Fractal Working Group [1]

Coordinators: Gerald Edgar, Larry Lindsay

Notes LaTeX-ed in real-time by Steven Miller

Participants: Bruce Adcock, Youri Dimitrov,
Gerald Edgar, Dean Eiger, Charles Estill,
Larry Lindsay, Steven Miller, James Tyne

[1]Ohio State, Autumn 2003, Wed. and Friday: $1:30-2:30$pm, MA 417

Abstract

The purpose of this working group was to introduce participants to various areas of fractal research. Meetings were held twice a week, Wednesday and Friday. The format of the meetings was usually an informal lecture by one of the participants. After a brief introduction to a variety of topics by the organizers, the other participants were encouraged to investigate specific areas of research and report back to the group. The following report records the lectures that were presented. Topics included definitions of dimension, iterated function systems, Julia sets, L-systems, fractal curves, and applications. Originally, the notes were taken in real-time during the lecture, and at the end of the quarter participants were allowed to edit the notes for errors and to enhance clarity.

Contents

Chapter 1

Introductory Lecture

Lecturer: Larry Lindsay (Wednesday, October 8, 2003)

The following is the last of three informal lectures on various topics in fractals. The purpose of this lecture (as well as the previous lectures, not included in this report) is to give participants in the Fractal Working Group an introduction to fractals. The wide assortment of topics is intended to demonstrate the diversity of research in the study of fractals. In the previous lectures we briefly touched on the history of fractal research, different ways to make a fractal, the intuitive idea of dimension, dynamical systems and chaos, analysis on fractals, the chaos game, and applications of fractals. In this lecture we will discuss probabilities on fractals, box-counting and Hausdorff dimension, cellular automata as fractal generators, graph-directed fractals, and quantization dimension.

1.1 Probabilities on Fractals

Much attention has been given to studying probabilities on fractals. One reason for this is that it leads to research on "multifractals," which requires that a probability be defined on the fractal. (We will talk about multifractals in the next lecture.) One way of assigning probabilities is similar to a chaos game. I will briefly present a visual example here. Start with a big circle, and consider four non-overlapping circles inscribed inside. Think of these as similarity maps which shrink the larger circle by $\frac{1}{2}$ for two (the right and left), and $\frac{1}{3}$ for the other two (top and bottom). From a seed point, say the center, pick these maps at random and apply the sequence of maps for several iterations. Plot the resulting point at each

iteration. If we give each map probability $\frac{1}{4}$, then this is like symbolic dynamics where we generate a sequence of four symbols $(1, 2, 3, 4)$, with each symbol having the same probability of being chosen. We are interested in what shapes we get after a large number of iterations – in this example, after 40 iterations, it looks like four "speckled" diamonds.

But the two smaller diamonds look a little darker, and this gives us an intuitive picture of a probability supported on a fractal. In this example, a more "natural" measure would give each equal darkness, with the smaller diamonds having less than $\frac{1}{4}$ probability and the larger diamonds having more than $\frac{1}{4}$ probability.

What exactly are the probabilities which make the diamonds equally dark? In our example the map scalings are $\frac{1}{2}, \frac{1}{2}, \frac{1}{3}, \frac{1}{3}$, so we would need to solve for d such that

$$\left(\frac{1}{2}\right)^d + \left(\frac{1}{2}\right)^d + \left(\frac{1}{3}\right)^d + \left(\frac{1}{3}\right)^d = 1. \tag{1.1}$$

Then the four numbers on the left of the equation are the probabilities, say p_1, p_2, p_3, p_4, needed to make the diamonds equally dark.

The idea of assigning arbitrary probabilities to the basic maps of an iterated function system is a generalization of the above example, and the above procedure would lead to an intuitive physical picture of the probability.

1.2 Box-counting Dimension

Upper and lower box-counting dimension. Given a set E, let $N(r)$ be the minimum number of balls of diameter r needed to cover E. Obviously, E needs to be bounded. The upper box-counting dimension is defined as

$$\overline{\dim}_B(E) = \overline{\lim}_{r \searrow 0} \frac{\log N(r)}{\log \frac{1}{r}}. \tag{1.2}$$

We also define the lower box-counting dimension using liminf, and if the two agree then we get the box-counting dimension of E. The idea is that $N(r) \sim \left(\frac{1}{r}\right)^d$.

As an example, let $E = [0, 1]$. Clearly we get approximately 1 each time in the calculation, using various values of r, and in the limit as $r \to 0$ we definitely get 1. So the box-counting dimension of $[0, 1]$ is 1, which is a good thing because an interval is normally thought of as a 1-dimensional object!

As our next example, take the rationals $\mathbb{Q}$ in $[0, 1]$. This is a countable set. What should the box-counting dimension be? We get a dimension of 1 again, since

we always need the same number of boxes to cover the rationals as we needed to cover all of $[0, 1]$. From a mathematician's standpoint this actually makes us a bit uneasy, because we have here a countable set (something relatively insignificant) with positive dimension. Other definitions of dimension give any countable set a dimension of 0. Despite this drawback box-counting dimension remains popular because it is relatively easy to calculate and understand. Here is a theorem which applies to the above example:

Theorem 1.2.1. *$dim_B(E) = dim_B(\overline{E})$, where $\overline{E}$ is the closure of E.*

Let me reiterate that we need a bounded set to do box counting. If we tried to do the calculation on an unbounded set, say $E = \mathbb{R}$, then we would get $\dim_B(E) = \infty$. This is not good, since $\mathbb{R}$ should intuitively have dimension 1.

1.3 Hausdorff dimension

Hausdorff dimension doesn't have this problem. We get

$$\dim_H(\mathbb{R}) = 1. \tag{1.3}$$

The reals are a countable union of intervals of length 1, and there is a theorem which states the following: a countable union of sets has Hausdorff dimension equal to the supremum of the Hausdorff dimensions of the individual sets.

Without using the theorem, though, suppose we want to calculate the Hausdoff dimension of the entire real line. We begin with the following quantity, where $\{A_i\}$ represents a certain covering of E:

$$H^{\delta}_{\epsilon}(\mathbb{R}) = \inf\{\sum |A_i|^{\delta} : \mathbb{R} \subset \cup A_i : |A_i| \leq \epsilon\}. \tag{1.4}$$

Given n, we could use $2n$ intervals of length $\frac{1}{n}$ to cover $[-1, 1]$. We could cover $[-2, -1]$ and $[1, 2]$ by intervals of length $\frac{1}{2n}$, and we would need $4n$ such intervals. We could use $8n$ intervals of size $\frac{1}{4n}$ to cover $[-3, -2]$ and $[2, 3]$. Continue in this way until we cover all of $\mathbb{R}$.

For this cover, we get

$$\begin{aligned}\sum |A_i|^\delta &= 2n\left(\frac{1}{n}\right)^\delta + 4n\left(\frac{1}{2n}\right)^\delta + \cdots + 2^k n\left(\frac{1}{2^{k-1}n}\right)^\delta + \cdots \\ &= \sum_{k=1}^{\infty} 2^k n \frac{1}{(2^{k-1}n)^\delta} \\ &= \frac{n}{n^\delta} 2^\delta \sum_{k=1}^{\infty} 2^{(1-\delta)k}. \end{aligned} \tag{1.5}$$

Now, to define Hausdorff dimension we first define the δ-dimensional Hausdorff measure

$$H^\delta(\mathbb{R}) = \lim_{\epsilon \searrow 0} H^\delta_\epsilon(\mathbb{R}), \tag{1.6}$$

and the Hausdorff dimension $\dim_H(\mathbb{R})$ is the value of δ where the δ-dimensional Hausdorff measure jumps from infinity to zero. In this particular example if $\delta > 1$, no matter what ϵ we use, the infimum defining $H^\delta_\epsilon(\mathbb{R})$ is 0 (just take larger and larger values of n in the above equation). If $\delta < 1$, it is infinity. If $\delta = 1$, it is infinity. We conclude that the Hausdorff dimension of $\mathbb{R}$ is 1, and the 1-dimensional Hausdorff measure of $\mathbb{R}$ is infinity.

1.4 Fun Exercise

Let $E = \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\}$. We get the Box Counting dimension is $\frac{1}{2}$! (But, of course, the Hausdorff dimension is 0, because E is a countable union of points, and a single point clearly has Hausdorff dimension 0.)

1.5 Example

Here is an example in which

$$\underline{\dim}_B(E) < \overline{\dim}_B(E). \tag{1.7}$$

For this example we will construct a special Cantor set. Let $0 < a < b \leq 1/3$. In the construction of the traditional middle-third Cantor set we start with $[0, 1]$ and take out the middle third, and then take out the middle third of the remaining intervals, and continue in this manner. In this example we have two choices:

Either take out $1-2a$ from the middle of the remaining intervals in one scenario, or take out $1-2b$ from the middle of the remaining intervals in the second scenario (do it symmetrically so the sets on left/right are both of length a or b).

At Level n in the construction of the Cantor set there will be 2^n intervals, and in this example they each have length $a^{s_n}b^{t_n}$, where $s_n + t_n = n$. (s_n is the total number of times we used a, and t_n is the total number of times we used b.)

At each level we have a choice between a and b. To make these choices we consider the calculation involved for upper box-counting dimension:

$$\begin{aligned}
\overline{\dim}_B(E) &= \overline{\lim}_{r\searrow 0}\frac{\log N(r)}{\log\frac{1}{r}} \\
&\geq \overline{\lim}_{n\to\infty}\frac{\log 2^n}{\log\frac{1}{a^{s_n}b^{t_n}}} \\
&= \overline{\lim}_{n\to\infty}\frac{n\log 2}{s_n\log\frac{1}{a}+t_n\log\frac{1}{b}}. \qquad (1.8)
\end{aligned}$$

By consistently following a string of a's with an appropriately long sequence of b's we can get this quantity close to $\log 2/\log\frac{1}{b}$. Similarly, for lower box-counting dimension, by consistently following a string of b's with an appropriately long sequence of a's we can get $\underline{\dim}_B(E)$ close to $\log 2/\log\frac{1}{a}$. In this way we can make $\underline{\dim}_B(E) < \overline{\dim}_B(E)$.

Question: What if we "randomly" choose a and b? (Many people study randomly generated fractals.) In this case, with probability one, the box-counting dimension would exist, and the "expected" dimension would be the same as the dimension we would get by simply alternating between a and b.

1.6 Cellular Automata

A classic example in the study of cellular automata is the Game of life (invented by John H. Conway). Here is a brief description. Start with a 2-dimensional grid in which some boxes are shaded and some are not, indicating if a box is "alive" or "dead." We start with some configuration at time $t = 0$. At time $t = 1, 2, 3, \ldots$ we update the configuration of shaded boxes by a certain rule, which determines whether or not a given box is alive or dead based on the condition of its neighbors at the previous stage. The general idea is this: You die from having too few living neighbors (die from loneliness), or you die from having too many living neighbors

(die from overcrowding). When conditions are just right, you live (or you are born if you weren't previously alive).

Now consider a 1-dimensional example. The boxes are in a single line. Again suppose there is some rule, which determines whether you are alive or dead at each point in time $t = 0, 1, 2, \ldots$. For visual purposes we align the evolution of our configurations, with configuration 0 above configuration 1 above configuration 2, etc.

Here is a simple rule: You remain alive (or come alive) if your neighbors on your left and right are different, and you die if they are the same.

Starting with a single live cell, we can think of this system as the evolution of polynomials $(1 + x)^n$ mod 2. (Or Pascal's triangle mod 2.) A shaded box corresponds to a coefficient of 1 for the appropriate power of x. We get

$$\begin{aligned} &1 \\ &1 + x \\ &1 + x^2 \\ &1 + x + x^2 + x^3 \end{aligned} \tag{1.9}$$

If we do this for a very long time and look at the resulting figure, we see that it's a Sierpinski triangle, oriented differently of course (right triangle).

More generally, we can take a prime p, and for our polynomial coefficients use the finite field $\mathbb{Z}/p\mathbb{Z}$. We then take a given polynomial and raise it to higher and higher powers mod p (or actually p^k should work, too). A surprising fact is that the resulting fractal turns out to be a graph-directed fractal, which is a generalized version of an iterated function system.

Question 1.6.1. *What about using Gaussian integers instead of primes?*

1.7 Graph-directed Fractal

In an iterated function system we start with a seed set and two or more maps of the set into itself. We then iterate these maps in all possible ways, ad infinitum. This can be generalized to several seed sets, and various maps between these sets, in which case we get several limit sets and call the resulting collection of sets a graph-directed fractal. In the particular example above, with cellular automata, to see the graph-directed fractal we have to keep zooming back and rescaling

appropriately (as we take higher powers of our polynomial), and in this context it is called the evolution set.

1.8 Quantization Dimension

My research has been in quantization dimension, and I will talk more about this later. For now I will give a brief introduction. In order to talk about quantization dimension we need a probability, and in the interesting case it will be a measure on a fractal. Therefore, one might suspect it is related to multifractals (it is, but we won't get into that here). The idea of quantization is useful for a variety of applications, and in fact it comes from electrical engineering.

For an intuitive example, say you have a signal coming to you, with data that you want to store. This signal might be for the colors of the pixels for some picture. But you can only store a finite number of colors, not the entire spectrum. Say the spectrum runs through Red - Yellow - Green - Blue.

It might work well to "uniformly" choose colors in the spectrum. If we are allowed three bits, we can assign colors using numbers ranging from 000 to 111. This is overly simplified, but of course, on a computer you can only work with finitely many things.

Now suppose that the colors in our picture are not uniformly distributed. Perhaps there are several shades of green, but in our original choice of quantization levels there may only be 2 possible shades of green (not good enough). If we really need to distinguish greens better, then we might want several greens. We would have to give up something somewhere else, so maybe we don't care about red as much. Skew the color distribution (of our quantizing levels), so that we can distinguish greens well, but reds not as well. (Example: a military analyst may need to distinguish between subtle shades of green in a satellite photo.)

Given a random picture from a random space of pictures (where there are some biases perhaps in the space it's drawn from), the engineer would want to figure out the best way of coding data so as to minimize distortion in the reproduced image.

Suppose we have a spectrum of colors (with a continuum of possible colors). Given n we are only allowed to quantize with n levels, so we need to pick the n most representative colors. If this choice has been made, then a given color in the signal gets assigned to one our representatives according to a Voronoi diagram. That is, it is assigned to the closest representative color. Effectively, we have replaced our spectrum of colors with a finite number of regions of constant color.

How should we choose points? This is done so as to minimize expected dis-

tortion. Without getting into specific definitions, let e_n be the minimum expected distortion (there are several ways of doing this). The quantization dimension is (if it exists)

$$d = \lim_{n\to\infty} \frac{\log n}{-\log e_n}. \tag{1.10}$$

The idea is that the quantization dimension gives us some idea as to how fast the distortion goes to zero as we quantize with more and more points: $e_n \approx \left(\frac{1}{n}\right)^{\frac{1}{d}}$.

As an example take the middle-third Cantor set. We need a measure, so consider the uniform distribution (probability take left or right is $\frac{1}{2}$). The following argument will be non-rigorous but will suggest what should happen. We will run through a sub-sequence of n's in order to calculate the limit in the definition of quantization dimension. When $n = 1$, we can use only 1 point, so we pick it intuitively in the middle, and we get some error e_1.

When $n = 2$, we can use two points, so we put each in the middle of the left and right halves of the Cantor set:

$$[--] * [--] \qquad\qquad [--] * [--] \tag{1.11}$$

We intuitively know what to do for $n = 1, 2, 4, 8, 16, \ldots$, and by self-similarity we get

$$\begin{aligned} n &= 1 \qquad e_1 \\ n &= 2 \qquad e_2 = \frac{1}{3}e_1 \\ n &= 4 \qquad e_4 = \frac{1}{9}e_1 \end{aligned} \tag{1.12}$$

In the calculation of quantization dimension we get

$$\begin{aligned} d &= \lim_{n\to\infty} \frac{\log n}{-\log e_n} \\ &= \lim_{k\to\infty} \frac{\log 2^k}{-\log(\frac{1}{3^k}e_1)} \\ &= \frac{\log 2}{\log 3}. \end{aligned} \tag{1.13}$$

This should be reassuring, because this number agrees with other definitions of dimension.

The above reasoning is certainly not rigorous, but the result is true. In fact, research into optimal configurations for $n \neq 2^k$ has led to some interesting questions which researchers are still trying to answer.

Chapter 2

Multifractals

Lecturer: Larry Lindsay (Friday, October 10, 2003)

The following lecture is motivated by the David Harte book, *Multifractals: Theory and Applications*. I will discuss the basics of multifractal analysis and then talk about how the spatial distribution of earthquake locations can be modelled as a multifractal. A chapter of Harte's book is devoted to earthquake analysis. Since much of the earthquake discussion in this lecture involves looking at figures from the book and the figures are not shown here, the following will mainly address the basic ideas of multifractal analysis with an occasional mention of the application to earthquakes.

2.1 Multifractal Measures

The term multifractal actually refers to a measure on a space. The set having "full measure," of course, tends to be a fractal, but in a natural way we are able to obtain fractals within fractals, and hence we use the term multifractal. Therefore, we start with a measure m, usually a probability measure, so that $m(A)$ is the probability of a randomly chosen point belonging to A.

We have seen examples of how we can visualize a measure on a fractal (chaos game). Recall that we got a static picture, which may have thousands of points, showing roughly, say, the Sierpinski triangle. By looking at the relative density of points in an area we get a picture of the bias of a randomly chosen point being in one part over another.

We will be looking at pictures of earthquake events, and seeing where they

clump. This gives an indication of those areas where there is a higher probability of having an earthquake. Because of the intuitive nature of fault zones, which have an apparent self-similarity in their crack systems, it seems appropriate to assume that a multifractal structure exists. With this in mind one collects data on earthquake locations and then tries to analyze that data using ideas from multifractal analysis.

