CHAPTER 4

Dimensions

4.1. Motivating examples, or, Fun with length and area.

Let's consider this question: How big is the Koch curve? We can estimate its length by computing the lengths of the curves K_n introduced in Example 1.2 and shown in Figure 2.

Recall that we began with K_0 as a line segment of length 1, for concreteness suppose it is on the x-axis from $x = 0$ to $x = 1$. To construct K_1, four affine transformations with contraction factor 1/3 are used. That means K_1 is the union of four segments of length 1/3 and thus has length 4/3. The union of four copies of K_1, scaled by 1/3, make up K_2. Each scaled copy has length 4/9, but there are four of them end-to-end and so K_2 has length 16/9. At the next stage, four copies of K_2, which have been scaled by 1/3 and thus have length 16/27, union together to become K_3, which then has length 64/27. In general we see that K_n will have length $4^n/3^n$. Since the Koch curve is $\lim\limits_{n\to\infty} K_n$ (where the limit of course is in the Hausdorff metric), we can see that its length is $\lim\limits_{n\to\infty} 4^n/3^n = \infty$.

We can consider this result to mean that we are not really measuring the size of the Koch curve correctly. It's a nice compact set, and to call it's size infinite seems not to give a lot of geometric information about the curve. We'll be able to get finer information using the idea of fractal dimension.

EXERCISE 4.1. Consider the middle-thirds Cantor set introduced in Example 1.1 and shown in Figure 1. Compute its length with the method used for the Koch curve, i.e. by finding the lengths of $C_0, C_1, C_2, \ldots$ and taking their limit.

EXERCISE 4.2. Consider the Sierpinski triangle introduced in Example 1.3 and shown in Figure 3. Compute its area with the method used for the Koch curve, i.e. by finding the areas of $S_0, S_1, S_2, \ldots$ and taking their limit.

Since the Koch curve has infinite length in a finite area, maybe we should have tried to compute its area rather than its length. However, it seems fairly clear that the area is zero. You might try to argue that point as follows: The area of any K_n is clearly 0, so it stands to reason that the area of the limit is 0 also. However, in a moment we will see an example of a "space-filling curve", where each approximating set has zero area but the limit is a set with nonzero area.[1]

[1] Notice that this calls into question the length and area computations we've done so far.

A different way to approximate the area of the Koch curve, and one which will generalize to our study of fractal dimension, involves covering each K_n with rectangles, or 'boxes', and compute those areas instead. Of course this will overestimate the area, but we're going to get 0 anyway and the example is instructive.

So, let's be ridiculous and cover K_0 with the box $[0,1]\times[-.5,.5]$, which is shown on the left of figure 1. The area of the box is one and that is an upper bound on the area of K_0.

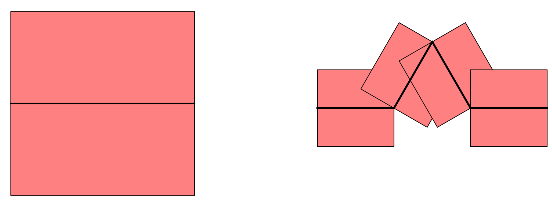


FIGURE 1. The first and second approximations of the Koch curve by boxes.

To construct a box-covering of K_1, simply apply the collage to the box that covers K_1. We see in the figure that K_1 is covered by four copies of this box, scaled *in length* by $1/3$, but that means that the *area* of each rescaled box is $1/9$. Again we shall be ridiculous and fail to account for the overlap between the boxes, getting an overestimate of the area of K_1 by adding the areas of the four boxes together. This gives us that the area of K_1 is less than $4/9$.

EXERCISE 4.3.
(1) Using the same technique, give an upper bound on the area of K_2.
(2) If necessary, repeat for K_3, K_4, etc. until you can generalize your answer to an upper bound on the area of K_n.
(3) Determine the area of the Koch curve.

Example 4.4. Here is an iterated function system called the Heighway Dragon (also known as the dragon curve). We begin with the vertical line segment connecting the origin to $(0, 1)$, and use the affine maps

$$T_1\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{pmatrix} .5 & -.5 \\ .5 & .5 \end{pmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \qquad T_2\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{pmatrix} -.5 & -.5 \\ .5 & -.5 \end{pmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

This IFS may seem simpler than that of the Koch curve, but the result looks really different. The reason is that this curve folds back to touch itself repeatedly. Figure 2 shows the first several applications of $\mathcal{T}$, and figure 3 shows H_{13} and the fixed set.

Suppose we were to overestimate the area of the Heighway dragon by covering H_0 with a rectangle. We would not have a sequence of covers that converge to 0. What would happen instead is more subtle, and we delay discussing it for now.

Side note: You can make iterations of the Heighway Dragon using a long, thin strip of paper. If you fold the paper in half and open it up again, you have a scaled version of H_1. If you fold it twice, being careful to fold the same way both times, then when you open it up so that the folds are at right angles you get H_2. Doing a bunch of iterations is a fun party trick.

