2013

Fractal dimension and Julia sets

Robert DeLorto
Eastern Washington University

FRACTAL DIMENSION AND JULIA SETS

A Thesis

Presented To

Eastern Washington University

Cheney, Washington

In Partial Fulfillment of the Requirements

for the Degree

Master of Science

By

Robert DeLorto

Spring 2013

THESIS OF ROBERT DELORTO APPROVED BY

____ DATE:________

W. DALE GARRAWAY, GRADUATE STUDY COMMITTEE

____ DATE:________

RONALD GENTLE, GRADUATE STUDY COMMITTEE

____ DATE:________

HANI SAAD, GRADUATE STUDY COMMITTEE

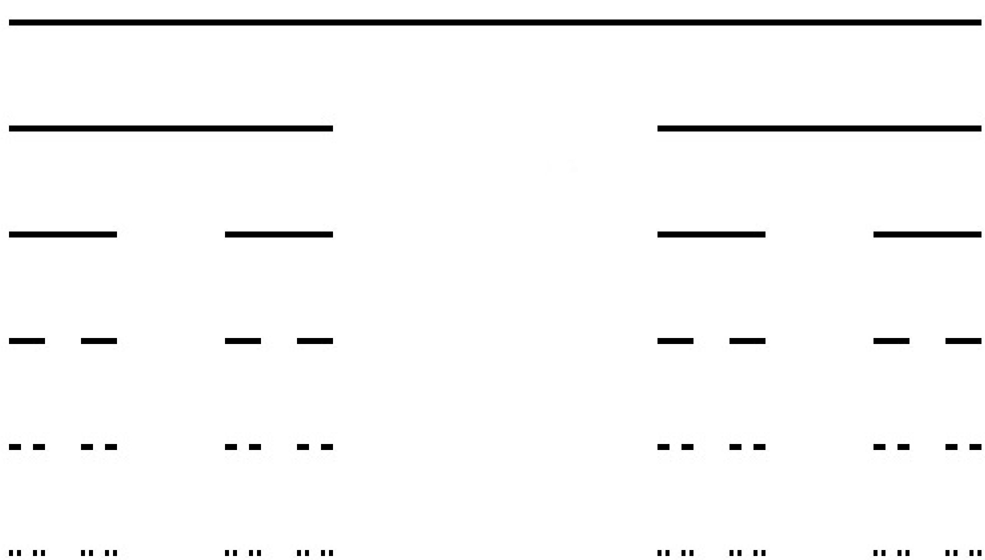
Contents

Chapter 1

Introduction

There are objects in our world defined by their own complexity, whose shape has so much detail that it cannot be observed by the human eye, nor by any instrument that man has contrived. These objects are observed from the intricate details of our bodies to the great expanse of the universe. They are images which cannot be described in the normal notion of viewing dimensions. These mysterious images are called fractals.

One of the first fractals studied was the Cantor set in 1874. At this time, however, the term "fractal" had not been formed and the properties of them had not been discovered. These things did not come until many decades later. It was Benoit Mandelbrot, the French mathematician, who coined the term "fractal" from the Latin "fractus," meaning "broken" or "fractured" in 1975. Mandelbrot was working with engineers at IBM, studying the noise in telephone lines. Information was carried from one computer to another by an electric current. However, given an interval of time there was sometimes some interference noise that would be transmitted, causing errors. This noise occurred seemingly at random, and the engineers wanted to find out why there were errors so that they could eradicate them. Mandelbrot discovered that, although the noise seemed to occur randomly, it happened at all scales during the transmission, and the proportion of error noise to clean transmission with no noise was constant. So given a period of time when error existed, the ratio of error to clean transmission in a second was the same as the ratio of error to clean transmission in an hour. What Mandelbrot found was that this is an abstraction of the Cantor set, which was developed many decades earlier. The Cantor set is a type of fractal constructed by first drawing a line with length one. Now remove the middle third of the line so that there are two line segments, each of length one third remaining. Then remove the middle thirds of each of these segments, and so on, resulting in the following figure.



Although the error was present in the transmissions, there was no way for the engineers to remove it. They were forced to settle with trying to produce a transmission with the least error possible. It is not surprising that it was Mandelbrot who discovered the relationship between the Cantor set and these transmissions. In school, Mandelbrot excelled at geometry, in fact, when he was in school this is how he came up with the answers to most of the questions in his math classes (much to the dismay of his superiors). Looking at the world with a geometric eye, he would often say things like, “Clouds are not spheres,” “Mountains are not cones,” or “Lightning does not travel in a straight line.” These natural objects have a jaggedness to them that makes them unique. In fact, they are very close to fractals in shape (though not exactly).

It was these jagged objects in nature that piqued Mandelbrot’s curiosity. He wanted a way of measuring these objects, but couldn’t do it with the existing ways of measuring. It was then that he posed his famous question, “How long is the coastline of Britain?” Say someone has a “ruler” that is a kilometer long. This person walks along the coast of Britain and measures the perimeter. Then someone else comes along with a “ruler” that is only a meter long and measures the coastline of Britain. The person with the meter long ruler will get a longer result than the person who used the ruler that was a kilometer long because the meter can go into some of the jagged areas that the kilometer ruler couldn’t reach. If someone else comes along with an even smaller ruler, the result would be even longer. So it seems as if the coastline of Britain has an infinite perimeter (though, of course, Mandelbrot knew that this is not the case since nothing in the real world can have infinite perimeter). To account for this, Mandelbrot decided to measure fractals using what is called *fractal dimension*. Dimension will be discussed in more detail in chapter 4.

Other examples of objects that have a fractal nature include trees, snowflakes, mountains, and clouds. Even our the lungs in our bodies are fractal. This should make sense because the goal of the lungs is to get as much air to the body, but it is contained in a finite cavity. So if the lungs can maximize its surface area, what better shape is there than a fractal? In the three dimensional sense, a fractal can have infinite surface area and a finite volume - perfect for what the lungs need[7].

In this human age of cinematography, technology made from fractals is used to create realistic special effects. For example, to make a mountain an artist might start off with a cone, then apply some fractal generator to the cone to make it look more mountainlike. In Star Wars episode III, fractals were used to create the splashing lava towards the end of the movie. The layers required to create the realistic lava were generated by a fractal. This method for creating functions that describe a fractal image will be discussed in chapter 5.

The mathematical structure of fractals are deeply tied to definitions from topology and analysis. So in this chapter we will explore definitions and concepts that will be needed. Note that definitions of topology are incredibly general so that they can be applied to various situations. The definitions may seem more complicated than the practice.

Topology is an area of mathematics that describes the abstract structure of certain sets or objects. It is one of the most general disciplines of math, so if a theorem is proved in topology, the theorem will hold for sets in other areas of math, such as algebra and analysis. The axioms of topology are defined in terms of open and closed sets.

Definition 1.1 Let X be a set. Then a *topology* $\mathcal{T}$ is a collection of subsets of X, called *open sets*, which satisfy the following:

1. X and $\emptyset$ are open sets.

2. Any union of open sets is open.

3. Any finite intersection of open sets is open.

These three statements are known as the axioms of topology. The pair $(X, \mathcal{T})$ is called a *topological space*.

Definition 1.2 A set is called *closed* if its complement is open.

We get the following proposition, which is proved in any book on topology. In fact, this proposition can be taken as an alternate definition of a topology.

Proposition 1.3 Let $(X, \mathcal{T})$ be a topological space. Then

1. X and $\emptyset$ are closed sets.

2. The finite union of closed sets is closed.

3. Any intersection of closed sets is closed.

Definition 1.4 Let $(X, \mathcal{T})$ be a topological space. Let $x \in X$. An *open neighborhood* around x is an open set that contains x. Let $A \subseteq X$. The *closure* of A, denoted $\overline{A}$, is the smallest closed subset of X containing A. The *interior* $A \subseteq X$, denoted $\text{int}(A)$ is the set of points of $x \in A$ such that any open neighborhood around x is also contained in A. The *boundary* of A is denoted ∂A and is the set $\overline{A} \backslash \text{int}(A)$.

A point x is in ∂A if and only of a neighborhood centered around x contains points both in A and in the complement of A. In our study of fractals we will mainly be using special types of topological spaces called *metric spaces*. These are topological spaces with some distance defined on the set.

Definition 1.5 A *metric* on a set X is a function $d : X \times X \to [0, \infty)$ such that for all $x, y, z \in X$, then

1. $d(x, y) > 0$ if $x \neq y$
2. $d(x, y) = 0$ if and only if $x = y$
3. $d(x, y) = d(y, x)$ (symmetry)
4. $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality)

The pair (X, d) is called a *metric space*.

Example 1.6 For an example of a metric space, let $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be defined by $d(x, y) = |x - y|$. This is a metric defined on the real line, measuring the distance between two points. There are similar metrics defined for $\mathbb{R}^n$, described by the appropriate distance formula. In $\mathbb{R}^2$, for two points (x_1, x_2) and (y_1, y_2) the distance between these two points is defined as $\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$. In general, for two points $(x_1, x_2, ..., x_n)$ and $(y_1, y_2, ..., y_n)$ in $\mathbb{R}^n$ we distance is defined as $\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + ... + (x_n - y_n)^2}$.

□

Definition 1.7 Let (X, d) be a metric space. If $x \in X$ and $\epsilon > 0$, then an *ϵ-ball* $B_\epsilon(x)$ around x is $B_\epsilon(x) = \{y \in X : d(x, y) < \epsilon\}$.

Example 1.8 Let $\epsilon = 2$ and $x = 4$. Then for d as described above, $B_\epsilon(x) = B_2(x)$ is the open interval $(2, 6)$.

In $\mathbb{R}^2$, take $\epsilon = 2$ and the point $A = (2, 4)$. Then $B_\epsilon(A)$ is the open circle (circle with no boundary) with radius 2 centered at $(2, 4)$.

Similarly, in $\mathbb{R}^3$, if we have $\epsilon = 2$ and the point $C = (2, 4, 6)$, then $B_\epsilon(C)$ is the open sphere with radius 2 centered at $(2, 4, 6)$ □

Given a metric there is an associated topology. Open sets in a metric space are countable unions of balls and finite intersections of balls. So if $U \subseteq X$, then U is an open set if for each $x \in U$, there exists $\delta > 0$, depending on x, such that $B_\delta(x) \subset U$. If (X, d) is a metric space and $x \in X$, then $B_\epsilon(x)$ is an open set. This is because for any $y \in B_\epsilon(x)$ we can find a $\delta > 0$ such that $B_\delta(y) \subseteq B_\epsilon(x)$, so $B_\epsilon(x)$ is open.

Example 1.9 The interval $(2, 4)$ is an open set in $\mathbb{R}$ since for any number x in that interval, we can find an ϵ such that the $\epsilon-$ball around x is in $(2, 4)$.

In $\mathbb{R}^2$, the circle C with no boundary centered at the point $(2, 4)$ with radius 2 is an open set since for any point $a \in C$ we can find an $\epsilon > 0$ such that $B_\epsilon(a) \subseteq C$. Note that we exclude the boundary of the circle because if we choose a point on the circle, then a ball centered around that point will contain points on the outside of the circle. □

Example 1.10 The closure of $(2, 4)$ is $[2, 4]$ since $[2, 4]$ is the smallest closed set containing $(2, 4)$.

