wisconsin, Madison, Wisc., 1987).
92. A. Namajūnas, J. Pozžela, and A. Tamaševičius, "An electronic technique for measuring phase space dimension from chaotic time series," Phys. Lett. A **131,** 85 (1988).
93. A. Destexhe, J. A. Sepulchre, and A. Babloyantz, "A comparative study of the experimental quantification of deterministic chaos," Phys. Lett. A **132,** 101 (1988).
94. J. D. Farmer and J. J. Sidorowich, "Predicting chaotic time series," Phys. Rev. Lett. **59,** 845 (1987).
95. D. A. Russell, J. D. Hanson, and E. Ott, "Dimension of strange attractors," Phys. Rev. Lett. **45,** 1175 (1980).
96. P. Grassberger, "On the fractal dimension of the Hénon attractor," Phys. Lett. A **97,** 224 (1983).
97. M. J. McGuinness, "A computation of the limit capacity of the Lorenz attractor," Physica **16D,** 265 (1985).
98. W. E. Caswell and J. A. Yorke, "Invisible errors in dimension calculations: geometric and systematic effects," in *Dimensions and Entropies in Chaotic Systems—Quantification of Complex Behavior,* G. Mayer-Kress, ed., Vol. 32 of Springer Series in Synergetics (Springer-Verlag, Berlin, 1986), p. 123.
99. A. Chhabra and R. V. Jensen, "Direct determination of the $f(\alpha)$ singularity spectrum," Phys. Rev. Lett. **69,** 1327 (1989).
100. F. Takens, "On the numerical determination of the dimension of an attractor," in *Dynamical Systems and Bifurcations, Groningen, 1984,* B. L. J. Braaksma, H. W. Broer, and F. Takens, eds., Vol. 1125 of Lecture Notes in Mathematics (Springer-Verlag, Berlin, 1985).
101. J. Theiler, "Quantifying chaos: practical estimation of the correlation dimension," Ph.D. dissertation (California Institute of Technology, Pasadena, Calif., 1988).
102. J. Theiler, "Statistical precision of dimension estimators," Phys. Rev. A (to be published).
103. R. Cawley and A. L. Licht, "Maximum likelihood method for evaluating correlation dimension," in *The Physics of Phase Space,* Y. S. Kim and W. W. Zachary, eds., Vol. 278 of Lecture Notes in Physics (Springer-Verlag, Berlin, 1986), p. 90.
104. S. Ellner, "Estimating attractor dimensions for limited data: a new method, with error estimates," Phys. Lett. A **113,** 128 (1988).

105. J. Theiler, "Lacunarity in a best estimator of fractal dimension," Phys. Lett. A **133,** 195 (1988).
106. J. Theiler, "Efficient algorithm for estimating the correlation dimension from a set of discrete points," Phys. Rev. A **36,** 4456 (1987).
107. S. Bingham and M. Kot, "Multidimensional trees, range searching, and a correlation dimension algorithm of reduced complexity," Phys. Lett. A **140,** 327 (1989).
108. F. Hunt and F. Sullivan, "Efficient algorithms for computing fractal dimensions," in *Dimensions and Entropies in Chaotic Systems—Quantification of Complex Behavior,* G. Mayer-Kress, ed., Vol. 32 of Springer Series in Synergetics (Springer-Verlag, Berlin, 1986), p. 74.
109. M. Franaszek, "Optimized algorithm for the calculation of correlation integrals," Phys. Rev. A **39,** 5540 (1989).
110. C.-K. and F. C. Moon, "An optical technique for measuring fractal dimensions of planar Poincaré maps," Phys. Lett. A **114,** 222 (1986).
111. P. Grassberger, "Do climatic attractors exist?" Nature (London) **323,** 609 (1986).
112. A. R. Osborne and A. Provenzale, "Finite correlation dimension for stochastic systems with power-law spectra," Physica **35D,** 357 (1989).
113. A. Brandstater and H. L. Swinney, "Strange attractors in weakly turbulent Couette–Taylor flow," Phys. Rev. A **35,** 2207 (1987).
114. N. B. Abraham, A. M. Albano, B. Das, G. DeGuzman, S. Young, R. S. Gioggia, G. P. Puccioni, and J. R. Tredicce, "Calculating the dimension of attractors from small data sets," Phys. Lett. A **114,** 217 (1986).
115. J. W. Havstad and C. L. Ehlers, "Attractor dimension of nonstationary dynamical systems from small data sets," Phys. Rev. A **39,** 845 (1989).
116. J. B. Ramsey and H.-J. Yuan, "Bias and error bars in dimension calculations and their evaluation in some simple models," Phys. Lett. A **134,** 287 (1989).
117. R. Badii, G. Broggi, B. Derighetti, M. Ravani, S. Ciliberto, A. Politi, and M. A. Rubio, "Dimension increase in filtered chaotic signals," Phys. Rev. Lett. **60,** 979 (1988).
