



ELSEVIER

Physica D 139 (2000) 276–300

PHYSICA D

www.elsevier.com/locate/physd

Computing connectedness: disconnectedness and discreteness

V. Robins^{a,*}, J.D. Meiss^a, E. Bradley^b

^a *Department of Applied Mathematics, University of Colorado, Boulder, CO 80309, USA*

^b *Department of Computer Science, University of Colorado, Boulder, CO 80309, USA*

Received 12 January 1999; received in revised form 15 October 1999; accepted 8 November 1999

Communicated by C.K.R.T. Jones

Abstract

We consider finite point-set approximations of a manifold or fractal with the goal of determining topological properties of the underlying set. We use the minimal spanning tree of the finite set of points to compute the number and size of its ϵ -connected components. By extrapolating the limiting behavior of these quantities as $\epsilon \rightarrow 0$ we can say whether the underlying set appears to be connected, totally disconnected, or perfect. We demonstrate the effectiveness of our techniques for a number of examples, including a family of fractals related to the Sierpinski triangle, Cantor subsets of the plane, the Hénon attractor, and cantori from four-dimensional symplectic sawtooth maps. For zero-measure Cantor sets, we conjecture that the growth rate of the number of ϵ -components as $\epsilon \rightarrow 0$ is equivalent to the box-counting dimension. © 2000 Elsevier Science B.V. All rights reserved.

PACS: 07.05.Kf; 02.40.Pc; 47.53.+n

Keywords: Computational topology; Fractal geometry; Minimal spanning tree

1. Introduction

One of the primary tasks of experimental science is the extraction of qualitative and quantitative information from data. Data are typically represented as subsets of metric spaces, either naturally, or through the process of phase space embedding [1–3]. The qualitative information obtained from data is most often geometric; common examples are fractal dimensions [4] or Lyapunov exponents [5]. Topological properties, although more fundamental than metric ones, are more difficult to determine computationally. Examples of topological techniques in the study of dynamical systems include the calculation of symbolic dynamics of a flow (when the attractor can be embedded in \mathbb{R}^3) by using knot theory to create a template [6–8], the computation of homology from flows that lie on smooth manifolds [9], and the application of Conley index theory to time-series data [10,11].

There is a growing literature on the formalization and representation of topological questions for computer applications and on the study of appropriate algorithms; see [12] for a survey of the field. By “computational

* Corresponding author.

E-mail address: robinsv@colorado.edu (V. Robins)

topology” we mean the study of topological properties of an object (such as the number of connected components) that can be computed to some finite accuracy. Our computational techniques are tailored to finite point-set data, but as we derived in [13], the underlying concepts apply in general to compact subsets of metric spaces. Often we can improve the accuracy of the computations at the expense of increased computer or experimental effort, and our algorithms lead to topology by extrapolation.

This notion can be distinguished from “digital topology” [14], which deals with the topological features of digital images, given by discrete values (typically binary) on a regular grid (typically two- or three-dimensional). This field has many applications, including algorithmic pattern recognition, which plays an important role in computer vision (e.g., determining whether a robot-width corridor exists between two obstacles [15]) and remote sensing (e.g., computing the boundaries of a drainage basin from satellite data [16]). The fundamental concept in this field is that of adjacency, the definition of which depends upon the lattice. Much work in this area has focused on algorithms for the labeling of components [14], boundaries [17], and other features of digital images. Basic results include consistent notions for connectedness [14], simple connectedness [18], a digital Jordan curve theorem [19], and algorithms for the Euler characteristic [20,21] of digital sets.

Section 2 briefly reviews the results we first described in [13]. The fundamental step of that paper is to formulate definitions of connectedness and disconnectedness at a resolution ϵ . We can then deduce the topological properties of connectedness, total disconnectedness, and perfectness by examining the limiting behavior of the number, $C(\epsilon)$, and size, $D(\epsilon)$, of ϵ -connected components of the set as $\epsilon \rightarrow 0$. We characterize the limiting behavior of these two quantities by a power law, and compute the corresponding disconnectedness and discreteness indices. In Section 3, we give a new algorithm based on the *minimal spanning tree* (MST) that implements these ideas numerically for arbitrary finite point-set data. An important step is to determine a “cutoff” resolution, ρ , so that the computed results are a good representation of the true set for $\epsilon > \rho$. When the underlying set is perfect, ρ is simply the resolution at which the data first appears to have an isolated point. We show that the MST is an ideal data structure for representing ϵ -components of a finite point-set. Essentially, this is because the MST always joins two subsets by the smallest possible edge.

The main purpose of this paper is to demonstrate the effectiveness of our techniques by applying them to a variety of examples. We present data that exhibit different types of scaling in the number and size of their ϵ -components. We also investigate the dependence of the cutoff resolution ρ on the number of data points and on the uniformity of their distribution over the attractor. The first three examples are fractals generated by closely related iterated function systems. Each has a distinct topology but all have the same Hausdorff dimension. We show that the disconnectedness and discreteness indices classify the sets according to their topology. We next analyze five Cantor sets to demonstrate the different types of scaling that can occur in the functions $C(\epsilon)$ and $D(\epsilon)$ as $\epsilon \rightarrow 0$. These examples lead us to conjecture that Cantor sets with zero Lebesgue measure have disconnectedness index equal to their box-counting dimension. Finally, we examine data from two dynamical systems. Our analysis of the Hénon attractor gives strong support to the common intuitive description that it has Cantor set cross-sections. We find that two cantori from symplectic sawtooth maps have subpolynomial growth of the number of ϵ -components, demonstrating yet another type of scaling. We anticipate that our techniques will be of even greater utility in identifying the structure of sets in higher dimensions where visualization is more difficult.

2. Review of theory

In [13], we reformulated the classical definition of connectedness in a way that relies on extrapolation, making it possible to implement the ideas numerically. The basic approach is to look at the set with a finite resolution ϵ , see how certain properties change as $\epsilon \rightarrow 0$, and infer information about the topology.

The standard definition is that a set X is *connected* if it is in one piece, i.e., if there do not exist two closed disjoint subsets, U and V with $U \cup V = X$. When such subsets do exist, the set is disconnected. Given a subset X of a metric space, we say it is ϵ -*disconnected* if it can be written as the union of two sets that are separated by a distance of at least ϵ , i.e., there are two closed subsets, U and V with $U \cup V = X$ and $d(U, V) \equiv \inf_{x \in U, y \in V} d(x, y) \geq \epsilon$. Otherwise, X is ϵ -*connected*. Our definition is equivalent to an early formulation of connectedness, due to Cantor [22], which uses the concept of an ϵ -*chain*: a finite sequence of points x_0, \dots, x_n that satisfy $d(x_i, x_{i+1}) < \epsilon$ for $i = 1, \dots, n$. A set X is Cantor-connected if and only if given any $\epsilon > 0$, for all $x, y \in X$, there is an ϵ -chain with $x_0 = x$ and $x_n = y$. Cantor's definition agrees with the now standard one only in the special case of compact metric spaces. For example, the rational numbers are Cantor-connected but disconnected in the regular sense. Restricting our attention to compact sets is not unreasonable, since we are primarily interested in sets that are well approximated on a computer as finite point-sets.

The objects in which we are most interested are the connected components of X . In ϵ -resolution terms, these are defined as follows: A is an ϵ -*component* of X if A is ϵ -connected and $d(A, X \setminus A) \geq \epsilon$. Given a resolution, ϵ , X has a natural decomposition as the disjoint union of its ϵ -components. We can exploit this decomposition to deduce topological properties of the set. For example, if the only ϵ -component is X itself for all ϵ , then we can conclude that X is connected. We also know that at any fixed resolution ϵ , a compact set has a finite number of ϵ -components. This motivates the introduction of a function $C(\epsilon)$ that measures the number of ϵ -components at resolution ϵ . This function is monotonic: if $\epsilon_1 < \epsilon_2$ then $C(\epsilon_1) \geq C(\epsilon_2)$.

We can also deduce the topological properties of total disconnectedness and perfectness by looking at the size of the ϵ -components. There are a number of ways to measure the size of a set; the quantity used depends on the context. In [13] we used the diameter,¹ since this is defined in any metric space. At a resolution ϵ , $\mathcal{D}(\epsilon)$ represents the set of diameter measurements of the ϵ -components. For notational convenience, we write $D(\epsilon) = \max \mathcal{D}(\epsilon)$ for the function that describes how the largest diameter changes with resolution. A set is *totally disconnected* if the connected component of each point is only the point itself. So, a compact set X is totally disconnected if and only if $\lim_{\epsilon \rightarrow 0} D(\epsilon) = 0$. Any finite set of points is totally disconnected; so are all Cantor sets.

Any compact set that is both totally disconnected and perfect is a Cantor set. Recall that a set is *perfect* if it is equal to the set of its accumulation points. In other words, every point has arbitrarily small neighborhoods containing infinitely many other points, so no point is isolated. In ϵ -resolution terms, a compact set X is perfect if and only if $\min \mathcal{D}(\epsilon) > 0$ for all $\epsilon > 0$. Another way to determine whether or not a set is perfect is to look for isolated points. These are straightforward to detect numerically, since a point is isolated at resolution ϵ if and only if its nearest neighbor is at least distance ϵ away. The number of isolated points as a function of resolution, $I(\epsilon)$, must be zero for all $\epsilon > 0$ if the set is perfect.

Given the above results, it is clear that the limiting behavior of $C(\epsilon)$ and $D(\epsilon)$ will tell us about the connectedness properties of a set. In some situations of interest, e.g., Cantor sets, it is expected that $C(\epsilon) \rightarrow \infty$ and $D(\epsilon) \rightarrow 0$ as $\epsilon \rightarrow 0$. An effective way to evaluate this behavior is to assume an asymptotic form and compute appropriate indices. We use a general power law: near $\epsilon = 0$, $C(\epsilon) \sim \epsilon^{-\gamma}$ and $D(\epsilon) \sim \epsilon^\delta$. The exponents may be found as the following limits:

$$\gamma = \liminf_{\epsilon \rightarrow 0} \frac{\log C(\epsilon)}{\log(1/\epsilon)},$$

$$\delta = \liminf_{\epsilon \rightarrow 0} \frac{\log D(\epsilon)}{\log \epsilon}.$$

¹ Recall that the diameter of a set A is the largest distance between any two points in the set: $\text{diam}(A) = \sup_{x, y \in A} d(x, y)$.

The component growth rate, γ , is called the *disconnectedness* index. A positive value of γ implies that the set has infinitely many components. We call δ the *discreteness* index. If δ is positive, the set must be totally disconnected. The function $D(\epsilon)$ relates the size of the ϵ -components to the distance between them, so δ can be thought of as measuring how sparse the set is. If the set is connected or has a finite number of components, then γ and δ are both zero.

For many of the Cantor set examples presented in Section 4, we find that γ is approximately equal to the box-counting dimension, \dim_B , and that δ is close to one. This leads us to conjecture that

Conjecture 1. *Cantor sets with zero Lebesgue measure have $\gamma = \dim_B$ and $\delta = 1$.*

We can prove this in the special case that the Cantor set is generated by an iterated function system of similarity transformations. These results are the topic of a forthcoming paper.

3. Implementation

The type of data we want to analyze are finite sets of points in a finite dimensional metric space (usually \mathbb{R}^n). This is a more general setting than digital data, which are restricted to a predefined, finite grid, so we need a more general implementation. Although many of the examples in Section 4 are attractors of iterated function systems, our algorithm is also more general than those developed specifically for that context [23,24]. We are most interested in the situation where the data represent an underlying infinite set, similar to the way floating point numbers in a computer approximate the real numbers. This situation occurs naturally in the study of dynamical systems, where a finite sequence of points (an orbit) on an attractor is used to represent the true attractor. Similar data sets arise in time-series analysis when sample measurements are recorded from a continuous process. If the data are embedded using time-delay techniques [1,2], the result is a finite point-set in \mathbb{R}^n . Knowing the topological and geometric properties of this set assists qualitative understanding of the dynamics and such properties have been used in mathematical modeling [7–9,25].

The theory outlined in Section 2 applies to arbitrary compact sets. Recall that compactness is the property that any covering of the set by open subsets can also be accomplished by a finite number of those subsets. This makes it possible to represent a compact set by a finite number of points. One way to do this is to take a covering of the set by open balls, choose a finite subcover, and let the centers of the balls be the finite point-set approximation. Suppose each point in the compact set, X , is within a maximum distance $\rho/2$ of some point in the approximating set, S . The original set and its finite point-set approximation therefore exhibit essentially the same behavior of $C(\epsilon)$, $D(\epsilon)$ and $I(\epsilon)$ when $\epsilon > \rho$. For example, if the original set, X , is connected and perfect, then examining S with resolution $\epsilon > \rho$ we should see $C_S(\epsilon) = 1$, $D_S(\epsilon)$ close to the diameter of X , and $I_S(\epsilon) = 0$. For $\epsilon < \rho$, S has many components and the majority of these are isolated points. In general, we have the bound $C_S(\epsilon - \rho) \geq C_X(\epsilon) \geq C_S(\epsilon + \rho)$ for $\epsilon > \rho$.