Here are some of the ideas involved. We begin with the set of points having local dimension α:

$$E_\alpha \;=\; \{x : \lim_{r \searrow 0} \frac{\log m(B_r(x))}{\log r} \;=\; \alpha\}. \tag{2.1}$$

This is how we are able to obtain, as mentioned above, fractals within fractals. Typically we have

$$a_{\min} \;\leq\; \alpha \;\leq\; a_{\max}. \tag{2.2}$$

Definition 2.1.1 (Multifractal spectrum). $f(\alpha) \;=\; \dim_H(E_\alpha)$.

2.2 Generalized Rényi Dimensions

See also the paper by Hentschel and Procaccia, *The infinite number of generalized dimensions of fractals and strange attractors.*

Define

$$\Theta(q) \;=\; \lim_{r \searrow 0} \frac{\log \int_{X_r} m(B_r(x))^{q-1} dm(x)}{\log r}, \tag{2.3}$$

where

$$X_r \;=\; \{x : m(B_r(x)) > 0\}. \tag{2.4}$$

$$D_q \;=\; \begin{cases} \frac{\Theta(q)}{q-1} & \text{if } q \neq 1 \\ \lim_{r \searrow 0} \frac{\int \log m(B_r(x)) dm(x)}{\log r} & \text{if } q = 1 \end{cases} \tag{2.5}$$

where in the case $q = 1$, $\log 0 = 0$.

Keep in mind that $\int f(x) dm(x)$ is the expected value of f with respect to the measure m. Also, $\left(\int f^{q-1}(x) dm(x)\right)^{\frac{1}{q-1}}$ is the $q-1$ norm of f with respect to the measure m. We have

$$D_q \;=\; \lim_{r \searrow 0} \frac{\log \left(\int m(B_r(x))^{q-1} dm(x)\right)^{\frac{1}{q-1}}}{\log r}. \tag{2.6}$$

Why the $q-1$ norm? Notice that the integral

$$\int m(B_r(x))^{q-1} dm(x) \tag{2.7}$$

can be estimated with the sum

$$\sum_i m(B_i)^{q-1} m(B_i), \tag{2.8}$$

where we have divided a region into a grid of small sub-boxes B_i.

Alternatively, Hentschel-Procaccia define

$$D_q = \frac{1}{q-1} \lim_{r \searrow 0} \frac{\log \sum_i m(B_i)^q}{\log r}, \tag{2.9}$$

where r is the length of a side of a box B_i.

2.3 Relations

We have defined the Multifractal spectrum $f(\alpha)$ and the generalized Renyi dimensions $\Theta(q)$. So how are they related? Typically, $f(\alpha)$ and $\Theta(q)$ form a Legendre transform pair. See the book by Kenneth Falconer, *Techniques in Fractal Geometry*, pages 194 - 195.

Consider again the Hentschel-Procaccia definition:

$$D_q = \frac{1}{q-1} \lim_{r \searrow 0} \frac{\log \sum_i m(B_i)^q}{\log r}. \tag{2.10}$$

What if $q = 0$? (Assume $0^0 = 0$.) Then we get D_0 is the box-counting dimension. The case $q = 1$ is called the information dimension, which is related to $\sum_i p_i \log p_i$. The case $q = 2$ is called the correlation dimension.

In general, for $q = 2, 3, \ldots$, we get the so-called q^{th} order interpoint distance. For these integer values of q, there exists an algorithm which is helpful for computations, as suggested by the following. Given a probability distribution, pick a sequence of points from this distribution: $X_1, X_2, X_3, \ldots, X_q$. Suppose the choices are independent, in that the choice of one doesn't affect another. We call this a sequence of i.i.d. random variables (i.i.d. = independent, identically distributed).

Let

$$Y \;=\; \max\{||X_1 - X_q||, ||X_2 - X_q||, \ldots, ||X_{q-1} - X_q||\}. \tag{2.11}$$

Take the cumulative distribution function of the above:

$$F_Y(\epsilon) \;=\; \text{Prob}(Y \leq \epsilon) \;=\; \int m(B_\epsilon(x))^{q-1} dm(x). \tag{2.12}$$

Note this is the same integral as we've seen in the definition of $\Theta(q)$. Therefore, we can rewrite the definition as

$$D_q \;=\; \frac{1}{q-1} \lim_{\epsilon \searrow 0} \frac{\log F_Y(\epsilon)}{\log \epsilon}. \tag{2.13}$$

2.4 Bootstrapped Hill Estimate

The actual calculations for earthquake analysis in Harte's book were done using the bootstrapped Hill estimate, which is described as follows. Consider a data set $\{x_1, x_2, \ldots, x_N\}$. Let i be the current bootstrap number, $1 \leq i \leq k$.

Let $\{y_s : y_s = 1, 2, \ldots, n\}$ be the bootstrap sample. For all s, choose $x_{s_1}, \ldots, x_{s_q}$. Let

$$y_s \;=\; \max\{||x_{s_1} - x_{s_q}||, \ldots, ||x_{s_{q-1}} - x_{s_q}||\}. \tag{2.14}$$

Sort:

$$y_{(1)} \;\leq\; y_{(2)} \;\leq\; \cdots \;\leq\; y_{(n)}. \tag{2.15}$$

Next calculate

$$z_{i,n} \;=\; \frac{-1}{m-1} \sum_{j=1}^{m-1} \log \frac{y_{(j)}}{y_{(m)}}. \tag{2.16}$$

We ultimately get an estimate for $\Theta(q)$ (which shouldn't depend on m):

$$\widehat{\Theta}_m \;=\; \frac{k}{\sum_{i=1}^{k} z_{i,m}}. \tag{2.17}$$

Let $\overline{y}_{(m)}$ be the average of the m^{th} order statistic over all k bootstrap samples.

Chapter 3

Iterated Function Systems

Lecturer: Gerald Edgar (Wednesday, October 15, 2003)

3.1 Iterated Function Systems

When Mandelbroit started writing about fractals, came up with the notion that they should be self-similar – what does this mean? Lots of examples from previous mathematics wi' this property. One way to formalize is thru iterated function systems (due to Barnsley), analogous to dynamical systems. Have lots of similar properties, where the name iterated function system comes from, reminds us of dynamical systems.

Cantor Set (1880); Henry Smith is sometimes said to have written about this before Cantor, but unclear if he had it. Before Cantor, notion of countable / uncountable wasn't clear, and what Smith talked about wasn't clear. Smith was interested in Riemann Integral. If you want to tell if a function is integrable, we nowadays say it must be continuous except on a set of measure 0. Back then, didn't have measure (which dates from around 1900). They came up with examples of sets that were really small such that you could still be integrable. Some of his examples looked like Cantor sets.

3.2 Cantor Set

3.2.1 Construction

The Cantor set is a subset of the real line. We'll consider subsets of $[0, 1]$:

$[0]---[1]$

Will look at a sequence of approximations to the Cantor set.

Let $C_0 = [0, 1]$.

Remove the middle third, and get $C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. Is a closed interval, keep endpoints.

$[0]---------[1/3]$ $\qquad\qquad$ $[2/3]---------[1]$

Continue the process. Next is

$$C_2 \;=\; \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]. \tag{3.1}$$

In general, C_n is the union of 2^n closed intervals, each of size 3^{-n}. Note

$$C_0 \;\supset\; C_1 \;\supset\; C_2 \;\supset\; \cdots \tag{3.2}$$

Definition 3.2.1 (Cantor Set). *The Cantor set C is define by*

$$C \;=\; \bigcap_{n=1}^{\infty} C_n \;=\; \{x \in \mathbb{R} : \forall n, x \in C_n\}. \tag{3.3}$$

Note that $0, 1 \in C$. In fact, once we find and endpoint, we never remove the endpoint, thus all the endpoints are in C. One might first think that only the endpoints are left, but not the case. In fact, we'll see C is uncountable later.

3.2.2 Non-Trivial Point in Cantor Set

Example 3.2.2. *$\frac{1}{4} \in C$, but $\frac{1}{4}$ is not an endpoint.*

The endpoints are always of the form $\frac{m}{3^n}$, $m \in \mathbb{N}$. By unique factorization of integers, cannot write $\frac{1}{4}$ as an integer divided by a power of 3.

Must show $\frac{1}{4} \in C_n$ for all n; infinitely many things to check, checking one at a time won't be useful. Need to do a more clever job of checking. Will proceed by induction on n.

Will proceed by Induction, showing that $\frac{1}{4}$ and $\frac{3}{4}$ are in C_n for all n.

Clearly, both points are in C_0, and the base case holds. We now assume that $\frac{1}{4}, \frac{3}{4} \in C_n$, and show they are in C_{n+1}.

How do we go from C_n to C_{n+1}? We remove the middle third of sets. We take an interval, remove the middle third, and what is left for each sub-interval looks like the union of two pieces, each one-third the length of the previous.

Thus, we have shrinking maps fixing the left and right parts $L, R : \mathbb{R} \to \mathbb{R}$ given by

$$\begin{aligned} L(x) &= \frac{x}{3} \\ R(x) &= \frac{x+2}{3}. \end{aligned} \tag{3.4}$$

Exercise 3.2.3. *Prove that*

$$C_{n+1} \ = \ L(C_n) + R(C_n). \tag{3.5}$$

Thus, each step is related to the previous step. The maps L and R are nice in that the two images of $[0, 1]$ are disjoint, so all future subintervals will be disjoint and it will be easy to count.

What happens to $\frac{1}{4}$ and $\frac{3}{4}$?

Note

$$\begin{aligned} L\left(\frac{3}{4}\right) &= \frac{1}{4} \\ R\left(\frac{1}{4}\right) &= \frac{3}{4}. \end{aligned} \tag{3.6}$$

We now have the inductive step: If $\frac{1}{4}, \frac{3}{4} \in C_n$, then $\frac{1}{4}, \frac{3}{4} \in C_{n+1}$.

Proof. Clearly,

$$\frac{1}{4} \ = \ L\left(\frac{3}{4}\right) \ \in \ L(C_n) \ \subset \ C_{n+1} \tag{3.7}$$

and

$$\frac{3}{4} \ = \ R\left(\frac{1}{4}\right) \ \in \ R(C_n) \ \subset \ C_{n+1}, \tag{3.8}$$

which completes the proof of the inductive claim.

□

Remark 3.2.4. *Note that the Induction was easier by working with* both $\frac{3}{4}$ *and* $\frac{1}{4}$ *and not just* $\frac{1}{4}$.

3.2.3 Alternate Formulation of C

Note we have proved that

$$C \;=\; R(C) \;\cup\; L(C). \tag{3.9}$$

Thus,

$$C \;=\; \bigcap_{n=0}^{\infty} C_n \;=\; \bigcap_{n=0}^{\infty} C_{n+1} \;=\; \bigcap_{n=0}^{\infty} \left(R(C_n) \cup L(C_n)\right). \tag{3.10}$$

Therefore, we find

$$\begin{aligned} C &= \bigcap_{n=0}^{\infty} R(C_n) \cup \bigcap_{n=0}^{\infty} L(C_n) \\ &= R(\bigcap_{n=0}^{\infty} C_n) \cup L(\bigcap_{n=0}^{\infty} C_n) \\ &= R(C) \cup L(C). \end{aligned} \tag{3.11}$$

This is what we mean by C being self-similar. See is self-similar under shrinking by $\frac{1}{3}$.

3.2.4 Another Formulation of the Cantor Set

Let $x \in [0,1]$, we may write x in base 3. In other words, we can write

$$x \;=\; \sum_{i=1}^{3} \frac{a_i}{3^i}, \;\; a_i \in \{0,1,2\}. \tag{3.12}$$

Note we are not claiming each number has a unique representation. Consider the string $.122222222222\cdots$ and $.2$.

Note that $C_1 = \{x \in [0,1] : a_1 \neq 1\}$.

Continuing, we find $C_2 = \{x \in [0,1] : a_1 \neq 1, a_2 \neq 1\}$, and in general

$$\begin{aligned} C_n &= \{x \in [0,1] : a_1, \ldots, a_n \in \{0,2\}\} \\ C &= \{x \in [0,1] : a_1, a_2, \cdots \in \{0,2\}\}. \end{aligned} \tag{3.13}$$

There are problems, however, As remarked, numbers need not have a unique base 3 expansion. The example given shows that we may replace a number with repeating block with a terminating set, and thus these numbers are rationals. If a number can be written in two ways, one way using 1s and one way not, then it is in the Cantor set (as it can be written in base 3 without using any 1s).

Remark 3.2.5. *The Cantor Set is uncountable.*

This follows from the fact that C is equivalent to numbers in base 3 without 1 as a digit. Formally, one could map any such $x = \sum \frac{a_i}{3^i}$, $a_i \in \{0,2\}$, to $y = \sum \frac{b_i}{2^i}$, where $b_i = 0$ if a_i is 0 and $b_i = 1$ if a_i is 2. Thus, C has as many points as all of $[0,1]$ (consider base 2 expansions of real numbers).

3.2.5 non-Cantor Sets

Let

$$\begin{aligned} A_0 &= \{0\} \\ A_1 &= A_0 \cup \left(A_0 + \frac{2}{3}\right) \\ A_2 &= A_1 \cup \left(A_1 + \frac{2}{9}\right) \\ A_3 &= A_2 \cup \left(A_2 + \frac{2}{27}\right) \\ &\vdots \\ A_{n+1} &= A_n \cup \left(A_n + \frac{2}{3^{n+1}}\right) \\ A &= \bigcup_{n=0}^{\infty} A_n. \end{aligned} \tag{3.14}$$

This is an increasing sequence of sets, its union is not the complete Cantor Set, but on the computer, cannot tell the difference between this and the Cantor set.

Similarly, we have

$$A_{n+1} \;=\; L(A_n) \cup R(A_n). \tag{3.15}$$

Therefore, the union of these sets, A, just like the Cantor Set, satisfies

$$A \;=\; L(A) \cup R(A). \tag{3.16}$$

Remark 3.2.6. *Note A is* not *the Cantor Set! A is a countable set, the Cantor Set is uncountable. We do have, however, that*

$$C \;=\; \overline{A}, \tag{3.17}$$

namely, C is the closure of A, and A approximates C as well as we want.

3.3 Uniqueness of sets under such constructions

Question 3.3.1. *Consider the maps L and R. Are there any other sets X such that*

$$X \;=\; R(X) \cup L(X), \tag{3.18}$$

maybe if we want the sets to be disjoint, and not the empty set?

Take $\frac{1}{4}$ and $\frac{3}{4}$ and keep applying these operators. This will generate a countable example. There are lots of examples of sets satisfying this relation.

Theorem 3.3.2 (Characterization of the Cantor Set). *Let X be a closed, bounded, non-empty set such that $X = L(X) \cup R(X)$. Then X is the Cantor set.*

3.4 Sierpinski Gasket (or Triangle)

Two-dimensional example. Start with an equilateral triangle S_0 in the plane. Subdivide into smaller triangles by taking the midpoints of the three sides, and joining them to form an equilateral triangle. There are now four equilateral triangles, remove the middle one (the one that doesn't touch any of the original vertices). Call this S_1.

Continue by removing the middle equilateral triangle from each of the remaining triangles, and call the resultant S_2.

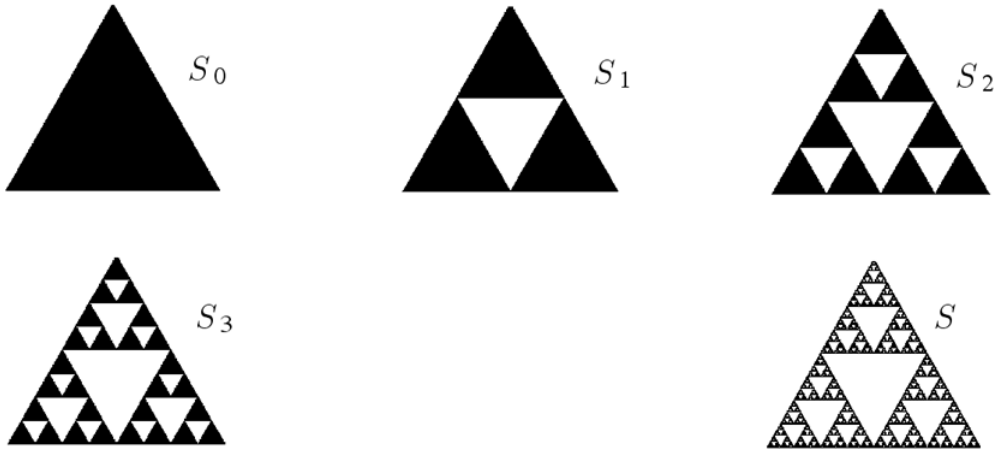
We obtain a sequence

$$S_0 \supset S_1 \supset S_2 \supset S_3 \supset \cdots, \tag{3.19}$$

and let

$$S = \bigcap_{n=0}^{\infty} S_n. \tag{3.20}$$

Exercise 3.4.1. *Analogous to the expansion using just 0s and 2s from the Cantor set, how can we characterize this?*



Chapter 4

Sierpinski Gaskets, Metric Spaces and IFS

Lecturer: Gerald Edgar (Friday, October 17, 2003)

4.1 Sierpinski Gasket

We saw last time that the Sierpinski Gasket is a generalization to the plane of the Cantor Set. How can we generalize the base-three expansion of the Cantor Set?

We will look at an expansion in base $\frac{1}{2}$. The digits are the vectors $\overrightarrow{0}, \overrightarrow{a}, \overrightarrow{b}$, where $\overrightarrow{0}$ is the 0-vector, $\overrightarrow{a}$ is the vector from the origin along the x-axis, going half-way the first side, and $\overrightarrow{b}$ is the vector along the positively sloped line going up from the origin, going half-way up.