For $\mathbb{R}^2$, the closure of the open circle centered at $(2, 4)$ with radius 2 is the circle of radius 2 with its boundary. The circle with its boundary is a closed set since its complement is and open set.

Likewise, in $\mathbb{R}^3$, the open sphere centered at $(2, 4, 6)$ with radius 2 is the same sphere with its boundary. □

Definition 1.11 A collection $\{U_i\}$ of subsets of a topological space X is a *cover* of X if $X \subseteq \bigcup U_i$. For a metric space, if each of the sets U_i has diameter at most δ for some $\delta > 0$, then the collection is set to be a δ-*cover* of X. Furthermore, if each of the sets U_i is an open set, then we call the collection an *open cover* of X.

Example 1.12 Take the interval (0,2). This set could be covered by two open balls: The intervals (0,$\frac{3}{2}$) and (1,2).

In $\mathbb{R}^2$, take the circle centered at (0,0) with radius 2. Take the 4 circles of radius 1 centered around each of the points (2,0), (0,2), (-2,0), and (0,-2). These 4 circles form a cover of the original circle.

□

Definition 1.13 Let $(X, \mathcal{T})$ be a topological space. If $A \subseteq X$, then A is *compact* if for any cover of A there is a finite subcover of open sets from X that also covers A.

Example 1.14 $\mathbb{R}$ is not a compact set. Consider the set of open intervals $\{(x, x+2) : x \in \mathbb{R}\}$. This set covers all of $\mathbb{R}$, but it does not have a finite subcover that also covers $\mathbb{R}$. Therefore $\mathbb{R}$ is not compact. □

Example 1.15 $U = \{0\} \cup \{\frac{1}{n} : n \in \mathbb{N}\}$ is a compact set, even though there are infinitely many elements. Let $\mathcal{A}$ be a cover of U. Note that any open set that contains 0 in this cover will cover an infinite number of points. Take one of these sets containing 0 to be an element of the subcover. Then since there are finitely many points left, we can pick finitely many of the covers remaining. So $\mathcal{A}$ has a finite subcover that covers U. Thus U is compact since $\mathcal{A}$ was arbitrary.

□

Definition 1.16 Let $H \subseteq X$. A subset A of a metric space X is said to be *connected* if there are no non-empty sets G and H such that $A = G \cup H$ and $\overline{H} \cap G = H \cap \overline{G} = \emptyset$. The space X is *totally disconnected* if the only non-empty connected subsets of X are sets containing a single element.

Example 1.17 The integers are a totally disconnected subset of $\mathbb{R}$ since the only connected sets of $\mathbb{Z}$ are the integer singletons.

$\square$

Functions that are not very well behaved may have differing limits, depending on how you approach a point. This can happen if the function is discontinous or if the function fluctuates very rapidly. Suppose we want to find the limit as $x \to 0$.

Definition 1.18 The *upper limit* is defined as

$$\overline{\lim_{x\to 0}} f(x) = \lim_{r\to 0}(\sup\{f(x) : 0 < x < r\}).$$

Likewise, the *lower limit* is defined as

$$\underline{\lim_{x\to 0}}\, f(x) = \lim_{r\to 0}(\inf\{f(x) : 0 < x < r\}).$$

If the upper limit and the lower limit agree, we may use our usual method of writing the limit: $\lim\limits_{x\to 0} f(x)$.

There are two cases to watch out for when determining whether a limit exists. Either the function will have a jump discontinuity, or the function will have an asymptote.

Example 1.19 Let

$$f(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases}.$$

Clearly, $\lim\limits_{x\to 0^+} f(x) = 1$ and $\lim\limits_{x\to 0^-} f(x) = -1$. So the upper limit, $\overline{\lim\limits_{x\to 0}} f(x) = 1$, since it is the supremum of the set $\{1, -1\}$, and for the lower limit we have $\underline{\lim\limits_{x\to 0}}\, f(x) = -1$. Since the upper and lower limits do not agree, the limit $\lim_{x\to 0} f(x)$ does not exist.

$\square$

Example 1.20 Let $f(x) = \sin\left(\frac{1}{x}\right)$. Here, as $x \to 0$, the function oscillates rapidly between 1 and -1. So the upper limit as $x \to 0$ is 1, and the lower limit is -1. Again, since the upper and lower limits are different values, the $\lim_{x\to 0} f(x)$ does not exist.

$\square$

Definition 1.21 There are different ways of measuring the size of a subset of $\mathbb{R}^n$. A measure will assign a positive number, 0, or ∞ to the set. A *measure*, μ, on $\mathbb{R}^n$ must satisfy the following three requirements:

1. $\mu(\emptyset) = 0$
2. $\mu(A) \leq \mu(B)$ if $A \subseteq B$
3. If $A_1, A_2, ...$ is a countable or finite sequence of sets, then
$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu(A_i).$$

For the purposes of this thesis, we will only talk about Hausdorff measure (to be mentioned later) and Lebesgue measure. Let $A \subset \mathbb{R}$ be an interval (a, b) (this interval could be closed or half open, and the measure would be the same). Then for Lebesgue measure, $\mu(A) = |b - a|$. This is just the length of the interval. Note that for a single point, a, we can think of this as the interval $[a, a]$. So the measure of this point is $\mu(a) = |a - a| = 0$. Hence by 3 above, if A is a countable or a finite set of points, then A has Lebesgue measure 0. A countable union of countable sets is countable, so the measure of such a union would also be measure 0. Lebesgue measure can also be applied to $\mathbb{R}^n$. Let $A = \{x_1, x_2, , ..., x_n) \in \mathbb{R}^n : a_i \leq x_i \leq b_i\}$. Then the n-dimensional Lebesgue measure is the *n-dimensional volume.* If S is a parallelepiped (or solid box) in $\mathbb{R}^n$, then the *n-dimensional* volume is defined as $\text{vol}^n = (b_1 - a_1)(b_2 - a_2)...(b_n - a_n)$. So for $n = 2$, the Lebesgue measure would just be the area of the rectangle and for $n = 3$, it would just be the 3-dimensional volume. Now if A is a more abstract set, the Lebesgue measure is defined to be the following:

$$\mathcal{L}^n(A) = \inf \left\{ \sum_{i=1}^{\infty} \text{vol}^n A_i : A \subset \bigcup_{i=1}^{\infty} A_i \text{ and where each } A_i \text{ are parallelepipeds} \right\}.$$

Again, if A is a "nice" geometric shape, like a line segment, a triangle, or a cube, the Lebesgue measure is just the length, area, or volume respectively.

The following chapter outlines the basics of chaos theory. This thesis is basically an extension of a thesis by Carrie Rose Gibson. Chapter 2 outlines the work she did, with the proofs of the theorems omitted and much fewer examples. In chapter 3 we will explore what it means to be a fractal, what they look like, and some differing defintions. Chapter 4 explains the calculation of different methods of fractal dimension and many examples that go with them. As stated earlier, chapter 5 gives us a way of constructing a fractal using difference equations and gives us another way of calculating the dimension. Chapter 6 deals with special kinds of fractals called *Julia sets*. A Julia set is a graph of a complex *difference equation*. Though not all such graphs are fractals, the ones we will be exploring are. All theorems and proofs come originally from either [4] or [5] unless otherwise noted.

Chapter 2

Chaos Theory

Ed Lorenz discovered chaos through weather systems in 1960. For his weather analysis, Lorenz wrote a program using twelve equations on his computer, describing variables such as pressure, temperature, and wind speed. One day as he was researching, Lorenz ran the program but wanted to look at a particular section of the output more thoroughly. So he ran the program again, but to save time he started from the middle of the code instead of the beginning. Also, when he input values, he rounded them to three decimal places, while the original had six decimal places. He ran the program and left the room to get some coffee. When he came back an hour later, he was shocked. The new printout looked nothing like the original. The two graphs looked close at first, but as they went on, the values grew further and further apart. At first he thought there was something wrong with his machine, but then he noticed his input values. Even though the numbers differed by only a few tenthousanths, this is what caused the output to be so drastically different. This is called "sensitive dependence on initial conditions." A very small change in the input can cause a very large change in the output. Any chaotic system will have this sensitive dependence. For the case of weather, Lorenz called this sensitive dependence “the Butterfly Effect,” and was the first to use the term. He used it as the title of a presentation, and the name stuck and became popular, even in disciplines other than mathematics. Basically, the Butterfly Effect asks the question, “Can a butterfly flapping its wings in Brazil cause a tornado in Texas a few weeks later?” The butterfly flaps its wings, which affects the air by such a small amount that no technology would pick it up. So in predicting the weather, the program would predict that no tornado will occur in Texas because it did not pick up the butterfly’s disturbance. However, in reality, the small effect of the butterfly’s wings set in motion small changes in the weather, which over a large amount of time creates a large changes. So a tornado occurs because the butterfly flapped its wings, even though the computer’s predicted otherwise.

Lorenz was not the only scientist studying chaotic systems. Physicists were trying to solve problems arising in dynamical systems. It seemed impossible for them to predict the trajectory of a projectile traveling through an area of turbulence, whether in air or in water. In fact, chaos was mostly discovered and developed by physicists rather than mathematicians. As time progressed, it could be seen by those who studied other disciplines. In economics, the seemingly random order of the stock market was discovered to have sensitive dependence on initial conditions. Even biologists studying populations of species saw chaos in their work. So throughout recent years, chaotic behavior has been discovered in several disciplines. Chaos was seen in the past as well, but because of the strong calculation that computers can do, it is now able to be researched more clearly. Yet despite its recent research and exploration, chaos still has no set definition. An important characteristic of a chaotic system, however, is a system that seems to act randomly, but is actually governed by a set of rules.

A particularly intriguing question throughout the centuries has been, "What causes the Great Red Spot of Jupiter?" As technology has advanced and telescopes have become more distinguished, several scientists have come up with different answers. The early astronomers thought it was lava flow from a great volcano. Others thought a moon might be forming from the surface of Jupiter. But Voyager saw something different. The pictures of voyager showed that the surface of Jupiter was gaseous, thus the Great Red Spot was surrounded by storms swirling around it. However, the Great Red Spot never changed. Every once in a while it will move slightly, but overall it remains constant. This demonstrates that in the same chaotic system there will be both order and disorder. Amidst the disordered storms, the Great Red Spot stays practically the same. Essentially, the idea of chaos theory is that many chaotic systems will be governed by a certain set of rules, even though disorder is present.

Mathematically, chaotic systems are frequently defined using difference equations (not to be confused with differential equations). A difference equation is merely a function that is defined recursively and certain points in them can have special properties. In order to describe them fully, we will need a definition.