In practice, we are given only the finite point-set approximation and wish to determine the connectedness properties of the underlying set. To do this, we need to estimate the minimum resolution, ρ , from the point-set itself. We can then attempt to deduce the limiting behavior of $C(\epsilon)$ and $D(\epsilon)$ from the data with $\epsilon > \rho$. The question that naturally arises is: how confident can we be that the limiting topology is that implied by the data for $\epsilon > \rho$? It is possible to construct examples that appear to be connected down to a given resolution, but are in fact Cantor sets. Conversely, there are connected sets whose finite point-set approximations appear to be totally disconnected. Both these problems can be addressed to some extent by checking the effect on $C(\epsilon)$, $D(\epsilon)$ and $I(\epsilon)$ of increasing the number of points approximating the underlying set. Ultimately, though, we are restricted by machine precision.

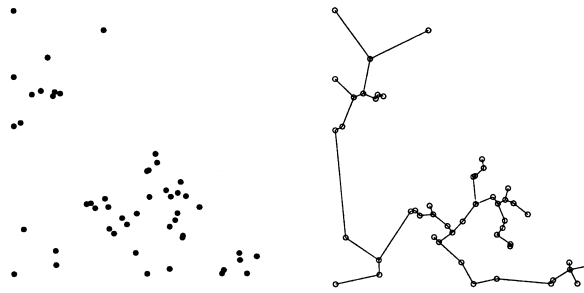


Fig. 1. A finite set of points and its minimal spanning tree. The weight of an edge is the Euclidean distance between the points it joins.

In order to compute $C(\epsilon)$, $D(\epsilon)$ and $I(\epsilon)$ numerically, we need an appropriate way to organize the finite point-set data. The structure we use is a graph — the MST [26]. This choice was inspired by Yip’s [27] work on computer recognition of orbit structures in two-dimensional area-preserving maps. Section 3.1 describes why the edge lengths of the MST naturally define the resolutions at which one should see a change in the number of components. In fact, the MST of a data set contains all the information we need about the ϵ -components.

3.1. Minimal spanning trees

A *graph* is a structure consisting of a finite set of points called *vertices* and a list of pairs of these points called *edges*. A *weighted graph* has a weight or cost assigned to each edge. A graph is *connected* if there is a sequence of edges (a *path*) joining any point to any other point. When it contains no closed paths, a connected graph is called a *tree*. Given a graph, a *spanning tree* is a subgraph that is a tree and contains all the vertices from the original graph. For our purposes, the vertices are data points in a metric space (usually \mathbb{R}^n), the edges are lines joining two points, and the weight of an edge is just the metric distance between the two points joined by that edge.

The MST of a graph is a spanning tree of that graph that has minimal total weight. It is unique when all the edge weights are distinct. For our application, the original graph has an edge between every pair of points, i.e., the *complete graph*. The intuitive way to construct the MST from the complete graph of a set of points is to delete edges as follows. Start with the longest edge (largest weight), and then remove successively shorter ones, provided that doing so leaves the subgraph connected. The algorithm ceases when removing any edge would leave a disconnected graph. An alternative constructive algorithm, due to Prim [28], is more readily implemented on a computer. The initial subtree consists of any point, its nearest neighbor, and the edge between them. The subtree grows by adding the point that has the shortest edge connecting it to the subtree. This step is repeated until all points are in the tree. The cost of this algorithm is $O(n^2)$ where n is the number of points in the set. It is also possible to construct the MST from a special graph called the Delaunay triangulation. For sets in the plane, this results in an algorithm with cost $O(n \log n)$ (see [26] for details).

The property of connecting closest points makes the MST a natural structure for our analysis. This is illustrated by the data set and MST presented in Fig. 1. Notice that the MST bridges the “gap” between two subsets with the shortest edge possible, organizing the point-set in a way that corresponds closely to ϵ -connected components of the set. The following results formalize this intuition.

First, we show that edge lengths in the MST have a close correspondence with ϵ -connectedness and the ϵ -components.

Lemma 2. *If X is a finite set of points that is ϵ -connected for all $\epsilon \geq \epsilon_0 > 0$, then the longest edge in the MST has length $l < \epsilon_0$.*

Conversely, if the MST of a finite set of points, X , has largest edge length $l < \epsilon_0$, then X is ϵ -connected for all $\epsilon \geq \epsilon_0$.

Proof. Recall from Section 2 that X is ϵ -connected if and only if there is an ϵ -chain joining every pair of points in X . This implies that in the complete graph over the points of X , it is possible to move from one vertex to any other along edges of length less than ϵ . Therefore, in the intuitive construction of the MST, all edges of length greater than ϵ_0 are removed, since they could not disconnect the graph. It follows that the MST of X can have no edge longer than ϵ_0 .

Conversely, the MST is a connected graph, so given any points $x, y \in X$, we can find an ϵ_0 -chain: $x_0 = x, \dots, x_n = y$, with $d(x_i, x_{i+1}) < \epsilon_0$. This implies that X is ϵ -connected for all $\epsilon \geq \epsilon_0$. \square

The next result shows that the MST is a suitable structure for representing ϵ -components. First we note that every edge in a MST defines a partition of X (the point-set) into two subsets P and Q containing the points from the two subtrees generated by removing the edge [29]. This results in the following lemma.

Lemma 3. *Removing an edge from an MST generates two subgraphs, each of which is an MST of its points.*

Suppose that we remove the longest edge, of length l , and that this edge length is unique.² Then we are left with two subsets P and Q that are ϵ -connected for some $\epsilon < l$. These subsets genuinely are ϵ -components since $d(P, Q) = l > \epsilon$.

We now have a straightforward way to find the number of ϵ -components by counting edges of the MST. Clearly $C(\epsilon)$ is just one more than the number of edges with lengths, $l \geq \epsilon$. The ϵ -components of the set X correspond directly to the connected components of the subgraph that is generated by removing edges from the MST of lengths, $l \geq \epsilon$. Also, the edge lengths of the MST automatically give the resolutions at which one should see a change in the number of components.

3.2. Some practicalities

The MST, once constructed, holds all the information we need to deduce connectedness properties from ϵ -components. The next task is to extract this information. Trees are frequently used in computation for organizing data effectively. While the MST is a tree, its root is arbitrary and little can be said in general about its branching structure, so it does not lend itself to efficient searches. There is, however, a natural binary structure to the MST, since each edge defines a partition of the MST into two components. This fact can be used to construct a binary tree from the MST, which is then much faster to search.

The binary tree represents information about the MST in the following way. Nodes (vertices) of the binary tree represent edges from the MST, ordered by length; the root is the longest edge. The two children of an edge node are the longest edges of the two sub-MSTs generated by removing that edge. The leaves (nodes with no children) of the binary tree represent the data points; the parent node of a leaf is the shortest edge incident to that point in the MST. In essence, each edge node in the binary tree represents a connected component of the MST. Given a value of ϵ , each ϵ -connected component is represented by an edge node with length less than ϵ , but whose parent has length greater than ϵ . The points in an ϵ -component can be found by listing the leaves “under” its representative node. This binary tree can be built so that each edge node has information about the component it represents and can therefore be searched more quickly than the MST, although it is not necessarily balanced.

As mentioned above, the number of ϵ -components, $C(\epsilon)$, is found by adding one to the number of edges that are longer than ϵ . The diameters of the ϵ -components are computed by listing the elements of an ϵ -component and then

² If there are two or more edges of length l , then the ideas follow through in the obvious way.

finding the diameter of that set of points. Finding the diameter of a set of n points in the plane is an order $O(n \log n)$ algorithm, since the computations can be restricted to points lying on the boundary of the convex hull. For subsets of higher-dimensional spaces, this restriction does not necessarily help and the algorithm we use is the brute-force comparison of distances between all pairs of points, which is $O(n^2)$ [26].

Finally, we must address the problem of how to determine the finest appropriate resolution, ρ , as discussed at the beginning of Section 3. To do this, we examine how the number of isolated points in the ϵ -decomposition of the set X varies with resolution, i.e., the function $I(\epsilon)$. A point, x , is isolated at resolution ϵ if $d(x, X \setminus x) \geq \epsilon$. In terms of the MST, a point is isolated at resolution ϵ if all edges incident to it have length longer than ϵ . In all of the examples below, the underlying sets are perfect, so the finite point-set approximation is “bad” at any resolution for which there are isolated points. It follows that the resolution at which we start to see isolated points is one way to estimate ρ , i.e., $\rho = \inf\{\epsilon : I(\epsilon) = 0\}$. The validity of this approach is supported by the numerical evidence given in Section 4; the data for $C(\epsilon)$ and $D(\epsilon)$ blur at the resolution at which isolated points are first detected.

4. Examples

In this section we present some examples that illustrate the behavior of the number of ϵ -components, $C(\epsilon)$, the largest diameter, $D(\epsilon)$, and the number of isolated points, $I(\epsilon)$, for fractals with different topology. The goal is to show that these quantities give useful information about the topology. The first examples are relatives of the Sierpinski triangle. These sets are generated from a family of iterated function systems (IFS). The Hausdorff dimension of each set is identical, even though they have different topological structure, as the disconnectedness and discreteness indices highlight. For these fractals, we show that the cutoff resolution decreases when the number of data points is increased, which is not surprising since more points sampled from an attractor constitute a better approximation of the underlying set. We also vary the way in which the data cover the set and find that for a fixed number of points the cutoff resolution is a minimum when the data is uniformly distributed. Again, this is exactly what we expect, since isolated points appear at larger values of ϵ when points are not evenly spaced.

We next present a number of Cantor set examples to illustrate different types of scaling behavior in $C(\epsilon)$ and $D(\epsilon)$. We observe that, for Cantor sets with zero Lebesgue measure, the computed value of γ is approximately equal to the dimension of the set. For a Cantor set with positive measure, though, γ and the dimension are significantly different. Cantor sets and other fractal structures are frequently observed in dynamical systems. We present $C(\epsilon)$ and $D(\epsilon)$ results for some representative examples, including the Hénon attractor and “cantori” from four-dimensional symplectic sawtooth maps. We investigate three cross-sections of the Hénon attractor, and present numerical evidence that the sections are Cantor sets, as is widely believed. Finally, we show that the cantori exhibit logarithmic rather than polynomial growth in the number of components, implying that even though $C(\epsilon) \rightarrow \infty$ as $\epsilon \rightarrow 0$, $\gamma = 0$. In all cases, we use the Euclidean metric.

4.1. Relatives of the Sierpinski triangle

The Sierpinski triangle relatives are attracting fixed sets of a family of iterated function systems

$$S = f[S] = f_1[S] \cup f_2[S] \cup f_3[S].$$

In each case, the functions f_i are similarity transformations³ of the unit square with contraction ratio $\frac{1}{2}$, as illustrated by the template in Fig. 2. The functions that generate the Sierpinski triangle, Fig. 3, are simple contractions composed

³ A similarity transformation, S , is an affine transformation that contracts or dilates distance uniformly, i.e., for all x and y , there is a positive number r such that $|S(x) - S(y)| = r|x - y|$.

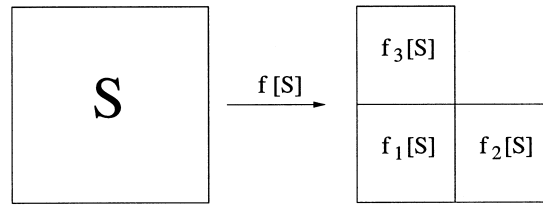


Fig. 2. Template for the iterated function system that generates the Sierpinski triangle relatives.

with a translation; the generators of the examples in Figs. 8 and 10 involve additional rotation or reflection symmetries of the square. There are 232 different fractals in this family [30]; their topology ranges from simply connected to totally disconnected to a class of examples with infinitely many connected components of nonzero diameter [13]. This range of topologies makes them ideal test cases for our techniques.

