# Fractals and Dimension

Rich Schwartz

November 3, 2014

## 1 What is a Fractal?

Asking what a fractal is in mathematics is a bit like asking: What is an object of art? It is hard to answer this question precisely. You might say that clearly a painting is an object of art, but suppose that someone has smeared toothpaste all over a canvas? Is it still art? Well, there might be some very beautiful toothpaste paintings. Modern painting is pretty broad, and a toothpaste composition is tame stuff for e.g. the folks over at R.I.S.D. Is a basketball placed in an aquarium an object of art? The Chicago Museum of Contemporary Art thinks so, but I'm a little bit skeptical. And so on.

The term *fractal* was popularized by Benoit Mandelbrot. There are certain classic sets, like the middle-third Cantor set and the Sierpinski triangle (described below) which one would certainly say are fractals. These sets exhibit *self-similarity*, meaning that the set breaks into a finite union of smaller copies of itself. However, this definition doesn't strictly work, because a line segment also has this property.

Another definition of a fractal is that it is a set whose dimension is not an integer. (Below I will explain what this might mean.) Technically, there isn't really a definition of dimension for sets; you need to have a distance relation on the set for the calculations to make sense. A set with a distance relation on it (which satisfies certain agreed-upon axioms) is called a *metric space*. The reason why people usually speak of fractal *sets* is that the sets are subsets of Euclidean space (e.g. the plane) and they automatically inherit a distance relation from the space in which they sit. So, all along, people mean *fractal metric space* when they say *fractal set*.

I would say that the most popular kinds of dimension are packing dimension, covering dimension, box dimension, and Hausdorff dimension. Box dimension is a notion which only works for subsets of Euclidean space and the other notions work for any metric space. Hausdorff dimension is probably the most "mathematically serious" kind of dimension used, but it is difficult to give a definition without a buildup of terminology. In these notes I'll focus on packing dimension, but I will also define covering demension. For the sets which are classically considered fractals – the household fractals, so to speak – all these dimensions coincide.

It is also problematic to say that a fractal is a set which has fractional dimension. For instance, you might have a set $A$ which has dimension $3/2$ and then you could define a new set $B$ which is the union of $A$ with a finite number of line segments. The line segments have dimension 1, and it turns out that $B$ will still have dimension $3/2$. However, you might feel like the presence of these line segments mars the original set, much like scratches on the canvas might mar a painting. You could do a little bit worse. For instance, you could take a countable union of fractals, $F_1 \cup F_2 \cup F_3...$ where $F_n$ has dimension $3/2 - 1/n$. The union will have dimension $3/2$, but is it still a fractal? If you say yes, then suppose that $F_n$ has dimension $2 - 1/n$. Then the union has dimension 2, but maybe you would still want to call it a fractal. And so on.

So, in short, I'm not really going to answer the question *What is a fractal*, but I will say that with practice you will know one when you see it.

## 2 Some Household Fractals

The most common kinds of (traditionally called) fractals are constructed using infinite processes, though many of them have other kinds of descriptions as well. The most classic example is the *middle third Cantor set*. This is the subset of the interval $[0, 1]$ consisting of real numbers whose base 3 expansion has no 1s in it. Alternatively, start with $S_0 = [0, 1]$ and cut out the middle third, leaving two intervals

$$S_1 = [0, 1/3] \cup [2/3, 1].$$

In general, given $S_n$, let $S_{n+1}$ denote the set obtained by chopping out the middle third of each of the intervals of $S_n$. The middle-third Cantor set is the intersection $S_1 \cap S_2 \cap S_3....$

Another example is the *Koch snowflake*. Start with an equilateral triangle $T_0$ of side length 1 and place 3 smaller equilateral triangles on the centers of the edges of $T_0$, as shown in the middle part of Figure 1. The new triangles should have side length 1/3. Call the result $T_1$. In general, place small equilateral triangles at the centers of the edges of $T_n$ to produce $T_{n+1}$. These triangles should have size $(1/3)^{n+1}$. Figure 1 shows a few steps of this construction. Let $T = T_1 \cup T_2 \cup T_3 ...$ be the union. The boundary of $T$ is the Koch snowflake.