Definition 2.1 Let $f : D \to D$ be a function from some domain D to itself and let $x \in D$. Then the *orbit* of x is the set $\{x, f(x), f^2(x), f^3(x), ...\}$, where

$$f^2(x) = f(f(x)), f^3(x) = f(f(f(x))), \text{etc.}$$

Consider the function called the tent map:

$$T(x) = \begin{cases} 2x & \text{if } 0 \le x \le \frac{1}{2} \\ 2(1-x) & \text{if } \frac{1}{2} < x \le 1 \end{cases}.$$

Let $x = \frac{1}{4}$. Let's calculate the orbit of x. $T(\frac{1}{4}) = \frac{1}{2}$, so $\frac{1}{2}$ is in the orbit of $\frac{1}{4}$. Then $T(\frac{1}{2}) = 1$, so 1 is in the orbit of $\frac{1}{4}$. $T(1) = 0$, so 0 is in the orbit. We are now done since $T(0) = 0$. So the orbit of x is the set $\{\frac{1}{4}, \frac{1}{2}, 1, 0\}$. Observe that 0 is a *fixed point.*

Definition 2.2 Let f be a difference equation and let $x^* \in D$. Then x^* is called a *fixed point* if $f(x^*) = x^*$. The point $\bar{x} \in D$ is a *k-periodic point* of f if $f^k(\bar{x}) = \bar{x}$ for some $k \in \mathbb{Z}^+$ (i.e. $\bar{x}$ a fixed point of f^k).

So according to this definition, the orbit of x^* is just $\{x^*\}$. Note also that fixed points of a real valued function f always lie on the intersection of the graph of the function and the line $y = x$ since the fixed point must satisfy $f(x^*) = x^*$.

Periodic points will bounce back and forth between different values and will stay in this cycle forever. We call the whole set of points that is hit a cycle.

Definition 2.3 A *k-cycle* is the orbit of a k-periodic point, denoted

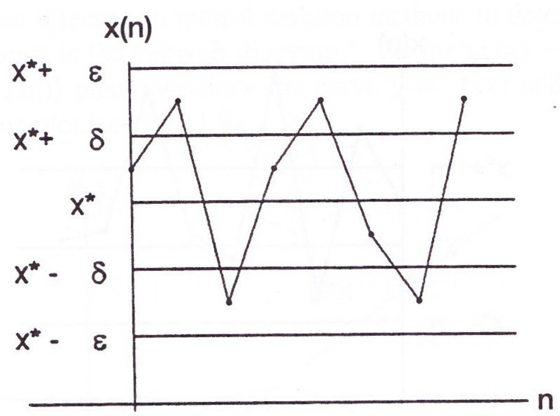
$$O(\bar{x}) = \{\bar{x}, f(\bar{x}), f^2(\bar{x}), ..., f^{k-1}(\bar{x})\}.$$

As an example for periodic points and k-cycles, let's look at the tent map again. Let $x = \frac{2}{5}$. Then $T(\frac{2}{5}) = \frac{4}{5}$ and $T(\frac{4}{5}) = \frac{2}{5}$. So since $\frac{2}{5}$ gets mapped to itself after two iterations, we say that $\frac{2}{5}$ is a periodic point of order 2, as is $\frac{4}{5}$. Fixed points and periodic points can still be classified further. These points can be stable, unstable, hyperbolic, or nonhyperbolic.

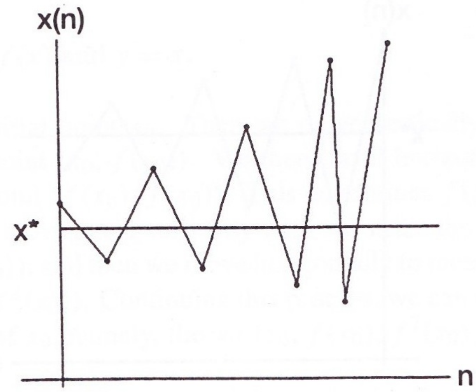
Definition 2.4 For a differentiable function f, a fixed point x^* is *hyperbolic* if $|f'(x^*)| \neq 1$. Otherwise, it is *nonhyperbolic.*

We need to know whether a point is hyperbolic or not since this will determine how to calculate the point's stability.

Definition 2.5 Let $f : D \to D$ be a mapping and x^* be a fixed (or periodic) point. Then x^* is *stable* if for any $\epsilon > 0$ there exists $\delta > 0$ such that if $|x_0 - x^*| < \delta$ then $|f^n(x_0) - x^*| < \epsilon$ for all $n \in \mathbb{Z}^+$ and all $x_0 \in D$. If a the point x^* is not stable, then it is said to be *unstable.* See the following diagrams.

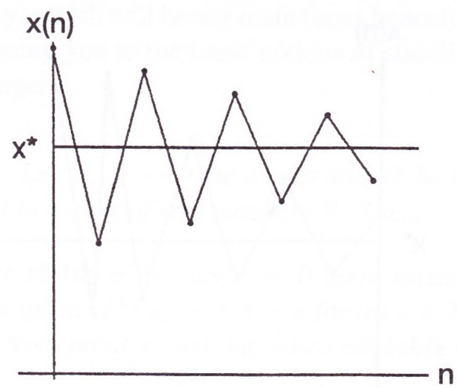


Stable fixed point x^*.



Unstable fixed point x^*.

Definition 2.6 A point (or orbit) is *attracting* if nearby points become close to the point (or orbit) under iteration of f. Let x^* be a stable fixed (or periodic) point. Then x^* is *asymptotically stable* if x^* is attracting (see diagram)[4].



Asymptotically stable fixed point x^*.

Stability of a point can also be thought of intuitively. Think of a marble at the bottom of a bowl. If the bowl is disturbed, the marble will roll around for a bit, but will eventually settle back down at the bottom of the bowl. The bottom of the bowl can be thought of as a "stable fixed point" since the marble returns to its original state. Now think of a pen standing up on its point. Any sort of disturbance will cause it to topple, and it will not return to its original state. The original position of the pen can be thought of as an "unstable fixed point" since any deviation from the original state will cause it to go somewhere else. In order to determine whether a fixed point is stable or unstable, we will use the following propositions.

Proposition 2.7 Let x^* be a hyperbolic fixed point of a function f, where f is continuously differentiable at x^*. Then

1. If $|f'(x^*)| < 1$, then x^* is stable

2. If $|f'(x^*)| > 1$, then x^* is unstable

Proposition 2.8 Let x^* be a fixed point of a map f such that $|f'(x^*)| = 1$ (so it's nonhyperbolic). Suppose $f'''(x^*) \neq 0$ and is continuous. Then

1. If $f''(x^*) \neq 0$, then x^* is stable.

2. If $f''(x^*) = 0$ and $f'''(x^*) > 0$, then x^* is unstable.

3. If $f''(x^*) = 0$ and $f'''(x^*) < 0$, then x^* is stable.

The proofs of the propositions in this chapter are left out since they are proved in another thesis, *Chaos Theory* by Carrie Gibson. If the fixed point in question is nonhyperbolic, that is, if $|f'(x^*)| = 1$, then we will need a different way of calculating its stability. In order to do this, we will need to compute the Schwartzian derivative.

Definition 2.9 The *Schwartian derivative*, Sf of a function f is defined by

$$Sf(x) = \frac{f'''(x)}{f'(x)} - \frac{3}{2}\left[\frac{f''(x)}{f'(x)}\right]^2.$$

Proposition 2.10 Let x^* be a fixed point of a function f, where $f'(x^*) = -1$. If $f'''(x^*)$ is continuous, then

1. If $\mathrm{S}f(x^*) < 0$, then x^* is stable.

2. If $\mathrm{S}f(x^*) > 0$, then x^* is unstable.

Proposition 2.11 Let $\bar{x}$ be a k-periodic point of a function f (so $\bar{x}$ is a fixed point of $f^k(x)$). Then $\bar{x}$ is a stable periodic point of f if

$$|f'(\bar{x})f'(f(\bar{x}))...f'(f^{k-1}(\bar{x}))| < 1.$$

In the following chapter there are examples of all of these definitions and propositions at work. We will be seeing these definitions in a setting where a fractal is generated. Fractals emerge from chaos in strange way.

An *attractor* M is another name for an asymptotically stable fixed point or cycle. An attractor will attract a certain set of values. Let x^* be an attractor of a continuous map $f : D \to D$. Then the *basin of attraction* $A(x^*)$ of x^* is defined as the maximal interval I containing x^* such that if $x \in I$, then $f^n(x) \to x^*$ as $n \to \infty$. A *repellor* is another name for an unstable fixed point. As the name suggests, a repellor will repel a certain set of values.

Example 2.12 For a simple example, take the function $f(x) = x^2$. The fixed points of f are $x_1 = 0$ and $x_2 = 1$ since these are the solutions to $f(x) = x$. Note that $x_1 = 0$ is asymptotically stable, so it is an attractor. The basin of attraction for x_1 is the interval $(-1, 1)$ since as points in this interval are iterated, they get closer and closer to 0. Also note that $x_2 = 1$ is a repellor. Values of x greater than 1 (in absolute value) are iterated off to ∞ and values less than 1 (in absolute value) are iterated towards 0. The only values that go to 1 are -1 and 1.

□

All these definitions are defined for difference equations, which is a way of describing chaos. There is no set definition of chaos, but many mathematicians have attempted to create one and some stuck. One definition of chaos relies on some properties of a difference equation. In order to define chaos in this way, we will need three more definitions.

Definition 2.13 Let $H \subseteq X$. Then H is said to be *dense* in X if $\overline{H} = X$.

Let's briefly look at an example of a dense set. Consider $\mathbb{Q}$, the set of rational numbers. Then $\mathbb{Q}$ is dense in $\mathbb{R}$ since an $\epsilon-$ball centered at any point in $\mathbb{R}$ will contain a rational number. Thus, the closure of $\mathbb{Q}$ is $\mathbb{R}$.

Definition 2.14 Let f be a map on a metric space (X, d). Then f is said to be *transitive* if for any pair of non-empty, open sets U and V, such that $U, V \subseteq X$ there exists $k \in \mathbb{Z}^+$ such that $f^k(U) \cap V \neq \emptyset$.

Definition 2.15 The map f on a metric space X is said to possess *sensitive dependence on initial conditions* if there exists $\epsilon > 0$ such that for any $x_0 \in X$ and for any open set $U \subseteq X$ containing x_0, there exists $y_0 \in U$ and $k \in \mathbb{Z}^+$ such that $d\left(f^k(x_0), f^k(y_0)\right) > \epsilon$.

This definition says that if we choose a point $x_o \in U \subseteq X$, then there is an $\epsilon > 0$ and a $y_0 \in U$ such that after some number of iterations, x_0 and y_0 will be a distance of more than ϵ apart.