It is easy to generate a finite number of points on the attractor of an iterated function system. One way is to choose an initial point x_0 in the domain of the IFS and then record its trajectory under the iteration $x_{n+1} = f_{i_n}(x_n)$ setting $i_n = 1, 2$ or 3 with probability p_1, p_2 and p_3 , respectively. If x_0 is in the attractor then its entire orbit is in the attractor; if not, the iterates converge to it. In the examples below we choose $x_0 = (0, 0)$ which is in each of the attractors. Thus, the orbit can be viewed as a random sampling of the attractor by a finite number of points. When $p_1 = p_2 = p_3 = \frac{1}{3}$ the data cover the fractal uniformly; if the probabilities are not equal the distribution of points is nonuniform and their density approximates a multifractal measure.

4.1.1. The Sierpinski triangle

The generating functions for the Sierpinski triangle are

$$\begin{aligned}
 f_1(x, y) &= \frac{1}{2}(x, y), \\
 f_2(x, y) &= \frac{1}{2}(x + 1, y), \\
 f_3(x, y) &= \frac{1}{2}(x, y + 1).
 \end{aligned}
 \tag{1}$$

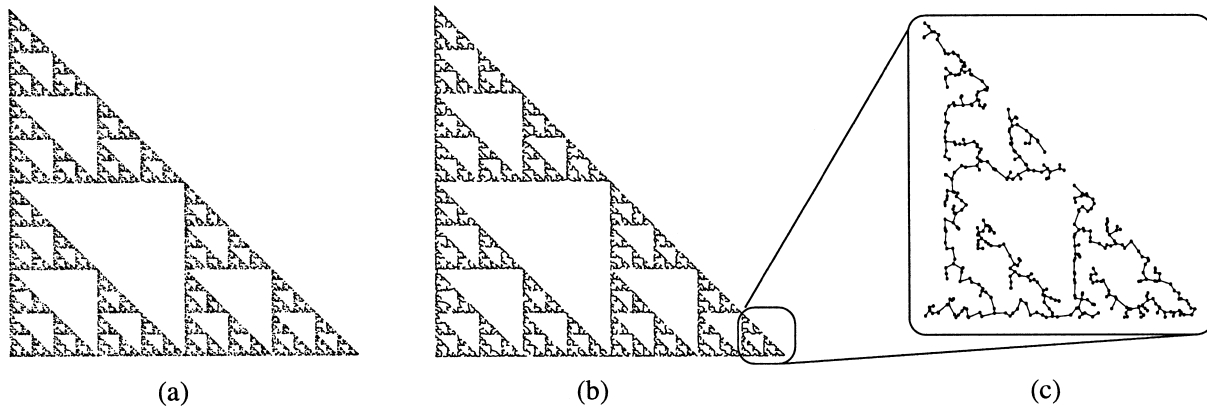


Fig. 3. (a) 10^4 points uniformly distributed on the Sierpinski triangle, (b) the corresponding MST, and (c) a close-up of the bottom right corner of the MST.

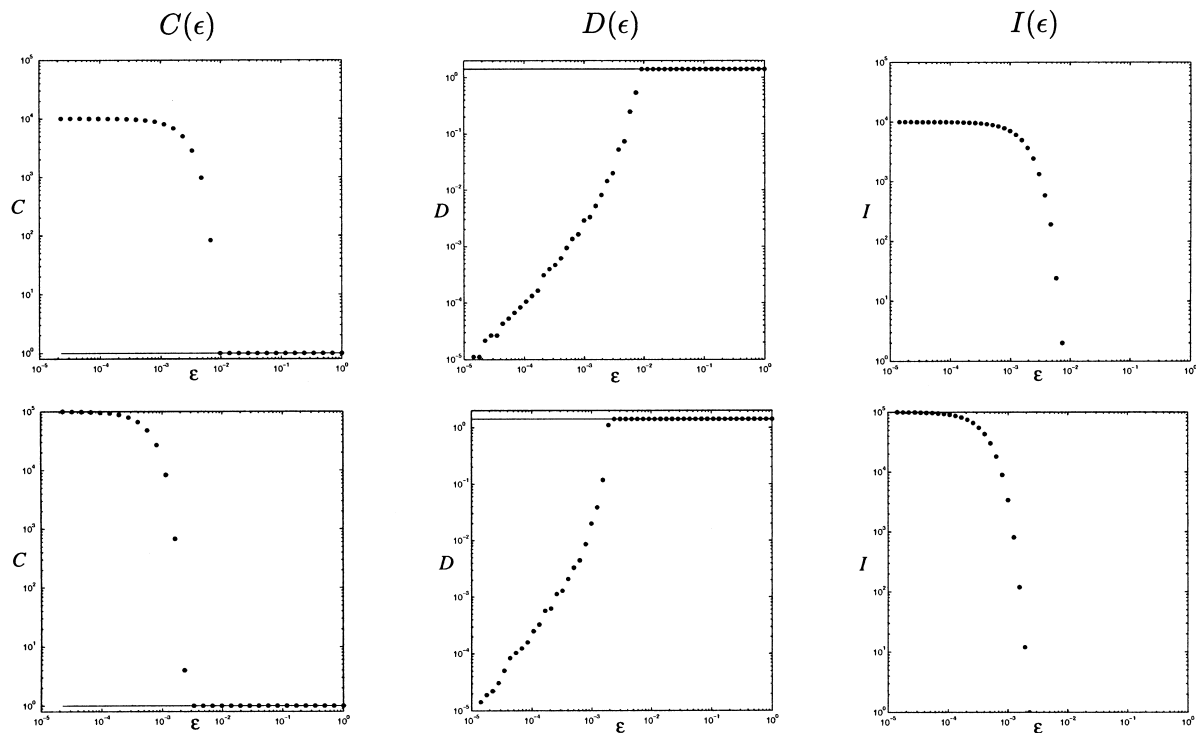


Fig. 4. $C(\epsilon)$, $D(\epsilon)$ and $I(\epsilon)$ for the Sierpinski triangle. The top row gives results for 10^4 uniformly distributed points on the fractal and the bottom row for 10^5 points. All axes are logarithmic. The horizontal axis range is $10^{-5} < \epsilon < 1$. The solid lines represent $C(\epsilon)$ and $D(\epsilon)$ for ideal data; the dots are the computed values.

A finite point-set approximation to the triangle and the corresponding minimal spanning tree are shown in Fig. 3. The underlying set is connected and perfect, so we expect to see $C(\epsilon) = 1$, $D(\epsilon) = \sqrt{2}$, and $I(\epsilon) = 0$ for $\epsilon > \rho$. This is reflected by the calculations of $C(\epsilon)$ and $D(\epsilon)$ for 10^4 and 10^5 point approximations to the triangle, as shown in Fig. 4. We see that for ϵ above a threshold value, the computed values of $C(\epsilon)$ and $D(\epsilon)$ are in exact agreement with our expectations. The point at which $C(\epsilon)$ and $D(\epsilon)$ deviate from the ideal values is the value of ϵ at which the number of isolated points, $I(\epsilon)$, becomes positive. This ϵ value is, of course, the cutoff resolution ρ discussed in Section 3.2. At finer resolutions, i.e., $\epsilon < \rho$, we see a sharp transition in the number of connected components from one to the number of points in the set; the diameters show a correspondingly sharp decrease. Both these effects are due to the narrow distribution of edge lengths of the MST. Clearly, the value of ρ depends on the number of points, N , covering the set. For the 10^4 point approximation, $\rho \approx 0.008$ and for $N = 10^5$, $\rho \approx 0.0022$; We expect the relationship to be $\rho \approx 1/\sqrt{N}$, since the data is homogeneously distributed on a subset of \mathbb{R}^2 . This is supported by the data in Fig. 5a. Here, we plot cutoff resolution versus the number of points for $10^3 \leq N \leq 10^5$; the slope of the least-squares fit line is -0.58 .

The results discussed so far are for uniformly distributed data; we now look at nonuniformly distributed data. As described earlier, we change the way an orbit covers the IFS attractor by choosing the functions f_1 , f_2 , and f_3 with different probabilities. To generate Fig. 6a, we set $p_1 = 0.05$ and $p_2 = p_3 = 0.475$. This highly nonuniform distribution of points induces perceptible changes in the $C(\epsilon)$, $D(\epsilon)$ and $I(\epsilon)$ data, Fig. 7, but the graphs remain qualitatively similar to those in Fig. 4. The cutoff resolution is significantly larger: $\rho \approx 0.04$ compared with 0.008 for the uniform distribution with the same number of points. The growth in the number of ϵ -components for $\epsilon < \rho$

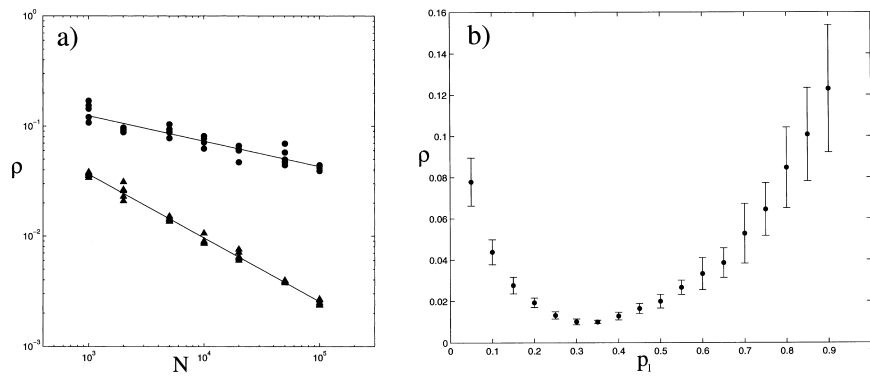


Fig. 5. (a) Cutoff resolution, ρ , as a function of the number of points, $10^3 \leq N \leq 10^5$, covering the Sierpinski triangle for two values of p_1 : (●) marks data for the nonuniform distribution with $p_1 = 0.05$; (▲) marks data for $p_1 = \frac{1}{3}$, i.e., a uniform distribution, and (b) cutoff resolution as a function of p_1 for 10^4 data points on the Sierpinski triangle. The error bars are the standard deviation about the mean of 20 calculations of ρ for each value of p_1 .

is also less rapid than that for the uniform data. Both these changes are due to a greater spread in the edge lengths of the MST. The geometry of the distribution is reflected in the graph of $D(\epsilon)$; the densely covered diagonal means $D(\epsilon) = \sqrt{2}$ for ϵ -values significantly less than ρ . We can lower the cutoff resolution by increasing the number of data points but, as shown in Fig. 5a, the rate at which ρ decreases is now $\rho \approx N^{-0.23}$, significantly slower than that for the uniform data.

Finally, the graph in Fig. 5b summarizes the data from a systematic study of the way cutoff resolution varies with the distribution of the data. The measure of nonuniformity in the data is p_1 , the probability of choosing f_1 ; we set $p_2 = p_3 = \frac{1}{2}(1 - p_1)$. Twenty data sets of 10^4 points were generated for values of p_1 in the range 0.05–0.9. The cutoff resolution reaches a minimum at about $p_1 = \frac{1}{3}$, i.e., for uniformly distributed data, as we expect. The other feature to note is that the standard deviation also depends on the distribution, and is greatest for highly nonuniform data. This is because when one function in the IFS is chosen with very low probability, there is greater variability in the way an orbit fills out the attractor; this in turn leads to greater variation in the edge lengths of the MST.

We conclude that for moderate amounts of nonuniformity (e.g., $0.2 \leq p_1 \leq 0.5$) the cutoff resolution is at a level comparable to that for perfectly uniform data and that our techniques are not adversely affected. For highly nonuniform coverings of an attractor, significantly more data points are needed to reach the same cutoff resolutions as for uniform data. The only effect this has is to generate inconclusive, rather than incorrect, diagnoses of the topology of the underlying set.