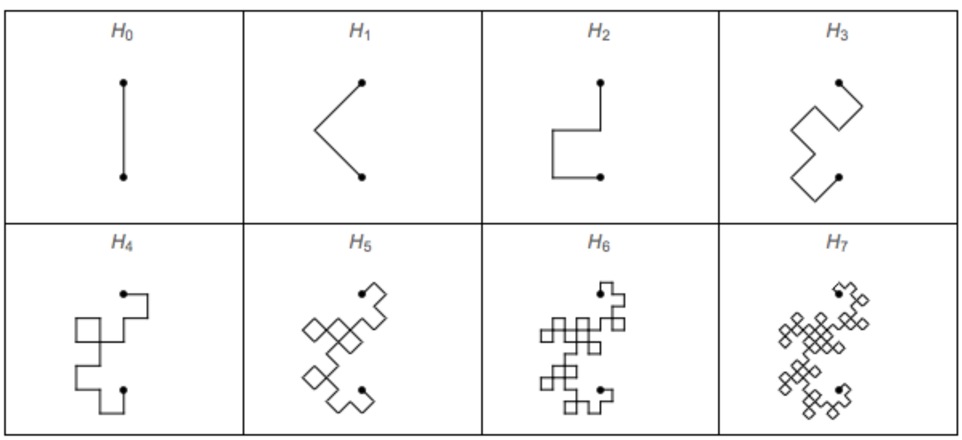


Figure 2. The first several iterations of the collage map. Points at the origin and at $(0, 1)$ are marked for reference.

Exercise 4.5. Make a really good paper Heighway dragon to show the class. Explain the questions you had or problems you solved to do it.

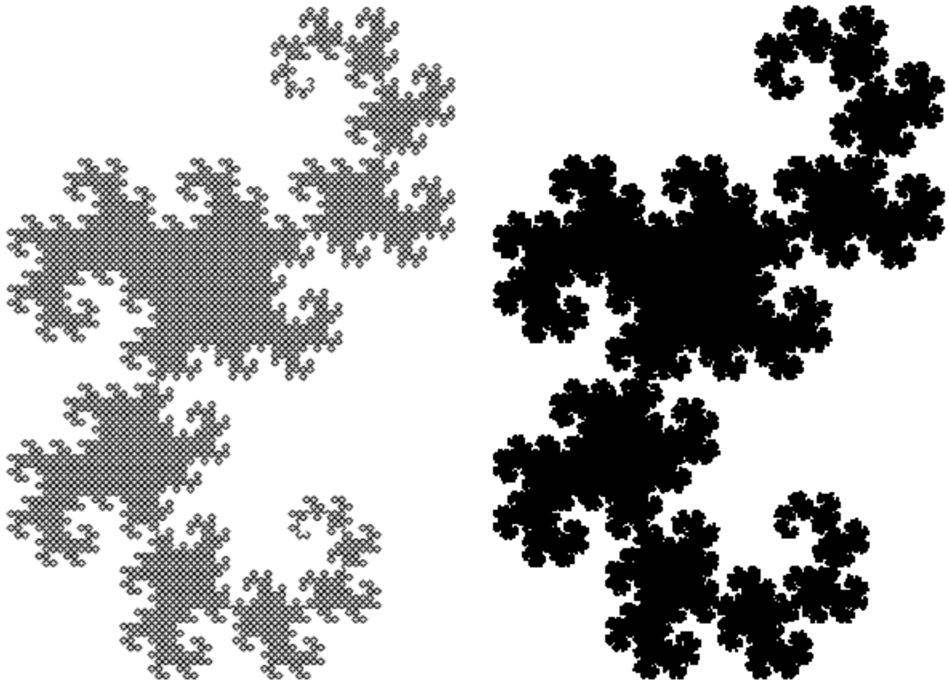


FIGURE 3. The left is H_{13}; the right is the fixed set of the IFS.

4.2. The idea of fractal dimension

Everyone has an intuitive idea of dimension: one-dimensional objects look like (possibly deformed) line segments; two-dimensional objects look like (possibly deformed) pieces of planes; three-dimensional space is the space we live in, where there are three independent directions one can move. In linear algebra a mathematical definition is given for vector spaces: it is the number of basis vectors required to define the space. In a topology class, objects are n-dimensional if they are locally homeomorphic images of $\mathbb{R}^n$ in a technical sense.

In calculus we learn an excellent strategy for calculating the length, area, or volume of an object we will call A. To compute the length $\mathcal{L}(A)$, we approximate it with line segments of length Δx; the length of A is approximately the sum of those segments, i.e. $\mathcal{L}(A) \approx \sum_{\#\text{ segments}} (\Delta x)^1$. We let $\Delta x \to 0$ and if the limit exists in a certain sense it becomes an integral that represents the length of the object. This is where the arclength formulae that you learn in calculus come from.

If we want to estimate the area of A, we can cover it with squares of side length Δx; the area of each square is $(\Delta x)^2$ and so the area $\mathcal{A}(A)$ is approximately $\mathcal{A}(A) \approx \sum_{\#\text{ squares}} (\Delta x)^2$. Again in a calculus class we would let $\Delta x \to 0$, and if the limit existed in a certain sense we'd get the area of A.

The same process works to estimate the volume of A, covering it with cubes of side length Δx. The volume of such a cube is $(\Delta x)^3$, so the volume $\mathcal{V}(A)$ is approximately $\mathcal{V}(A) \approx \sum_{\#\text{ cubes}} (\Delta x)^3$.