Definition 2.16 A map $f : X \to X$ is said to be *chaotic* if:

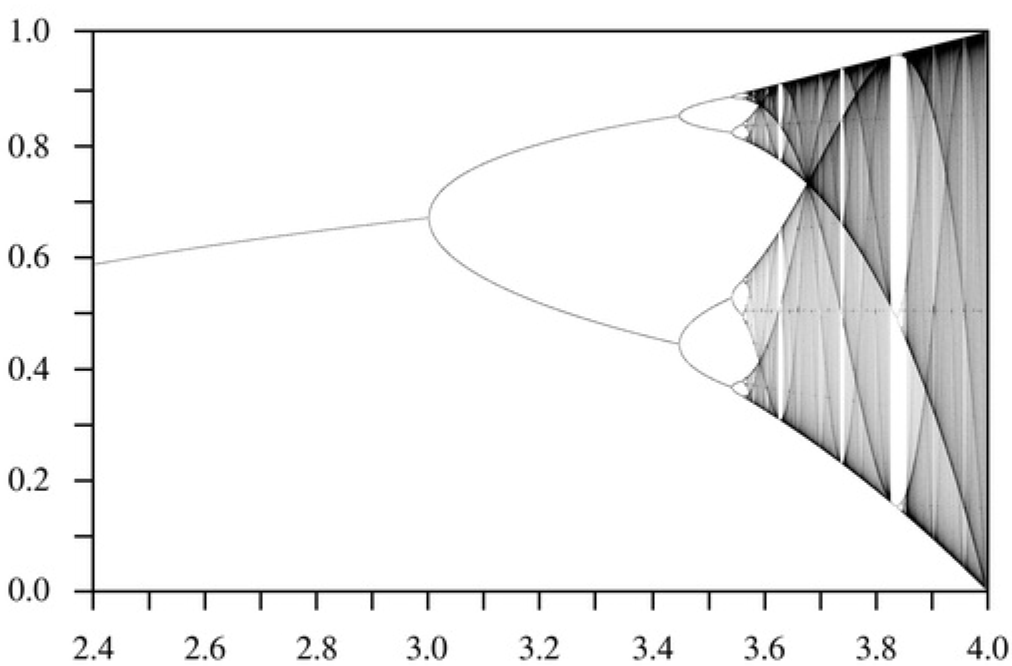
1. f is transitive.

2. The set of periodic points P of f is dense in X.

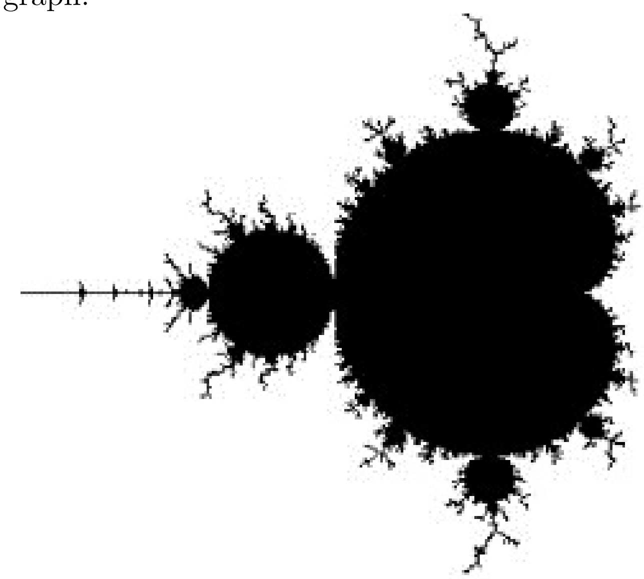
3. f has sensitive dependence on initial conditions.

Furthermore, Li and Yorke proved that a system is chaotic if there is a cycle with period 3 in the difference equation. From this, they wrote their famous article "Period three implies chaos"[9]. In this article they proved that if a map has a point with period 3, then there will be points for any period. Let's look at an example of this.

Consider the logistic map: $f(x) = \mu x(1-x)$, where the parameter μ is a positive constant. This mapping is used for population modeling. See the graph of this mapping (this graph is called a bifurcation diagram). Note that there is an orbit of order 3 at about $\mu = 3.8$. This implies that the logistic function is chaotic. On the graph, the chaotic sections are the fuzzy parts of the graph where we can't tell what the periodicity of the points are [3].



Let's briefly look at a chaos theory at work. Consider the Mandelbrot set. The Mandelbrot set is a Julia set, which we will explore more thoroughly in chapter 6. It is a famous fractal and is defined by the complex difference equation $f(z) = z^2 + c$, where $z \in \mathbb{C}$ and c is a complex constant. For the graph of this function, color each point on the complex plane according to whether the point is bounded under iteration or escapes to infinity. We get the following graph:



The set of black points forms the Mandelbrot set. These are the points that do not iterate to infinity, they are bounded. The other points are usually colored according to how quickly they escape to infinity, but in the above image they are just colored white. Now consider the boundary of the Mandelbrot set. It just so happens that a point on the boundary will be mapped to another point on the boundary. Now let's look at a point close to the edge. Here we can see sensitive dependence on initial conditions. Say we have one point on the boundary and one point very close outside the boundary, on the outside of the Mandelbrot set (so the first point is black, but the second point is white). They will end up arbitrarily far apart since the first point will stay bounded and the other will iterate to infinity.

The Mandelbrot set is just one of many great images of chaos. We call these wonderful images fractals and will study these in more detail in the following chapters.

Chapter 3

Fractals

As the theory chaos was developed and understood in the 1960s, difference equations were giving rise to some curious graphical images. In the 1970s, these images were named "fractals," creating a new area of intrigue for scientists and mathematicians. Fractals can be constructed in a number of different ways, but what are they exactly? Some say that it is an image that is self-similar. A self-similar image is a picture that repeats itself at a decreasing scale. So inside of the big picture, there are copies of itself that get smaller and smaller. Zooming in on a certain part of a fracal at a certain ratio will yield the same picture again. This could be done indefinitely since we could keep zooming in on the same section.

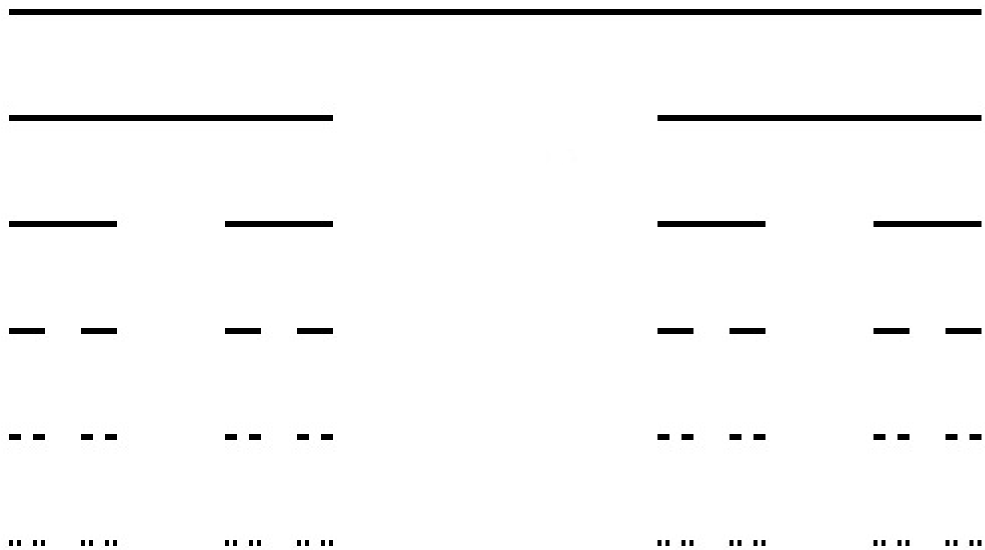
This definition of self-similarity alone is too broad since a line segment is considered self-similar, yet a line segment is not a fractal. It is difficult to define exactly what a fractal is, thus there is no set definition of a fractal because no definition is all encompassing. Note also that chaos has no set definition. Since fractals are the result of chaos, it makes sense that we can't create a set definition for fractals either. Despite the difficulty of defining a fractal, we can at least describe some properties that fractals have. A fractal should have the following properties:

1. They are infinitely detailed, that is, they have detail at arbitrarily small scales.
2. They cannot be described in the classical terms of geometry because their shapes are too irregular.
3. They usually have some degree of self-similarity (though not always).
4. The fractal dimension is usually larger than the topological dimension, which we will talk about later.
5. Usually, a fractal can be described in a recursive manner or some other simple way.

Though fractals were first studied in the 1970s, some fractals have been known since before then. Mathematicians have discovered and created them throughout history without realizing it. The reason that it took so long to study them in depth is that their detailed analysis relies on technology. With the ideas of chaos theory in mind, fractals can be generated using difference equations. In order to get a good picture of what the orbits are doing, many iterations for many points need to be calculated. The discovery of these images came about from inputting values into a computer program. The program could then create a graphical representation of the orbits used much more quickly than a human could. Since computers started becoming more popular in the 1960s, this is the reason that chaos theory became more developed during this time. As mathematicians studied chaos theory by inputting values on a computer and graphing the results, images of *strange attractors* and fractals were discovered.

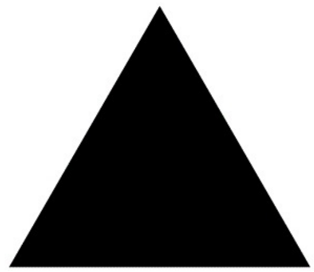
In order to describe exactly what a fractal is, we will demonstrate their properties with some examples and constructions. When constructing a fractal, usually we start with some base set and apply some transformation on the set. The base set E_0 is called the initiator. The first transformation on the set, E_1, is called the generator, and the nth transformation of the set is called E_n. The limiting set $\lim\limits_{n\to\infty} E_n$ is the fractal F. Let's look at some examples.

Example 3.1 One of the classic examples of a fractal is the Cantor set. It was explored first by Georg Cantor in the late 1800s - before the term "fractal" had even been coined. For the purpose of this example, in order to construct the Cantor set, start off with a line segment of length 1. This line segment is E_0, the initiator. Now remove the middle third of this line, so we are left with 2 segments each of length $\frac{1}{3}$, and call this set E_1, the generator. Again, remove the middle thirds of each of the remaining segments. So we are left with 4 segments each of length $\frac{1}{9}$, and call this set of segments E_2. Keep doing this forever. Then E_n will have 2^n segments each of length 3^{-n}. Taking the limit as $n \to \infty$ results in the Cantor set. (See the figure). This set is self similar and is totally disconnected.

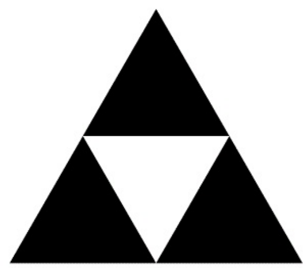


□

Example 3.2 Another well known fractal is the Sierpinski Triangle (also known as the Sierpinski gasket). The Polish mathemtician Waclaw Sierpinski created the Sierpinski triangle in 1906. The initiator for the Sierpinski triangle is a solid equilateral triangle with sides of length 1. Now connect the midpoints of each of the sides to create 4 smaller triangles. Remove the middle triangle, leaving three triangles each with one forth the area of the original. This is the generator E_1. Now for each of the three triangles, do the same thing and remove the middle triangle in each, leaving 9 triangles. Then E_n will have 3^n triangles with area 4^{-n}. Taking the limit as $n \to \infty$ gives us the Sierpinski triangle.