We can write a general point $\overrightarrow{p}$ in the gasket by

$$\overrightarrow{p} \;=\; \sum_{i=1}^{\infty} \left(\frac{1}{2}\right)^{i} \overrightarrow{d}_i, \quad \overrightarrow{d_i} \in D, \tag{4.1}$$

where

$$D \;=\; \{\overrightarrow{0}, \overrightarrow{a}, \overrightarrow{b}\}. \tag{4.2}$$

4.2 Metric Space

Have a set of "points" S and a distance function $p(x, y) \geq 0$ defined on pairs of points in S satisfying:

For all $x, y \in S$,

1. $p(x, y) = 0$ if and only if $x = y$;
2. $p(x, y) = p(y, x)$;
3. $p(x, z) \leq p(x, z) + p(y, z)$.

We say the space is a Complete Metric Space if every Cauchy sequence converges. This is true for Euclidean spaces.

Definition 4.2.1 (Fixed Point). *Let $f : S \to S$. A point $x \in S$ is a fixed point of f if $f(x) = x$.*

Definition 4.2.2 (Contraction Map). *A continuous function $f : S \to S$ is a contraction map if $\exists \alpha \in (0, 1)$ such that $\forall x, y \in S$,*

$$p\left(f(x), f(y)\right) \ \leq \ \alpha p(x, y). \tag{4.3}$$

Remark 4.2.3. *We do not need to assume a contraction map is continuous; this follows from the definition.*

Example 4.2.4. *For the Sierpinski Gasket, we can define maps by*

$$\begin{aligned} f_{\overrightarrow{0}}(\overrightarrow{x}) &= \frac{1}{2}\overrightarrow{x} \\ f_{\overrightarrow{a}}(\overrightarrow{x}) &= \frac{1}{2}\overrightarrow{x} + \overrightarrow{a} \\ f_{\overrightarrow{b}}(\overrightarrow{x}) &= \frac{1}{2}\overrightarrow{x} + \overrightarrow{b}\,. \end{aligned} \tag{4.4}$$

Theorem 4.2.5 (Contraction Mappings). *Let S be a non-empty complete metric space, with $f : S \to S$ a contraction map. Then there exists a* unique *fixed point.*

Remark 4.2.6. *If x_1 and x_2 were both fixed points, then*

$$p(x_1, x_2) \ = \ p\left(f(x_1), f(x_2)\right) \ \leq \ \alpha p(x_1, x_2) \ < \ p(x_1, x_2). \tag{4.5}$$

Thus, $p(x_1, x_2) = 0$, and the two points are the same.

Proof. Choose $x_0 \in S$ (possible as S is non-empty). Then let $x_1 = f(x_0)$, $x_2 = f(x_1)$, and in general, $x_n = f(x_{n-1})$.

Then $\{x_n\}$ is a Cauchy sequence, and by the completeness of S (every Cauchy sequence converges), we know $x_n \to y$, say.

As f is a continuous function, $f(x_n) \to f(y)$. But as $f(x_n) = x_{n+1}$, the left hand side is just the original sequence shifted by one term. Thus, $x_{n+1} \to f(y)$, or, recalling the definition of y, we have $y = f(y)$, and y is a fixed point.

□

4.3 Iterated Function Systems

Let S be a complete metric space, and let $f_i : S \to S$ be contraction maps for $i \in \{1, \ldots, n\}$ with constants $\alpha_1, \ldots, \alpha_n$. Let $\alpha = \max_{i=1}^{n} \alpha_i$; note all f_i are contraction maps with parameter α.

Explicitly,

$$\forall x, y \in S, \;\; p\left(f_i(x), f_i(y)\right) \;\leq\; \alpha p(x, y). \tag{4.6}$$

We call the above an *Iterated Function System* with parameter α.

Definition 4.3.1 (Attractor). *An attractor for the Iterated Function System $\{f_i\}$ s a non-empty compact set E such that*

$$E \;=\; f_1(E) \cup f_2(E) \cup \cdots \cup f_n(E). \tag{4.7}$$

Theorem 4.3.2 (Hutchinson). *A complete metric space with a finite number of contraction maps has a unique attractor E.*

Proof. Consider the set

$$K(S) \;=\; \text{all non-empty compact subsets of } S. \tag{4.8}$$

$K(S)$ is non-empty as S is non-empty (consider the compact sets consisting of just single points in S).

We introduce the *Hausdorff distance* on $K(S)$ by: for $K_1, K_2 \in K(S)$, $D(K_1, K_2) < r$ if every element of K_1 has distance less than r from some element of K_2 and vice versa.

Then $K(S)$ is a compact metric space. Define $F : K(S) \to K(S)$ by

$$F(K) \;=\; f_1(K) \cup f_2(K) \cup \cdots \cup f_n(K). \tag{4.9}$$

Also, one has

$$D\left(F(K_1), F(K_2)\right) \;\leq\; \alpha D(K_1, K_2). \tag{4.10}$$

This follows from the fact that each $x_1 \in K_1$ is close to some point $x_2 \in K_2$. Thus,

$$p\left(f(x_1), f(x_2)\right) \;\leq\; \alpha p(x_1, s_2). \tag{4.11}$$

Book-keeping yields the claim.

Therefore, the Contraction Map Theorem applied here implies that *F* has a unique fixed point.

□

Moreover, we have an explicit *construction*! Let K_0 be any non-empty compact set in S. Let

$$\begin{array}{rcl} K_1 & = & F(K_0) \;=\; f_1(K_0) \cup \cdots \cup f_n(K_0) \\ K_2 & = & F(K_1) \;=\; f_1(K_1) \cup \cdots \cup f_n(K_1) \\ & \vdots & \end{array} \tag{4.12}$$

and so on. Then $K_n \to E$ in D, where E is the unique attractor set.

One can use this to draw images on the computer. Break the plane into small rectangles (introduce a grid), and color the grid if any point of K_n is in that grid. As n grows, this becomes a good approximation.

Since on the computer you are using floating points and do not have exact numbers, one can get a cluster of points representing the fixed point; however, if you stand far enough back you see a fixed point.

Chapter 5

Interpolation and Fractals

Lecturer: Bruce Adcock (Wednesday, October 22, 2003)

5.1 Interpolating Functions and IFS

Say one has a set of data, and wants to interpolate and find a function that fits it. We will arrive at an IFS that interpolates the data, as opposed to more traditional functions. From the book *Fractals Everywhere* by Michael Barnsley.

Assume we have a collection of data

$$\Big\{(x_i, y_i) \in \mathbb{R}^2 : i \in \{0, \ldots, N\}\Big\}, \quad N \in \mathbb{Z}^+, x_0 < \cdots < x_N. \tag{5.1}$$

Definition 5.1.1 (Interpolating Function). *An* interpolating function *(for the data set above) is a continuous function* $f{:}[x_0, x_N]{\to}\mathbb{R}$*,* $f(x_i) = y_i$ *for* $i \in \{0, \ldots, N\}$*. The points* (x_i, y_i) *are the* interpolation points*, and* f interpolates *the points.*

Example 5.1.2. *Consider the data set consisting of* $(0, 1)$ *and* $(1, 2)$*. Let*

$$w_1 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix} \tag{5.2}$$

and

$$w_2 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix}. \tag{5.3}$$

Letting $y = x + 1$, we find

$$w_1 \begin{pmatrix} x \\ x+1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}x \\ \frac{1}{2}x + 1 \end{pmatrix} \quad \textit{and} \quad w_2 \begin{pmatrix} x \\ x+1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}x + \frac{1}{2} \\ \frac{1}{2}x + \frac{3}{2} \end{pmatrix}. \tag{5.4}$$

meaning the graph of $y = x + 1$ is the attractor.

We want an IFS: $\{\mathbb{R}^2; w_n, n \in \{1, \ldots, N\}\}$. Will limit ourselves to simple affine functions such as

$$w_n \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_n & 0 \\ c_n & {}_n \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e_n \\ f_n \end{pmatrix}. \tag{5.5}$$

Also, we want

$$w_n \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix} \quad \text{and} \quad w_n \begin{pmatrix} x_N \\ y_N \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix}. \tag{5.6}$$

We get

$$\begin{aligned} a_n x_0 + e_n &= x_{n-1} \\ a_n x_N + e_n &= x_n \\ c_n x_0 +_n y_0 + f_n &= y_{n-1} \\ c_n x_N + d_n y_N + f_n &= y_n. \end{aligned}$$

We have five unknowns, four equations – and so one free variable. Will pick and choose d_n. Looking at d_n, that acts vertically – stretches or compresses vertical line segments, acts like a sheer.

We will study $0 \leq d_n < 1$, and show that we have contraction maps.

5.2 Contraction Maps

Theorem 5.2.1. *Let $N > 1$, $\{\mathbb{R}^2, w_n, n \in \{1, \ldots, N\}\}$ be as earlier, with data set $\{(x_i, y_i) \in \mathbb{R}^2, i \in \{0, \ldots, N\}\}$. Let $d_n \in [0, 1)$. Then there exists a metric d on $\mathbb{R}^2$ equivalent to the Euclidean metric, such that the IFS is hyperbolic (ie, contracts) with respect to d. Therefore, there is a unique nonempty compact set $G \subset \mathbb{R}^2$ such that $G = \cup_{n=1}^{N} w_n(G)$.*

Proof. Let

$$d\left((p_1,q_1),(p_2,q_2)\right) \;=\; |p_1-p_2|+\theta|q_1-q_2|, \tag{5.7}$$

where θ is defined at the end. This is equivalent to the Euclidean metric for $\theta > 0$. Note this is equivalent to the taxicab metric (your distance is how much you go north/south, then how much east/west; if $\theta \neq 1$, one direction is more expensive to travel than the other).

Then, noting that $a_n > 0$ (which comes from the fact that we've arranged the points in increasing order; c_n might also be positive), so if we set $D = d\left(w_n(x_1,y_1), w_n(x_2,y_2)\right)$ then

$$\begin{aligned}
D &= d\left((a_nx_1+e_n, c_nx_1+d_ny_1+f_n),(a_nx_2+e_n, c_nx_2+d_ny_2+f_n)\right)\\
&= |a_n|\cdot|x_1-x_2|+\theta|c_n(x_1-x_2)+d_n(y_1-y_2)|\\
&\leq (|a_n|+\theta|c_n|)|x_1-x_2|+\theta|d_n|\cdot|y_1-y_2|\\
&\leq a|x_1-x_2|+\theta\delta|y_1-y_2|\\
&\leq \max\{a,\delta\}\cdot d((x_1,y_1),(x_2,y_2)).
\end{aligned}$$

In the above, as $N \geq 2$, we have

$$|a_n| \;=\; \frac{|x_n-x_{n+1}|}{|x_N-x_0|} < 1.$$

If $\exists c_i \neq 0$ then

$$\theta \;=\; \frac{\min_{1\leq i\leq N}\{1-|c_n|\}}{\max_{1\leq i\leq N}\{2|x_n|\}}; \tag{5.8}$$

otherwise take $\theta = 1$. Further, we took

$$\begin{aligned}
a &= \frac{1}{2}\left(1+\max_{1\leq n\leq N}\{|a_n|\}\right) \;<\; 1\\
\delta &= \max_{1\leq n\leq N}\{|d_n|\} \;<\; 1.
\end{aligned}$$

As a and δ are less than 1, we obtain a contraction map.

□

5.3 Fractal Interpolation Function

Theorem 5.3.1. *With the same conditions as above, if G is the attractor of the IFS, then G is the graph of a continuous function $f : [x_0, x_N] \to \mathbb{R}$ that interpolates the interpolation points.*

Proof. The proof is several pages. The basic idea is that we have created a system of interpolating points. Doesn't look normal by any means, but it does work. □

We call the above a *fractal interpolation function*.

Chapter 6

Self-Similar Curves

Lecturer: Gerald Edgar (Wednesday, October 29, 2003)

6.1 Snowflake Curve

Start with an equilateral triangle. On each edge, put an equilateral triangle in the middle, one-third of the size of the original. Iterate. Koch snowflake.

In the limit, what is the perimeter, what is the area? One can see the area is finite (contained in a large disk), but what about the perimeter? What is the curve's length (in the limit).

One can obtain this by using an Iterated Function System. Isn't so successful if you start with the initial set – better to start with the *boundary* of the initial set, the three sides of the initial equilateral triangle.

In this case, we obtain the boundary of the Koch snowflake. If we look at just one edge, obtain something which is self-similar. Whatever the dimension of this piece is, that is also the dimension of the boundary.

If we divide a line-segment into 5 equal parts, each part is equal to the other, and we are looking for an exponent d such that

$$5 \cdot \left(\frac{1}{5}\right)^d = 1. \tag{6.1}$$

We end up with $d = 1$. If we have a solid square, and divide into 5 equal pieces, we get 25 equal pieces, and now we have

$$25 \cdot \left(\frac{1}{5}\right)^d = 1. \tag{6.2}$$

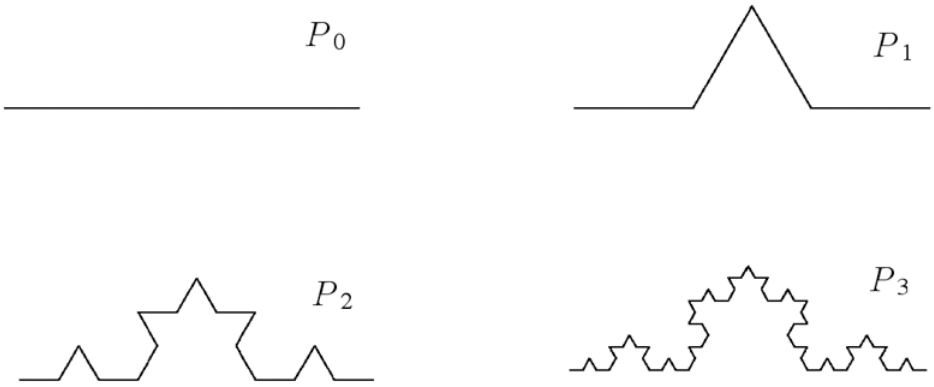
In this case, we end up with $d = 2$.

So, for *similarity dimension*, we obtain a line is dimension 1, and a square is dimension 2.

For the Koch boundary, when we shrink by a third, what happens? When we so divide we see there are four equal parts, each shrunk by a factor of $\frac{1}{3}$. We now have

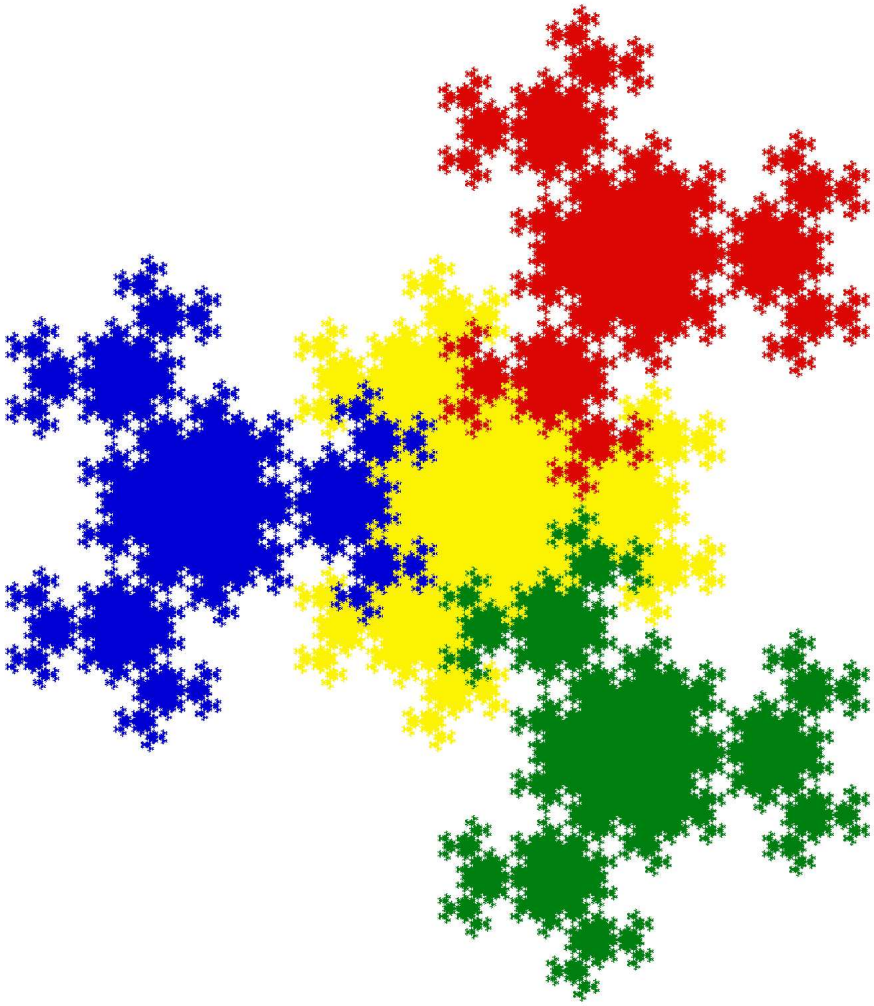
$$4 \cdot \left(\frac{1}{3}\right)^d = 1. \tag{6.3}$$

Therefore, $d = \log_3 4$; this is the self-similarity dimension of the snowflake – it is higher than 1, but lower than 2.