The general term we will use for a line segment, square, cube, or an n-dimensional analogue thereof will be a *box*. The n-dimensional *volume* of such a box will be it's side length to the nth power. Thus to estimate the n-dimensional volume of an object A, we approximate the set with boxes with n-volume $(\Delta x)^n$, arriving at the formula

$$\mathcal{V}_n(A) \approx \sum_{\#\text{ boxes}} (\Delta x)^n$$

Clearly, **the dimension of a box appears as the exponent of its volume calculation**, with length corresponding to exponent 1, area to exponent 2, ordinary volume to exponent 3, and n-dimensional volume to exponent n.

There is nothing to stop us from trying to use the calculus procedure to compute the n-dimensional volume of an object whether or not it is n-dimensional. Indeed, calculating the n-volume of an object that isn't fundamentally n-dimensional yields predictable results that we are beginning to understand.

EXAMPLE 4.6. Let A be the circle $x^2 + y^2 = 1$. We use the above process to compute the length of A. Let us estimate A using an inscribed regular k-gon, so that the circle is approximated by k line segments. As we let $k \to \infty$, the approximation will converge in Hausdorff metric to the circle. A k-gon is pictured for $k = 11$ on the left of figure 4.

Each edge of the k-gon has length Δx_k and can be seen as the short side of an isosceles triangle with long sides 1 and angle $2\pi/k$, which means by that $\Delta x_k = 2\sin(\pi/k)$ and so the circumference is approximately $\displaystyle\sum_{\#\text{ segments}} (\Delta x_k)^1 = 2k\sin(\pi/k)$, and from calculus we know (and you will verify in a homework exercise) that the limit as $k \to \infty$ is 2π.

An alternative trick to determine Δx_k, perhaps a bit circular, is to use the fact that the arc subtended by an angle $2\pi/k$ in the unit circle has length $2\pi/k$. Since Δx_k is approximately the length of that arc it follows that $\Delta x_k \approx 2\pi/k$, with the approximation becoming more and more accurate as $k \to \infty$. There are k sides to the k-gon, so the length of the circle is approximately $\displaystyle\sum_{\#\text{ segments}} (\Delta x_k)^1 = k\Delta x_k \approx k(2\pi/k) = 2\pi$. Either way, as $k \to \infty$, our approximation approaches 2π, as it should.

EXERCISE 4.7. Let's now cover the circle with squares whose diagonals are the sides of the regular k-gons above. We picture the covering for $k = 11$ on the right of Figure 4.

(1) Compute the approximate length of the side of each square using the second approximation of the diagonal, which was $2\pi/k$. Use this to approximate the area of each square.
(2) Compute the approximation of the area of A using your squares.
(3) Letting $k \to \infty$, show that the area of A is 0.

There was no particular reason to choose the covering of the circle by squares to look like the left side of figure 4. We could instead have had each edge of the k-gon be a side of the square, or the midline of the square like we did for the Koch curve.

EXERCISE 4.8. Suppose we did the covering of the circle using the midline version of the square covering. Compute the area of A again using this covering.

When doing box-related calculations like these it is very important that the answer be independent of the placement of the boxes. There exist bizarre examples that fail to have this independence, but we will not encounter any. It is also important that the shape of the boxes not matter; that is, if we approximated the shape circles, or rectangles, or sets with different dimensions, we'd always get the same answer as the diameters of the sets went to 0. In the wild woolly world of analysis there exist examples for which this is not true. Such examples will not concern us in this course.

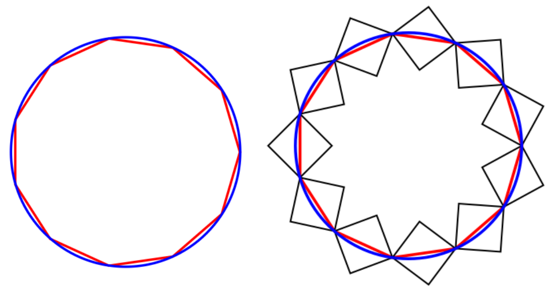


FIGURE 4. Approximating the circle with boxes of dimension 1 and 2.

EXERCISE 4.9. Consider the line segment L connecting the origin to the point $(1,1,1)$ in $\mathbb{R}^3$. For integers k let $\Delta x_k = \sqrt{3}/k$. Calculate the length, area, and volume of L using the calculus method. Namely, for $n = 1,2,3$, cover L with n-boxes of side length $\sqrt{3}/k$, compute the estimate $\sum_{\#\text{ boxes}} (\Delta x_k)^n$, and take the limit as $k \to \infty$.

EXERCISE 4.10. Now let A be the unit square lying in the $x-y$ plane of $\mathbb{R}^3$. Compute the length, area, and volume of A using boxes of side length $\Delta x_k = 1/k$. Note: your estimate of the square using line segments will leave a lot of the square uncovered, but in the Hausdorff metric limit the approximation by line segments will converge to the square.

In each one of the examples we have seen there were three possible outcomes: 0, a finite number, or ∞. We got 0 when we overestimated the dimension of A, we got ∞ when we underestimated the dimension of A, and we got a finite number when we got the right dimension. Except for the Koch curve: we got 0 for 2 dimensions and ∞ for one dimension, suggesting that the correct dimension is somewhere between 1 and 2.

Exercise 4.11. Suppose we are covering the approximations K_j of the Koch curve with boxes of side length Δx_k, except that we imagine the boxes to be s-dimensional in the sense that they have volume $(\Delta x_k)^s$.