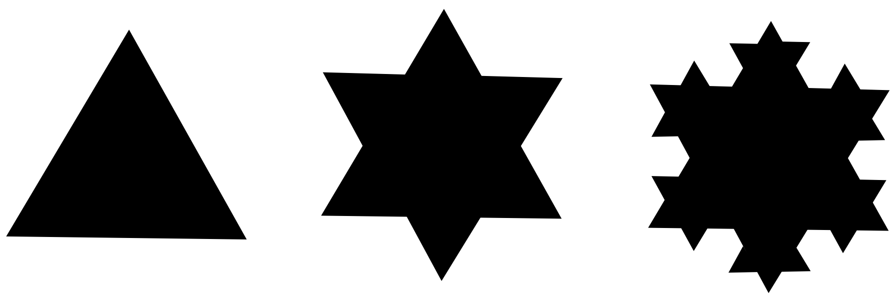

**Figure 1:** Three steps in the construction of the Koch snowflake.

Another example is the Sierpinski Triangle. Start with an equilateral triangle $T_0$, break the triangle into 4 equal piaces, and cut out the (interior of the) middle one. This leaves $T_1$, as shown in Figure 2. In general, cut out the middle of all the triangles in $T_n$ to produce $T_{n+1}$. The intersection $T_1 \cap T_2 \cap T_3 ...$ is called the Sierpinski triangle.

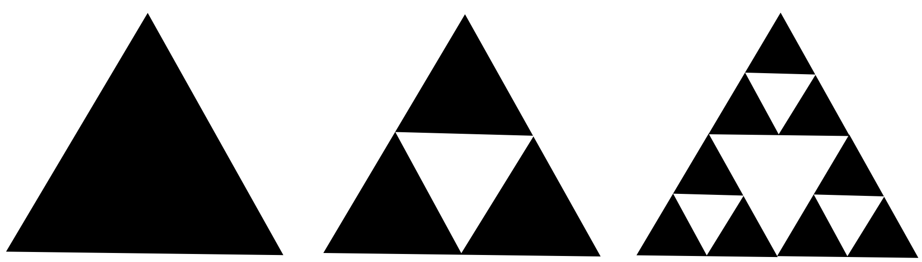

**Figure 2:** Three steps in the construction of the sierpinski triangle

There are some other examples too, such as the Sierpinski carpet, the Menger sponge, the dragon curves. The internet has some great pictures of these sets.

You could imagine modifying these constructions in various ways. For instance, you could cut out randomly chosen intervals when making the Cantor set, or you could add on triangles of random shapes and sizes when making the Koch snowflake. Are the results still fractals? Yeah, sort of.

# 3 Metric Spaces

On your first pass through mathematics, you encounter sets contained in Euclidean space (e.g. the plane), like the ones drawn above. For this reason, you might not think too much about the distinction between sets and metric spaces. The notion of Euclidean distance is so ingrained that it is practically invisible. In this section, I'm going to define what is meant by a metric space. Such a space need not be a subset of Euclidean space.

A *metric space* is a set $X$ together with what is called a distance relation $d$. The distance relation satisfies the following axioms.

1. (Positivity) $d(x,y) \geq 0$ for all $x,y \in X$.
2. (Nontriviality) $d(x,y) = 0$ if and only if $x = y$.
3. (Symmetry) $d(x,y) = d(y,x)$ for all $x,y \in X$.
4. (Triangle Inequality) $d(x,y) + d(y,z) \geq d(x,z)$ for all $x,y,z \in X$.

These properties are abstract versions of the usual properties that hold for distances in Euclidean space. The first three are pretty obvious. The last one expresses the idea that the sum of the lengths of two sides of a triangle is at least as large as the length of the third side. This property might seem a bit arbitrary at first, but it turns out to be very important.

A metric space is usually denoted by $(X,d)$. In case $X$ is already a subset of Euclidean space, the quantity $d(x,y)$ is just the ordinary Euclidean distance between two points $x,y \in X$. In this case, the axioms are certainly satisfied.