4.1.2. A Cantor set relative

Fig. 8 shows the attractor for the iterated function system generated by

$$\begin{aligned}
 f_1(x, y) &= \frac{1}{2}(-y + 1, x), \\
 f_2(x, y) &= \frac{1}{2}(y + 1, x), \\
 f_3(x, y) &= \frac{1}{2}(y, -x + 2).
 \end{aligned}
 \tag{2}$$

This fractal is a Cantor set, so we should see $C(\epsilon) \rightarrow \infty$ and $D(\epsilon) \rightarrow 0$ as $\epsilon \rightarrow 0$. As derived in [13], the functional relationships are

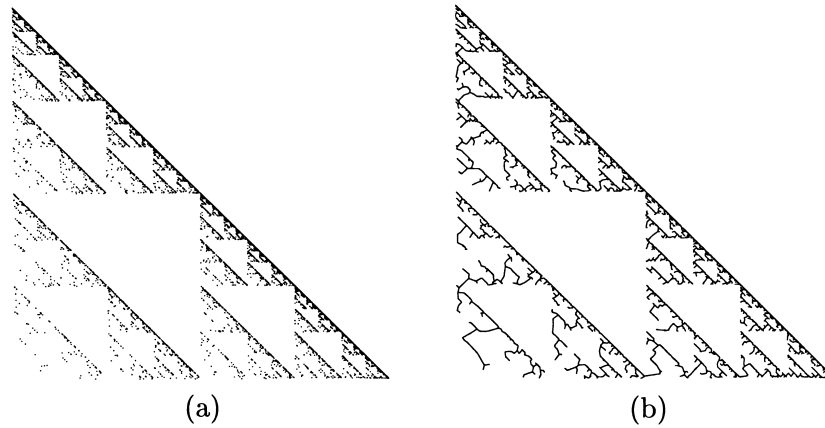


Fig. 6. (a) 10^4 points on the Sierpinski triangle generated by setting $p_1 = 0.05$ and $p_2 = p_3 = 0.475$, and (b) the corresponding MST.

$$\epsilon_n = \frac{\epsilon_0}{2^n}, \quad C(\epsilon_n) = \begin{cases} 3^n + 2 \cdot 3^{(n-1)/2} & \text{if } n \text{ is odd,} \\ 3^n + 3^{n/2} & \text{if } n \text{ is even,} \end{cases} \quad D(\epsilon_n) = \frac{D_0}{2^{n-1}} \quad \text{for } n \geq 3.$$

These expressions give $\gamma = \log 3 / \log 2 \approx 1.585$ and $\delta = 1$.

We can see in Fig. 9 that the numerical calculations agree very well with the theory down to the cutoff resolution $\rho \approx 0.003$. When $\epsilon < \rho$, the computed values of $C(\epsilon)$ are larger than the predicted values because isolated points are counted as extra components. For still smaller values of ϵ , every point is resolved as an isolated point and the $C(\epsilon)$ curve levels off. The meaningful portion of the data — between these extremes — shows a staircase periodicity about a linear trend. The slope of the linear trend is an estimate of γ . We determine γ numerically, using a least-squares fit, to be 1.41 ± 0.05 . This is lower than the true limiting value given above (1.58) because of second-order effects at the relatively large values of ϵ for which the $C(\epsilon)$ data are valid. We estimate the slope of the true curve over the same range to be 1.48, which is closer to the value computed above.

The numerically calculated values of $D(\epsilon)$ also show a staircase periodicity about a linear trend. The data have a systematic bias for the jumps at slightly larger values of ϵ than predicted by theory. This is due to finite data

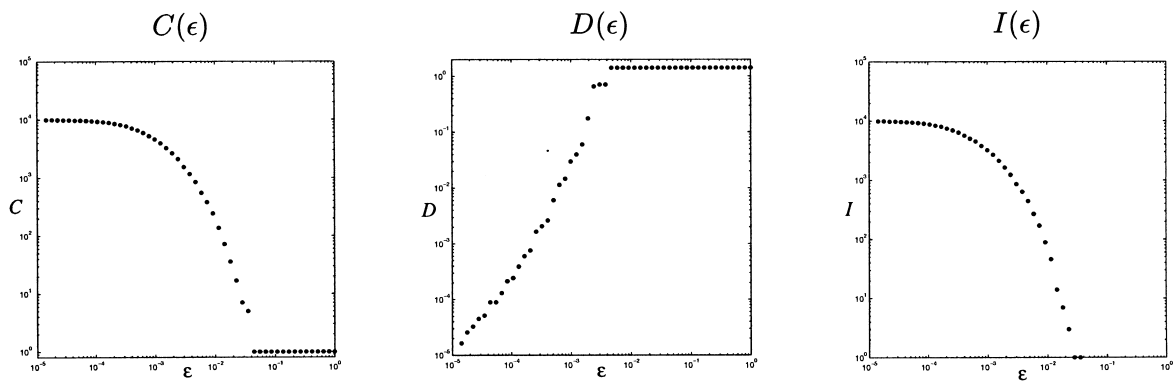


Fig. 7. $C(\epsilon)$, $D(\epsilon)$ and $I(\epsilon)$ for the nonuniformly distributed data set in Fig. 6. All axes are logarithmic. The horizontal axis range is $10^{-5} < \epsilon < 1$.

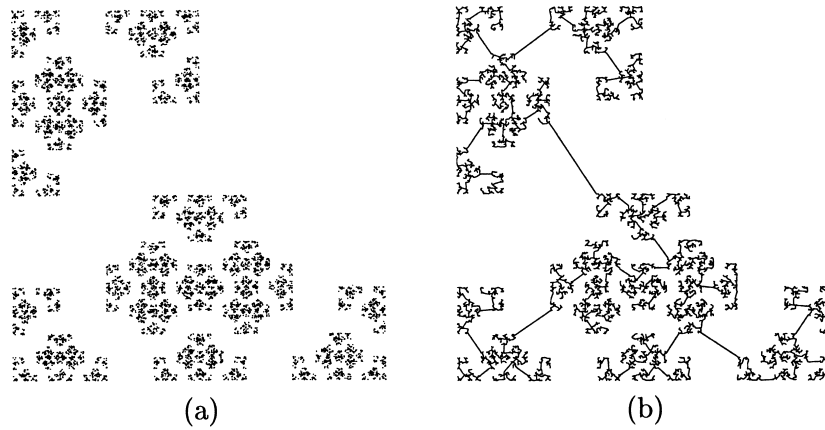


Fig. 8. (a) 10^4 points on the Cantor set generated by (2), and (b) the corresponding MST.

effects: the points are not filling out the corners, so edges in the MST are a little longer than the “true” gaps and the diameters of the ϵ -components are a little less. The data points fall below the true curve when $\epsilon < \rho$ because interpoint distances are comparable to the intercomponent distances at these resolutions. Since $D(\epsilon)$ measures the largest diameter, the flat tails of the $D(\epsilon)$ data in Fig. 9 are due to the presence of at least one triple/pair combination within a distance ϵ of each other, whilst almost all the other points have become isolated. The slope of the linear trend of $D(\epsilon)$ is an estimate of δ . We estimate the slope over the range $\rho < \epsilon < 0.06$ because the first few steps are shallower than the limiting trend. Using a least-squares fit, again, we calculate δ to be 1.00 ± 0.03 , as predicted above.

The cutoff resolution for data from this IFS varies with the number of data points and their distribution in exactly the same manner as the data from the Sierpinski triangle. For moderately nonuniform data, estimates of γ and δ are the same as those given above. When the data is very unevenly distributed ρ increases significantly and there may be too small a range of ϵ to make an estimate of the slope. Again, this leads to inconclusive results rather than incorrect ones.

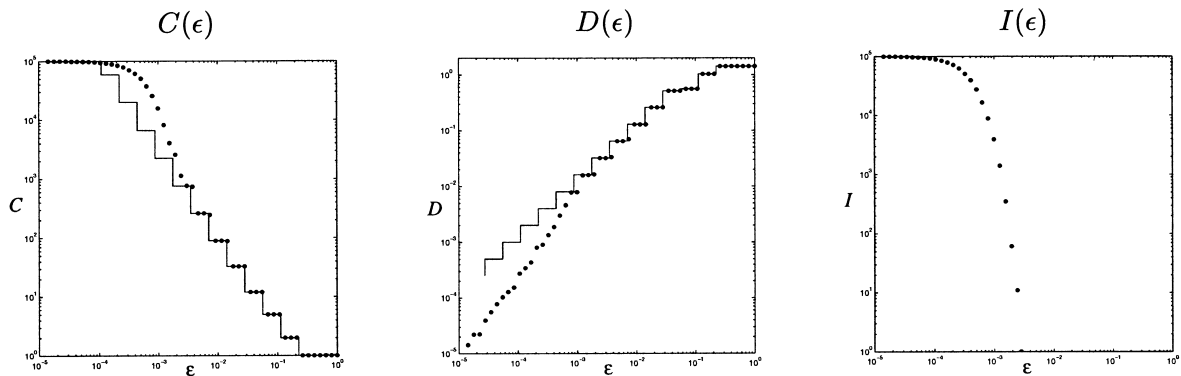


Fig. 9. $C(\epsilon)$, $D(\epsilon)$ and $I(\epsilon)$ for 10^5 points uniformly distributed over the Cantor set triangle relative. All axes are logarithmic. The horizontal axis range is $10^{-5} < \epsilon < 1$. The solid lines represent $C(\epsilon)$ and $D(\epsilon)$ for ideal data; the dots are the computed values.

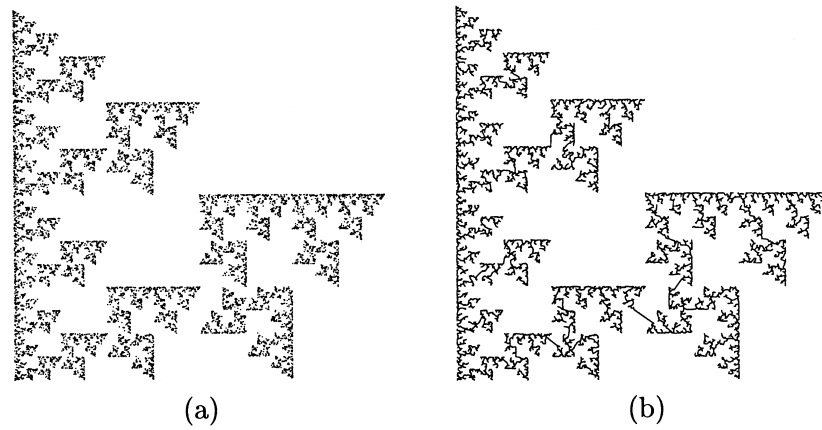


Fig. 10. (a) 10^4 points on the fractal generated by (3), and (b) the corresponding MST.

4.1.3. A relative with infinitely many connected components

A third triangle relative, shown in Fig. 10, is generated by the following similarities:

$$\begin{aligned}
 f_1(x, y) &= \frac{1}{2}(x, y), \\
 f_2(x, y) &= \frac{1}{2}(y + 1, -x + 1), \\
 f_3(x, y) &= \frac{1}{2}(x, y + 1).
 \end{aligned}
 \tag{3}$$

The attractor for this system has infinitely many connected components, yet is not totally disconnected because the components have positive diameters. We, therefore, expect to see $C(\epsilon) \rightarrow \infty$ and $D(\epsilon) \rightarrow 1$ as $\epsilon \rightarrow 0$. As we derive in [13]

$$\epsilon_n = \frac{\epsilon_0}{2^n}, \quad C(\epsilon_n) = \frac{1}{2}(3^{n+1} + 1), \quad D(\epsilon_n) \rightarrow 1.$$

The graphs of $C(\epsilon)$ and $D(\epsilon)$ in Fig. 11 reflect this; the former has characteristics similar to those of the Cantor set above, but $D(\epsilon)$ looks like that for the Sierpinski triangle. The slope of the $C(\epsilon)$ staircase is estimated from the

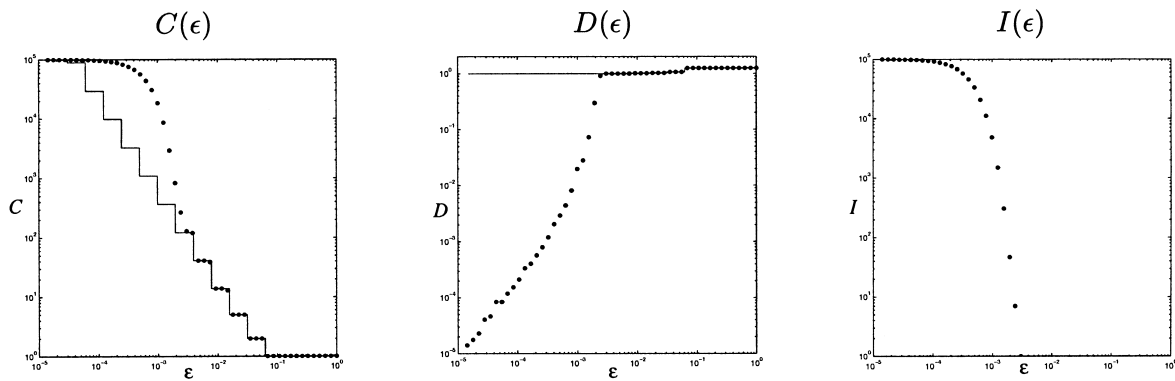


Fig. 11. $C(\epsilon)$, $D(\epsilon)$ and $I(\epsilon)$ for a triangle relative with infinitely many components. Again, the data is for 10^5 points uniformly distributed on the set. All axes are logarithmic. The horizontal axis is $10^{-5} < \epsilon < 1$. The solid line represents $C(\epsilon)$ and $D(\epsilon)$ for ideal data; the dots are the computed values.