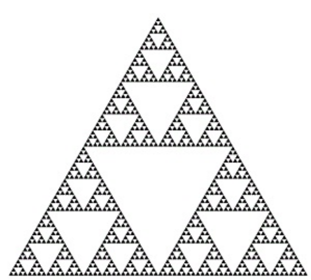


Stage E_0



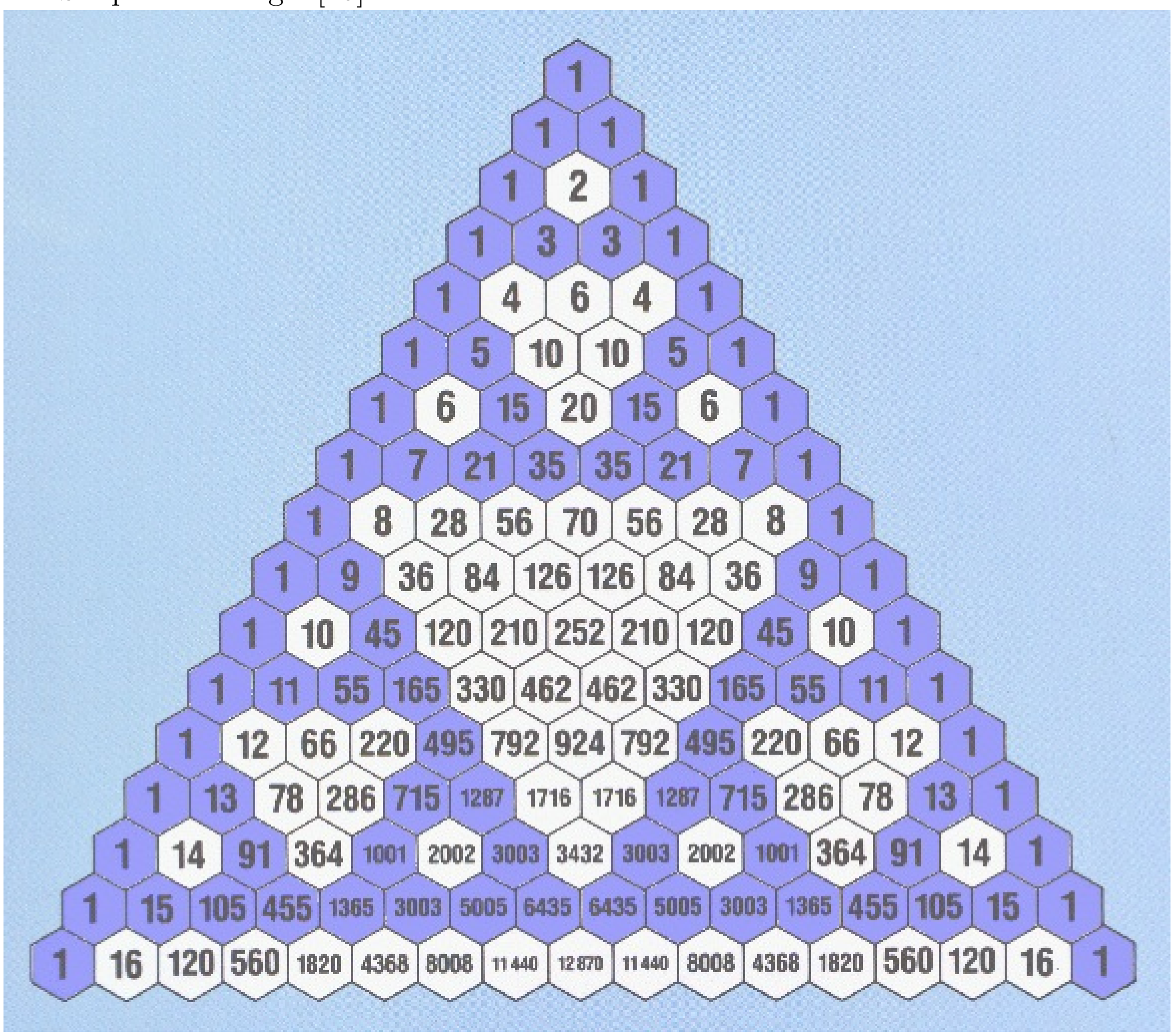
Stage E_1

Stage E_2



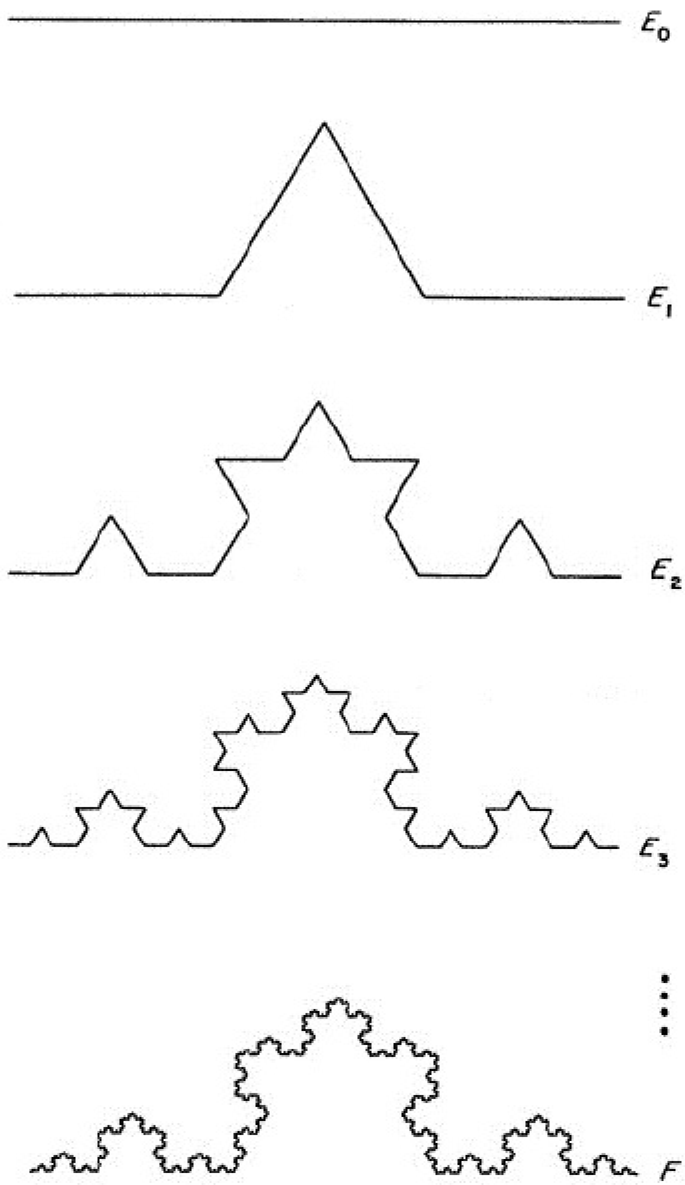
F

As an interesting side note, the Sierpinski triangle looks like Pascal's triangle modulo 2. If we look at Pascal's triangle mod 2, the 1's create the outline of Sierpinski's triangle and the 0's create the solid triangles. In the following figure, the odd numbers are shaded, while the even numbers are not. So the even numbers are 0 mod 2 and the odd numbers are 1 mod 2. The white region can be compared to the regions that are removed when constructing the Sierpinski triangle [15].

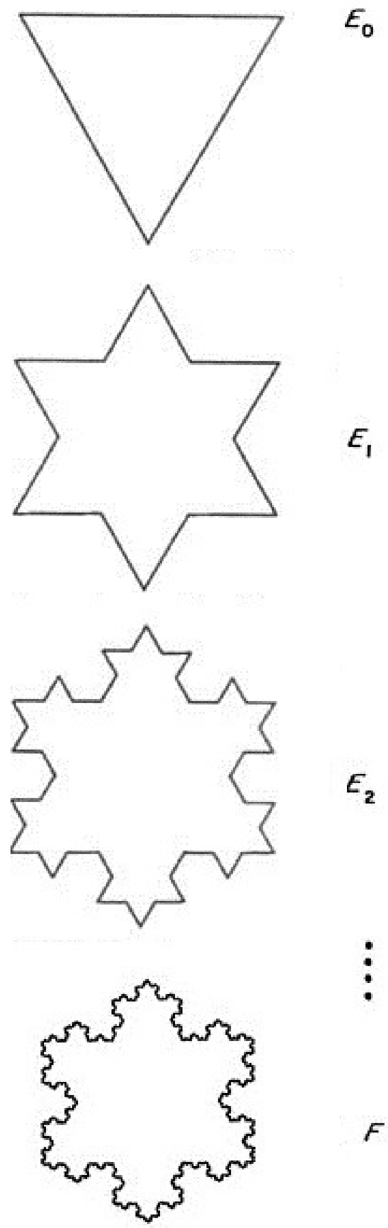


□

Example 3.3 The Koch curve was founded by Helge von Koch in 1904. The construction of the Koch curve is similar to that of the Cantor set, except it does not live in just one dimension. The initiator of the Koch curve is a straight line, say of length 1. For the generator, remove the middle third of the line and replace it with an equilateral triangle with sides of length $\frac{1}{3}$, but with no base. Now remove the middle thirds of each segment and replace those with equilateral triangles and so on. It is interesting to note that the length of the initiator E_1 is $\frac{4}{3}$, E_2 has length $\frac{16}{9} = \left(\frac{4}{3}\right)^2$, and so on. So for the nth iteration, the length of the curve is $\left(\frac{4}{3}\right)^n$. Then the fractal itself, as $n \to \infty$ has infinite length, an important feature of fractals that live in more than the first dimension. See the figure on the following page [16].



An extension of the Koch curve is the Koch Snowflake. The initiator for the Koch Snowflake is an equilateral triangle instead of a straight line segment. If we apply the same transformation to each side of the triangle as we did to the Koch curve, we end up with the Koch snowflake. Since each side of the figure has infinite length, the resulting figure will have infinite perimeter, yet finite area. See the figure.



□

This notion of a fractal having infinite perimeter is another way of saying that the figure is infinitely detailed. If we zoom in on a part of the figure, it will always have some sort of detail, or jaggedness, no matter how far we zoom in. This is what Mandelbrot was explaining when he asked the question, “How long is the coastline of Britain?”

We have looked at some examples of fractals that we can construct geometrically. As stated earlier, however, these are not the only fractals that exist. They can also arise from chaos. Let’s examine a more complicated example of a fractal: a bifurcation diagram. A bifurcation diagram is essentially the graph of a chaotic function. So let’s consider the difference equation of the logistic map, defined as $f_\mu(x) = \mu x(1-x)$, where $x \in [0,1]$ and $\mu \in (0,4]$. Note also that $f'(x) = \mu - 2\mu x$. The goal is to determine how many periodic points a certain μ will have, and to determine the stability of each periodic point. To find the fixed points, set the equation equal to x and solve. So, solve $\mu x(1-x) = x$ for x. This gives us two fixed points: $x_1^* = 0$ and $x_2^* = \frac{\mu-1}{\mu}$. So no matter what value for μ we choose these will always be the two fixed points, but we care about stability. The point $x_1^* = 0$ is only stable for $0 < \mu < 1$ by evaluating the derivative f' at $x = 0$. The point x_2^* is stable for $1 < \mu < 3$. But what about the periodic points?