Many metric spaces do not live in Euclidean space. Let $X$ be the set of integers. Define $d(x,y) = 0$ if $x = y$ and $d(x,y) = 1$ if $x \neq y$. So, in other words, every two integers are defined to be 1 apart. Let's think about this metric space geometrically. If we just pick 3 points in $X$, we can think of them as the vertices of an equilateral triangle. If we pick 4 points, we can think of them as the vertices of a regular tetrahedron. If we pick 5 points, we can think of them as the vertices of a regular 4-dimensional tetrahedron. There is nothing wrong with trying to picture $(X,d)$ this way, but you can see that you have to go up a dimension every time you add a point to the image you are building. For this reason, there is no subset of a finite dimensional Euclidean space which has the same distance relation as $X$. The space $(X,d)$ is the set of vertices of an infinite dimensional regular tetrahedron!

Here is another example. Suppose that $X = [0,1]$, the interval from 0 to 1. We define a new distance on $X$ using the formula

$$d(x,y) = \sqrt{|x-y|}.$$

It is pretty obvious that $d$ satisfies the first 3 axioms above. It is also not too hard to show that $d$ satisfies the last property: It boils down to the fact that $\sqrt{t} + \sqrt{1-t} \geq 1$ for all $t \in (0,1)$.

Here is another wierd example. Again let $X$ be the set of integers. Define $d(m,n) = 3^{-k}$ where $k$ is the largest integer such that $3^k$ divides $m-n$. For example $d(2,11) = 1/9$ because $11 - 2 = 9$ and hence 9 divides $11 - 2$. Here are some other distances in $X$:

- $d(2,8) = 1/3$ because 3 divides $6 = 8 - 2$ but 9 does not divide 6.
- $d(2,83) = 1/81$.
- $d(11,18) = 1$.

It turns out that $(X,d)$ is a metric space. Geometrically $(X,d)$ looks a lot like a Cantor set. Also, $(X,d)$ is very closely related to the 3-adic numbers. In fact, $d$ is called the 3-adic metric.

**Metric Balls:** Let $(X,d)$ be a metric space, let $x \in X$, and let $r > 0$ be some real number. The *ball of radius* $r$ *about* $x$ is the set of all $y \in X$ such that $d(x,y) \leq 1/r$. If you try this definition out in Euclidean space, you get the usual notion of a ball.

Let's consider our last example again. What is the ball of radius 1 about the point 5? It is the set of all integers $n$ such that $n - 5$ is divisible by 3. For instance 8 is in this ball and 7 is not. What is the ball of radius 1/3 about 5? It is the set of all integers $n$ such that $n - 5$ is divisible by 9. So, for instance, 14 is in the ball of radius 1/3 about 5 but 8 is not.

Here is one property I will talk a lot about later. If two points in $X$ are at least $2r$ apart, then the balls of radius $r$ about each of these points are disjoint from each other: they have no points in common. Why does this work? Well, the triangle inequality guarantees it. This is one hint that the triangle inequality is important. If you have a bunch of points that are all separated by $2r$ than all the balls of radius $r$ around these points are disjoint from each other.

## 4 Growth of a Sequence

Before I talk about the notion of dimension, I want to talk about how fast sequences grow. Suppose that we have a sequence $A_1, A_2, A_3, ...$ and we suspect that the sequence grows like some polynomial. That is $A_n \approx n^d$ for some exponent $d$. Don't worry too much now what the symbol $\approx$ means. One way to capture what we want to is to computer the quantity

$$\lim_{n \to \infty} \frac{\log(A_n)}{\log(n)}. \tag{1}$$

This expression looks a bit forbidding at first, but let's try it out.

Let's consider the sequence where $A_n = n^2$. So, this sequence starts out $1, 4, 9, 16, 25, ....$ We compute

$$\frac{\log(A_n)}{\log n} = \frac{\log(n^2)}{\log n} = \frac{2 \log n}{\log n} = 2.$$

In this case, the limit in Equation 1 is 2. Let's try a messier example. Suppose that $A_n = 365n^2$. We have

$$\frac{\log(A_n)}{\log n} = \frac{\log(365n^2)}{\log n} = \frac{2 \log n}{\log n} + \frac{\log 365}{\log n}.$$

The second term goes away in the limit, and again the answer is 2.

Suppose we have $A_n = 111n^2 + 17n$. In this case, we have

$$n^2 < A_n < 365n^2.$$

If we apply Equation 1 to either the left sequence or the right sequence, we get 2. Therefore, we get 2 for the middle sequence as well.