Table 1

This table summarizes values of γ and δ for Cantor subsets of the plane^a

Data set	\dim_B	γ	δ
Fig. 12a	1.262	1.23 ± 0.02	0.96 ± 0.04
Fig. 12b	1.126	1.11 ± 0.02	1.00 ± 0.03
Fig. 14a	1.131	1.13 ± 0.01	0.98 ± 0.03
Fig. 14b	2	0.80 ± 0.05	0.46 ± 0.05
Fig. 16	$1.21 < \dim_B < 1.34$	1.36 ± 0.03	0.95 ± 0.05

^aThe numbers are estimated using a least-squares linear fit to logarithmic plots of $C(\epsilon)$ and $D(\epsilon)$, respectively; the error margins are estimated by varying the scaling range; the second column gives the box-counting dimension, \dim_B , for each set; these numbers are computed using formulas from Falconer [32].

data for a 10^5 point approximation, giving a value of $\gamma = 1.55 \pm 0.03$. This is in very close agreement with the theoretical value of $\gamma = \log 3 / \log 2 \approx 1.585$.

4.2. Cantor sets in the plane

One of our objectives is to use our techniques to identify and characterize phase space structures in dynamical systems. Cantor sets are often present in chaotic dynamical systems, so it is useful to examine some simple Cantor set examples to gain a better understanding of the different types of scaling that can occur in the $C(\epsilon)$ and $D(\epsilon)$ graphs. In Figs. 12, 14 and 16, we show five Cantor sets in the plane. In each case, the orbit has 50 000 points. Four of these have zero Lebesgue measure and one (Fig. 14b) has positive measure, so it is termed a fat Cantor set (this is analogous to the term “fat fractal” for fractals with positive measure [31]). All are attractors of iterated function systems of the form

$$S = f[S] = f_1[S] \cup f_2[S] \cup f_3[S] \cup f_4[S].$$

The generating functions, f_i , become increasingly complex in this series of examples. The three simplest involve only affine transformations and another uses conformal functions; the functions that generate the fat Cantor set cannot be written in closed form. The geometric structure of each set is reflected in the type of staircase seen in the graphs of $C(\epsilon)$ and $D(\epsilon)$. For the four examples with zero Lebesgue measure, we expect to see $\gamma = \dim_B$, the box-counting dimension; this is supported by our results, summarized in Table 1. For the Cantor set with positive measure, the value of γ is significantly different from the dimension. We again observe, for all of the examples, that the cutoff resolution, ρ , is well approximated by the ϵ -value where the number of isolated points ceases to be zero. Again, this is because all of the underlying sets are perfect.

We start with a simple example where each f_i is a similarity transformation with contraction ratio $\frac{1}{3}$, as shown in Fig. 12a. The numerical calculations of $C(\epsilon)$, $D(\epsilon)$, and $I(\epsilon)$ are presented in the top row of Fig. 13. These graphs show staircase scaling behavior similar to the triangle relative Cantor set presented in Section 4; this makes sense because both are generated by iterated function systems of similarity transformations. For this example, the jumps in $C(\epsilon)$ and $D(\epsilon)$ are at $\epsilon_n = 1/3^n$; because there are four self-similar copies at one-third the size, we see $C(\epsilon_n) = 4^n$ and $D(\epsilon_n) = 1/3^n$, which gives the theoretically determined limits of $\gamma = \log 4 / \log 3 \approx 1.262$ and $\delta = 1$. This is in close agreement with our numerical estimate of $\gamma = 1.23 \pm 0.05$ and $\delta = 0.97 \pm 0.02$.

The second example, Fig. 12b, is also generated by similarities. This time, the lower two have a contraction ratio of $\frac{1}{3}$ and the upper two have a ratio of $\frac{1}{4}$. As can be seen from the second row of Fig. 13, this leads to a more complicated staircase pattern in the $C(\epsilon)$ and $D(\epsilon)$ graphs. Jumps in these graphs occur at values of ϵ corresponding to edge lengths in the MST. The structure from the IFS means these edge lengths are of the form $l(\frac{1}{3})^m(\frac{1}{4})^n$, for all

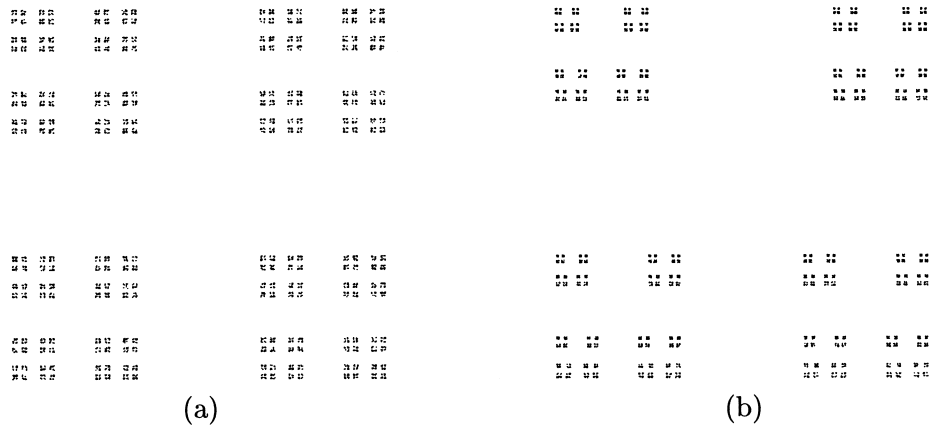


Fig. 12. Cantor sets generated by iterated function systems of four similarity transformations. Both sets have 50 000 points: (a) similarities with contraction ratio $\frac{1}{3}$, and (b) the upper two similarities have ratio $\frac{1}{4}$ and the lower two have ratio $\frac{1}{3}$.

integers m and n , where l is one of the two longest edges. Values of γ and δ , presented in Table 1, are again very close to the expected values.

To generate the set in Fig. 14a, more general affine transformations are used, each contracting by $\frac{1}{3}$ horizontally and $\frac{1}{4}$ vertically. The corresponding graphs of $C(\epsilon)$ and $D(\epsilon)$ in Fig. 15 show the now familiar staircase scaling

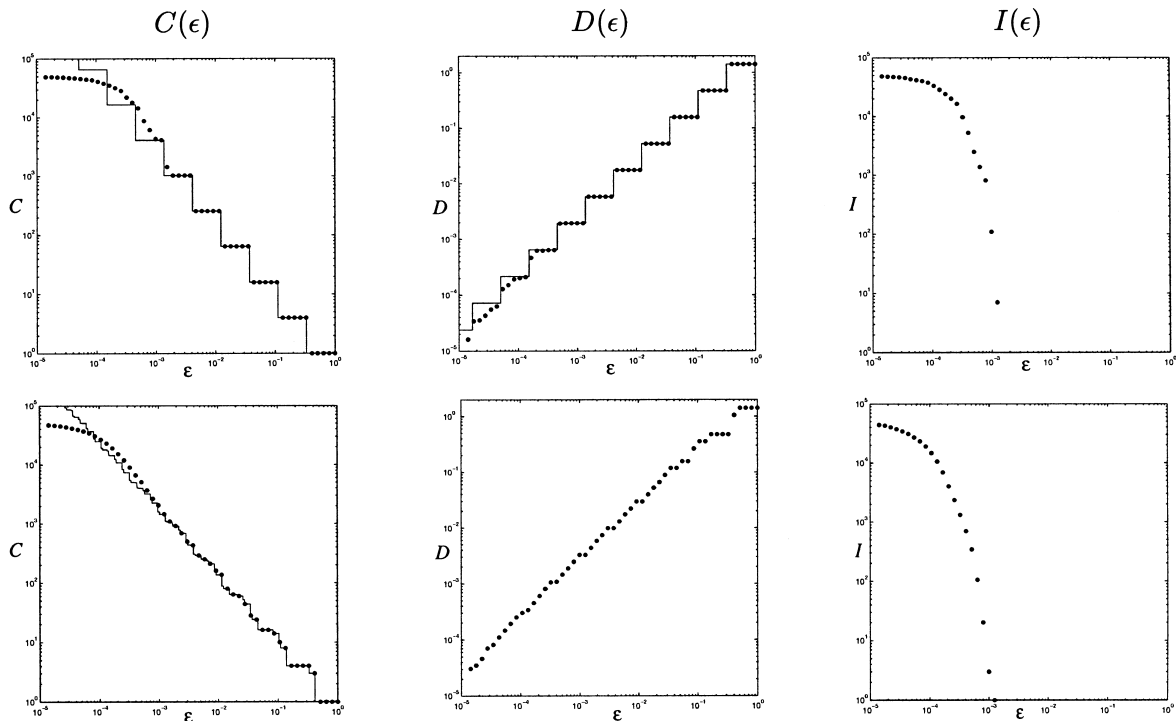


Fig. 13. $C(\epsilon)$, $D(\epsilon)$ and $I(\epsilon)$ for the Cantor sets in Fig. 12. The top row is data for Fig. 12a; the second row is for Fig. 12b. All axes are logarithmic. The horizontal axis range is $10^{-5} < \epsilon < 1$. The solid lines represent $C(\epsilon)$ and $D(\epsilon)$ for ideal data; the dots are the computed values.

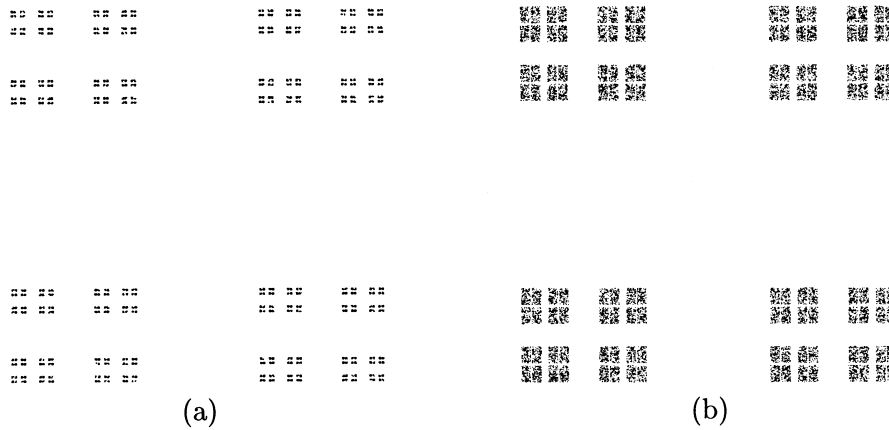


Fig. 14. Two Cantor sets with largest gaps of $\frac{1}{2}$ and $\frac{1}{3}$: (a) a set generated by an IFS of four affine transformations with horizontal contraction of $\frac{1}{3}$ and vertical contraction of $\frac{1}{4}$, and (b) a fat Cantor set, generated as the cross product of two Cantor sets of positive measure in the real line.

pattern. Compared to the second Cantor set example, the larger steps in these graphs reflect the more regular geometric structure of the set.

The fourth example is a Cantor set with positive Lebesgue measure and therefore a dimension of two. It is possible to represent this set as the attractor of an iterated function system of the general form above. The functions involved,

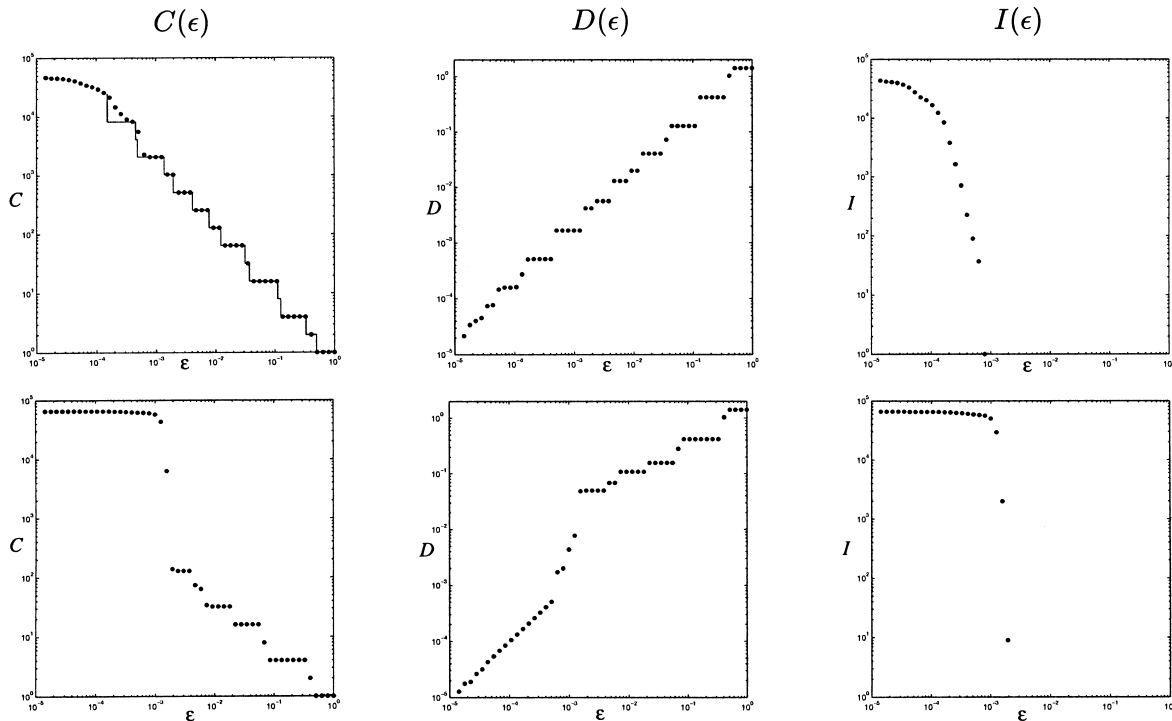


Fig. 15. $C(\epsilon)$, $D(\epsilon)$ and $I(\epsilon)$ for the 2D Cantor sets in Fig. 14. The top row is data for Fig. 14a; the second row for Fig. 14b, the fat Cantor set. All axes are logarithmic. The horizontal axis range is $10^{-5} < \epsilon < 1$.