(1) We cover K_0 by a box with side length 1 as before. The s-volume of that box is then $1^s = 1$. When we apply $\mathcal{T}$ to this box, we get four boxes, where the side length of each is now $\Delta x_1 = 1/3$. Thus our s-dimensional volume approximation for K_1 is then $\sum_{\#\text{ boxes}} (\Delta x_1)^s = 4(1/3)^s = 4/3^s$.
(2) The side length of each box making up the cover for K_2 has length $1/9$. Compute $\sum_{\#\text{ boxes}} (\Delta x_2)^s$.
(3) Find the formula for the s-volume of the cover of K_j, for $j \geq 2$.
(4) We want the limit of that s-volume to be a finite positive number C. Solve for s. (Hint: set your answer from (3) approximately equal to C and solve for s.)

4.2.1. Four tenets of a good notion of dimension. A really good definition of fractal dimension ought to have a number of properties that make sense. The two definitions we will make will satisfy some or all of them in some or all types of examples. These four tenets are described in the lecture notes [**FN**], p. 102. Let's suppose we have called our candidate for fractal dimension Dim.

- (Familiarity) The dimension of $\mathbb{R}^n$ should be n. Moreover, line segments and vector spaces of linear algebraic or topological dimension 1 should also have a fractal dimension of 1. Similarly a square should have fractal dimension 2, a cube dimension 3, and so on.
- (Monotonicity) If $A \subset B$, then $Dim(A) \leq Dim(B)$. In particular this means that if $A \subset \mathbb{R}^n$, then $Dim(A) \leq n$.
- (Stability) The process of taking unions of sets should not affect the fractal dimension unpredictably. Thus we prefer our definition of dimension to satisfy $Dim(A \cup B) = \max\{Dim(A), Dim(B)\}$.
- (Invariance) Suppose $T : X \to X$ is an isometry like rotation, reflection, or translation, or some other 'nice' map like a similarity. Then although T might affect the *size* of A, it should not affect its fractal dimension. Thus we prefer that our definition of fractal dimension satisfy, for such a map T, that $Dim(T(A)) = Dim(A)$.

4.3. Similarity dimension

We begin with the simplest type of fractal dimension to compute: the similarity dimension. An advantage is that it can be computed for any IFS once the contraction factors of its transformations are known, but a disadvantage is that it is accurate[2] only in some of those cases. Fortunately, a many of the examples we have been considering are ones for which the similarity dimension makes sense.

[2] Ideally, a set's similarity dimension would agree with its other kinds of dimensions in a wide swath of examples.

4.3.1. Similarity dimension: one scaling factor. We know that if we scale a figure by $1/2$, then its length scales by $1/2$, but its area scales by $1/4$, and its volume scales by $1/8$, and by extension its volume in dimension s scales by $(1/2)^s$. If we know that a set such as the Koch curve is made up of a certain number of copies of itself, all scaled by the same amount, we can use that fact to solve for s exactly. This is the foundation for our definition of similarity dimension.

EXAMPLE 4.12. Let us denote the Koch curve by $\mathcal{K}$ and its s-volume by $V_s(\mathcal{K})$. We know that $\mathcal{K}$ is the fixed point of a collage map $\mathcal{T}$ composed of four maps, each with contraction factor $1/3$, i.e., that $\mathcal{T}(\mathcal{K}) = T_1(\mathcal{K}) \cup T_2(\mathcal{K}) \cup T_3(\mathcal{K}) \cup T_4(\mathcal{K})$. We also know that $V_s(T_i(\mathcal{K})) = V_s(1/3(\mathcal{K})) = (1/3)^s V_s(\mathcal{K})$. This implies that

$$V_s(\mathcal{K}) = 4(1/3)^s V_s(\mathcal{K})$$

This is an equation we can solve for s. We see quickly that $3^s = 4$, which means that $s \ln 3 = \ln 4$, and so $s = \ln 4/\ln 3$.

We have now arrived at the dimension $\ln 4/\ln 3$ for the Koch curve using both parts of the Existence Theorem: in Exercise 4.11 we did it by seeing $\mathcal{K}$ as the limit of sets under repeated iteration of $\mathcal{T}$, and in the previous example we did it using the fact that $\mathcal{K}$ is the fixed point of $\mathcal{T}$. We can be fairly confident that this number is representative of the fundamental 'size' of the Koch curve.

EXERCISE 4.13. Use the technique of example 4.12 to compute the similarity dimension of the middle-thirds Cantor set $\mathcal{C}$.

EXERCISE 4.14. Use the technique of example 4.12 to compute the similarity dimension of the Sierpinski triangle $\mathcal{S}$.

Computation of similarity dimension is appropriate for sets $A \in \mathcal{H}(X)$ that are unions of rescaled copies of themselves, and the scaling factors don't all have to be the same. However, the transformations have to be *similarities* in the sense of your high school geometry course. Let us take a moment to review them.

4.3.2. Similarity transformations. Although in this course we focus on compact subsets of Euclidean space, the following definition of similarity holds in any metric space.

DEFINITION 4.15. A transformation $T : X \to X$ is a *similarity* if there is a positive $c \in \mathbb{R}$ such that for every $x, y \in X$, $d(T(x), T(y)) = c\, d(x, y)$. We call c the *similarity ratio* of T.