All k-periodic points are fixed points of the function $f_\mu^k(x)$. We can solve for 2^n-cycles by solving $f_\mu^{2^n}(x) = x$ for x. So to find the 2-cycles, solve the equation $f_\mu^2(x) = x$ and to find the 2^2cycles, solve $f_\mu^4(x) = x$, and so on. When looking for cycles for the powers of 2, it is called period doubling. Let’s find the 2-periodic points by solving for x in $f_\mu^2(x) = x$. Recall that $f_\mu^2(x) = f_m u(f_\mu(x)) = \mu(\mu x(1-x)(1-\mu(\mu x(1-x)))$. Set this equal to x to get $\mu(\mu x(1-x)(1-\mu(\mu x(1-x))) - x = 0$. After we expand and simplify, we can factor out x and $(x - \frac{\mu-1}{\mu})$ since these are fixed points. So we obtain $x(x-\frac{\mu-1}{\mu})(\mu^2x^2-(\mu^2+\mu)x+(\mu+1)) = 0$. Using the quadratic formula yields two more periodic points:

$$\bar{x}_1 = \frac{\mu+1+\sqrt{(\mu+1)(\mu-3)}}{2\mu}$$

and

$$\bar{x}_2 = \frac{\mu+1-\sqrt{(\mu+1)(\mu-3)}}{2\mu}.$$

So there are two periodic points of order 2 (remember that the other two solutions were fixed points, or order 1). Note that these values for x only make sense for $\mu > 3$. Now what is the stability of these points? Since $f(\bar{x}_1) = \bar{x}_2$, from Proposition 2.11, we know that a cycle is stable if

$$|f'_\mu(\bar{x}_1)f'_\mu(f_\mu(\bar{x}_1))| = |f'_\mu(\bar{x}_1)f'_\mu(\bar{x}_2)| < 1.$$

So

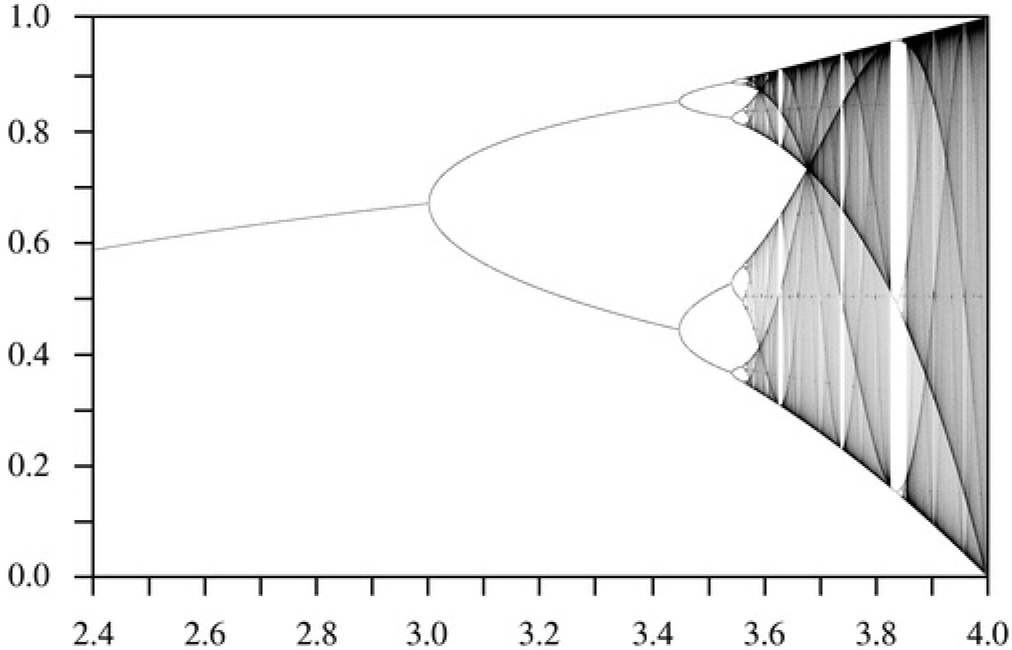
$$-1 < (\mu - 2\mu\bar{x}_1)(\mu - 2\mu\bar{x}_2) < 1$$
$$-1 < \mu^2(1-2\bar{x}_1)(1-2\bar{x}_2) < 1$$
$$-1 < \quad \mu^2(1 - \tfrac{\mu+1+\sqrt{(\mu+1)(\mu-3)}}{\mu})(1 - \tfrac{\mu+1-\sqrt{(\mu+1)(\mu-3)}}{\mu}) \quad < 1$$
$$-1 < \mu^2 - 2\mu - 4 < 1$$

Now separate the inequality:

$$0 < \mu^2 - 2\mu - 3 < 2 \text{ or} - 2 < \mu^2 - 2\mu - 5 < 0$$

Solving the first one yields $\mu < -1$ or $\mu > 3$. But $\mu > 0$, so only take the solution $\mu > 3$. Solving the second inequality yields $\mu < 1+\sqrt{6}$ or $\mu < 1-\sqrt{6}$. Again since $\mu > 0$, we only regard $\mu < 1+\sqrt{6}$. Thus μ is stable for $3 < \mu < 1+\sqrt{6}$. Now we need to determine the stability of $\mu = 1+\sqrt{6}$, but $f'_\mu(\bar{x}_1)f'_\mu(\bar{x}_2) = -1$. Therefore, we will need to use the Schwartzian derivative on f^2_μ. The computations are left out since they are so messy. In the end, however, we end up with the fact that $S(f(\bar{x}_1)) < 0$ and $S(f(\bar{x}_2)) < 0$. This indicates that $\bar{x}_1$ and $\bar{x}_2$ are indeed asymptotically stable.

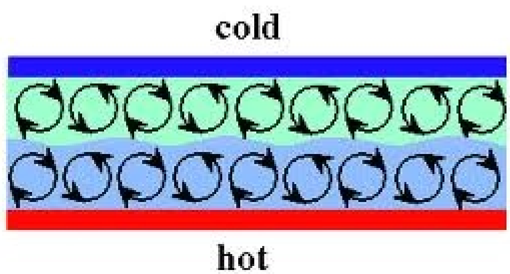
Now to solve for the 4-cycles, we would have to solve $f^4_\mu(x)$, which is an order 12 equation. With some numerical analysis, or by looking at the graph produced, we see that we get a stable 4-cycle for $1+\sqrt{6} < \mu \leq 3.54409$. Let μ_n be the value of μ corresponding to the 2^n-cycle. If we kept doing this, we would see that for a 2^n-cycle, $\mu_n \to 3.570$ as $n \to \infty$. However, there are other cycles passed 3.570. See the diagram on the next page. This diagram that describes the behavior of the difference equation is called a bifurcation diagram. The x-axis represents the values for μ, and the y-axis reperesents the values of x that are periodic points. Earlier, we only calculated 2^n-cycles. There could be other cycles, such as 3-cycles or 6-cycles that we haven't computed yet. Notice on the diagram that at about $\mu = 3.829$ there is a 3-cycle. What does this indicate? Remember that period 3 implies chaos. The chaos in the diagram are the fuzzy grey areas. The x-values in these areas bounce around all over the place. If we think about sensitive dependence on initial conditions, then if we start with some values very close together, say x_α and x_β, then at some point in the iterations, $f^n(x_\alpha)$ and $f^n(x_\beta)$ will be arbitrarily far apart for some natural number n. The chaotic portions of this diagram are fractals, however, these fractals do not seem to have any self-similarity, yet they are still infinitely detailed.



Mitchell Feigenbaum was a physicist who studied turbulence - a dynamical system. It just so happens that many dynamical systems are chaotic systems, so they have a fractal structure. He was looking at the bifurcations of difference equations - specifically period doubling functions - studying when the bifurcation changed from periodic to chaotic. As he was examining his numerous calculations, he discovered something fascinating. Had he been using a fast computer like we have today, he could have completely missed this, but since he had to do many of his calculations by hand, he was able to see this pattern. He discovered that the rate at which a function converges to chaos is always the same. The ratio of the interval between bifurcation points approaches this constant. Using the notation we used for describing the values for μ in the logistic equation, this constant is $\delta = \lim_{n\to\infty} \dfrac{\mu_n - \mu_{n-1}}{\mu_{n+1} - \mu_n} \approx 4.6692016090$, but is most likely irrational. To visualize this ratio, compare the ratio between the intervals with chaos of the logistic function with the areas of attracting values of the Mandelbrot set. In the figure below, it is seen that these intervals are the same. This is meant to demonstrate that Feigenbaum's number is indeed a constant, not relying on the chaotic function being used [11].

Another kind of fractal that arises from chaos is a strange attractor. Recall that an attractor is a point or set of points that a function approaches. Sometimes a set of iterated points will seem to be plotted randomly, but will stay close to a certain set. This set is called a *strange attractor*. The chaotic pieces of the bifurcation diagram for the logistic function above can be thought of as a strange attractor. However, the more well known strange attractors usually have some interesting shape that defines them.

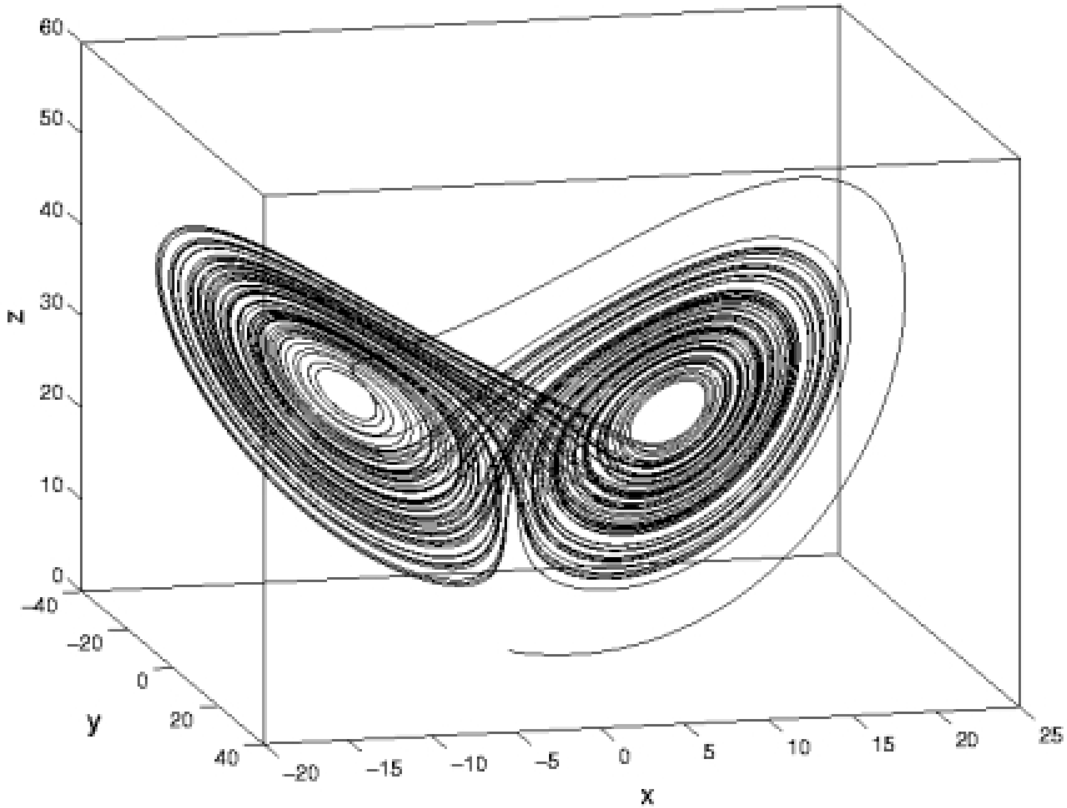
A special kind of strange attractor is the Lorenz attractor. Lorenz the meteorologist discovered this strange attractor while studying the flow of a fluid trapped between two plates with different temperatures (like the figure). This figure is a model of the Raleigh-Benard convection experiment.