118. F. Mitschke, M. Moller, and W. Lange, "Measuring filtered chaotic signals," Phys. Rev. A **37,** 4518 (1988).
119. E. J. Kostelich and J. A. Yorke, "Noise reduction in dynamical systems," Phys. Rev. A **38,** 1649 (1988).
120. M. Möller, W. Lange, F. Mitschke, N. B. Abraham, and U. Hübner, "Errors from digitizing and noise in estimating attractor dimensions," Phys. Lett. A **138,** 176 (1989).
121. L. A. Smith, "Intrinsic limits on dimension calculations," Phys. Lett. A **133,** 283 (1988).
122. B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, Calif., 1982).
123. B. B. Mandlebrot, "Corrélations et texture dans un nouveau modéle d'univers hiérarchisé, basé sur les ensembles trémas," C. R. Acad. Sci. A **288,** 81 (1979).
124. Y. Gefen, Y. Meir, A. Aharony, and B. B. Mandelbrot, "Geometric implementation of hypercubic lattices with noninteger dimension," Phys. Rev. Lett. **50,** 145 (1983).
125. Y. Gefen, A. Aharony, and B. B. Mandelbrot, "Phase transitions on fractals: III. Infinitely ramified lattices," J. Phys. A **17,** 1277 (1984).
126. R. Badii and A. Politi, "Intrinsic oscillations in measuring the fractal dimensions," Phys. Lett. A **104,** 303 (1984).
127. L. A. Smith, J.-D. Fournier, and E. A. Spiegel, "Lacunarity and intermittency in fluid turbulence," Phys. Lett. A **114,** 465 (1986).
128. A. Arneodo, G. Grasseau, and E. J. Kostelich, "Fractal dimensions and $f(\alpha)$ spectrum of the Hénon attractor," Phys. Lett. A **124,** 426 (1987).
129. D. Bessis, J.-D. Fournier, G. Servizi, G. Turchetti, and S. Vaienti, "Mellin transforms of correlation integrals and generalized dimension of strange sets," Phys. Rev. A **36,** 20 (1987).
130. P. Atten, J. G. Caputo, B. Malraison, and Y. Gagne, "Détermination de dimension d'attracteurs pour différents écoulements," J. Mec. Theor. Appl. **133** (Suppl.) (1984).
131. J. Gleick, *Chaos: Making a New Science* (Viking, New York, 1987).
132. H.-O. Peitgen and P. H. Richter, *The Beauty of Fractals* (Springer-Verlag, Berlin, 1986).
133. H.-O. Peitgen and D. Saupe, eds., *The Science of Fractal Images* (Springer-Verlag, New York, 1988).
134. T. S. Parker and L. O. Chua, "Chaos: a tutorial for engineers," Proc. IEEE **75,** 982 (1987).
135. T. S. Parker and L. O. Chua, *Practical Numerical Algorithms for Chaotic Systems* (Springer-Verlag, New York, 1989).
136. N. B. Abraham, A. M. Albano, B. Das, T. Mello, M. F. H. Tarroja, N. Tufillaro, and R. S. Gioggia, "Definitions of chaos and measuring its characteristics," in *Optical Chaos,* J. Chrostowski and N. B. Abraham, eds. Proc. Soc. Photo-Opt. Instrum. Eng. **667,** 2 (1986).
137. R. S. Shaw, "Strange attractors, chaotic behavior, and information flow," Z. Naturforsch. **36a,** 80 (1981).
138. J. P. Crutchfield, J. D. Farmer, N. H. Packard, and R. Shaw, "Chaos," Sci. Am. **255,** 46 (1986).
139. E. Ott, "Strange attractors and chaotic motions of dynamical systems," Rev. Mod. Phys. **53,** 655 (1981).
140. A. V. Holden, ed., *Chaos* (Princeton U. Press, Princeton, N.J., 1986).
141. H. G. Schuster, *Deterministic Chaos: An Introduction* (VCH, Weinheim, Federal Republic of Germany, 1988).
142. H. Bai-Lin, *Chaos* (World Scientific, Singapore, 1984).
143. P. Cvitanović, *Universality in Chaos* (Hilger, Bristol, UK, 1986).
144. P. Grassberger, "Estimating the fractal dimensions and entropies of strange attractors," in *Chaos,* A. V. Holden, ed. (Princeton U. Press, Princeton, N.J. 1986), Chap. 14.
145. G. Mayer-Kress, ed., *Dimensions and Entropies in Chaotic Systems—Quantification of Complex Behavior,* Vol. 32 of Springer Series in Synergetics (Springer-Verlag, Berlin, 1986).
146. N. B. Abraham, J. P. Gollub, and H. L. Swinney, "Testing nonlinear dynamics," Physica **11D,** 252 (1984).