6.2 Other Examples

Eisenstein Fractions:

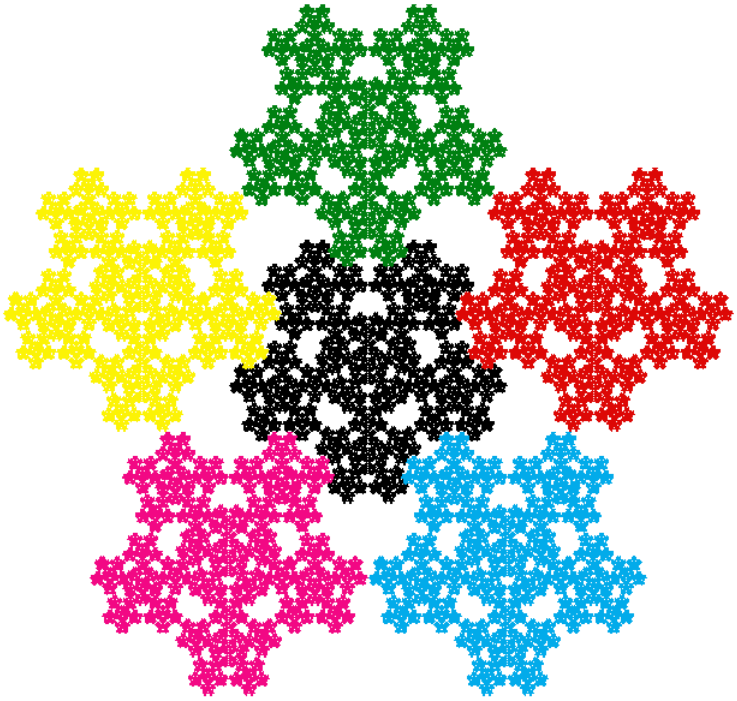


$$4 \cdot \left(\frac{1}{2}\right)^d = 1. \tag{6.4}$$

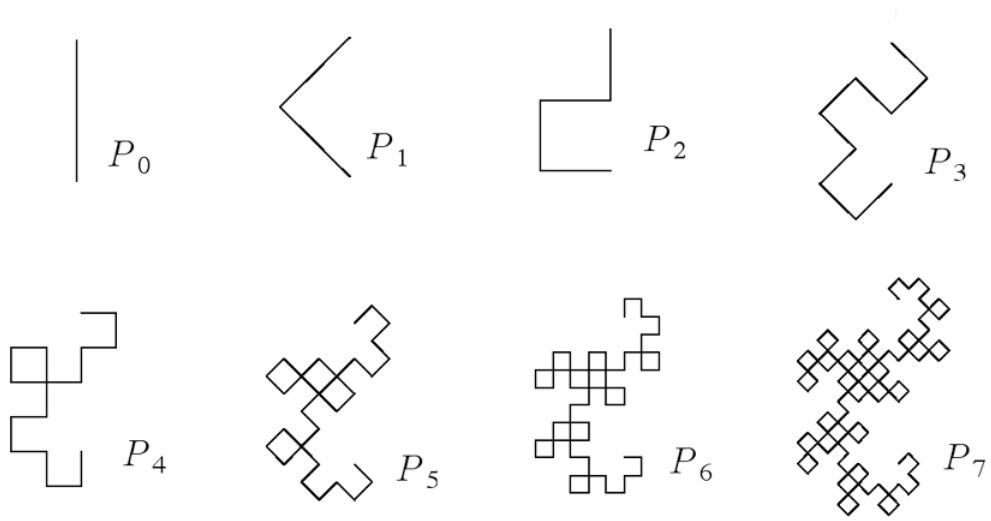
So the Eisenstein fractions are dimension 2. They tile the plane. For the McWorter pentigree, we have

$$6 \cdot \left(\frac{3-\sqrt{5}}{2}\right)^d = 1. \tag{6.5}$$

Hence, d is about 1.86.



Another example: Heighway’s dragon:



$$2 \cdot \left(\frac{1}{\sqrt{2}}\right)^d = 1. \tag{6.6}$$

Start with a line segment. Split and replace with two line segments of half the length at right angles. Continue this process. As you continue, you get choices as to which side to do things – alternate.

One can show the sides never cross; the limit is Heighway's dragon. This will have non-empty interior, tiles the plane.

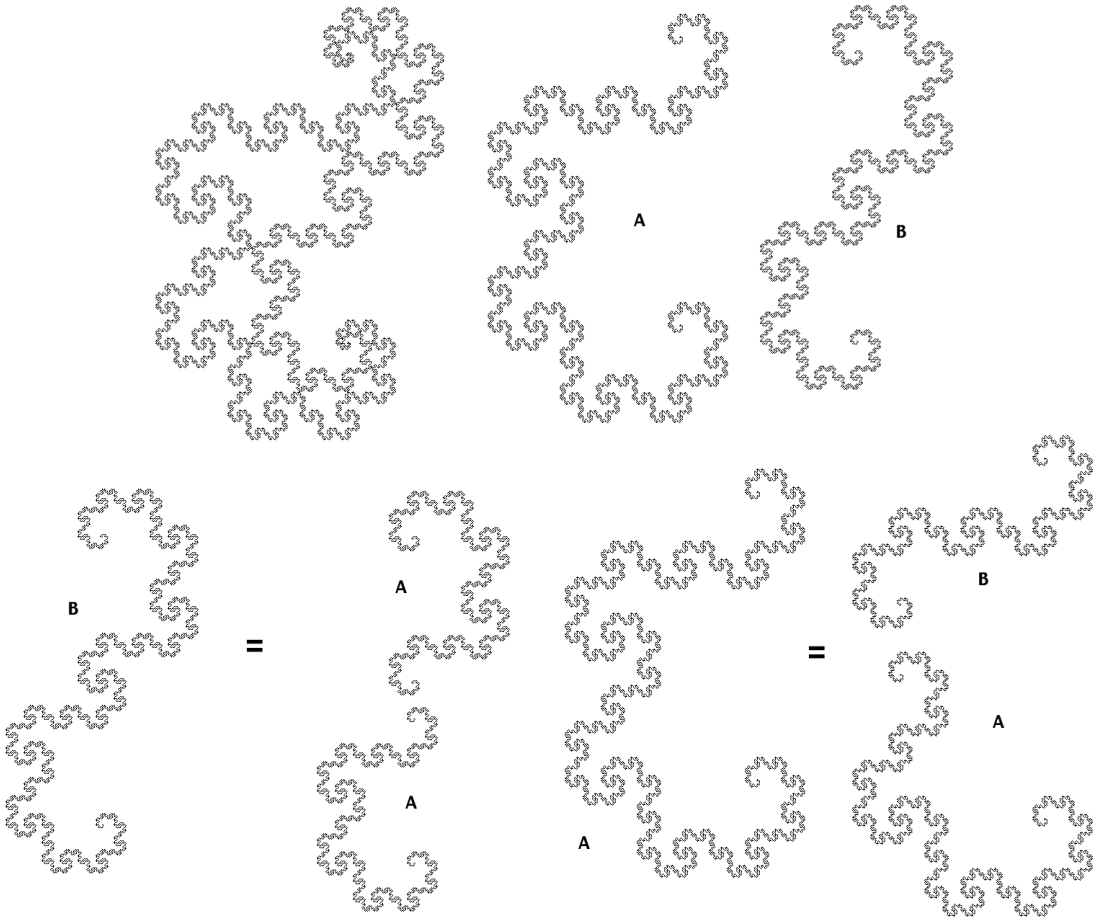
Can also look at as an attractor for an IFS. Can also think of the approximations as not just polygons in the plane, but ranges of function defined on $[0, 1]$. First is just a linear function. For the second, divide interval in half. Map each half linearly into the two segments (respectively).

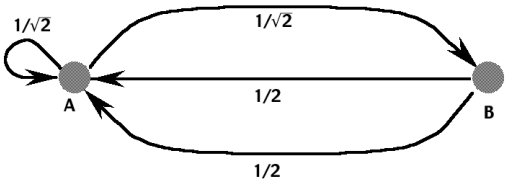
For next stage, map in quarters, and so on. Once we have a quarter, keep the corner vertex. Stays fixed in all future steps. The distances go down geometrically (and the geometric series converge), so the images of any point converges. This sequence of functions with values in the plane converges uniformly to a new function in the plane. The range of that limit function is Heighway's dragon. This is a space-filling curve: the domain is a line segment, but the range has nonempty interior in the plane! Other examples include Polya's and Hilbert's space-filling curves.

What about the boundary of Heighway's dragon? What is its dimension? The boundary looks like a nice curve, with necks where sides meet.

6.3 Graph Self-Similar

We can analyze this in a way like self-similar. Think of the boundary as being made up of a left (A) and a right (B) part. The boundary is a closed curve, but not simple closed curve (necks). We can find the dimensions of A and B.





For B, it turns out that it is made up of two shrunken As, rotated, shrunk by a half. Looked at in another way, can get that B is made up of a B and an A, each shrunk by $\frac{1}{\sqrt{2}}$.

Can look as a multi-graph: Have nodes at A and B, have an arrow from A to itself of weight $\frac{1}{\sqrt{2}}$; an arrow from A to B of weight $\frac{1}{\sqrt{2}}$, and then arrows from B

to A of weight $\frac{1}{2}$.

Not self-similar in the usual sense; often called *graph self-similar*.

Our matrix has two rows and two columns:

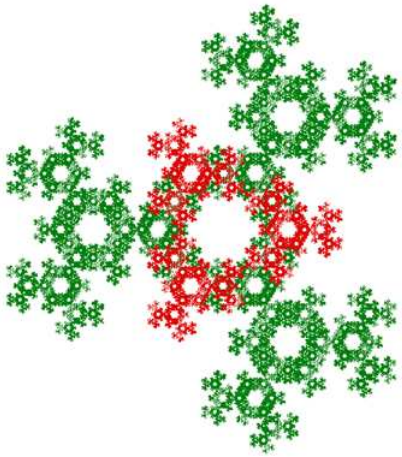
$$A(d) \;=\; \begin{pmatrix} \left(\frac{1}{\sqrt{2}}\right)^d & \left(\frac{1}{\sqrt{2}}\right)^d \\ 2\left(\frac{1}{2}\right)^d & 0 \end{pmatrix} \tag{6.7}$$

here $A(d)$ is the adjacency matrix, dimension d; do not confuse $A(d)$ with node A. In the above, $0 < d$.

In some sense, we want the above matrix to equal 1. This is a non-negative matrix. To such matrices, there is a nice theory due to Perron-Frobenius, which describes how the eigenvalues of such a matrix can behave. We have a positive eigenvalue large that the absolute value of everything else. Often the largest eigenvalue is called the *spectral radius*; all the eigenvalues are often called the *spectrum*.

We want to find d so that the spectral radius is 1; thus, the other eigenvalues will be less than 1.

The similarity dimension is always greater than or equal to the Hausdorff dimension For the Barnsley Wreath, the self-similar dimension is 1.9227 (approx), and the Hausdorff is 1.8459 (approx).



Chapter 7

Logistics Curve, Julia Sets, and Efficient Algorithms

Lecturer: Dean Eiger (Friday, October 31, 2003)

7.1 Notation

By $f^n(x)$ we mean $f(f(\cdots f(x))\cdots)$, ie, the n-fold composition. By the orbit of a point x under f we mean

$$O(x) \;=\; \{n \in Z \cap [1,\infty) : f^n(x)\}. \tag{7.1}$$

It is possible for a variety of behaviors: the points in the orbit could escape to infinity, or cycle.

For example, in the complex plane consider $f(z) = z^2$. If $z = a + ib$, we find $f(a+ib) = (a^2-b^2)+(2ab)i$. We can create a graph using the real and imaginary parts.

7.2 Logistics Curve

Let L be the limiting population of a system, let x_n be the population at the n^{th} generation.

$$L(x_{n+1}) \;=\; cx_n(1-x_n) \tag{7.2}$$

govern the population: the x_n means that the population growth is proportional to the size of the population, while the $1 - x_n$ encodes information on how the population competes with itself for resources. Often $c \in [0, 4]$.

7.3 Julia Set

Consider an arbitrary point in the domain of the Logistics curve, and use that as an input (starting point) for iteration. The Logistics Curve looks like a parabola, going through $(0, 0)$ and $(1, 0)$. Look at the intersection with the line $y = x$. Starting at any x on the real axis, go up till you hit the Logistics curve, then over till you hit the line, and repeat. In this case, we obtain a cycle. If we started at a different point, we would have obtained different behavior.

We define the Julia set as follows: let i refer to the i^{th} element of the set.

$$J(O_i) \; = \; \{\forall O_i \ni f(O_i) = O_i \cap f(O_{i\pm1}) \neq O_{i\pm1}\} \tag{7.3}$$

Take $f(z) = z^2$. If $|z| < 1$, $|f(z)|^2 = |z|^2$. Note that $|f^{n+1}(z)| < |f^n(z)|$ in this case; in fact, ratios will also tend to zero if $|z|$ is strictly less than 1.

If the modulus is greater than 1, the point will go off to infinity.

If the modulus equals 1, then all these points will be points in the Julia set.

1. If $|z| < 1$, then $f^n(z_0) = z_0$ (this is notation for there exists a fixed point).
2. If $|z| > 1$, then $f^n(z_0) \neq z_0$ (this is notation for there is no fixed point).
3. If $|z| = 1$, then $f^n(z_0) \in J$, where J is the Julia set.

7.4 Efficient Algorithms

Lecturer: Steven Miller (Friday, October 31, 2003)

For computational purposes, often having an algorithm to compute a quantity is not enough; we need an algorithm which will compute *quickly*. Below we study three standard problems, and show how to either rearrange the operations more efficiently, or give a more efficient algorithm than the obvious candidate.

7.4.1 Polynomial Evaluation

Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$. The obvious way to evaluate is to calculate x^n and multiply by a_n (n multiplications), calculate x^{n-1} and multiply by a_{n-1} ($n-1$ multiplications) and add, et cetera. There are n additions and $\sum_{k=0}^{n} k$ multiplications, for a total of $n + \frac{n(n+1)}{2}$ operations. Thus, the standard method leads to $O(n^2)$ computations.

Instead, consider the following:

$$\left(\Big((a_n x + a_{n-1})x + a_{n-2}\Big)x + \cdots + a_1\right)x + a_0. \tag{7.4}$$

For example,

$$7x^4 + 4x^3 - 3x^2 - 11x + 2 = \left(\Big((7x+4)x - 3\Big)x - 11\right)x + 2. \tag{7.5}$$

Evaluating the long way takes 14 steps; cleverly rearranging takes 8 steps.

Exercise 7.4.1. *Prove that the second method takes at most $2n$ steps to evaluate $a_n x^n + \cdots a_0$.*

7.4.2 Exponentiation

Consider x^n. The obvious way to evaluate involves $n-1$ multiplications. By writing n in base two, we can evaluate x^n in at most $2\log_2 n$ steps.

Let k be the largest integer such that $2^k \leq n$. Then $\exists a_i \in \{0,1\}$ such that

$$n \;=\; a_k 2^k + a_{k-1}2^{k-1} + \cdots + a_1 2 + a_0. \tag{7.6}$$

It costs k multiplications to evaluate x^{2^i}, $i \leq k$. How? Consider $y_0 = x^{2^0}$, $y_1 = y_0 \cdot y_0 = x^{2^0} \cdot x^{2^0} = x^{2^1}$, $y_2 = y_1 \cdot y_1 = x^{2^2}$, $\ldots$, $y_k = y^{k-1} \cdot y^{k-1} = x^{2^k}$.

Then

$$\begin{aligned} x^n &= x^{a_k 2^k + a_{k-1}2^{k-1} + \cdots + a_1 2 + a_0} \\ &= x^{a_k 2^k} \cdot x^{a_{k-1}2^{k-1}} \cdots x^{a_1 2} \cdot x^{a_0} \\ &= \left(x^{2^k}\right)^{a_k} \cdot \left(x^{2^{k-1}}\right)^{a_{k-1}} \cdots \left(x^2\right)^{a_1} \cdot \left(x^1\right)^{a_0} \\ &= y_k^{a_k} \cdot y_{k-1}^{a_{k-1}} \cdots y_1^{a_1} \cdot y_0^{a_0}. \end{aligned} \tag{7.7}$$

As each $a_i \in \{0, 1\}$, we have at most $k + 1$ multiplications above (if $a_i = 1$ we have the term y_i in the product, if $a_i = 0$ we don't).

Thus, it costs k multiplications to evaluate the x^{2^i} ($i \leq k$), and at most another k multiplications to finish calculating x^n. As $k \leq \log_2 n$, we see that x^n can be determined in at most $2 \log_2 n$ steps.

Note, however, that we do need more storage space for this method, as we need to store the values $y_i = x^{2^i}$, $i \leq \log_2 n$.

Exercise 7.4.2. *Instead of expanding n in base two, expand n in base three. How many calculations are needed to evaluate x^n this way? Why is it preferable to expand in base two rather than any other base?*

7.4.3 Euclidean Algorithm

The Euclidean Algorithm is an efficient way to determine the greatest common divisor of x and y, denoted $\gcd(x, y)$ or (x, y). Without loss of generality, assume $1 < x < y$.

The obvious way to determine $\gcd(x, y)$ is to divide x and y by all positive integers up to x. This takes at most $2x$ steps.