In Euclidean space, transformations are similarities if they are certain kinds of affine maps. In one dimension, any affine map will do as long as it is invertible.

EXERCISE 4.16. Prove that $T : \mathbb{R} \to \mathbb{R}$ is a similarity if and only if there are constants $a, b \in \mathbb{R}$ with $a \neq 0$ such that $T(x) = ax + b$.

In two dimensions, similarity transformations are the ones that you actually learned about in high school geometry: they take triangles to similar triangles (that is, they preserve angles). There are several ways to identify when an affine transformation is a similarity, of which we give three. The first two are geometric and the third comes from linear algebra.

One way is to look at what the transformation does to the standard basis vectors $\vec{e_1}$ and $\vec{e_2}$. If it sends them to vectors that are orthogonal and that have the same length, then the transformation is a similarity. This trick works to identify a similarity in any dimension.

Another way is to know all possible *isometries* of $\mathbb{R}^2$, that is, all possible maps that don't change distance between points. Those are simply the translations, rotations, and reflections. If your affine map is a scaling factor times one of these isometries, then it is a similarity.

A linear algebraic way to determine if a transformation in any $\mathbb{R}^d$ is a similarity is to determine the eigenvalues and eigenvectors of the underlying linear transformation. The matrix must be diagonalizable in that it must have a basis of (possibly complex) eigenvectors, and its eigenvalues must all have the same magnitude. Under those conditions the eigenvalues stretche the basis vectors by the same amount, and thus stretches every vector in $\mathbb{R}^d$ by the same amount.

EXERCISE 4.17. Consider the transformations in the collage for Barnsley's fern. Determine which, if any, are similarities. If it is a similarity, determine the scaling factor.

Notice that a similarity transformation is a contraction whenever $c \in [0, 1)$.

4.3.3. Similarity dimension: multiple scaling factors. Now let us consider the fractal dimension of the attractor of the IFS $(X; T_1, ..., T_n)$. If T_i is a similarity of X for $i = 1, 2, ...n$ then it is an appropriate system on which to make a definition of similarity dimension. The similarity dimension $Dim_S(A)$ will be defined via the scaling factors of its contraction maps in a manner quite similar to our discussion of the Koch curve.

When we computed the dimension of the Koch curve in example 4.12, we needed to solve the equation

$$(1/3)^s + (1/3)^s + (1/3)^s + (1/3)^s = 1,$$

which we of course simplified to solving $4(1/3)^s = 1$. In exercise 1.30 we saw a Cantor-set-like example made of transformations of different scaling factors that makes a good test case for similarity dimension.

EXAMPLE 4.18. Let $X = [0,1]$, $T_1(x) = x/4$, $T_2(x) = x/3 + 1/3$, and $T_3(x) = x/4 + 3/4$. The attractor A of the IFS $(X; T_1, T_2, T_3)$ is pictured in exercise 1.30. If we have the correct dimension s of A calculated, then its volume $V_s(A)$ must satisfy

$$V_s(A) = V_s(1/4A) + V_s(1/3A) + V_s(1/4A),$$

giving rise to the equation

$$V_s(A) = (1/4)^s V_s(A) + (1/3)^s V_s(A) + (1/4)^s V_s(A).$$

Thus the dimension s must be the solution to

$$(1/4)^s + (1/3)^s + (1/4)^s = 1.$$

Mathematica finds the approximation $s \approx 0.856738$ using the expression[a]

FindRoot[(1/4)^s + (1/3)^s + (1/4)^s - 1, {s, 1}]

[a] I told it to look for the root near $s = 1$ since s should be somewhere between 0 and 1.

So we see that when the scaling factors are different we arrive at the equation

$$c_1^s + c_2^s + \cdots + c_n^s = 1.$$

We need to know that that equation has a unique solution, which is the content of this lemma that we state without proof.

LEMMA 4.19 (see [**Edg90**], p. 105). *Suppose $c_1, c_2, ..., c_n$ are numbers in $[0,1)$ for all i. Then there is a unique number $s \geq 0$ such that $c_1^s + c_2^s + ... + c_n^s = 1$. The number s is 0 if and only if there is only one nonzero value for c_i.*

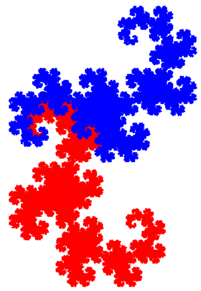
An unfortunate but unavoidable fact is that in general there is no way to solve for s unless the c_i's are related to each other in some way. However it is very easy to get your computer or calculator to give you an approximation that is accurate to as many digits as you like. Here is the official definition of similarity dimension.

DEFINITION 4.20. Let A be the attractor for an iterated function system $(X, T_1, T_2, ..., T_n)$ for which each T_i is a similarity with similarity ratio $c_i \in [0,1)$. The *similarity dimension* of A, denoted $Dim_S(A)$, is defined to be the solution s to the equation

$$c_1^s + c_2^s + ... + c_n^s = 1.$$

EXERCISE 4.21. Compute the similarity dimension of the spiral fractal from exercise 1.36. Give the equation that it solves and obtain a numerical estimate of its value.

EXERCISE 4.22. The Heighway dragon is the union of two copies of itself under the two transformations that make up its collage map, given in example 4.4. Compute its similarity dimension.