To model the behavior of the fluid, he used the three-dimensional differential equations

$$\begin{aligned}
\dot{x} &= -\sigma x + \sigma y \\
\dot{y} &= -xz + rx - y \\
\dot{z} &= xy - bz
\end{aligned}$$

where σ, r, and b are parameters. In these equations, σ is called the Prandtl number, r is called the Raleigh number, and b is related to the height of the fluid layer. Also, "x is proportional to the circulatory fluid flow velocity, y is proportional to the temperature differenece between the ascending and descending fluid elements, and z is proportional to the distortion of the vertical temperature profile from its equilibrium."[4] If the fluid is flowing clockwise, $x > 0$, and if the fluid is flowing counterclockwise, $x < 0$. In his calculations, Lorenz used $\sigma = 10$, $r = 28$, and $b = \frac{8}{3}$. Plotting $x(t)$ versus $z(t)$ yielded the following butterfly shaped graph.



Unseen by many, fractals make their appearance many places in nature. Although fractals cannot exist in the purest sense of going on indefinitely, many objects come very close. In the words of Michael Barnsley: "Fractal geometry will make you see everything differently. There is danger in reading further. You risk the loss of your childhood vision of clounds, forests, galaxies, leaves, feathers, flowers, rocks, mountains, torrents of water, carpets, bricks, and much else besides. Never again will your interpretation of these things be quite the same."[4] To extend his list, there have also been discoveries of fractals in coastlines, lungs, blood vessels. A healthy heart beat is chaotic, so even that could be described using a fractal. In fact, studies in a hospital have shown that dying hearts tend to have a steady rhythm. In 1914 at the age of 28, the scientist George Mines hooked himself up to a machine that sent out electrical impulses to his heart at a constant rate. The same evening, the janitor found him dead in his laboratory. These incidents indicate that a heart beat with a steady rhythm is not only unhealthy, but also lethal.

Notice that all the examples that have physical shape are either incredibly jagged or incredibly complex. It makes sense that our lungs and blood vessels are fractal since when transferring oxygen to our bodies, our lungs need the largest possible surface area to acquire it from the air. However, the lungs are enclosed in a finite cavity in our chest. The solution? A fractal of course! Remember that in the 2-dimensional sense, a fractal has infinite perimeter and finite area. In the 3-dimensional sense, a fractal would be an object that has infinite surface area contained in a finite volume. Using this same idea, our blood vessels branch out for the same reason.

As they are becoming more well understood, fractals are beginning to be more and more used in the development of technology. They are currently being used in movie editing. For example, in Star Wars episode III Revenge of the Sith, the lava towards the end of the movie was generated using fractals. They used layers of red and orange to create a fractal that resembled lava. This is also how they create computer generated mountans and trees - starting off with an initial shape and then scaling.

Another place that fractals are used are in cell phones. The newer models of cell phones have fractal antennas. Since a fractal has supposedly infinite perimeter, these antennas are able to pick up a larger variety of wave frequencies than a standard antenna can.

Chapter 4

Fractal Dimension

Broccoli is not three-dimensional. It's not two-dimensional either. This is because broccoli is a fractal, and fractals almost always have a dimension that is not an integer. The dimension of broccoli is approximately 2.66 [10].

Now that we have seen a broad overview of what fractals are and what they look like, it is time to talk about how to measure these obscure objects. Thinking of a fractal in 2-space, it is not enough to only calculate the perimeter or area because as we have already seen, fractals could have infinite length and it may not enclose an area (think of the Koch curve for example). While Mandelbrot was studying shapes at IBM, he asked his famous question, "How long is the coastline of Britain?" Though it seems like it would not be difficult to measure the coastline of Britain, this question is much more difficult and deep than it first appears. Clearly, the edges of a coastline are not straight, but jagged. Say someone measures the coastline using just a yard stick. This person would get a smaller value than someone using a ruler because a ruler could fit in some of the cracks that a yard stick couldn't. But someone else may come along with an even smaller tool for measuring and get an even larger value. Going on like this, it seems as if the coastline may have an infinite length, though Britain has a finite area. So since the coastline has seemingly infinite length, Mandelbrot proposed a new way of measure such objects: dimension.

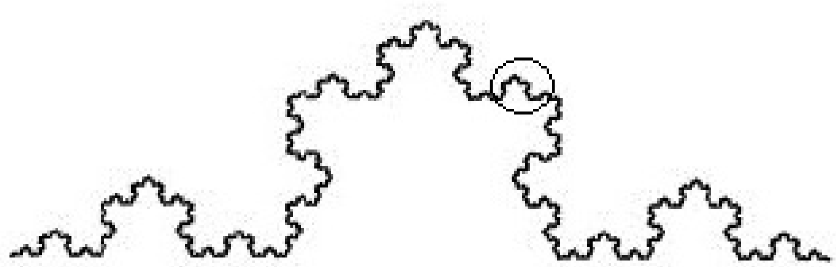
We know that a line is of dimension one, and a square is of dimension two, but what would this jagged, infinitely detailed object have for a dimension? It seems to fill up more space than line, but less space than a two dimensional square, so the dimension of this particular fractal will be something between one and two. There are two classes of dimension that we will discuss: topological dimension and fractal dimension. Mandelbrot originally defined a fractal to be a set whose fractal dimension is strictly larger than its topological dimension. But what is a fractal or topological dimension? There are many methods of calculating fractal dimension, and unfortunately different definitions will sometimes give differing values. Topological dimension is our normal way of thinking about dimension. For this thesis, when shapes such as such as circles, squares, etc. are mentioned, only the boundary of such shapes is to be considered. If we want to consider the inside of the circle or square, we will call the circle a "disc" or "solid circle" and the square a "solid square." In order to define topological dimension formally, we have the following definition.

Definition 4.1 A set $F \subseteq \mathbb{R}^n$ has *topological dimension 0* if for any point $x \in F$, an open ball around x does not intersect any point of F. F has *topological dimension* $k > 0$ if any open ball around x intersects F at a set of points of topological dimension $k - 1$ on the ball's boundary, and k is the smallest integer for which this holds.

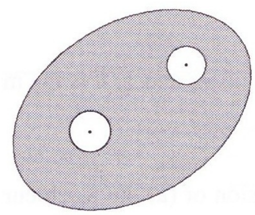
Let's look at some examples of sets with certain topological dimension.

Sets of topological dimension 0 include the Cantor set, the set of integers, and the set of rational numbers. These are sets of scattered or isolated points. Consider the set of integers. Clearly, the boundary of a sufficiently small ball centered around an integer will not hit another integer. For another example, consider the set of rational numbers. A ball centered around a rational number could have an irrational radius. In this case, the boundary of the ball "misses" a rational number, so the rational numbers have topological dimension 0.

Lines, circles, and the Koch snowflake are examples of sets with topological dimension 1. Consider a line for a moment. For any point x on the line, a circle centered around x will intersect the line at two points. These two points have topological dimension 0, so the line will have topological dimension 1. Choose a point on the Koch snowflake. The boundary of a ball centered around this point may intersect the snowflake at more than two points, but the intersection could never be uncountable[4]. So since the intersection is at a finite number of points (which has topological dimension 0), the Koch curve has topological dimension 1.



Topological dimension 2 sets include discs, solid squares, and the like. Consider a solid square with a point on its interior. The boundary of a ball around this point will intersect the square in the form of a circle, which is a 1-dimensional object. This is illustrated in the following image.



As would be expected, topological dimension 3 sets are the classic Euclidean solids such as cubes and spheres. Consider a cube and a point in the cube not on the boundary. A ball centered around this point will be in the form of a sphere. The boundary of the sphere that intersects the cube is its surface area. Surface area is 2-dimensional, thus a cube is 3-dimensional.

Notice that the topological dimension is always an integer. Fractal dimension, however, will very rarely be an integer. The number associated with the dimension of a fractal indicates how much space the set "fills up." For instance, if a fractal has fractal dimension 1.2, the set will resemble a line with enough quirks to make it fill up more space. But if the fractal has dimension 1.9, the set would more closely resemble a surface with holes. There are too many definitions of fractal dimension to discuss in this thesis, so we will only explore the most used definitions.

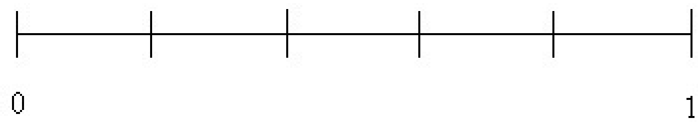
Recall from the previous chapter that scaling is an important characteristic of a fractal, and it is how we construct most fractals. Scaling is closely related to the dimension. To illustrate the idea of scaling with dimension, let's see what will happen to different objects when we double one of its sides. Let's start in the first dimension. Say we have a line segment with length 3. Then if we double it, we are left with a line segment of length 6. So the length of the scaled object increased by a factor of 2.

Now let's see what happens in two dimensions. Let's say we have a square with side length 3. The area of said square is 9. What will happen to the area of the square when we double this side length? When we double the length of one side, we must double the length of the other sides as well. Thus the area of the scaled square is $6 \times 6 = 36$. So the area of our square increased from 9 to 36. So this 2-dimensional object increased by a factor of 4, which is 2^2.

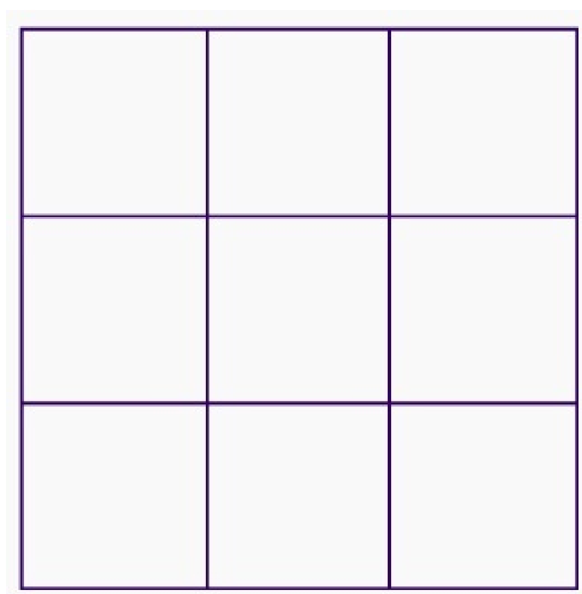
Finally, let's look at the 3-dimensional case. This time, say we have a cube with side length 3. The volume of this cube is 27. Again, if we double the side length of one side, we must double every side of the cube so the scaled volume is $6 \times 6 \times 6 = 216$. So the scaled volume increased by a factor of 8, which is 2^3.