Let $[z]$ denote the greatest integer less than or equal to z. We write

$$y = \left[\frac{y}{x}\right] \cdot x + r_1, \ 0 \leq r_1 < x. \tag{7.8}$$

Exercise 7.4.3. *Prove that $r_1 \in \{0, 1, \ldots, x - 1\}$.*

Exercise 7.4.4. *Prove $\gcd(x, y) = \gcd(r_1, x)$. Hint: $r_1 = y - \left[\frac{y}{x}\right] \cdot x$.*

We proceed in this manner until r_k equals zero or one. As each execution results in $r_i < r_{i-1}$, we proceed at most x times (although later we prove we need to apply these steps at most $2 \log_2 x$ times).

$$
\begin{aligned}
x &= \left[\frac{x}{r_1}\right] \cdot r_1 + r_2,\ 0 \leq r_2 < r_1 \\
r_1 &= \left[\frac{r_1}{r_2}\right] \cdot r_2 + r_3,\ 0 \leq r_3 < r_2 \\
r_2 &= \left[\frac{r_2}{r_3}\right] \cdot r_3 + r_4,\ 0 \leq r_4 < r_3 \\
&\vdots \\
r_{k-2} &= \left[\frac{r_{k-2}}{r_{k-1}}\right] \cdot r_{k-1} + r_k,\ 0 \leq r_k < r_{k-1}. \qquad (7.9)
\end{aligned}
$$

Exercise 7.4.5. *Prove that if $r_k = 0$, then $\gcd(x, y) = r_{k-1}$, while if $r_k = 1$, then $\gcd(x, y) = 1$.*

We now analyze how large k can be. The key observation is the following:

Lemma 7.4.6. *Consider three adjacent remainders in the expansion: r_{i-1}, r_i and r_{i+1} (where $y = r_{-1}$ and $x = r_0$). Then $\gcd(r_i, r_{i-1}) = \gcd(r_{i+1}, r_i)$, and $r_{i+1} < \frac{r_{i-1}}{2}$.*

Proof. We have the following relation:

$$
r_{i-1} = \left[\frac{r_{i-1}}{r_i}\right] \cdot r_i + r_{i+1},\ 0 \leq r_{i+1} < r_i. \qquad (7.10)
$$

If $r_i \leq \frac{r_{i-1}}{2}$, then as $r_{i+1} < r_i$, we immediately conclude that $r_{i+1} < \frac{r_{i-1}}{2}$. If $r_i > \frac{r_{i-1}}{2}$, then we note that

$$
r_{i+1} = r_{i-1} - \left[\frac{r_{i-1}}{r_i}\right] \cdot r_i. \qquad (7.11)
$$

But $\left[\frac{r_{i-1}}{r_i}\right] = 1$ (easy exercise). Thus $r_{i+1} < \frac{r_{i-1}}{2}$.

□

We count how often we apply Euclid's Algorithm. Going from $(x, y) = (r_0, r_{-1})$ to (r_1, r_0) costs one application. Every two applications leads to the first entry in the last pair being at most half of the second entry of the first pair.

Thus, if k is the largest integer such that $2^k \leq x$, we see we apply Euclid's Algorithm at most $1 + 2k \leq 1 + 2\log_2 x$ times. Each application requires one integer division, where the remainder is the input for the next step.

We have proven

Lemma 7.4.7. *Euclid's Algorithm requires at most $1 + 2\log_2 x$ divisions to find the greatest common divisor of x and y.*

Let us assume that $r_i = \gcd(x, y)$. Thus, the last equation before Euclid's Algorithm terminated was

$$r_{i-2} = \left[\frac{r_{i-2}}{r_{i-1}}\right] \cdot r_{i-1} + r_i,\ 0 \leq r_i < r_{i-1}. \tag{7.12}$$

Therefore, we can find integers a_{i-1} and b_{i-2} such that

$$r_i = a_{i-1}r_{i-1} + b_{i-2}r_{i-2}. \tag{7.13}$$

Looking at the second to last application of Euclid's algorithm, we find that there are integers a'_{i-2} and b'_{i-3} such that

$$r_{i-1} = a'_{i-2}r_{i-2} + b'_{i-3}r_{i-3}. \tag{7.14}$$

Substituting for $r_{i-1} = r_{i-1}(r_{i-2}, r_{i-3})$ in the expansion of r_i yields that there are integers a_{i-2} and b_{i-3} such that

$$r_i = a_{i-2}r_{i-2} + b_{i-3}r_{i-3}. \tag{7.15}$$

Continuing by induction, and recalling $r_i = \gcd(x, y)$ yields

Lemma 7.4.8. *There exist integers a and b such that $\gcd(x, y) = ax + by$. Moreover, Euclid's Algorithm gives a constructive procedure to find a and b.*

Exercise 7.4.9. *Find a and b such that $a \cdot 244 + b \cdot 313 = \gcd(244, 313)$.*

Exercise 7.4.10. *Add details to complete an alternate proof of the existence of a and b with $ax + by = \gcd(x, y)$:*

1. *Let d be the smallest positive value attained by $ax + by$ as we vary $a, b \in \mathbb{Z}$. Such a d exists. Thus, $d = ax + by$. We now show $d = \gcd(x, y)$.*
2. *$\gcd(x, y)|d$.*

3. *Let $e = Ax + By > 0$. Then $d|e$. Therefore, for any choice of $A, B \in \mathbb{Z}$, $d|(Ax + By)$.*

4. *Consider $(a, b) = (1, 0)$ or $(0, 1)$, yielding $d|x$ and $d|y$. Therefore $d \leq \gcd(x, y)$. As we've shown $\gcd(x, y)|d$, this completes the proof.*

Note this is a non-constructive proof. By minimizing $ax + by$, we obtain $\gcd(x, y)$, but we have no idea how many steps is required. Prove that a solution will be found either among pairs (a, b) with $a \in \{1, \ldots, y - 1\}$ and $-b \in \{1, \ldots, x - 1\}$, or $-a \in \{1, \ldots, y - 1\}$ and $b \in \{1, \ldots, x - 1\}$.

Chapter 8

L-Systems

Lecturer: Charles Estill (Wednesday, November 5, 2003)

8.1 L-System

An L-System is a method of starting with a seed object, say a triangle, and replacing certain portions with more complicated themes.

We start with what Mandelbroit called an initiator and a generator. Each time, we replace a pre-defined object with the generator.

For example, take an equilateral triangle. Have the generator be the first part of a Koch snowflake (take the unit line segment, replace the middle third with two segments at angles of 60 degrees). We then go through the object, replacing sides with shrinked versions of the generator.

Usually we represent this by using an alphabet and a set of rules. For example, let us have characters F, $-$ and $+$. We can form words, say $F \to F - F + +F - F$. Here F is the initiator. The word above tells us how to replace parts of the generator to form the next stage.

Consider ab, with $a \to ab$ and $b \to ba$ (a set of rules). Here, we apply these rules in parallel. So, the next generation of ab is $abba$.

This gives us strings of characters. We now give rules on how to make figures from this. For example, let F be go forward by a certain distance d (will determine this distance later) and draw a line. Let $-$ be a right rotation by a certain angle (say δ), and $+$ a left rotation by a certain angle (say δ). Further, let f be go forward without drawing a line.

At the end of the 19th century, Peano found an odd phenomenon. There are

continuous functions from the closed interval $[0,1]$ onto the unit square $[0,1] \times [0,1]$.

Can have stochastic systems, where say one-third of the time we have $a \to abc$, one-third of the time to bac and one-third of the time to cba.

Another example is Conway's game of life, where the fate of a cell depends on the population of its neighbors and some fixed rules.

8.2 Peano Spacefilling Curve

Space-filling, Self-Avoiding Simple and Self-Similar curve; abbreviated by FASS. Doesn't intersect itself, a bijection from $[0,1]$ to $[0,1] \times [0,1]$.

8.3 Motivation

The reason for creating this formulation was to model plants / leaves. We add additional characters, the brackets [and]. Consider $FF + [+F - F - F]$. The FF in the beginning is clear: go forward, go forward, dropping lines. The plus rotates us. The left bracket does the following: push the state into a push-down stack: the direction we're facing and the direction we're at; we basically drop a marker at this point and record the direction we were going. The we do what the other symbols tell us until we hit the right bracket, at which point we return to the marked position and restore the original orientation.

Obviously, need the same number of left and right brackets.

8.4 IFS and L-Systems

Consider $F \to F[+F]F[-F]F$. Here, let the angle be $\frac{2\pi}{14}$. This will generate a nice leaf.

Chapter 9

Hausdorff Dimension

Lecturer: Gerald Edgar (Friday, November 7, 2003)

9.1 Hausdorff Dimension

Let A be a set in a Euclidean Space (although more generally one may take a metric space). Let $A \subset \cup_{i=1}^{\infty} C_i$ be a cover of A, where $|C_i| = \text{diam}(C_i) < \epsilon$. We call the above an ϵ-cover of A.

For $s > 0$, consider

$$H_\epsilon^s(A) \;=\; \inf_{\epsilon-covers} \sum_{i=1}^{\infty} |C_i|^s. \tag{9.1}$$

We are often interested in

$$\lim_{\epsilon \to 0} H_\epsilon^s(A) \;=\; H^s(A). \tag{9.2}$$

The limit is a supremum - as ϵ decreases, fewer covers are available. Typically it is infinite up to a certain point, then a jump, and past the jump point it is often zero; the value at the jump point can be anything between 0 and ∞.

Lemma 9.1.1. *If $s < t$ and $H^t(A) > 0$ then $H^s(A) = \infty$.*

Proof. Consider an ϵ-cover C_i, and compare $\sum_i |C_i|^t$ and $\sum_i |C_i|^s$. We may assume $\epsilon < 1$ as $\epsilon \to 0$. We find

$$\sum_i |C_i|^t \;\leq\; \epsilon^{t-s} \sum_i |C_i|^s. \tag{9.3}$$

This is true for all ϵ-covers, so we find

$$H^t_\epsilon(A) \;\leq\; \epsilon^{t-s} H^s_\epsilon(A). \tag{9.4}$$

Sending $\epsilon \to 0$, if the LHS is positive, as $\epsilon \to 0$, we must have $H^s_\epsilon(A) \to \infty$. □

Lemma 9.1.2. *If $s < t$ and $H^s(A) < \infty$, then $H^t(A) = 0$.*

Proof. Same as before. □

Let $s_0 = \sup\{s : H^s(A) > 0\}$.

Caratheodory: 1905: defined arc length using method like this (n-dimensional space). Worked for sets not given parametrically. Then Caratheodory showed that for exponent 2, up to constant, get surface area. Hausdorff read the paper, said the exponent didn't need to be an integer, could be arbitrary. In his paper, he then examined sets like the Cantor Set to show non-integral limits are possible.

9.2 Sierpinski Gasket

Equilateral triangle with sides of length 1 in the plane. Divide into four smaller triangles, remove the central one and keep the other three. Continue (repeat with the remaining three triangles).

Can do in terms of expansions. Consider two vectors $\overrightarrow{e_1}$ along one side and $\overrightarrow{e_2}$ along the other side, both sides emanating from the lower left vertex (at $\overrightarrow{0}$).

Then we have all points in the plane that can be written in the form

$$\sum_{i=1}^{\infty} \frac{1}{2^i} \overrightarrow{a_i}, \quad \overrightarrow{a_i} \in \{\overrightarrow{0}, \overrightarrow{e_1}, \overrightarrow{e_2}\}. \tag{9.5}$$

Note the vectors $\overrightarrow{e_1}$ and $\overrightarrow{e_2}$ have length $\frac{1}{2}$ if the initial triangle has sides of length 1.

We can consider it also as a sequence

$$S_0 \;\supset\; S_1 \;\supset\; S_2 \;\supset\; S_3 \;\supset\; \cdots \tag{9.6}$$

where S_n consists of 3^n triangles of sides of length 2^{-n}. Note the unit equilateral triangle has diameter 1.

If we find one ϵ-cover, we get a bound. Given $\epsilon > 0$, choose n so that $2^{-n} < \epsilon$. Then S is covered by 3^n triangles of side length $2^{-n} < \epsilon$. Thus,

$$H^s_\epsilon(A) \ \leq \ 3^n \cdot (2^{-n})^s. \tag{9.7}$$

For what s is the above bounded? Clearly, $s_0 = \frac{\log 3}{\log 2} = \log_2 3$. Anything smaller than s_0 will give infinity, so we know that $H^{s_0}_\epsilon(S) \leq 1$. Therefore, $H^{s_0}(S) \leq 1$, so $\dim(S) \leq s_0$.

If A is a set of diameter D, let k be such that $2^{-k} < D \leq 2^{-k+1}$. Without loss of generality, we may assume $D < 1$.

Claim 9.2.1. *Let*

$$S_k \ = \ \bigcup_{i=1}^{3^k} T_{k,i}. \tag{9.8}$$

Then

$$m \ = \ \#\{i : T_{k,i} \cap A \neq \phi\} \ \leq \ 100. \tag{9.9}$$

Proof. Except at singleton points, $t_{k,i} \cap T_{k,i'} = \phi$ and all are in a ball of radius $\leq D + 2^{-k}$. Therefore

$$m \cdot 4^{-k} \cdot \frac{\sqrt{3}}{4} \ \leq \ \pi(D + 2^{-k})^2. \tag{9.10}$$

Thus,

$$m \ \leq \ \frac{4}{\sqrt{3}} \cdot 4^k \cdot \pi(D + 2^{-k})^2 \ \leq \ \frac{4 \cdot 4^k \pi (2^{-k+2})^2}{\sqrt{3}} \ = \ \frac{4\pi}{\sqrt{3}} \ < \ 100. \tag{9.11}$$

Of course, we can take a smaller number than 100! □

If A is a set of diameter D, it intersects with at most $m \leq 100$ triangles of size 3^{-k}. Therefore,

$$\mu(A) \ \leq \ 100 \cdot 3^{-k} \ = \ 100(2^{-k})^{s_0} \ \leq \ 100\mu(A)^{s_0}. \tag{9.12}$$

Therefore, any ϵ-cover A of S gives

$$\sum_i |A_i|^{s_0} \;\geq\; \frac{1}{100}\sum_i \mu(A_i) \;\geq\; \frac{1}{100}\mu(S). \tag{9.13}$$

This gives the desired answer – here μ is not Lebesgue measure, it is the measure on the Sierpinski gasket. We use $|A|$ to denote the diameter.

Chapter 10

Koch Fractal and Non-Differentiable Curves

Lecturer Youri (Wednesday, November 12, 2003)

10.1 Koch Fractal

>From the paper by Koch (turn of the last century): it is a curve which is nowhere differentiable. He wanted to construct such a curve using *elementary* geometry, and not the Fourier Series techniques of Weierstrass and Hardy.

Start with a segment $[A, B]$. Divide into three equal parts. Remove the middle third, say at points $C < E$. Form an equilateral triangle with base CE going up to third vertex at D, and add the lines CD and DE.

Continue this process on the four segments; in the limit we get the Koch Fractal (also the Koch snowflake).

Let P_n be the curve after n steps (after the first step, there are four sides). In general, at the n^{th} step, there are 4^n segments of length 3^{-n}.

Let $k_n : [0, 1] \to \mathbb{R}^2$, a function to the n^{th} curve. If $x \in P_n \cap P_{n-1}$, then $k_n(x) = k_{n-1}(x)$. Note $k_1(x) = (x, 0)$ if x is in the first or last third, else go up from x till you hit the point. Note the Cantor set are on P_n for all n; in fact, there are lots of Cantor Sets!

These functions converge uniformly, as the heights of the triangles decrease:

$$||k_n(x) - k_{n-1}(x)|| \leq \frac{\sqrt{3}}{2 \cdot 3^n}. \tag{10.1}$$

We see $k_n \to k$, where k is a continuous fraction describing the fractal.

10.2 Properties of k

10.2.1 k is Bounded

Note k is bounded: for k_n, the height is bounded by the sum of the heights of n equilateral triangles, where each triangle is one-third the previous. Thus,

$$||k(x)|| \ \leq \ \sum h_n \ = \ \sum \frac{\sqrt{3}}{2 \cdot 3^n} \ = \ \frac{\sqrt{3}}{4}. \tag{10.2}$$

10.2.2 k is not Rectifiable

P_n has 4^n segments, each of length 3^n. Thus, the length of P_n is $\frac{4^n}{3^n}$. The length of k is the limit as $n \to \infty$, which is infinite.

However, the area between the curve and the x-axis is finite; one may show the area is bounded by $\frac{3}{20}$. In the n^{th} step, there are 4^n equilateral triangles of side length 3^{-n}, for a total of $4^n \cdot \frac{\sqrt{3}}{4 \cdot 3^{2n}}$. Summing over n gives the claimed answer; note we are using the area of an equilateral triangle of side length s is $\frac{s^2\sqrt{3}}{4}$.

10.2.3 k doesn't intersect itself

In the original picture with $ACDEB$, look at the line from B to D. Draw a line L with angle 60 degrees from B, parallel to DE. We want to show the curve is to the left of this.

The distance from D to the line L is $d(DE, L) \geq \frac{\sqrt{3}}{6}$. Similarly, $d(CE, L) \geq \frac{\sqrt{3}}{6}$.

If a point x is built on the segment DE, then $k_1(x) \in ED$ with say height y on DE, then

$$d\left(k_1(x), k(x)\right) \ \leq \ \frac{1}{3}\frac{\sqrt{3}}{4} \ = \ \frac{\sqrt{3}}{12}. \tag{10.3}$$

Thus, this point is to the left of the line.

Now consider $x \in EB$. Then if $k(x) = (x, 0)$, it is on the left side. If $k(x) \neq (x, 0)$, we have a similar situation as before, with everything smaller (self-similarity).

Basically, draw equilateral triangles with sides at AC, CD, DE and EB. We see all points on the curve built on AC are constrained to lie in the equilateral triangle built on AC; all the points built on CD are constrained to lie in the equilateral triangle built on CD; and so on.

Assume $x \neq y$, and choose n so that $\frac{1}{3^n} < y - x$. Then $k(x)$ and $k(y)$ belong to parts of the curve built on different lines. As they are built on two different segments, the curves cannot intersect.

As the curve doesn't intersect itself, we can order the points on the curve. If $x < y$, we say the point $k(x)$ is *before* (or *preceeds*) the point $k(y)$.

10.3 k is not Differentiable

For a curve to be differentiable at some point A, if we take any B also on the curve, we can consider the secant line AB. This secant line must approach the tangent line T_A as $B \to A$.

10.3.1 Vertex Points

One type of points on the curve are the vertices of the polygonal lines. Let us assume $k(x)$ is a vertex of P_n; thus, it is a vertex of all P_m for $m \geq n$. Clearly not differentiable there, as the two sides (before and after) meet at 120 degrees. Thus, look at secant lines from one side gives a different slope than secant lines from the other. We need that there are infinitely many vertices on each side approaching $k(x)$ – this follows from the presence of Cantor sets on each side.

10.3.2 Cantor Set but not Vertex Point

Assume $K = k(x) = k_m(x)$ for all $m \geq n$. Then it is a point in some Cantor set. We can find two vertices for every n that are close to K, one on each side. Say the points are $A < K < B$. Then we can find a vertex C going up from B. The angle from K to C is between 30 and 60 degrees, or between 60 and 90 degrees. This comes from either C to the right and above B, or from C being to the left and below B.