EXERCISE 4.23. Compute the equation for the similarity dimension of the fractal in exercise 1.37 and give an numerical estimate for its dimension.

There is a clever way to solve for the similarity dimension in the last exercise exactly that capitalizes on the fact that $1/4 = (1/2)^2$.

EXERCISE 4.24. In the equation obtained for the similarity dimension in the previous exercise, let $x = (1/2)^s$. Rewrite the equation in terms of x and solve using the quadratic formula. This gives a solution for x that can be solved for s exactly. Give the exact solution for s and approximate it with a calculator value.

Here is a simple example where the similarity dimension does not do so well.

EXAMPLE 4.25. Let $X = [0,1]$, $T_1(x) = 2x/3$, and $T_2(x) = 2x/3 + 1/3$. Each of T_1 and T_2 are similarities and so the definition of similarity dimension applies and we find that $Dim_S(A)$ is the solution s to $2(2/3)^s = 1$. Thus $Dim_S(A) = \ln 2/\ln 1.5 > 1$, a troubling result. Because of the amount of overlap, it turns out that $\mathcal{T}(X) = X$ and so the attractor A of $\mathcal{T}$ is all of $[0,1]$. By the tenet of familiarity, we should therefore have arrived at the dimension of A as being 1. We also failed the tenet of monotonicity: the dimension certainly should not have exceeded 1 since A is a subset of $\mathbb{R}$.

The problem evident in this example is that its IFS is *overlapping*. Later in this chapter we will define what if means for an IFS to be *totally disjoint* or *just-touching*. The similarity dimension of an IFS satisfying either of those two conditions will be seen to have an 'accurate' similarity dimension.

So we see some obvious benefits and obvious drawbacks to our definition of similarity dimension. The main benefits are that it is quite easy to compute, at least numerically, and that it is quite natural for the examples for which it is defined. There are a few drawbacks. One is that it is not defined for iterated function systems such as Barnsley's fern, where the collage contains transformations that are not similarities. Another is that even when it is defined, there are situations such as our previous example where the result is misleading. The good news is that it is possible to write down which situations give us trouble, and that the number obtained from the equation for similarity dimension has meaning even in cases like the fern. To discuss this further we need to think about a more general definition of dimension.

4.4. Box-counting dimension

Let $A \subset \mathbb{R}^d$ be a compact set. A key idea necessary to compute dimension of A is how many boxes of a given side length are needed to cover A. Let us co-opt the notation we used earlier for balls of radius ϵ and let

$$B(x, \epsilon) = \text{box of side length } \epsilon \text{ centered at } x$$

In this definition we mean that a box in $\mathbb{R}$ is an interval of length ϵ with x at its center; in $\mathbb{R}^2$ it is a square of side length ϵ with x at its center, in $\mathbb{R}^3$ it is a cube, and so on. So the boxes are easy to picture even when their s-volume is not.

Definition 4.26. Let $A \in \mathbb{R}^d$ be compact and let $\epsilon > 0$. The smallest positive integer N for which $A \subset \displaystyle\bigcup_{n=1}^{N} B(x_n, \epsilon)$ is called the *ϵ-covering number* of A and is denoted by $\mathcal{N}(A, \epsilon)$.

Let us take a moment to parse this definition. For any integer, say K, we can choose any K points in X to call $x_1, x_2, ..., x_K$. Those points can be taken to be the centers of closed boxes of radius ϵ; in that case the union of all of those boxes will be some subset $\displaystyle\bigcup_{n=1}^{K} B(x_n, \epsilon)$ of X. Maybe our set A is in that union and maybe it is not. Maybe we should have made a more judicious choice for our centers. What is clear is that there are infinitely many choices of K and then (mega-uncountably) infinitely many choices for $x_1, x_2, ..., x_K$ we could make for the centers. In this definition we consider all of them simultaneously, focusing in particular on choices where $A \subset \displaystyle\bigcup_{n=1}^{K} B(x_n, \epsilon)$. We ask ourselves the question, what is the smallest possible value of K for which $A \subset \displaystyle\bigcup_{n=1}^{K} B(x_n, \epsilon)$? The answer to that is $\mathcal{N}(A, \epsilon)$.

Example 4.27. Let $A \in \mathbb{R}^2$ be the line connecting the origin to $(1, 1)$ and suppose that $\epsilon = 1$. Then $\mathcal{N}(A, \epsilon) = 1$ because we can take x_1 to be the midpoint $(1/2, 1/2)$. The box of side length ϵ will equal the unit square and thus contains A.

Slightly more interesting might be to take $\epsilon = 1/2^n$. In this case we can space out our centers along A to see that we need 2^n boxes to cover A.

Exercise 4.28. Verify the example for $n = 2$ and $n = 3$. In both cases make a sketch of $\displaystyle\bigcup_{n=1}^{N} B(x_n, \epsilon)$. Write a formula for placing the centers for a general n.

Of course in the previous examples the ϵ was selected to be particularly nice relative to the set A. The result was that we were able to cover A without needing much overlap of the boxes. In general the fit won't be quite as nice, as you see in the next exercise.

Exercise 4.29. Continue with example 4.27, this time letting $\epsilon = e^{-1}$ and $\epsilon = \pi^{-1}$.