By now, you can probably guess that if we take

$$A_n = C_d n^d + C_{d-1} n^{d-1} + ... + C_2 n^2 + C_1 n + C_0,$$

then the limit in Equation 1 is $d$. (Here $C_0, C_1, ..$ are constants whose values aren't important.) At least for polynomial sequences like this, the limit in Equation 1 picks out the biggest exponent. The quantity in Equation 1, when it exists, is sometimes called the *polynomial growth* of the sequence.

## 5 Some Geometrically Defined Sequences

Now imagine that we start with a line segment $S$. Suppose we let $A_n$ be any of the following quantities:

- The minimum number of intervals of length $1/n$ needed to completely cover $S$.
- The maximum number of points you can place in $I$ so that every two points are at least $1/n$ apart.

Then both sequences have growth 1.

Now imagine that we replace the word *interval* by the word *square*. Then the new sequences have growth 2. That is, we can cover the square $S$ by about $n^2$ squares of side-length $1/n$, and we can place about $n^2$ points inside $S$ so that every two points are separated by at least $1/n$. The exact number might be a bit hard to figure out; it depends on the size of $S$. Whatever the answer, the growth is always 2.

Finally, imagine that we replace the word *square* by the word *cube*. Now we get sequences whose growth is 3. In all three cases, the growth of the sequences coincides with the dimension of the object.

The construction is pretty robust: If we used triangles or pentagons in place of squares, we would still get sequences of growth 2. Likewise, if we used solid polyhedra, or solid balls, or solid ellipsoids instead of cubes, we'd still get sequences of growth 3. So, using the above sequences are a pretty good way to compute dimension.

## 6 Packing Dimension

Let $(X, d)$ be a metric space. (If you don't like metric spaces, you can just think about sets in the plane.) Define $A_n$ to be the maximum number of points you can put in $X$ so that no two points are within $1/n$ of each other. If we call the set of our points $S$, then we are saying that $d(x, y) \geq 1/n$ for all $x, y \in S$.

If the growth of $A_n$ exists, number is defined to be the *packing dimension* of $S$. It might happen that the limit in Equation 1 does not exist (because not all sequences converge) but let's not worry about this. In case the limit doesn't exist, the packing dimension is not defined.

The reason for the name is as follows. Suppose that we have placed points in $X$ so that every two points are at least $1/n$ apart. Then, if we place balls of radius $1/(2n)$ about each of our points, all these balls are disjoint. So, when we are filling up $X$ with points like this, we are asking how many balls of a given size we can pack into $X$ without overlaps.

Let's try an example: The middle-third Cantor set. At the $k$th stage of the construction of the middle-third Cantor set, you are left with $2^k$ intervals, each of which has length $(1/3)^k$. Moreover, these intervals are all separated from each other by a distance of at least $(1/3)^k$.

So, if we take $n = 3^k$, then we can place one point in the center of each interval and get $2^n$ points which are all separated by at least $1/n$. On the other hand, we can't put more than one point in any interval. So, for this choice of $n$, we have $A_n = 2^k$.

Let's say it again: Whan $n = 3^k$, we have $A_n = 2^k$. This gives us

$$\frac{\log 2^k}{\log 3^k} = \frac{k \log 2}{k \log 3} = \frac{\log 2}{\log 3}.$$

We haven't defined $A_n$ for every value of $n$. We've only defined it for every value of $n$ which has the form $3^k$. If we believe that the full sequence $A_n$ does have a well-defined growth, then this growth must be $\log(2)/\log(3)$. So, if the middle third Cantor set has a well-defined packing dimension, then the dimension must be $\log 2/\log 3$.

This discussion was a bit unsatisfying, so let's prove that the middle-third Cantor set really does have a packing dimension. That is, let's deal with the values of $A_n$ at other values of $n$. We know that the sequence $A_1, A_2, A_3, ...$ is not decreasing. So, if we have $3^{k-1} < n < 3^k$, then we have $2^{k-1} \leq n \leq 2^k$. But this means that

$$\frac{k}{k-1} \times \frac{\log 2}{\log 3} \leq \frac{\log A_n}{\log n} \leq \frac{k-1}{k} \times \frac{\log 2}{\log 3}.$$

As $k \to \infty$, both sides converge to $\log 2/\log 3$, so the limit in Equation 1 really does exist. All in all, the middle third Cantor set has packing dimension $\log 2/\log 3$.