The above argument also works for the points $k(0) = (0, 0)$ and $k(1) = (1, 0)$.

10.3.3 Limit Point

If $K = k(x)$ is a limit point, then we have a sequence of distinct $k_n(x)$s converging to K (else we're in a previous case). Let $K_n = k_n(x)$. Choose a sub-sequence $\{n_i\}$ such that $K_{n_i}(x) \neq k_{n_i}(x)$. Let $K_n \subset [A_{n-1}, B_{n-1}]$. Then K_n is built on a triangle that is removed. Then take $[A_n, B_n]$. Note the angle between $A_{n-1}B_{n-1}$ and A_nB_n is 60 degrees. But $A_n, B_n \to K$, but the tangent line will have a different slope.

10.4 More Non-Differentiable Curves

Start with an interval $[A, B]$. Divide in thirds with $C < E$. Let M be the midpoint of $[C, E]$, and let D be at some angle from M, with DM length $\frac{\sqrt{3}CE}{2}$. We perform a similar process as before, and obtain another non-differentiable curve.

Chapter 11

Quantization Dimension

Lecturer: Larry Lindsay (Friday, November 14, 2003)

11.1 Quantization Dimension — Background

Quantization is the process of estimating a given probability P on $\mathbb{R}^d$ by a discrete probability supported on a finite set A.

This came from electrical engineering, for efficiently encoding information in a signal and minimizing distortion. Specifically, the original motivation was pulse code modulation (Reeves 1938).

We will need the following definitions:

$$\begin{aligned} V_{n,r}(P) &= n^{th} \text{ quantization of order } r \\ &= \inf_{|A|\leq n} \int d(x,A)^r dP(x) \\ e_{n,r}(P) &= V_{n,r}(P)^{1/r}. \end{aligned} \tag{11.1}$$

To estimate a given probability with a discrete probability supported on a finite set, we could form a Voronoi diagram. Then the mass of a cell is the probability of that cell, and we assign that mass to the centroid of the cell.

As an example consider the uniform distribution in some finite portion of the plane $\mathbb{R}^2$. What does the optimal Voronoi partition look like? Could be squares, hexagons, and so on. It turns out that hexagons are the best way to pack. This was discovered in economics in the context of service centers for customers. The goal is to minimize the average distance people need to travel to service centers.

But this is kind of a trick question. It depends in some sense on how we define distance in the plane. If we use the sup norm, we would use squares (which would correspond to balls). The taxi-cab metric would lead to diamonds (again, corresponding to balls).

There is a famous algorithm used to get the optimal set, called Lloyd's Algorithm I (also discovered by Steinhaus): Start with n points (randomly placed) and form a Voronoi partition. Within each area, move the point to the centroid of the area. Then start over, forming a new Voronoi partition. Continue — number of points stays fixed. Hopefully this converges to an optimal set, but it sometimes converges to a locally optimal set (not globally optimal) or to a saddle-point.

Open question: What are necessary conditions for the algorithm to converge to an optimal set?

11.2 Generalization to the r^{th} Power Distortion

Bucklew and Wise (1982): Let P_a be the absolutely continuous part of a probability P on $\mathbb{R}^d$. Suppose $E[\|X\|^{r+\delta}] < \infty$ for some δ. Then

$$\lim_{n\to\infty} n^{r/d} V_{n,r}(P) \;=\; Q(r,d) \left\| \frac{dP_a}{d\lambda^d} \right\|_{d/(d+r)} \tag{11.2}$$

where $Q(r,d) > 0$ depends only on r and d, and above we are taking a p-norm, $p = \frac{d}{d+r}$. In this case λ^d is d-dimensional Lebesgue Measure.

The formula for $Q(r,d)$ is

$$Q(r,d) \;=\; \inf_{n\geq 1} n^{r/d} V_{n,r}(U([0,1]^d)). \tag{11.3}$$

Open question: What specific values does $Q(r,d)$ take on for different values of r and d?

This leads us into quantization dimension. What if P is singular with respect to Lebesgue measure? Maybe there is a formula similar to above, but using a smaller d. What exactly happens? Some definitions and results are known, but there are still many open problems.

11.3 Definition of Quantization Dimension

Zador (1982):

$$\begin{aligned} D_r(P) &= \text{quantization dimension of order } r \\ &= \lim_{n\to\infty} \frac{\log n}{-\log e_{n,r}(P)}. \end{aligned} \tag{11.4}$$

Note that the limit may not exist. In that case we define upper quantization dimension $\overline{D}_r(P) = \limsup$ and lower quantization dimension $\underline{D}_r(P) = \liminf$.

Idea: $e_{n,r}(P)$ behaves like $(\frac{1}{n})^{1/d}$.

As an example consider the symmetric probability on the middle-third Cantor set. As we saw in a previous lecture we can intuitively put quantization points in the middle of the intervals at different levels in the construction of the Cantor set. We get $e_{2^k,r} = \frac{1}{3^k} e_{1,r}$.

$$D_r \;=\; \lim_{n\to\infty} \frac{\log n}{-\log e_{n,r}} \;=^{?}\; \lim_{k\to\infty} \frac{2^k}{-\log e_{2^k,r}} \;=\; \log_3 2. \tag{11.5}$$

If instead of using the symmetric probability, what if our probability is skewed to the left or right? The above intuitive argument may not help us much.

11.4 Self-Similar Example

Graf and Luschgy (1999): Let P be a self-similar probability generated by a system of contracting similarities $\phi_1, \phi_2, \ldots, \phi_N$ (satisfying the Open Set Condition) and a probability vector $(p_1, \ldots, p_N)$, then $D_r(P)$ is the solution D_r to the equation

$$\sum_{i=1}^{N} (p_i s_i^r)^{\frac{D_r}{r+D_r}} \;=\; 1, \tag{11.6}$$

where s_i is the contraction ratio of ϕ_i.

Comparison to Thermodynamic Formalism: Let $\beta(q)$ be the inverse temperature

$$\sum_{i=1}^{N} p_i^q s_i^{\beta(q)} \;=\; 1. \tag{11.7}$$

Then we get

$$D_r \;=\; \frac{\beta(q)}{1-q} \tag{11.8}$$

precisely when $\beta(q) = rq$.

Therefore,

$$D(r) \quad = \quad \frac{\beta(q(r))}{1-q(r)} \tag{11.9}$$

implies

$$\begin{aligned} D'(r) \;&=\; \frac{\beta'(q(r))q'(r)(1-q(r)) + \beta(q(r))q'(r)}{(1-q(r))^2} \\ &=\; q'(r)\frac{\beta'(q(r))(1-q(r)) + \beta(q(r))}{(1-q(r))^2}. \end{aligned} \tag{11.10}$$

The first derivative is non-negative. We are assuming $q(r)$ is differentiable. As r increases, $q(r)$ decreases. Thus, $q'(r) < 0$. Therefore, $D'(r) \geq 0$ if and only if

$$\beta'(q(r)) \;\leq\; \frac{-\beta(q(r))}{1-q(r)}, \tag{11.11}$$

which is easily seen to be true.

Open Question: In numerical examples, the second derivative appears to be negative (or non-positive). How can we prove this?

Lindsay and Mauldin (2000): The above can be extended to conformal iterated function systems. Given appropriate conditions on a finite conformal iterated function system (in particular, satisfying the Open Set Condition), one gets

$$D_r(m) \;=\; \frac{\beta(q)}{1-q} \tag{11.12}$$

precisely when $\beta(q) = rq$.

11.5 History of Results

The self-similar case (Graf and Luschgy) was first done for the Strong Separation case: for the IFS, there was a positive distance between the different parts of the

fractal (no touching boundaries). Allowing for boundaries to touch, we have the Open Set Condition, which states: there exists an open set U such that $\phi_i(U) \subset U$ and $\phi_i(U) \cap \phi_j(U) = \phi$ when $i \neq j$. Note for the Cantor Set we could take $U = (0, 1)$; but we could also use $[0, 1]\backslash C$, C the Cantor Set. This example is bad, as the open set U misses the actual fractal. This leads to the Strong Open Set Condition, which is the Open Set Condition *plus* U can be chosen such that $U \cap J \neq \phi$, i.e., there is an intersection of the open set and the fractal set. For the systems considered above the Open Set Condition and the Strong Open Set Condition are equivalent.

Open question: How do we prove the previous result for infinite iterated function systems?

Chapter 12

Iterated Function Systems

Lecturer: Bruce Adcock (Wednesday, November 17, 2003)

12.1 IFS with Escape Time Algorithms

Looking at IFS with escape time algorithms. From there, one can go on to looking at Julia sets.

Say we have an IFS $\{\mathbb{R}^2; w_1, \ldots, w_n\}$; each w_i is a contraction. Let S be the attractor for the IFS. Further, let

$$\begin{aligned} A_1 &= w_1(S) \\ &\vdots \\ A_n &= w_n(S) \setminus \bigcup_{i=1}^{n-1} A_i. \end{aligned} \tag{12.1}$$

We can define a dynamical system $\{\mathbb{R}^2; f\}$, where f is defined so that when we restrict $f : \mathbb{R}^2 \to \mathbb{R}^2$ to S, it satisfies

$$f(x,y) = \begin{cases} w_1^{-1}(x,y) & \text{if } (x,y) \in A_1 \\ \qquad\vdots & \\ w_n^{-1}(x,y) & \text{if } (x,y) \in A_n. \end{cases} \tag{12.2}$$

We have $f(S) \subset S$.

12.2 Example

A right-angle version of the Sierpinski Triangle: $\{\mathbb{R}^2; w_1, w_2, w_3\}$, where

$$\begin{aligned} w_1(x,y) &= (.5x, .5y + .5) \\ w_2(x,y) &= (.5x + .5, .5y) \\ w_3(x,y) &= (.5x, .5y) \end{aligned} \tag{12.3}$$

Consider

$$f(x,y) = \begin{cases} (2x, 2y-1) & \text{if } y \geq .5 \\ (2x-1, 2y) & \text{if } x \geq .5, y < .5 \\ (2x, 2y) & \text{otherwise.} \end{cases} \tag{12.4}$$

Assume $\overrightarrow{v}$, $\overrightarrow{z}$ are in the same case of definition of $f(x,y)$. Then

$$d(f(\overrightarrow{v}), f(\overrightarrow{z})) = 2d(\overrightarrow{v}, \overrightarrow{z}). \tag{12.5}$$

If we mix cases, the above formula is not true.

Exercise 12.2.1. *Assume $\overrightarrow{v} \in \mathbb{R}^2 \backslash S$, then*

$$\lim_{n\to\infty} d(\overrightarrow{0}, f^{\circ n}(\overrightarrow{v})) = \infty, \tag{12.6}$$

where in the above, we mean f composed with itself n times. Thus, anything outside of the attractor escapes to infinity.

Consider again the right Sierpinski Triangle. Let $\mathcal{W}$ be a rectangle containing the Sierpinski Triangle (the *window*). Let $\mathcal{V} = \{(x,y) \in \mathbb{R}^2 : \sqrt{x^2+y^2} > R\}$.

Let $\overline{n}$ be a cut-off parameter (called *numits*). Consider $(x,y) \in \mathcal{W}$. We assign a color based on how quickly its iterates tend to infinity. Explicitly,

$$\text{Color}(x,y) = \min\left(\overline{n}, \text{least } n \text{ such that} f^{\circ n}((x,y)) \in \mathcal{V}\right). \tag{12.7}$$

In other words, we iterate a point at most $\overline{n}$ times. If it hasn't escaped (made it into $\mathcal{V}$), we color in $\overline{n}$. The smaller the color, the faster it escapes. As $f(S) = S$, the points in S are always colored $\overline{n}$.

With the Sierpinski Triangle example, we can get the following:

12.3 Julia Sets

Let $f : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ be a polynomial of degree greater than one. Let F_f denote the points of $\mathbb{C}$ whose orbits don't converge to the Point at Infinity (note we are using $\widehat{\mathbb{C}}$, not $\mathbb{C}$). Explicitly,

$$F_f \;=\; \Big\{ z \in \mathbb{C} : \{|f^{\circ n}(z)|\}_{n=0}^{\infty} \text{ is bounded} \Big\}. \tag{12.8}$$

The above is the filled in Julia Set. The boundary of F_f is called the Julia Set of f, and is written as J_f.

Typically, when we talk about Julia sets, we mean $f_\lambda(z) = z^2 + \lambda$. This is tied in with the Mandelbrot set, which is $\{\lambda \in \mathbb{C} : \{z_n\}_{n=0}^{\infty} \,\mathrm{does\,not\,diverge}\}$ (where $z_0 = 0$ and $z_n = f(z_{n-1})$.)

Chapter 13

Fractals and Biology

Lecturer: Larry Lindsay (Friday, November 19, 2003)

In this lecture we will discuss the paper *Fractal Graphical Representation and analysis of DNA and protein sequences*, by Victor V. Solovyev (BioSystems, 30, 1993, pages 137 to 160). The discussion below, unfortunately, doesn't make much sense without the accompanying figures from the paper, but it does give an idea of how the study of fractals has been applied to research on DNA and protein sequences.

13.1 Introduction

13.1.1 DNA

There are two types of **purines:** *Adenine (A)* and *Guanine (G)*; there are two types of **pyrimidines**: *Cytosine (C)* and *Thymine (T)*. These code over 100,000 proteins by sequences of these four letters.

There is a place on a DNA where we start and end. Not everything codes – we have a coding region where proteins are coded (that's what start and stop refers to). We have **exons** (these are the coding) and the **introns** (these don't code, we may not know what they do; spliced out during some processes). There are also **intergenic sequences** (outside of the coding of a gene). For some strange reason, we call the starting point $5'$ and the stopping point $3'$.

13.1.2 Chaos Game

Take a Sierpinski triangle, play the Chaos game and see what is filled in. If instead of starting with three vertices we start with 4 vertices, we instead uniformly fill in the square. Now, give each vertex a probability p_i. Now if the p_i's are not all the same we won't fill in the square uniformly, but will get self-similar sets; p_i tells us how often we are in the quadrant near vertex i, then $p_1 p_4$ would be the square up near vertex 1, and then divide that quadrant in fourths, and take the one corresponding to vertex 4.

We can use this for proteins / DNA: label the corners of the square by A at $(0,0)$, C at $(1,0)$, G at $(1,1)$ and T at $(0,1)$. Given a sequence, go to the center of the quadrant with label the same as our first letter. Then, each letter tells us how we move. There is more self-similarity on the intron coding than on the exon coding (observation of figure).

We can talk about standard deviations. Going down to a certain level (say n^2 squares in our grid), then for a given sequence of length L, we expect each cell to have $\frac{L}{n^2}$ percent. Let

$$p_{ij} \;=\; \#\{\text{points in the } (i,j)\text{-cell}\}. \tag{13.1}$$

Let δ be the standard deviation. Then

$$\delta \;=\; \sqrt{\frac{1}{n^2-1}\sum_i \sum_j (p_{ij}-p)^2}. \tag{13.2}$$

A mask M_n^m ($m = 0, 1, 2, \dots$) is the collection of cells where p_{ij} is at least $p + m\delta$. In other words, these are the cells where we are much above p. This gets rid of cells where not much is going on.

13.1.3 Graphic Representation

Take a large grid, say 128 by 128. Take a sequence of length 50,000. Play our chaos game, and see how many points are in each cell. This is our p_{ij}. For a given sequence, say we are interested in GATATACC – this will correspond to a certain set of boxes that correspond to this. The number of boxes divided by the number of cells is the weight we give this sequence.

Define

$$F(subsequence) \;=\; \sum_i \sum_j p_{ij}, \tag{13.3}$$

where above the right sums are only over the boxes that have the sub-sequence; i.e., the number of cells containing this. Then we only look at strings of length 8 – there are 2^8 such strings.

In a histogram, the x-coordinate is the Fs, and the y is the number that have that F value.

13.2 Use of Fractal Representation for Analysis

13.2.1 Functional Regions / Gene Structure

Sliding window going along, say of length 50. Then move over 1, repeat, calculating F values. Then plot the F value. We've filtered first, so if it is relevant, it should have a spike there.

Idea: family of sequences that do same thing, not sure which genes do it. Make a family mask, test against another sequence that is known to do the same thing.

13.2.2 Proteins and Amino Acids

There are 20 amino acids. Look at a 4×5 grid. Label each cell with one of the 20 amino acids. Now map proteins in terms of amino acid sequences. This works well for non-primes; in fact, works best when a number factors as two comparable numbers.

Now let the amino acids be sectors on a circle. This works also for primes now. Say we have n sectors. Then in the subsequent stages, we divide each sector into $n - 1$ subsectors with a smaller radius, and then the remaining sector in the wedge is the n^{th}.

Chapter 14

Fractals and Random Walks

Lecturer: Dean Eiger (Wednesday, December 3, 2003)

14.1 Random Walks

A random walk is a chance process in which an initial point is selected, and advancements are made along the coordinate system with equal probability of movement in each possible direction.

We define coordinate systems as follows: in $\mathbb{R}^1 = \mathbb{R}$, let $\widehat{i}$ be the unit vector in the positive direction. Let $\overrightarrow{r}_0$ be a starting point. We often choose at each stage either $+\widehat{i}$or $-\widehat{i}$. Thus, after n steps, we are at

$$\overrightarrow{r_n} \;=\; \overrightarrow{r_0} + \sum_{j=1}^{n} (\pm\widehat{i}). \tag{14.1}$$

Let $p(x) = \frac{x}{n}$. This is the probability that we reach the right end point (n) before we reach the left endpoint (0); let $x = \overrightarrow{r_0}$ for simplicity. Now $p(0) = 0$, $p(n) = 1$, and $p(x) = \frac{p(x-1)+p(x+1)}{2}$.