Example 4.30. Let A be a square of side length 1 in $\mathbb{R}^2$. It doesn't matter which one, but for concreteness center it at $(1/2, 1/2)$. Consider $\epsilon = 1/4$. Then $\mathcal{N}(A, \epsilon)$ is 16 because we need 16 squares of that side length to completely cover the A.

Exercise 4.31. In the previous example, consider $\epsilon = 1/8$ and determine $\mathcal{N}(A, \epsilon)$. Compare and contrast to your answer for the same epsilon for the line connecting the origin to $(1, 1)$.

The fact that $\mathcal{N}(A, \epsilon)$ is finite for any $A \in \mathcal{H}(X)$ is a consequence of compactness. (Indeed, the formal definition of compactness is that any open cover contains a finite subcover.)

Notice that unless A is a finite set it will be true that $\mathcal{N}(A, \epsilon) \to \infty$ as $\epsilon \to 0$. The real question is, *how* does it go to infinity? To give this question meaning, compare and contrast examples 4.27 and 4.30. It is clear that they will go to infinity at quite different rates, and this is related to the fact that they are fundamentally of different dimensions. The main diagonal is one-dimensional, whereas the square is two-dimensional; this fact appears in the covering numbers for various epsilons.

Exercise 4.32. Let $\mathcal{C}$ be the middle-thirds Cantor set and consider the subset of $\mathbb{R}^2$ given by

$$A = \{(x, y) \text{ such that } x \in \mathcal{C} \text{ and } y \in [0, 1]\}$$

The set A is pictured in figure 5 below. Let $\epsilon_n = 1/3^n$. Compute $\mathcal{N}(A, \epsilon_n)$ for $n = 1, 2$ and give a formula for a general n.

Exercise 4.33. Let A be a compact subset of $\mathbb{R}^d$ and let $\epsilon > 0$. Suppose that you are given $\mathcal{N}(A, \epsilon)$ and let $V_s(A)$ denote the s-volume of A.

(1) Given an approximate equation of $V_s(A)$ in terms of $\mathcal{N}(A, \epsilon)$. That is, fill in the right side of the expression $V_s(A) \approx$ ____________
(2) Solve your approximate equation for s, treating your $\approx$ sign as an equals sign.
(3) Your approximate equation must be true for all $\epsilon > 0$, and in fact becomes increasingly accurate as $\epsilon \to 0$ since the approximation by boxes becomes more accurate. What expression do you get for s in the limit?

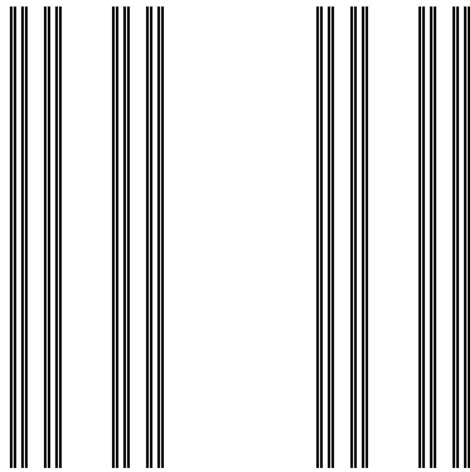


FIGURE 5. The set A for exercise 4.32.

This exercise leads us directly to another definition of dimension, simply called fractal dimension in [**Bar12**, p. 173] and box-counting dimension in [**FN**, p. 118].

DEFINITION 4.34. Let A be a compact subset of $\mathbb{R}^d$. The *box-counting dimension* of A is defined to be

$$Dim_B(A) = \lim_{\epsilon \to 0} \left(\frac{\ln(\mathcal{N}(A, \epsilon))}{\ln(1/\epsilon)} \right),$$

provided this limit exists.

The limit does not always exist, in which case it can be convenient to look at the "upper" and "lower" box dimensions instead, which are defined in terms of $\limsup$ and $\liminf$. We will not pursue that issue further here, but the interested reader can refer to [**PC09**, p. 87] or [**Fal06**, p. 41] for more details.

EXERCISE 4.35. Let X be the unit square and let $A = \{a, b, c\}$ be any set of three nonequal points in X. Compute $Dim_B(A)$.

It can be difficult to compute $\mathcal{N}(A, \epsilon)$ for general ϵ in many cases, making the limit in our definition of box-counting dimension intractable to use. If the limit exists, there are a number of equivalent ways of computing it, some of which are summarized in [**Fal06**, p. 43] and in the *Box Counting Theorem* of [**Bar12**, p. 175].[3] Of particular interest to us is the fact that instead of letting $\epsilon \to 0$ continuously it suffices to choose a constant $M > 1$ and let $\epsilon_n = 1/M^n$ in our computations of the box-counting dimension. We have the following lemma that makes the computation convenient.

[3] I should probably expand the boxes-in-a-grid idea that is used for approximation of the dimension for later versions of these notes.

LEMMA 4.36. *Let A be a compact subset of $\mathbb{R}^d$ and let $M > 1$. If the box-counting dimension of A exists, then*

$$Dim_B(A) = \lim_{n\to\infty}\left(\frac{\ln(\mathcal{N}(A, 1/M^n))}{n\ln(M)}\right)$$

EXERCISE 4.37. Fill in the blanks to prove the lemma.