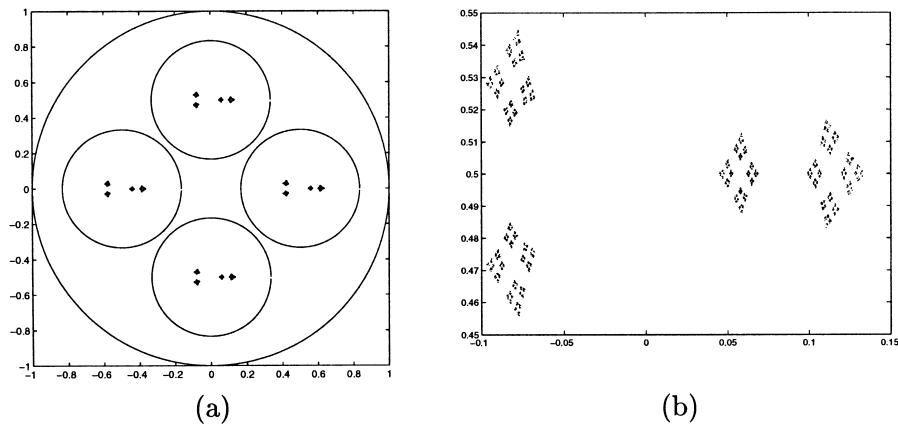


Fig. 16. A Cantor set generated by an IFS consisting of four nonlinear affine transformations, each mapping the unit circle into a circle of radius $\frac{1}{3}$: (a) the data set with circle boundaries, and (b) a close-up of one of the four clusters.

however, are limits of piecewise linear approximations and it is not possible to write them in closed form. Instead, we generate the set as the cross product of two positive measure Cantor subsets of the unit interval. These sets are constructed as follows: at each level, $n \geq 1$, 2^{n-1} gaps of length $a/2^{p(n-1)}$ are removed from the center of an interval remaining from level $n - 1$. The sum of the gap lengths is $a/(2^p - 1)$; choosing p and a to make this length less than one ensures the Cantor set has positive measure. It is easy to recursively generate the end points of the gaps (down to some level) and these points are used as the finite point-set approximation. For the set in Fig. 14b, we set $a = \frac{2}{3}$ and $p = 2$ for the horizontal cross-section, and $a = 2$, $p = 3$ for the vertical one. The behavior of $C(\epsilon)$ and $D(\epsilon)$, shown in the bottom row of Fig. 15, is not unlike that of the previous example. The slopes are significantly shallower, however, because the gaps are decreasing at a faster rate than the component diameters, giving $\delta \approx 0.46$. The component growth rate, γ , is approximately 0.80, which is clearly distinct from the box-counting (and Hausdorff) dimension of two.

The final IFS example uses nonlinear, conformal transformations. A function, F , is conformal if its derivative matrix at each point, $DF(x)$ is a similarity transformation. For the set illustrated in Fig. 16, the functions are

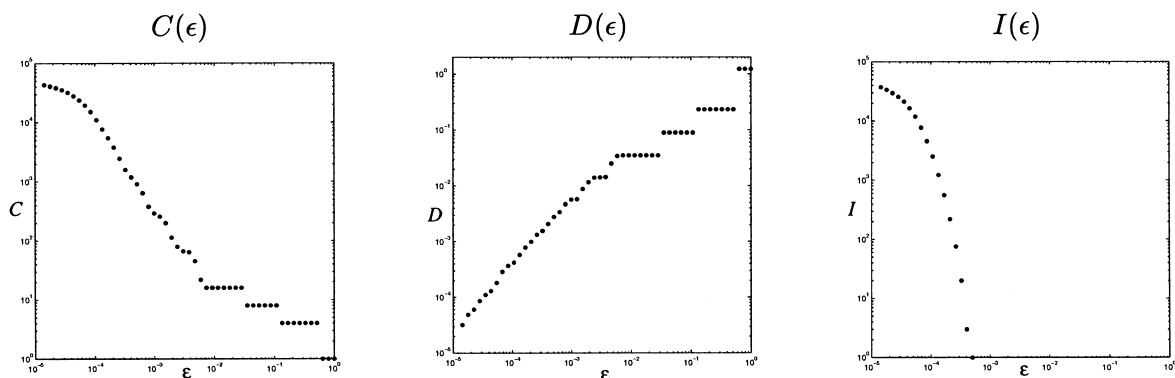


Fig. 17. $C(\epsilon)$, $D(\epsilon)$ and $I(\epsilon)$ for the nonlinear Cantor set of Fig. 16. All axes are logarithmic. The horizontal axis range is $10^{-5} < \epsilon < 1$.

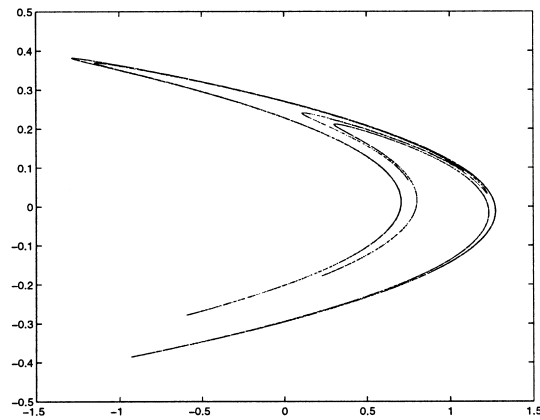


Fig. 18. An orbit on the Hénon attractor.

of the form $f_k(z) = \frac{1}{3}z^2 + c_k$, where $z = x + iy$, and the translations, c_k for $k = 1, \dots, 4$, take the values $\{\pm\frac{1}{2}, \pm\frac{1}{2}i\}$. Notice that although we choose f_i with equal probability, the nonlinearity introduces a nonuniformity to the distribution of points over the Cantor set. The cutoff resolution $\rho \approx 5 \times 10^{-4}$ is nevertheless comparable to the previous examples with uniformly distributed data. Scaling in the graphs of $C(\epsilon)$ and $D(\epsilon)$ occurs in two distinct ϵ intervals, see Fig. 17. For $0.005 < \epsilon < 1$, there are three shallow steps reflecting the large-scale structure that is visible in Fig. 16a. The second portion of the data, for $\rho < \epsilon < 0.005$, has a steeper slope, corresponding to the limiting small-scale structure of the set. The values of γ and δ given in Table 1 are slopes of the $C(\epsilon)$ and $D(\epsilon)$ over the interval $\rho < \epsilon < 0.005$. We find, as for the previous zero-measure Cantor sets, that γ is close to the box-counting dimension and $\delta \approx 1$.

4.3. Examples from dynamical systems

Our last two examples are Cantor sets from iterated maps.

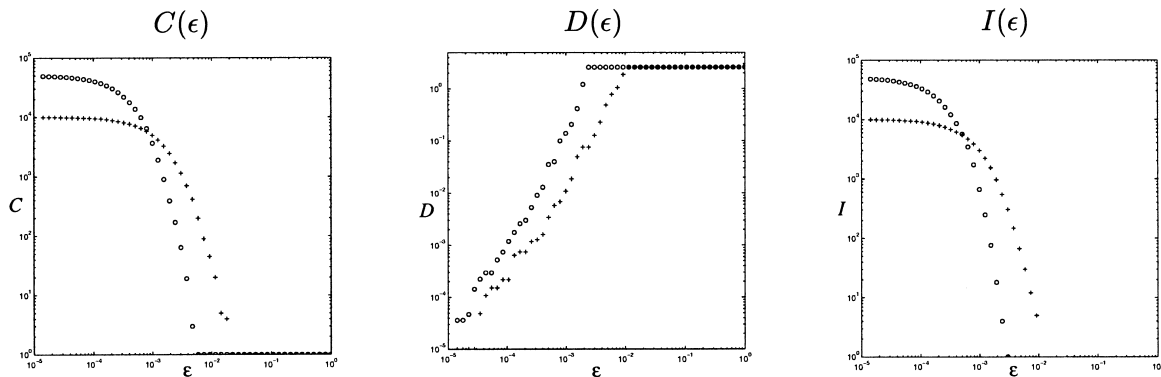


Fig. 19. $C(\epsilon)$, $D(\epsilon)$ and $I(\epsilon)$ data for two orbits on the Hénon attractor. The crosses (+) represent calculations for the orbit of 10^4 iterates and the circles (O) are for an orbit with 5×10^4 points. All axes are logarithmic. The horizontal axis range is $10^{-5} < \epsilon < 1$.

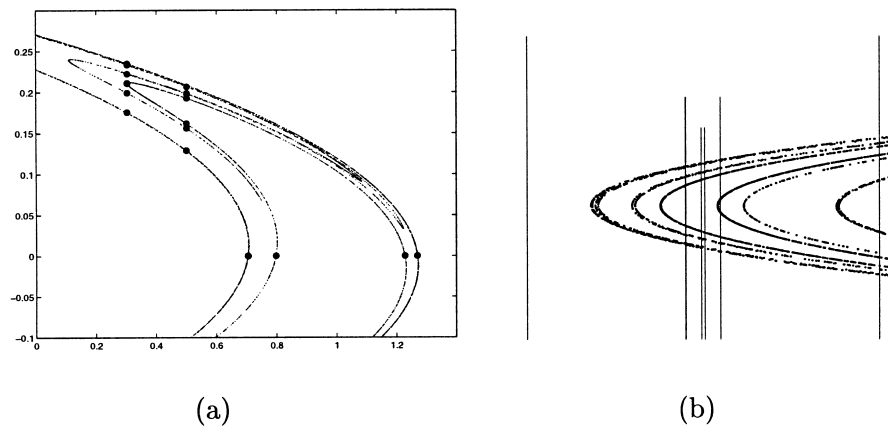


Fig. 20. (a) A close-up of the Hénon attractor. The dark spots are points in the three cross-sections considered in the text: slices at $x = 0.302435$, $x = 0.5$ and $y = 0$, and (b) a small part of the slice at $x = 0.302435$, $y = 0.22$ that shows the folding of the attractor. The pairs of vertical lines are the boundaries of the different subslices of widths 2×10^{-5} , 2×10^{-6} and 2×10^{-7} .

4.3.1. The Hénon attractor

Fig. 18 shows the well-known Hénon attractor, generated by iterating the map

$$\tilde{x} = y + 1 - ax^2, \quad \tilde{y} = bx, \quad (4)$$

with parameter values $a = 1.4$ and $b = 0.3$.

This attractor is the closure of the unstable manifold of a fixed point [33]. It follows that the set has a topological dimension of one and must be connected. This is exactly what the $C(\epsilon)$, $D(\epsilon)$ and $I(\epsilon)$ data in Fig. 19 reflects. The attractor has fractal structure nonetheless, and is often described as having a Cantor set cross-section [33]. To confirm this intuition, in Fig. 21 we present $C(\epsilon)$, $D(\epsilon)$ and $I(\epsilon)$ data for thin slices taken through the attractor at three different places: $x = 0.302435$, $x = 0.5$ and $y = 0$ (the dark spots in Fig. 20a).