14.2 Example: Circuit Analysis

Consider an electrical circuit. Let it have a voltage source of unit voltage, and an arbitrary number of resistors connected in series. We can define current by

$$i_{x,x'} \;=\; \frac{V(x) - V(x')}{R}, \tag{14.2}$$

where $V(0) = 0$, $V(n) = 1$. By Kirchoff's Law, we have

$$\frac{V(x-1) - V(x)}{R} + \frac{V(x+1) - V(x)}{R} \;=\; 0. \tag{14.3}$$

Therefore,

$$V(x) \;=\; \frac{V(x+1) - V(x-1)}{2}. \tag{14.4}$$

14.3 Harmonic Functions

Let $B = \{[0, n]\}$, and let $I = \{[1, n-1]\}$. We say a function is harmonic if, on the interior points I, we have $f(x) = \frac{f(x+1)+f(x-1)}{2}$; we say B is the boundary.

Dirichlet Problem: given the values of a harmonic function on the boundary, find the function in the interior.

Lemma 14.3.1 (Maximum Principle). *The largest value of a harmonic function occurs on the boundary.*

Proof. Assume not; thus, let f take its maximum value in the interior. For interior points $x \in I$, if $f(x) = M$ then $f(x \pm 1) = M$, implying the function is constant (and in this case, the maximum will also occur on the boundary). □

Lemma 14.3.2 (Uniqueness Principle). *If f and g are harmonic functions and $f(x) = g(x)$ for all $x \in I$, then $f(x) = g(x)$ for all $x \in B$.*

Proof. Let $h(x) = f(x) - g(x)$. For all $x \in I$,

$$\frac{h(x-1) + h(x+1)}{2} \;=\; \frac{f(x-1) + f(x+1)}{2} - \frac{g(x-1) + g(x+1)}{2}. \tag{14.5}$$

Therefore, if $f(x) - g(x) = h(x)$, then $h(x)$ is harmonic (because f and g are). By the maximum principle, the maximum of h occurs on the boundary. As $f(x) = g(x)$ on the boundary, this implies the maximum of h is zero, which then gives $h(x) = 0$ for $x \in I$. □

14.4 Random Walks in Infinite Spaces

Whereas the central problem of random walks in finite spaces is calculating the probability of reaching one terminal point before the other, the probability of a random walk escaping to infinity is the central problem in an infinite space. A random walk returning to the initial point is said to be recurrent, and one that does not is transient.

14.4.1 Escape Probability

The probability of a random walk on an infinite lattice escaping to infinity is determined by consideration of finite subsets of the lattice.

Let $G^{(r)}$ be any finite subset of L, where L is a lattice. Here $r < \sqrt{\sum_{i=1}^{d} x_i^2}$ is the integer-valued radius of the subset $G^{(r)}$. Let $S^{(r)}$ be the initial point.

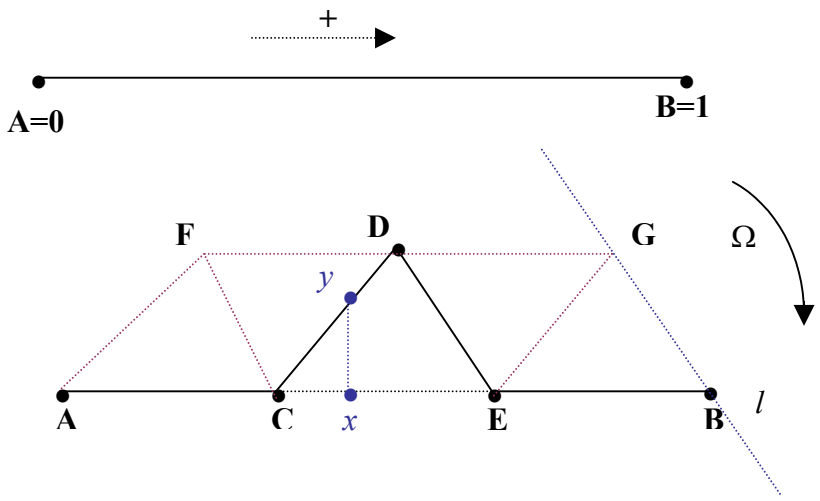
Again, consider the probability of a system being transient (the probability of escape). Thus, $p_{esc}^{(r)}$ is the probability of leaving $G^{(r)}$ before returning to $S^{(r)}$ (where we start). Therefore

$$p_{esc} \;=\; \lim_{r\to\infty} p_{esc}^{(r)}. \tag{14.6}$$

Helge von Koch: On a Continuous Curve without Tangents from Elementary Geometry

Koch Fractal

The segment $P_0=AB$ has a positive direction from A to B. The operation Ω replaces the segment AB with the polygonal line $P_1=ACDEB$.



The polygonal line P_n is obtained by applying n times the operation Ω starting from $AB=[0,1]$.

P_n has 4^n segments of length $\frac{1}{3^n}$. Let S be the set of all vertices of the polygonal lines P_1 ,P_2,,..., P_n,... and $S'=\overline{S}$ the set of the limit points of the points of S.

Define functions $k_n:[0,1]\to P_n$ in the following way

$k_0(x)=x$

$$k_1(x)=\begin{cases} x & \text{if } x\in AC \text{ or } EB \\ y & \text{if } x\in CE \text{ (as shown)} \end{cases}$$

The function k_n coincides with k_{n-1} on their common segments and $k_n(x)=y_n$ if $k_{n-1}(x)$ belongs to a segment ST of P_{n-1} which is removed in the construction of P_n and y_n is the point of intersection of the perpendicular to ST with P_n.

The functions k_n are continuous and converge uniformly to a continuous function k because

$$\| k_n(x)-k_{n-1}(x)\| < \frac{1}{2.3^n}\tan(\frac{\pi}{3})=\frac{\sqrt{3}}{2.3^n}$$

Let P be the curve obtained with the limiting process described above. The curve P is the continuous function $k:[0,1]\to S'$ and consists of all points of S'.

- ***The function k is bounded***

$$\| k_n(x)\| \le \sum_{m=1}^{n} \| k_m(x)-k_{m-1}(x)\| \le \sum_{m=1}^{n} \frac{\sqrt{3}}{23^m} = \frac{\sqrt{3}}{6}\frac{1-\frac{1}{3^n}}{1-\frac{1}{3}} = \frac{\sqrt{3}}{4}(1-\tfrac{1}{3^n})$$

$$\| k(x)\| = \lim_{n\to\infty} \| k_n(x)\| \le \frac{\sqrt{3}}{4} \quad \text{i.e. } \| k(x)\| < |AB| \frac{\sqrt{3}}{4}$$

- ***The curve k is not rectifiable***

The length L_n of the curve P_n is $\frac{4^n}{3^n}$. Then the length L of the curve P is

$$L = \lim_{n\to\infty} L_n = \lim_{n\to\infty} \frac{4^n}{3^n} = \infty$$

- ***The area between the curve P and the x – axis is*** $\mathbf{\frac{\sqrt{3}}{20}}$

The area of an equilateral triangle with sides of length x is $\frac{1}{2}x^2\sin(60°) = \frac{\sqrt{3}}{4}x^2$

The region between the curves P_n and P_{n-1} consists of 4^{n-1} equilateral triangles and has area $A_n = \frac{\sqrt{3}}{4}\frac{1}{3^{2n}}4^{n-1} = \frac{\sqrt{3}}{16}\frac{4^n}{9^n}$.

The area between the curve P and the $x-axis$ is

$$A = \sum_{n=1}^{\infty} A_n = \sum_{n=1}^{\infty} \frac{\sqrt{3}}{4}\frac{4^n}{9^n} = \frac{\sqrt{3}}{16}\frac{4}{9}\frac{1}{1-\frac{4}{9}} = \frac{\sqrt{3}}{20}$$

Now we want to show that the curve P doesn't intersect itself. Let l be the line that meets DE at the point E with an angle $60°$.

- ***The curve P lies on the left side of l***

Let $x\in[0,1]$

If $k(x)=x$ then $k(x)\in AB$ and AB is on the left side of l.

If $k(x)\neq x$ then there is an integer n such that $k_n(x)\neq x$. The distance between l and the segments of P_n (except the segment with vertex B) is $\ge \frac{1}{3^n}\cos(30°) = \frac{\sqrt{3}}{2.3^n}$. The point $k(x)$ lies on the part of P generated by a segment of length $\frac{1}{3^n}$. Then $\| k(x)-k_n(x)\| \le \frac{\sqrt{3}}{4.3^n}$ and so $k(x)$ is on the left side of l.

- ***The curve P doesn't intersect itself***

Suppose that $0 \le x < y \le 1$

Let n be an integer such that $\frac{1}{3^n} < y - x$. Then the polygonal line P_n has a vertex which lies between $k_n(x)$ and $k_n(y)$. Then the points $k(x)$ and $k(y)$ belong to parts of P defined by different segments of P_n. As we have shown in Proposition1, these parts of P don't intersect and so $k(x) \neq k(y)$.

The points of P_n (and P) have a natural ordering i.e. the point $X = k_n(x)$ *precedes* the point $Y = k_n(y)$ if $x < y$ and X *succeeds* Y if $x > y$.

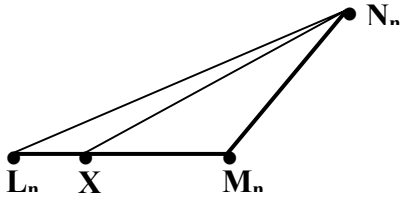
Theorem: ***The curve P doesn't have a well defined tangent at any point.***

Proof: Let X be a point on the curve P.

- If X is also a vertex of one of the polygonal lines P_n then both segments of P_n which have vertex X contain infinitely many vertices approaching X.

The angle between these segments is 60° or $120°$ and so P doesn't have a tangent line at X.

- If X is not a vertex and $k_n(x) = X$ for some integer n, then X belongs to all polygonal lines P_m, $m \ge n$. Then the lines XM_n and XN_n meet at an angle between $30°$ and $60°$ and so the sequence of lines $l_{2n-1} = XM_n$, $l_{2n} = XN_n$ do not converge i.e. the curve P doesn't have a tangent line at the point X.



- If X doesn't belong to any P_m, let L_n and M_n be two consecutive vertices of P_n such that $X_n = k_n(x)$ lies between them. If P has a tangent line at X, then the lines L_nM_n approach the tangent line at X. The angle between L_nM_n and $L_{n+1}M_{n+1}$ is $0°, 60°$ or $120°$. This angle is nonzero infinitely many times because $k(x) \neq k_n(x)$. Therefore the curve P doesn't have a tangent line at the point X. Since X is an arbitrary point of P the function k is not differentiable at any point in the interval $[0,1]$.

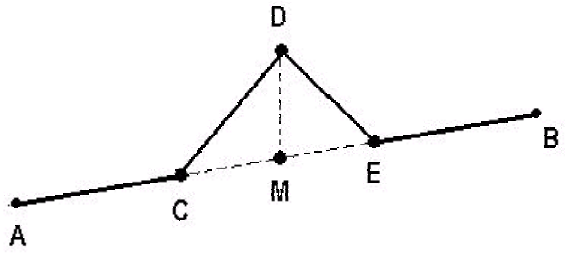
Interactive constructions of Koch fractal and Koch snowflake are implemented on

http://www.3rd-imperium.com/Java/Fractals/KF.html

Another Nondifferentiable Curve

The median DM of triangle CDE is vertical and has length

$$|DM| = \frac{\sqrt{3}}{2}|CE|$$

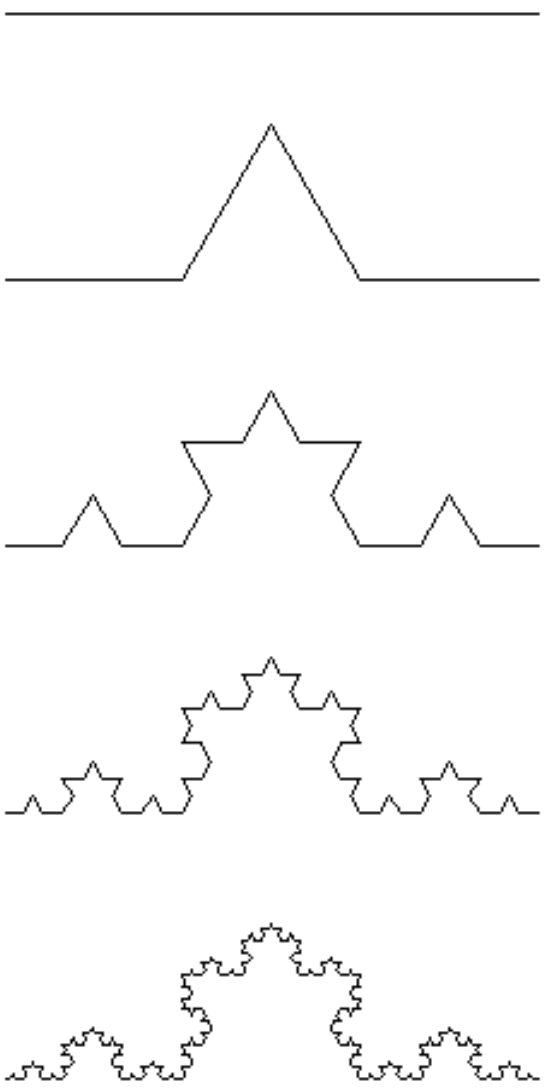


Let φ_n be the polygonal line obtained by applying the above operation n times starting with the unit interval. The functions φ_n converge uniformly to a continuous function φ.

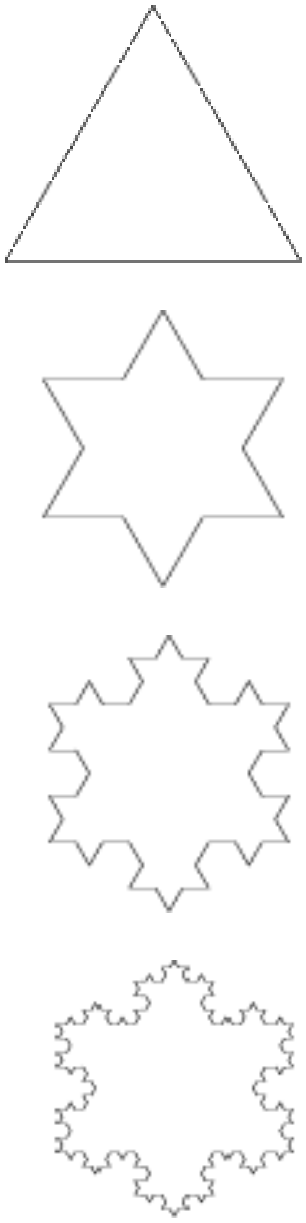
The function φ doesn't have derivative for any value of x

If $\varphi(x)$ belongs to one of the polygonal lines the proof is analogous to the Koch snowflake. The other possibility is that the points $\varphi_n(x)$ approach the point $\varphi(x)$ and the sequence $\{\varphi_n(x)\}$ contains infinitely many different points. If the line AB has a positive slope then the angle between CD and CE is less than $60°$ and the angle between EC and ED is greater than $60°$. Let's denote the unit interval as *left side* and the first side of a triangle of φ_n built on a left side of φ_{n-1} is *left* and the second is *right*. If the triangle is built on a *right* side of φ_{n-1}, then the first side is *right* and the second is *left*. Let $X_n = \varphi_n(x)$ and let's choose a subsequence $Y_k = X_{n_k}$ which doesn't have equal points. To each point Y_k we correspond the segment S_k with vertices s_{k_1} and s_{k_2} such that the point Y_k lies between s_{k_1} and s_{k_2}. If the sides S_k change from *left* to *right* (or from *right* to *left*) infinitely many times then the angle between every two consecutive sides is greater than $60°$. Then the sequence of secants $s_{k_1}\, s_{k_2}$ doesn't converge and so $\varphi(x)$ doesn't have a tangent line at x. Suppose that the segments $s_{k_1}\, s_{k_2}$ are left sides for all $k \geq k_0$. The sequences of left sides approach a vertical line and so the function φ cannot have a finite derivative at x. If the slope of AB is negative, the proof that the curve $\varphi(x)$ is nowhere differentiable is similar.

Koch Fractal



Koch Snowflake



A Simple Example of a Function, which is Everywhere Continuous and Nowhere Differentiable

Karl Kiesswetter

Let $f:[0,1]\to R$ be the function defined as follows

$f(x)=\sum(-1)^{N_\nu}\frac{X_\nu}{2^\nu}$ where

$x=\sum\frac{x_\nu}{4^\nu}$, $x_\nu=0,1,2,3$ is the base four expansion of $x\in[0,1]$ and

$$X_\nu=\begin{cases}x_\nu-2 & \text{if } x_\nu>0\\ 0 & \text{if } x_\nu=0\end{cases}$$

$N_\nu=$ number of x_k such that $x_k=0$ and $k<\nu$.

Then the values of X_ν are $-1,0,1$ and the series for $f(x)$ is absolutely convergent for every x. The values of f for $x=0,\frac{1}{4},\frac{1}{3},\frac{1}{2},1$ are

$f(0)=0$ $\qquad$ $f(\frac{1}{4})=-\frac{1}{2}$

$\frac{1}{3}=\sum_{k=1}^{\infty}\frac{1}{4^n}$ $\qquad$ $f(\frac{1}{3})=\sum_{k=1}^{\infty}-\frac{1}{2^k}=-1$

$\frac{1}{2}=\frac{2}{4}$ $\qquad$ $f(\frac{1}{2})=0$

$1=\sum_{k=1}^{\infty}\frac{3}{4^n}$ $\qquad$ $f(1)=\sum_{\nu=1}^{\infty}\frac{1}{2^\nu}=1$

The numbers $x=\frac{p}{4^n}$ in the interval $[0,1]$ have two different base four representations.

Now we show that the value of f is independent of the representation of x.