PROOF. Let us assume that $Dim_B(A)$ exists. If that is the case, then since the limit exists as ϵ goes to 0 it must also exist when we let $\epsilon = 1/M^n$ and let $n \to \infty$ because in that case $1/M^n$ ____________________. [a] Thus we see that by definition $Dim_B(A) = \lim_{\epsilon\to 0}\left(\rule{5cm}{0.4pt}\right)$, which by substituting $\epsilon = 1/M^n$ becomes $Dim_B(A) = \lim_{n\to\infty}\left(\rule{5cm}{0.4pt}\right)$. By applying ____________________, the denominator becomes ____________________, which finishes the proof. □

[a]The interplay between continuous and sequential limits requires rigorous treatment in an analysis class, but in this situation there is no logical problem. A logical problem could arise if we tried to argue that a limit existing for $\epsilon = 1/M^n$ as $n \to \infty$ implied that it existed for $\epsilon \to 0$. We're not doing that here. Our sequence of ϵs is going to 0 along with the rest of them.

EXERCISE 4.38. Compute the box-counting dimension of the Koch curve.

EXERCISE 4.39. Compute the box-counting dimension of the example in exercise 4.32.

4.5. When the dimensions are equivalent

Box-counting and similarity dimension agree when the set A is the attractor of an iterated function system that not only is made of similarities, but also satisfies one of the first two conditions given in this definition.

DEFINITION 4.40. Let $(X; T_1, ..., T_n)$ be an iterated function system and let A be its attractor. The iterated function system is said to be *totally disconnected* if $T_i(A) \cap T_j(A) = \emptyset$ for all $i \neq j$ with $i, j \in \{1, 2, ..., n\}$. It is said to be *just-touching* if there is an open set[a] $\mathcal{O} \subset A$ for which

(1) $T_i(\mathcal{O}) \cap T_j(\mathcal{O}) = \emptyset$ for all $i \neq j$ with $i, j \in \{1, 2, ..., n\}$, and
(2) $\mathcal{T}(\mathcal{O}) \subset \mathcal{O}$. If the IFS is neither totally disconnected nor just-touching it is said to be *overlapping*.

[a]A set is open if its complement $X/\mathcal{O}$ is closed.

EXAMPLE 4.41. The middle-thirds Cantor set satisfies the definition of totally disconnected since $T_1(\mathcal{C}) \cap T_2(\mathcal{C}) = \emptyset$.

EXERCISE 4.42. Determine whether the Sierpinski triangle and Koch curve are totally disconnected. If they are not, give elements of A that are in more than one image of A.

The just-touching definition is a bit more subtle but isn't so bad when you understand it. Basically what it is saying in a mathematically precise way is that the attractor is allowed to overlap just on its "boundary points". The points you gave in the last exercise are such boundary points. In such a situation, the set $\mathcal{O}$ can be taken to be the part of the attractor that is on the "inside" in a way we will see in the next example. When the collage is applied to it, it stays on the inside and the images of the individual maps in the collage do not overlap. That is enough, it turns out, to ensure that the dimension computation is accurate.

EXAMPLE 4.43. We construct the set $\mathcal{O}$ for the Sierpinski triangle as follows. Here X is the unit square and $\mathcal{S}$ denotes the Sierpinski triangle. Let $P \subset X$ denote the right triangle connecting the origin to $(1,0)$ and $(0,1)$, which is the boundary of $\mathcal{S}$. Then let $\mathcal{O} = \mathcal{S}/P$, i.e., $\mathcal{O}$ is the set of points in $\mathcal{S}$ that are not in this boundary triangle. Note that $\mathcal{O}$ is open since its complement in X is just P, which contains its limit points and is therefore closed.

Now we must verify that it satisfies the two conditions. The first condition seems relatively clear, since the only way for the images to overlap is on the boundary, which we have removed. The second condition also checks out, since $\mathcal{T}(\mathcal{O})$ does not contain any portion of P and thus must be inside $\mathcal{O}$.

EXERCISE 4.44. Determine a set $\mathcal{O}$ for the Koch curve to show that it is just-touching.

EXAMPLE 4.45. The iterated function system in example 4.25 is overlapping.

THEOREM 4.46. [**Bar12**, p. 183] *Let $(X; T_1, ..., T_n)$ be an iterated function system, let $c_1, c_2, ..., c_n$ be the similarity ratios of $T_1, T_2, ..., T_n$ respectively, and let A be the attractor of the IFS. If the IFS is totally disconnected or just-touching then the box-counting dimension is the similarity dimension of A. If the IFS is overlapping, then $Dim_B(A) \leq Dim_S(A)$.*

4.6. Exercises

EXERCISE 4.47. Go back through your introductory calculus notes or book and give the proof that $\lim_{k\to\infty} 2k\sin(\pi/k) = 2\pi$. You may assume the identity $\lim_{h\to 0} \frac{\sin(h)}{h} = 1$.

EXERCISE 4.48. Use the methods of section 4.2 and especially problem 4.11 to obtain a value s for the dimension of the middle-thirds Cantor set.

EXERCISE 4.49. Let $A \subset \mathbb{R}$ be the set given by $A = \{1/n,$ such that $n = 1, 2, 3, ...\}$. Find the box-counting dimension of A.

EXERCISE 4.50. Let $A \subset \mathbb{R}$ be the set given by $A = \{1/n^2,$ such that $n = 1, 2, 3, ...\}$. Find the box-counting dimension of A.