# 7 Covering Dimension

Now I'm going to define another kind of dimension. Given a metric space $X$, let $A_n$ be the smallest number of points you can place in $X$ so that every point of $X$ is within $1/n$ of the points you have placed. The reason for the name is that, if you place a ball of radius $r$ around each of the points, you will have covered all of $X$ with balls. When the sequence has a well defined growth, the growth of $A_n$ is called the *covering dimension* of $X$.

The covering dimension of the middle third Cantor set is $\log(2)/\log(3)$. To see this, we take $n = 3^k/2$. If we consider the $2^k$ center points of the intervals of size $(1/3)^k$, then every point of the Cantor set is within $1/n$ of one of the points. Moreover, this is the best we can do. This time we have

$$\frac{\log(A_n)}{\log(n)} = \frac{k\log 2}{k\log 3 - \log 2}.$$

As $k \to \infty$ the limit is $\log 2/\log 3$. The extra term does not matter. The same arguement as for the packing dimension shows that we would get the same answer if we compute the limit for all values of $n$, and not just for the special values we have considered.

Notice that the middle third Cantor set has the same packing dimension and covering dimension. For most metric spaces you encounter, the two dimensions exist and coincide. You have to work hard to find a set where this doesn't happen. There are many variants of the two kinds of dimensions I've talked about in these notes, but the two I've talked about are probably the easiest to define.

# 8 Completeness

The notions of packing and covering dimension work well for certain kinds of metric spaces, but they work terribly for others. For instance, let's take $X$ to be the set of rational numbers between 0 and 1, and let $d$ be the usual metric on the line, restricted to points of $X$. That is $d(x,y) = |x-y|$. In this case, the packing and covering dimension of $X$ will be 1, just as it would be for the whole interval $[0,1]$. On the other hand, $X$ is just a countable set of points, so any sane person would want to say that $X$ has dimension 0. What is going on?

The problem is that our space $X$ is has lots of holes in it, but these holes are not detected by our definitions above. The usual fix for this problem is to insist that $X$ must be a *complete* metric space. Informally, this means that $X$ has no holes the way that the rational numbers do. For the sake of completeness (of exposition) I will give a precise definition of what it means for $X$ to be a complete metric space.

An infinite sequence $\{x_n\}$ of points in $X$ is called *Cauchy* if, for every $\epsilon > 0$, there is some $N$ such that $d(x_m, x_n) < \epsilon$ provided that both $m$ and $n$ are greater than $N$. In other words, as you look farther and farther out the sequence, the points are settling down: they jump around less and less.

The sequence $\{x_n\}$ is called *convergent* if there is some $y \in X$ so that $d(x_n, y) \to 0$ as $n \to \infty$. The difference between a Cauchy sequence and a convergent sequence is that a Cauchy sequence is settling down and a convergent sequence is settling down *to a point of* $X$. Every convergent sequence is a Cauchy sequence, but it might happen that there are Cauchy sequences which are not convergent sequences.

Here's an example: If $X$ is the set of rationals between $(0, 1)$ we can take any sequence which converges to $\sqrt{2}/2$. This sequence is a Cauchy sequence in $[0, 1]$ because it converges to $\sqrt{2}/2$. However, it is not a convergent sequence in $X$.

The space $X$ is called *complete* if every Cauchy sequence in $X$ is a convergent sequence in $X$. This is one way to say that $X$ has no holes. The space $[0, 1]$ is complete, but the set of rationals in $[0, 1]$ is not complete.

Typically, the notions of packing and covering dimension are only applied to complete metric spaces. They can be applied to all metric spaces, but they will sometimes yield screwy results.

# 9 Problems

**1.** Let $A_n = 23n^3 + 17n^2 + 119n$. Show that this sequence has growth 3.

**2.** Draw a careful picture of the first 4 steps of the Koch snowflake.

**3.** Let $X = [0, 1]$ and let $d(x, y) = \sqrt{|x - y|}$. Show that $(X, d)$ has packing dimension 2.

**4.** Compute the packing dimension or the covering dimension of the Sierpinski triangle. Take your pick.