Let

$a=\sum_{\nu=1}^{n}\frac{a_\nu}{4^\nu}=a^*+\frac{a_n}{4^n}$ $\qquad$ with $a_n>0$

$b=\sum_{\nu=1}^{\infty}\frac{b_\nu}{4^\nu}$ with $b_\nu=a_\nu$ for $\nu<n$, $b_n=a_n-1$, and $b_\nu=3$ for $\nu>n$.

Then $a=b$ and $b=a^*+\frac{a_n-1}{4^n}+\frac{3}{4^{n+1}}+\frac{3}{4^{n+2}}+\cdots$

Case1: $a_n=3$: $\quad f(a)=f(a^*)+\frac{(-1)^{N_n}}{2^n}$

$$f(b) == f(a^*) + \sum_{\nu=n+1}^{\infty} \frac{(-1)^{N_n}}{2^{\nu}} = f(a^*) + \frac{(-1)^{N_n}}{2^n} = f(a)$$

Case2: $a_n = 2:\quad f(a) = f(a^*)$

$$f(b) == f(a^*) - \frac{(-1)^{N_n}}{2^n} + \sum_{\nu=n+1}^{\infty} \frac{(-1)^{N_n}}{2^{\nu}} = f(a^*)$$

Case3: $a_n = 1 \qquad f(a) = f(a^*) - \frac{(-1)^{N_n}}{2^n}$

$$f(b) == f(a^*) + \sum_{\nu=n+1}^{\infty} \frac{(-1)^{N_n+1}}{2^{\nu}} = f(a^*) - \frac{(-1)^{N_n}}{2^n} = f(a)$$

Now we show that f is a continuous nondifferentiable function

Theorem1:

The function f is Lipschitz continuous and $|f(x) - f(y)| \le 4|x-y|^{\frac{1}{2}}$

Proof: Let $0 \le x, y \le 1$ and

$|x - y| = \frac{t_1}{4^n} + \frac{t_2}{4^{n+1}} + \cdots$ where $t_1 > 0$

Then $|x - y| \ge \frac{1}{4^n}$ and

$|f(x) - f(y)| \le \frac{2}{2^n} + \frac{2}{2^{n+1}} + \cdots$

$|f(x) - f(y)| \le \frac{4}{2^n}$

Therefore $|f(x) - f(y)| \le 4|x-y|^{\frac{1}{2}}$ ■

Since $f(0) = 0$ and $f(1) = 1$, the function f can be extended continuously to $\mathbb{R}$ with $f(x) = [x] + f(x - [x])$. The fractal properties of f follow from the following lemma:

Lemma 1: *For every* $\frac{1}{4} \le x \le \frac{1}{2}$

(i) $f\left(x - \frac{1}{4}\right) + f(x) = \frac{1}{2}$

(ii) $f\left(x + \frac{1}{4}\right) - f(x) = \frac{1}{2}$

(iii) $f\left(x + \frac{1}{2}\right) + f(x) = 1$

Proof:

(i) Let $x = \frac{1}{4} + x_1$ where $x_1 < \frac{1}{4}$

Then $x-\frac{1}{4}=x_1$ and $f(x)=-\frac{1}{2}-f(x_1)=-\frac{1}{2}-f(x-\frac{1}{4})$

(ii) $x+\frac{1}{4}=\frac{2}{4}+x_1$ Then $f(x+\frac{1}{4})=-f(x_1)=f(x)+\frac{1}{2}$

(iii) $x+\frac{1}{2}=\frac{3}{4}+x_1$ Then $f(x+\frac{1}{2})=\frac{1}{2}-f(x_1)=1+f(x)$ ■

Let's denote by K the graph of f i.e. $K=\{(x,f(x))\mid x\in[0,1]\}$. The set K is compact because f is continuous on $[0,1]$.

Consider the following IFS $F=\{f_1,f_2,f_3,f_4\}$ in the plane

$$f_1\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}1/4 & 0\\0 & -1/2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}$$

$$f_2\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}1/4 & 0\\0 & 1/2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}+\begin{bmatrix}1/4\\-1/2\end{bmatrix}$$

$$f_3\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}1/4 & 0\\0 & 1/2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}+\begin{bmatrix}1/2\\0\end{bmatrix}$$

$$f_4\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}1/4 & 0\\0 & 1/2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}+\begin{bmatrix}3/4\\1/2\end{bmatrix}$$

From *Lemma1* follows that $K=f_1(K)\cup f_2(k)\cup f_3(K)\cup f_4(K)$
Then the set K is the attractor set of F because K is compact. Kiesswetter's fractal is self-affine and has Hausdorff dimension $\frac{3}{2}$.

Theorem2:
f is nondifferentiable at each point in the interval $[0,1]$.

In the proof of *Theorem2* we need the following two lemmas

Lemma 1:
Suppose that $0\le V_n=\frac{p}{4^n}$ *and* $W_n=V_n+\frac{1}{4^n}\le 1$. *Then* $|f(W_n)-f(V_n)|=\frac{1}{2^n}$.

Proof: Let $W_n=V_n+\sum_{\nu=n+1}^{\infty}\frac{3}{4^n}$. Then

$$f(W_n)=f(V_n)+\sum_{\nu=n+1}^{\infty}\frac{(-1)^{M_{n+1}}}{2^{\nu}}$$

$$| f(W_n) - f(V_n) | = \sum_{\nu=n+1}^{\infty} \frac{1}{2^{\nu}} = \frac{1}{2^n} \qquad \blacksquare$$

Lemma 2:

If $A \le x \le B$ with $A < B$, then at least one of the following inequalities hold

$$\left|\frac{f(B)-f(x)}{B-x}\right| \ge \left|\frac{f(B)-F(A)}{B-A}\right| \quad or \quad \left|\frac{f(x)-f(A)}{x-A}\right| \ge \left|\frac{f(B)-F(A)}{B-A}\right|$$

Proof:

$$(B-A)\frac{f(B)-f(A)}{B-A} = (B-x)\frac{f(B)-f(x)}{B-x} + (x-A)\frac{f(x)-f(A)}{x-A}$$

$$(B-A)\left|\frac{f(B)-f(A)}{B-A}\right| \le (B-x)\left|\frac{f(B)-f(x)}{B-x}\right| + (x-A)\left|\frac{f(x)-f(A)}{x-A}\right| \quad (*)$$

If both held simultaneously:

$$\left|\frac{f(B)-f(A)}{B-A}\right| > \left|\frac{f(B)-f(x)}{B-x}\right| \quad \text{and} \quad \left|\frac{f(B)-f(A)}{B-A}\right| > \left|\frac{f(x)-f(A)}{x-A}\right|$$

Then

$$(B-A)\left|\frac{f(B)-f(A)}{B-A}\right| > (B-x)\left|\frac{f(B)-f(x)}{B-x}\right| + (x-A)\left|\frac{f(x)-f(A)}{x-A}\right|$$

This is a contradiction to inequality $(*)$ $\blacksquare$

Proof of Theorem 2:

For each $0 \le x \le 1$ and for every natural number n there exists at lest one V_n such that

$$0 \le V_n \le x \le W_n = V_n + \frac{1}{4^n} \le 1$$

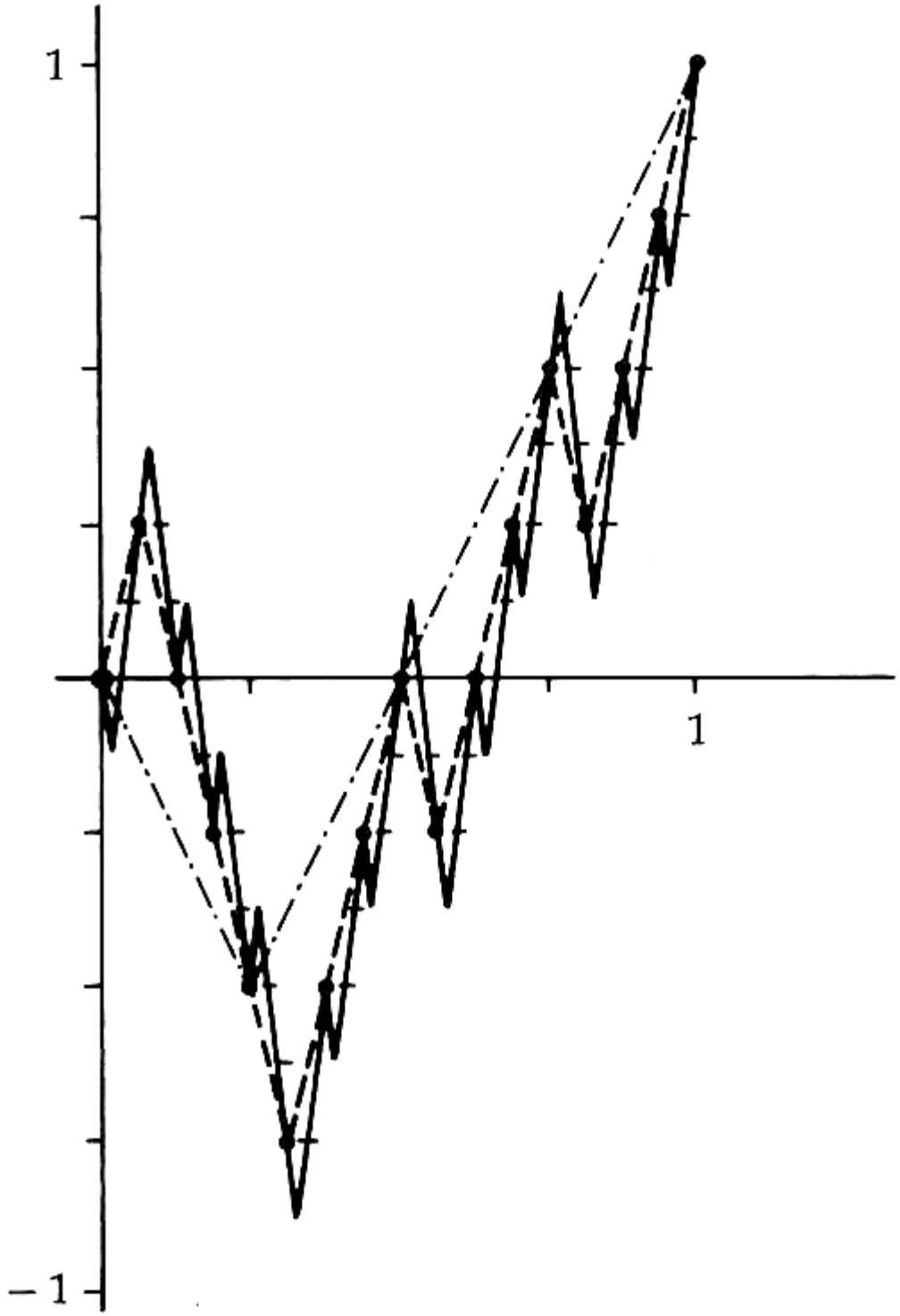
From *Lemma 2* it follows for at least one of V_n or W_n (denoted by K_n) that

$$\left|\frac{f(x)-f(K_n)}{x-K_n}\right| \ge \left|\frac{f(W_n)-f(V_n)}{W_n-V_n}\right| = \frac{1/2^n}{1/4^n} = 2^n$$

Therefore f is not differentiable at x because $\lim_{n\to\infty} K_n = x$ and the sequence of difference quotients $\dfrac{f(x)-f(K_n)}{x-K_n}$ diverges. $\blacksquare$

References.

G. Edgar, *Classics on Fractals*, Westview Press, 2004


1
1
−1

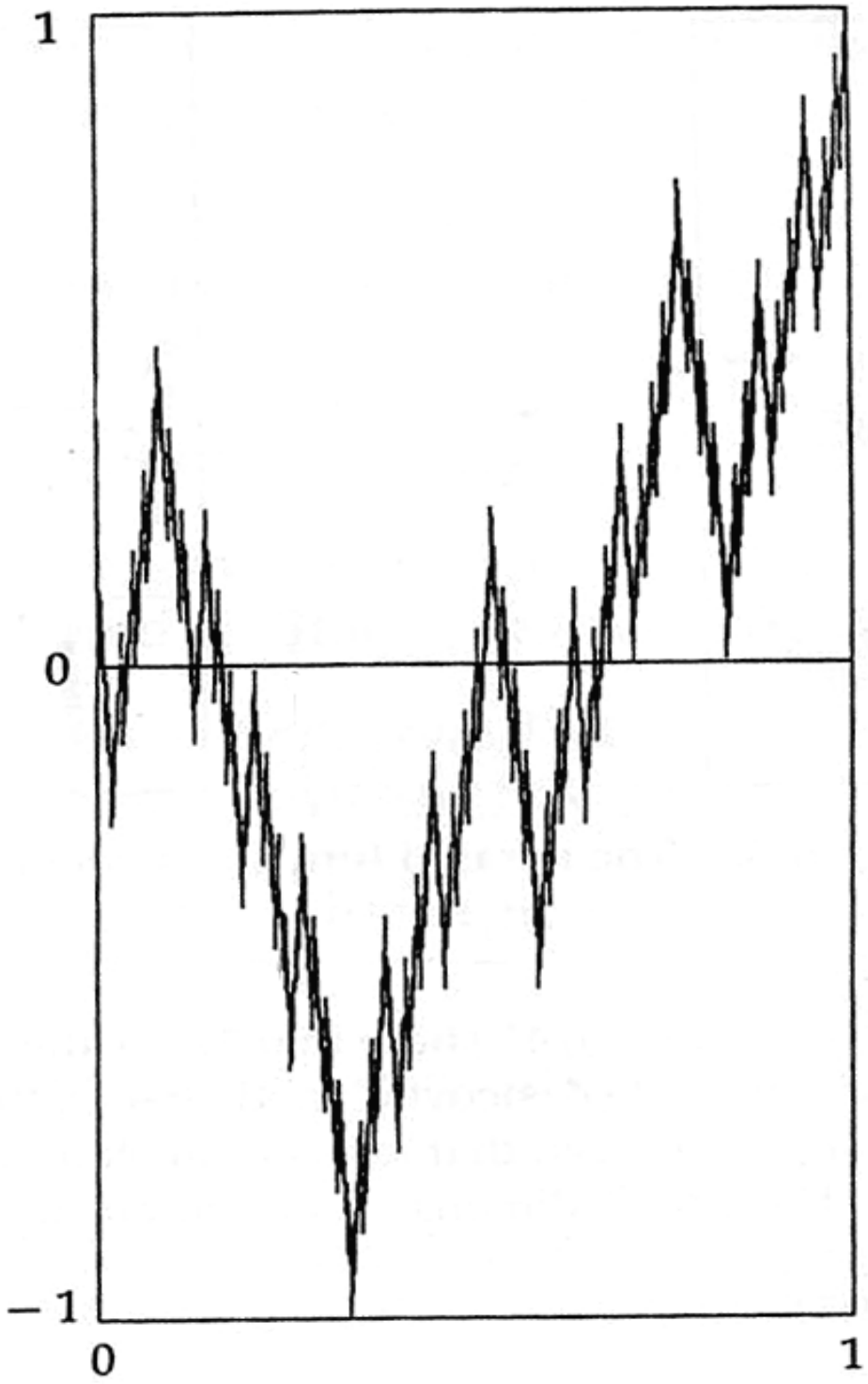


Figure 2. Kiesswetter's fractal.

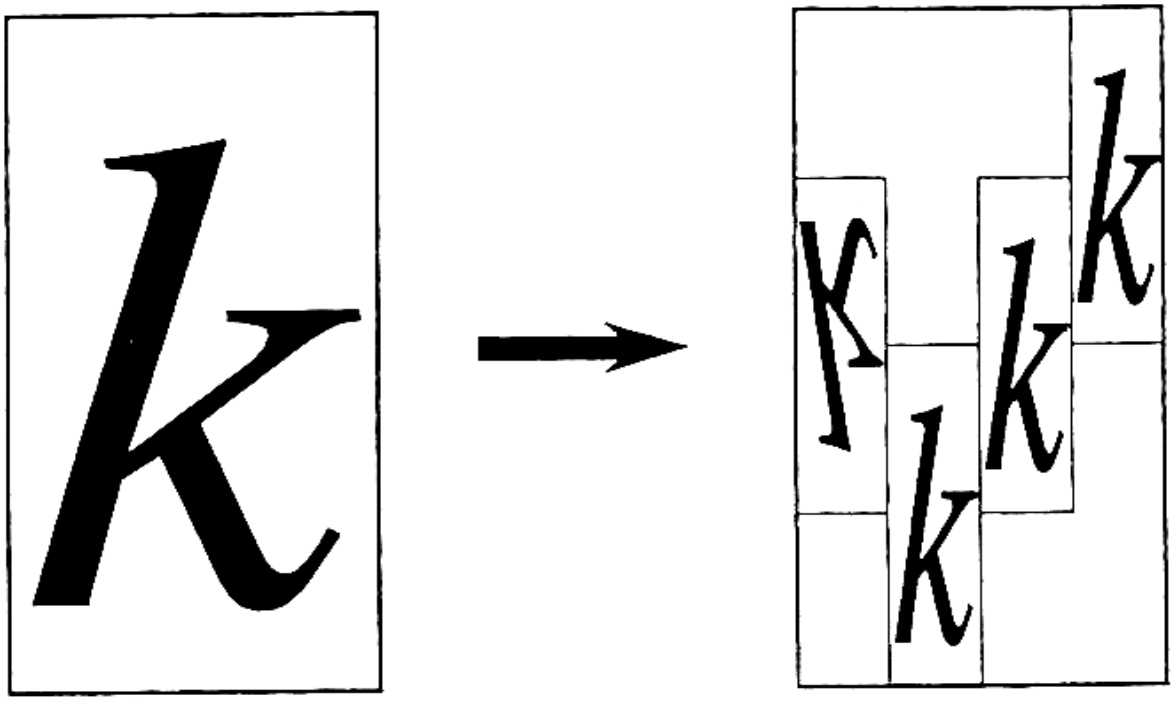


Figure 3. The iterated function system.