A common technique for visualizing the structure of attractors from *flows* is the Poincaré section. A surface of codimension one is chosen and points on the section are recorded whenever the trajectory pierces the surface. The Hénon attractor is generated by iterating a map, so it is not possible to find many points on a given section. Instead, the sections are generated by recording points that fall within an interval of the given section coordinate. This means that the slices have a finite width and the data are still two-dimensional. To be confident that the observed scaling behavior is approximating that of a one-dimensional Cantor set, we compute $C(\epsilon)$, $D(\epsilon)$ and $I(\epsilon)$ for four successively narrower slices at each cross-section. The thinnest slice in each case has a width of 2×10^{-7} .

The section at $x = 0.302435$ is interesting because it cuts through a fold in the attractor. The folding of the Hénon attractor is the source of its nonhyperbolicity and complexity. A close-up of the fold is shown in Fig. 20b. This figure shows that slices of different widths taken at this x value capture different folding structure. This is reflected in the $C(\epsilon)$ data in Fig. 21. The data for different slices at $x = 0.302435$ does not coincide exactly for $\epsilon > \rho$, as it does for the other two sections at $x = 0.5$ and $y = 0$, which show no folding at these resolutions. Note that we can reverse this observation and use it to detect cross-sections that touch a fold. The folding of the attractor presents a slight problem for the description of the set as having Cantor set cross-section. If a genuine one-dimensional cross-section touches a fold, then the point of tangency must be an isolated point. It follows that the

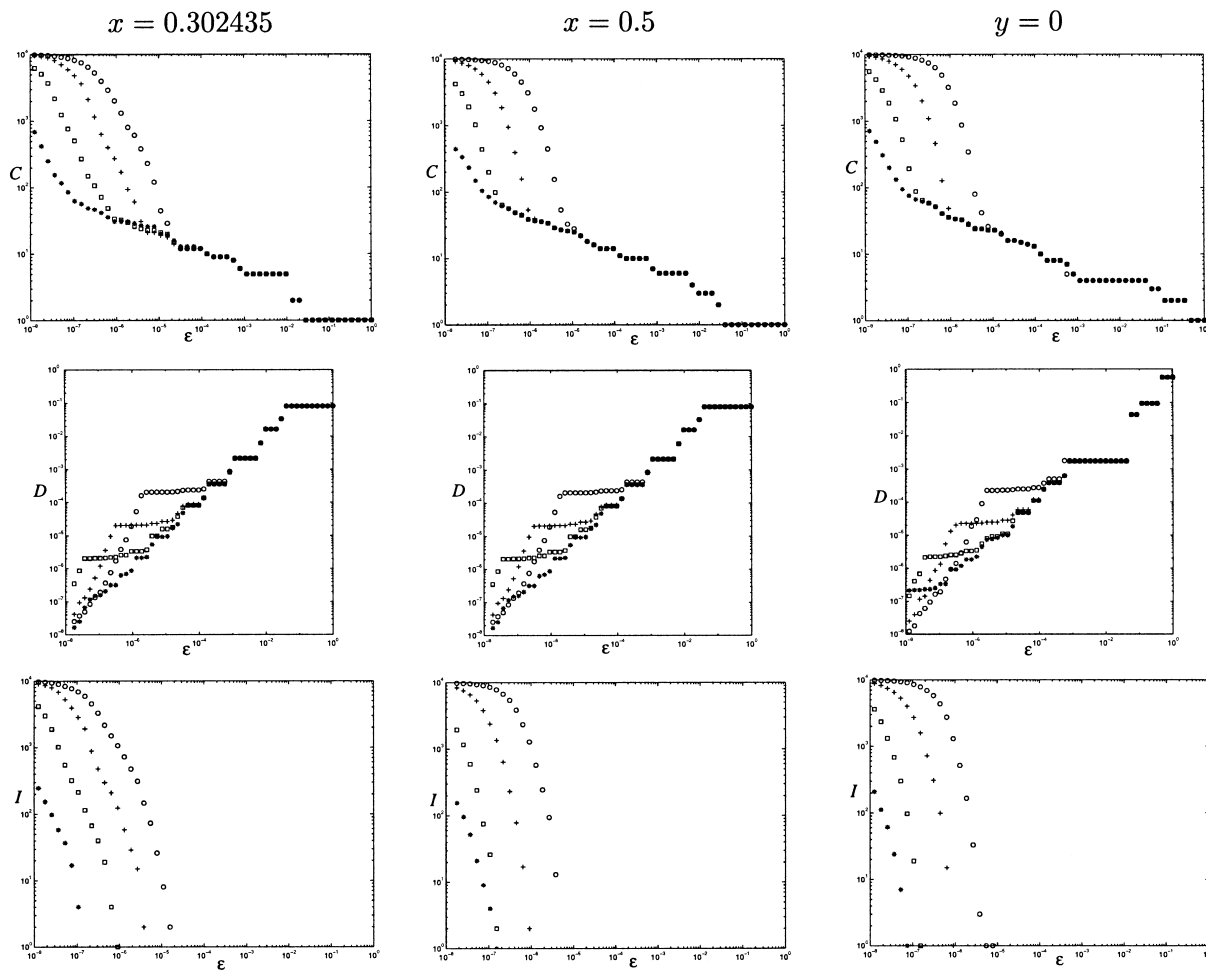


Fig. 21. Top row: $C(\epsilon)$, middle: $D(\epsilon)$, and bottom: $I(\epsilon)$ data for three sections of the Hénon attractor. The circles (\circ) represent calculations for a section of width 2×10^{-4} , the crosses (+) for one of width 2×10^{-5} , the squares (\square) of width 2×10^{-6} , and the stars ($*$) width 2×10^{-7} . All axes are logarithmic. The horizontal axis range is $10^{-8} < \epsilon < 1$.

cross-section can therefore not be a Cantor set; though of course, removing any fold-tangency points does leave a Cantor set.

The sections at $x = 0.5$ and $y = 0$ have simpler structure. The graphs of $D(\epsilon)$ show the now familiar staircase structure of a Cantor set. The flat segments in each graph of $D(\epsilon)$ are due to the finite width of each slice, making the data appear like a Cantor set of lines. Values of γ and δ are calculated from the $C(\epsilon)$ and $D(\epsilon)$ data for the thinnest slice at each section. The results are summarized in Table 2. The multifractal nature of the Hénon attractor [34] means that we expect to see the value of the dimension vary for different cross-sections. For the three examples given here, though, the variation is not significant.

The above results give strong numerical support for the common belief that cross-sections of the Hénon attractor are Cantor sets. The box-counting dimension of the Hénon attractor is estimated to be about 1.27 [34,35]. Results about the dimension of intersections of sets [32] imply that the dimension of a cross-section through the Hénon

Table 2
Values of γ and δ for the three sections of the Hénon attractor shown in Fig. 20

Section	γ	δ
$x = 0.302435$	0.25 ± 0.01	0.9 ± 0.1
$x = 0.5$	0.26 ± 0.01	0.85 ± 0.04
$y = 0$	0.27 ± 0.01	0.88 ± 0.02

attractor should be $1.27 - 1 = 0.27$. The values of γ given in Table 2 are in close agreement with this value, providing further support for our conjecture that Cantor sets of zero-measure have γ equal to the box-counting dimension.

4.3.2. Cantori

The final set of examples we examine in this paper are orbits from symplectic sawtooth maps. A cantorus is an invariant set of a symplectic mapping that is semi-conjugate to an incommensurate rotation, and is topologically a Cantor set. It is known that close enough to an “anti-integrable” limit, symplectic maps have cantori for all such rotations [36]. In the case of nearly-integrable area-preserving twist maps, cantori arise from the destruction of invariant circles, but it is not known, however, if this always occurs in higher dimensions. Thus, it is useful to develop numerical techniques that can distinguish the topological properties of such sets.

A simple model for which cantori are analytically computable is the following piecewise linear map, called a sawtooth map:

$$\tilde{p} = p + \nabla V(x), \quad \tilde{x} = x + \tilde{p} \bmod 1.$$

Here $x \in \mathbb{T}^2$ and $p \in \mathbb{R}^2$ and $V = \frac{1}{2}x^t Ax$ is a quadratic potential. Choosing an incommensurate pair of irrational rotation numbers, e.g., $\omega = \frac{1}{2}((\sqrt{5} - 1), \sqrt{2})$, we can find an orbit analytically by looking for the semi-conjugacy, $x_t = X(\theta + \omega t)$. In Fig. 22, we show the projection of the cantorus onto the configuration plane for two choices of ω : for Fig. 22a, $\omega = \frac{1}{2}((\sqrt{5} - 1), \sqrt{2})$, and for Fig. 22b, $\omega = (\tau^{-1}, \tau^{-2})$, where τ is the real root of $\tau^3 - \tau - 1$. In both cases, the matrix for the quadratic form is

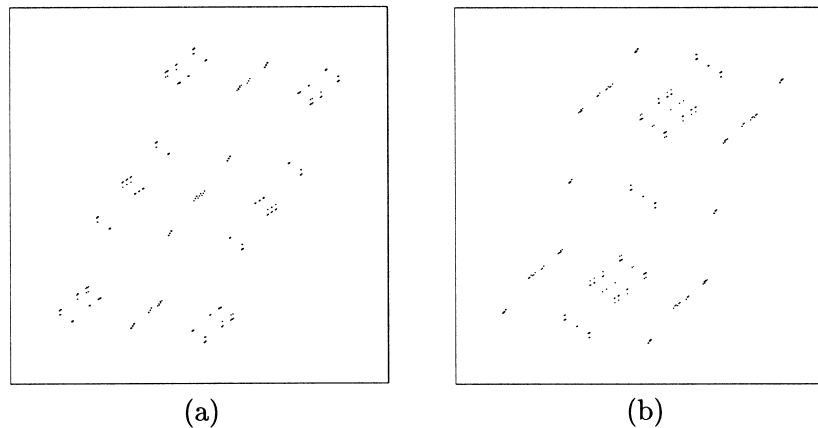


Fig. 22. Two examples of cantori generated by symplectic sawtooth maps. Each orbit has 10^4 points.

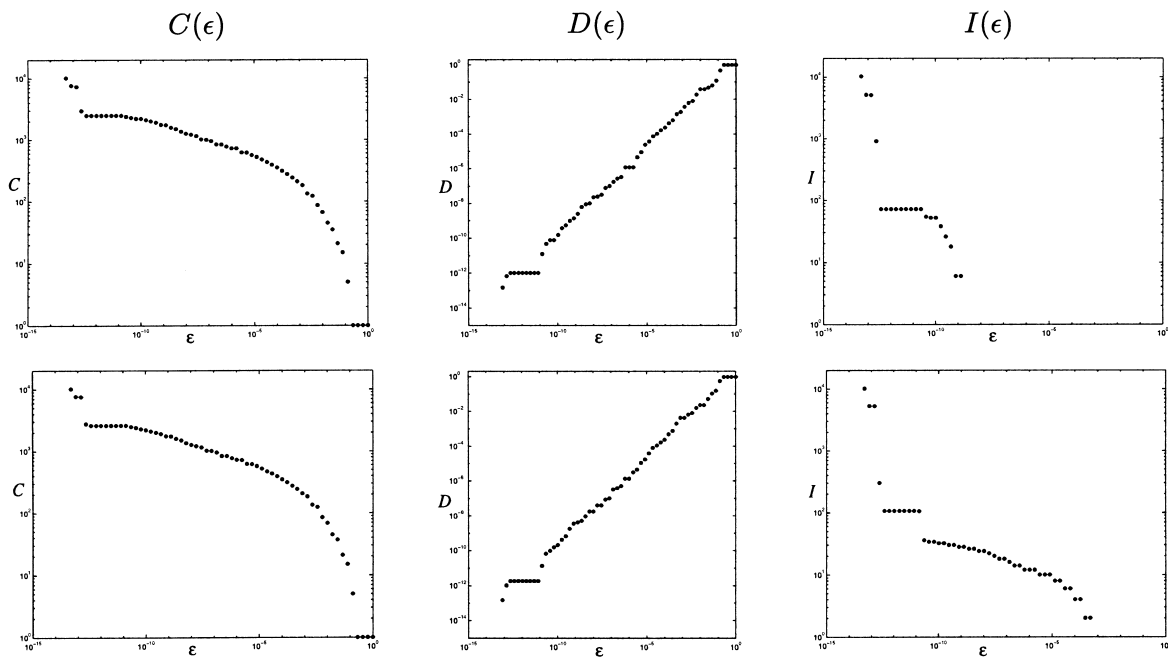


Fig. 23. $C(\epsilon)$, $D(\epsilon)$ and $I(\epsilon)$ data for the two cantori: (top row) data for the cantorus in Fig. 22a, and (bottom row) data for the cantorus in Fig. 22b. All axes are logarithmic. The horizontal axis range is $10^{-15} < \epsilon < 1$.

$$A \approx \begin{pmatrix} 1.9152 & -2.0358 \\ 0.5214 & 0.0847 \end{pmatrix}.$$

The graphs of $D(\epsilon)$ in Fig. 23 are similar to those for previous Cantor sets. These graphs tell us that the cantori are totally disconnected, because $D(\epsilon) \rightarrow 0$. Again, we estimate δ to be very close to one: for the cantorus of Fig. 22a $\delta = 1.09 \pm 0.05$; for that of Fig. 22b $\delta = 1.01 \pm 0.02$. The graphs of $C(\epsilon)$,

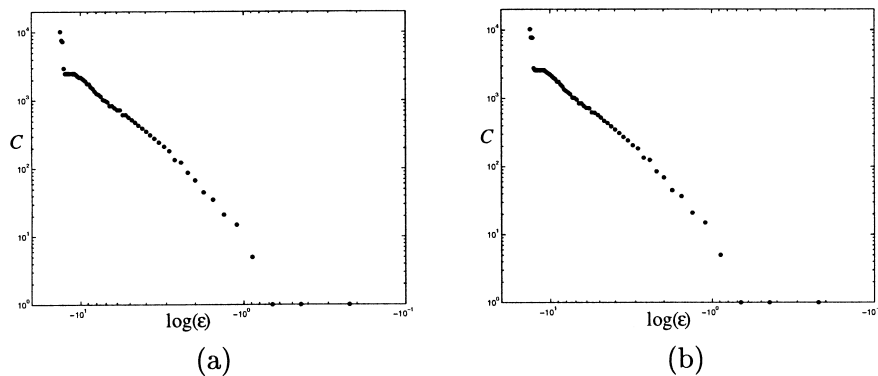


Fig. 24. $C(\epsilon)$ versus $\log(\epsilon)$ for the two cantori of Fig. 22a and 22b, respectively. All axes are logarithmic. The horizontal range is $-20 < \log(\epsilon) < -0.1$.

however, exhibit very different behavior from that of the other Cantor set examples. There is no linear scaling region from which to determine γ , and this suggests that our assumption that $C(\epsilon) \sim \epsilon^{-\gamma}$ does not hold here. Since the underlying sets are Cantor sets, we must still see $C(\epsilon) \rightarrow \infty$ as $\epsilon \rightarrow 0$, but the growth rate is possibly logarithmic, rather than polynomial, in ϵ . To test this hypothesis, we plot $C(\epsilon)$ versus $\log(\epsilon)$ on logarithmic axes in Fig. 24. The observed linear relationship tells us that $C(\epsilon) \sim (\log \epsilon)^{-\nu}$ near $\epsilon = 0$. We estimate the value of ν using a least-squares fit of the slope of the line and find that $\nu = 2.30 \pm 0.05$ for both cases. This also implies that $\gamma = 0$. A result due to Fathi [37] implies that these cantori have Hausdorff dimension equal to zero; so again we see that γ and the dimension are equal for a zero-measure Cantor set.

5. Conclusions

Our results demonstrate that the minimal spanning tree of a finite set of points can provide accurate information about the topology of the underlying set, down to a numerically computable resolution $\rho > 0$. In particular, we are able to identify sets that are connected, totally disconnected, or have infinitely many connected components with nonzero diameter. Confidence in the extrapolation can be increased by sampling more points in order to get a better approximation to the underlying set, but of course we are still ultimately restricted by the machine precision. Connected sets have disconnectedness index $\gamma = 0$, and discreteness index $\delta = 0$. We conjecture that Cantor sets with zero Lebesgue measure have γ equal to the box-counting dimension and $\delta = 1$. There are some results in this direction for subsets of the real line [38] and for Cantor sets generated by iterated function systems [39]. The fat Cantor set example shows that γ and dimension are not the same when the set has positive Lebesgue measure.

From the dynamical systems perspective, our techniques allow a quantitative analysis of the topology of attractors and other structures in phase space. Our analysis of some sections of the Hénon attractor provides strong evidence for the intuition that they have Cantor set structure. The cantori of Section 4.3 have subpolynomial growth in the number of components, i.e., $\gamma = 0$, even though they are Cantor sets. We resolved this problem by showing that the number of components grew logarithmically.

Most of the examples in this paper have points fairly evenly distributed over the underlying set. However, it is often the case that orbits cover an attractor in a highly nonuniform way. Previous work on characterizing fractals has dealt with this by introducing a concept of dimension for measures, and developing a theory of multifractals [38]. The results of Section 4.1 show that our techniques are most effective for uniformly distributed data because for a fixed number of data points, the cutoff resolution, ρ , is minimal when the points are evenly spread over the attractor. Our techniques still give valid results for highly nonuniform data; the difference is that ρ may be too high in these cases to make strong statements about the topology of the underlying set. It may be possible to weight the edges of the MST by the density of the data distribution and thereby lower ρ in these situations.

A natural question that arises from studying the Sierpinski triangle relatives is how to distinguish between simply connected sets and ones with holes. This involves reformulating concepts from homology theory by introducing a resolution parameter, similar to the way we treat the definition of connectedness. There are already a number of algorithms for computing homological invariants, such as Betti numbers, from data [9,40,41]. We are currently investigating ways to formalize the relationship between the homology of a set and the homology of a finite point-set approximation to it. The ability to characterize the connectedness properties of a set from finite data will give a deeper understanding of the structure of attractors, and therefore assist the modeling of the dynamics.

Acknowledgements

Thanks to A. Hogan for assistance with programming and the binary tree algorithm. JDM was supported in part by NSF grant number DMS-9971760. EB and VR were supported by NSF National Young Investigator award #CCR-9357740; EB was also supported by ONR #N00014-96-1-0720 and a Packard Fellowship in Science and Engineering.

References

- [1] N.H. Packard, J.P. Crutchfield, J.D. Farmer, R.S. Shaw, Geometry from a time series, *Phys. Rev. Lett.* 45 (1980) 712–716.
- [2] T. Sauer, J.A. Yorke, M. Casdagli, Embedology, *J. Stat. Phys.* 65 (1991) 95–116.
- [3] F. Takens, Detecting strange attractors in turbulence, in: D.A. Rand, L.S. Young (Eds.), *Dynamical Systems and Turbulence, Lecture Notes in Mathematics*, Vol. 898, Springer, Berlin, 1981, pp. 366–381.
- [4] P. Grassberger, I. Procaccia, Characterization of strange attractors, *Phys. Rev. Lett.* 50 (1983) 346–349.
- [5] A. Wolf, J.B. Swift, H.L. Swinney, J.A. Vastano, Determining Lyapunov exponents from a time series, *Physica D* 16 (1985) 285–317.
- [6] R.W. Ghrist, P.J. Holmes, M.C. Sullivan, *Knots and links in three-dimensional flows, Lecture Notes in Mathematics*, Vol. 1654, Springer, Berlin, 1997.
- [7] G. Mindlin, R. Gilmore, Topological analysis and synthesis of chaotic time series, *Physica D* 58 (1992) 229–242.
- [8] N.B. Tufillaro, P. Wyckoff, R. Brown, T. Schreiber, T. Molteno, Topological time-series analysis of a string experiment and its synchronized model, *Phys. Rev. E* 51 (1) (1995) 164–174.
- [9] M.R. Muldoon, R.S. MacKay, J.P. Huke, D.S. Broomhead, Topology from a time series, *Physica D* 65 (1993) 1–16.
- [10] K. Mischaikow, M. Mrozek, J. Reiss, A. Szymczak, Construction of symbolic dynamics from experimental time series, *Phys. Rev. Lett.* 82 (1999) 1144–1147.
- [11] T. Kaczynski, M. Mrozek, Conley index for discrete multivalued dynamical systems, *Topol. Appl.* 65 (1995) 83–96.
- [12] T.K. Dey, H. Edelsbrunner, S. Guha, Computational topology, in: B. Chazelle, J.E. Goodman, R. Pollack (Eds.), *Advances in Discrete and Computational Geometry, Contemporary Mathematics*, Vol. 223, Amer. Math. Soc., Providence, RI, 1999.
- [13] V. Robins, J.D. Meiss, E. Bradley, Computing connectedness: an exercise in computational topology, *Nonlinearity* 11 (1998) 913–922.
- [14] T.Y. Kong, A. Rosenfeld, Digital topology: introduction and survey, *Comput. Vis. Graph. Image Process.* 48 (1989) 357–393.
- [15] M. Brady, J.M. Hollerbach, T.L. Johnson, T. Lozano-Perez, M.T. Mason (Eds.), *Robot Motion: Planning and Control*, MIT Press, Cambridge, MA, 1984.
- [16] J. Williams (Ed.), *Geographic Information from Space: Processing and Applications of Geocoded Satellite Images*, Wiley, New York, 1995.
- [17] J.K. Udupa, Surface connectedness in digital spaces, in: A. Rosenfeld, T.Y. Kong (Eds.), *Topological Algorithms for Digital Image Processing*, Elsevier, Amsterdam, 1993.
- [18] G.T. Herman, E. Zhao, Jordan surfaces in simply connected digital spaces, *J. Math. Imaging Vis.* 6 (1996) 121–138.
- [19] A. Rosenfeld, Connectivity in digital pictures, *J. Assoc. Comput. Mach.* 17 (1970) 146–160.
- [20] T. Kong, A.W. Roscoe, A. Rosenfeld, Concepts of digital topology, *Topol. Appl.* 46 (1992) 219–262.
- [21] C.-N. Lee, T. Poston, A. Rosenfeld, Winding and Euler numbers for 2D and 3D digital images, *CVGIP: Graph. Models Image Process.* 53 (1991) 522–537.
- [22] J.G. Hocking, G.S. Young, *Topology*, Addison-Wesley, Reading, MA, 1961.
- [23] M.F. Barnsley, *Fractals Everywhere*, 2nd Edition, Academic Press, Boston, MA, 1993.
- [24] B. Burch, J.C. Hart, Linear fractal shape interpolation, *Proc. Graphics Interface '97*, May 1997, pp. 155–162.
- [25] A.I. Mees, Tessellations and dynamical systems, in: M. Casdagli, S. Eubank (Eds.), *Nonlinear Modeling and Forecasting, SFI Stud. Sci. Complexity Proc.*, Vol. XII, Addison-Wesley, Reading, MA, 1992, pp. 3–24.
- [26] F.P. Preparata, M.I. Shamos, *Computational Geometry: An Introduction*, Springer, Berlin, 1985.
- [27] K. Yip, *KAM: A System for Intelligently Guiding Numerical Experimentation by Computer*, MIT Press, Cambridge, MA, 1991.
- [28] R.S. Prim, Shortest connection networks and some generalizations, *Bell Sys. Technol. J.* (1957) 1389–1401.
- [29] C.T. Zahn, Graph-theoretical methods for detecting and describing Gestalt clusters, *IEEE Trans. Comput. C* 20 (1971) 68–86.
- [30] H.-O. Peitgen, S. Jürgens, D. Saupe, *Chaos and Fractals: New Frontiers of Science*, Springer, Berlin, 1992.
- [31] R. Eykholt, D.K. Ueberger, Extension of the fat-fractal exponent β to arbitrary sets in d dimensions, *Phys. Lett. A* 163 (1992) 409–414.
- [32] K.J. Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, Wiley, New York, 1990.
- [33] C. Robinson, *Dynamical Systems: Stability, Symbolic Dynamics, and Chaos*, CRC Press, Boca Raton, FL, 1995.
- [34] R. van de Weygaert, B.J.T. Jones, V.J. Martinez, The minimal spanning tree as an estimator for generalized dimensions, *Phys. Lett. A* 169 (1992) 145–150.
- [35] H.E. Nusse, J.A. Yorke, *Dynamics: Numerical Explorations*, 2nd Edition, Springer, Berlin, 1998.
- [36] R.S. MacKay, J.D. Meiss, Cantori for symplectic maps near the anti-integrable limit, *Nonlinearity* 5 (1992) 149–160.

- [37] A. Fathi, Expansiveness, hyperbolicity and Hausdorff dimension, *Commun. Math. Phys.* 126 (1989) 249–262.
- [38] K.J. Falconer, *Techniques in Fractal Geometry*, Wiley, New York, 1997.
- [39] V. Robins, *Computational Topology with Applications to Fractal Geometry*, Ph.D. Thesis, University of Colorado, Boulder, CO, in preparation.
- [40] C.J.A. Delfinado, H. Edelsbrunner, An incremental algorithm for Betti numbers of simplicial complexes on the 3-sphere, *Comput. Aided Geom. Design* 12 (1995) 771–784.
- [41] W.D. Kalies, K. Mischaikow, G. Watson, Cubical approximation and computation of homology, *Banach Center Publ.* 47 (1999) 115–131.