Let's recap. In the 1-dimensional case, the object increased by a factor of 2. In the 2-dimensional case, the object increased by a factor of $2^2 = 4$. In the 3-dimensional case, the object increased by a factor of $2^3 = 8$. Notice that the power of each of these factors is the dimension of the object. So in general, if we have an s-dimensional cube and double the sides, then the volume will increase by a factor of 2^s. We call 2 the scaling factor and s the dimension. Although, the objects that we are talking about are much more complicated than simple cubes, we can use the same way of thinking when calculating dimension.

The first method for calculating fractal dimension that we will explore is called similarity dimension and is the first definition that Mandelbrot used. Similarity dimension is the easiest way of calculating fractal dimension, however, this method will only work if the fractal is self-similar. For self-similar fractals, there are two things that we can look at. The first thing is the number of objects, N, produced after each iteration. The second thing is the scaling ratio, h. This is the ratio of the size of the original piece to the new pieces produced. Consider for a moment a line of length 1 divided up into N equal subsegments with scaling ratio h. Break the line up into 5 segments, each of length $\frac{1}{5}$. Then $N = 5$ and $h = \frac{1}{5}$, so $Nh = 1$.



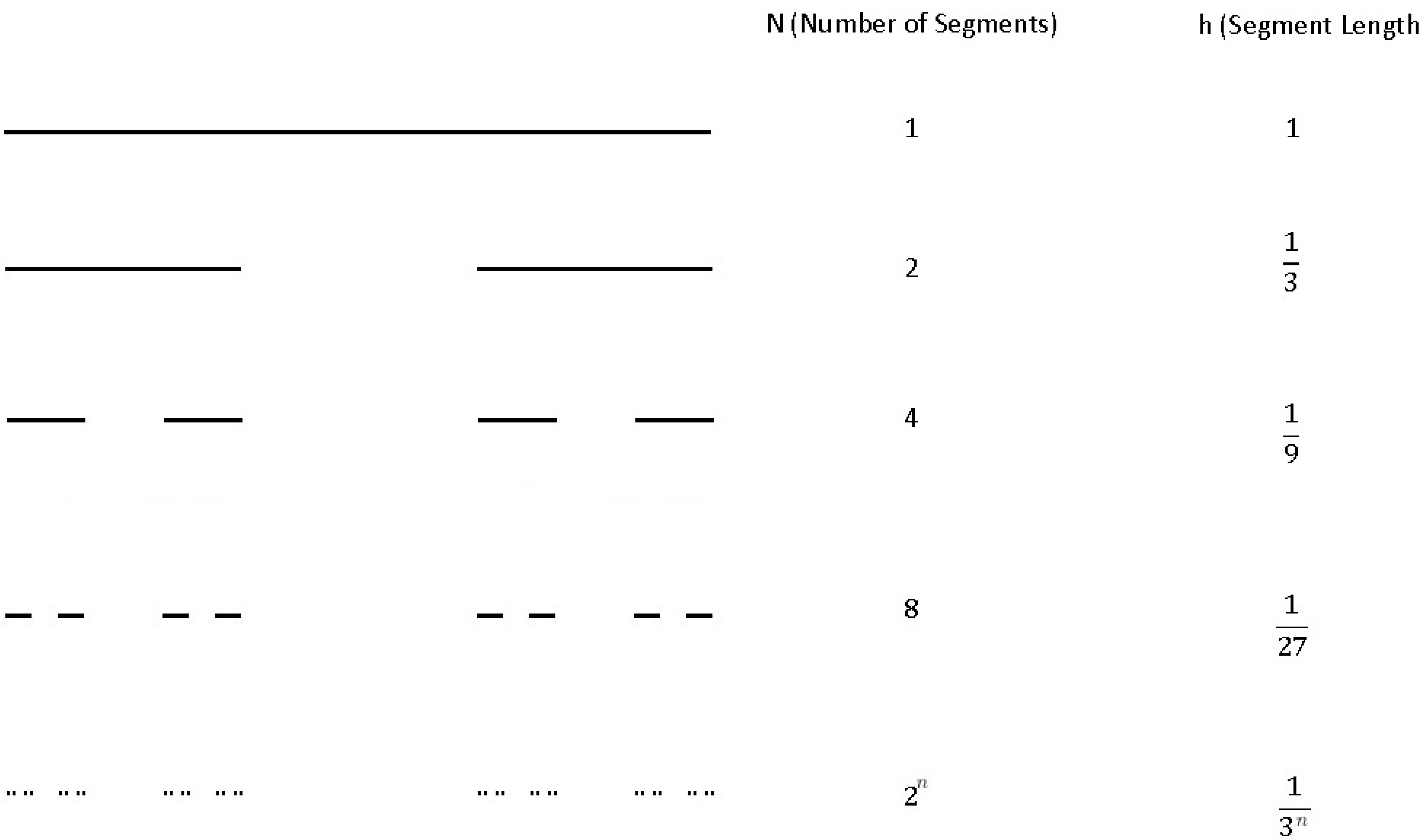
Now let's say we have a 1×1 square instead and divide the square into N equal subsquares with the sides scaled by ratio h. Let's say we divide the square into 9 equal subsquares, so all of their sides have length $\frac{1}{3}$. Then $N = 9$, $h = \frac{1}{3}$ and $Nh^2 = 1$.



Similarly, for a cube divided up into N equal cubes we would get $Nh^3 = 1$. Note that the power of h is the dimension of the set (the 1-dimensional line, the 2-dimensional square, and the 3-dimensional cube). So for any object with dimension d that is divided up into N equal parts with scaling ratio h, we have

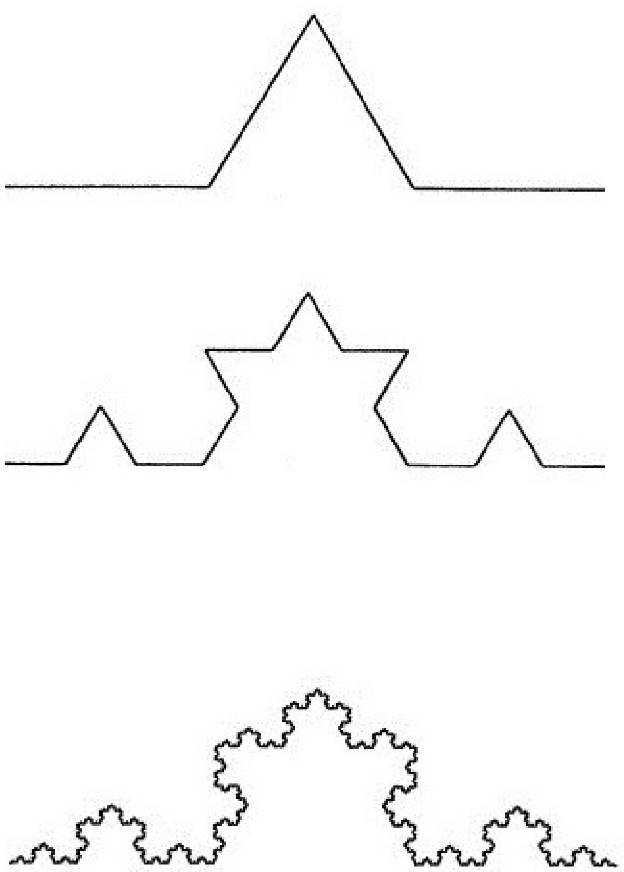
$Nh^d = 1$. Solving this equation for d yields $d = \dfrac{\ln N}{\ln \frac{1}{h}}$. So we will define the similarity dimension as $\dim_s(F) = \dfrac{\ln N}{\ln \frac{1}{h}}$. Let's look at a few examples.

Example 4.2 Let F be the Cantor Set. Recall that the Cantor Set is the line segment $[0, 1]$ with the middle thirds continuously removed and for each stage E_n we get 2^n intervals of length 3^{-n}.



So for the formula of similarity dimension, $N = 2$ and $h = \dfrac{1}{3}$. Thus $D_s(F) = \dfrac{\ln 2}{\ln \frac{1}{\frac{1}{3}}} = \dfrac{\ln 2}{\ln 3} \approx 0.6309$. Our answer makes sense because the Cantor set, a disconnected set of points, should have dimension less than one since a disconnected set of points fills up less space than a line. Also, our answer should be greater than zero since the Cantor set is an uncountable number of points. $\square$

Example 4.3 Now let F be the Koch curve. At each iteration E_n, we have 4^n segments each of length 3^{-n}. So for calculating the dimension, $N = 4$ and $h = \frac{1}{3}$. Thus $\dim_s(F) = \frac{\ln 4}{\ln 3} \approx 1.26186$.

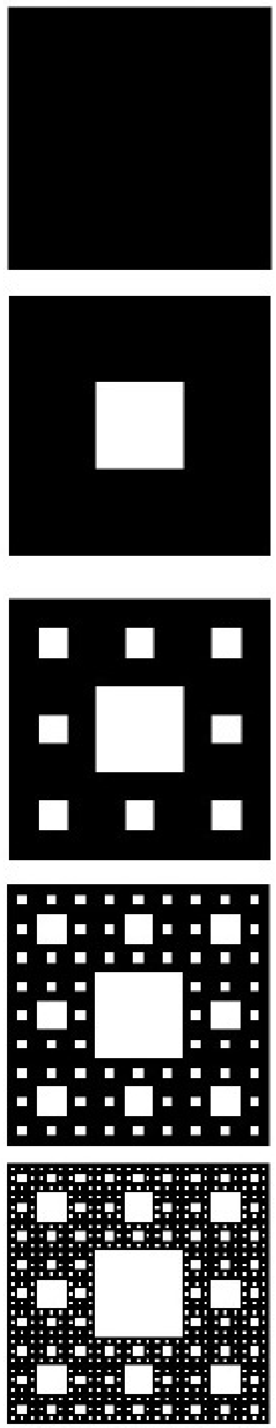


N (Number of segments)	h (segment length)
1	1
4	$\frac{1}{3}$
4^2	$\frac{1}{3^2}$
⋮	
4^n	$\frac{1}{3^n}$

$\square$

Example 4.4 Revised Cantor set: Consider a line segment with length 1. Separate this segment into fifths and remove the second and fourth fifth. Do the same thing to each of the remaining pieces, etc. Let F be the set at infinity. At the nth stage, there are 3^n pieces of length 5^{-n}, so $N = 3$ and $h = \dfrac{1}{5}$. Then $\dim_s(F) = \dfrac{\ln 3}{\ln \frac{1}{5}} = \dfrac{\ln 3}{\ln 5} \approx 0.682606$ $\square$

Example 4.5 Let F be the Sierpinski carpet. This is constructed as follows. Let E_0 be a solid square. The initiator E_1 is obtained by separating the square into nine equal subsquares and removing the middle one, and F is the limiting set. Do this again for each of the remaining eight squares. So at iteration E_n, we will have 8^n squares each with area 9^{-n}. However, let's recall what h is. h is the amount that just one side is scaled by. Since each side is only scaled by $\frac{1}{3}$, $h = \frac{1}{3}$. So $\dim_s F = \dfrac{\ln 8}{\ln 3} \approx 1.893$ [1].



	N	h
	1	1
	8	$\frac{1}{9}$
	8^2	$\frac{1}{9^2}$
	8^3	$\frac{1}{9^3}$
	$\vdots$	
	8^n	$\frac{1}{9^n}$

While the similarity dimension is nice because it is easy to use and calculate, it has its shortcomings when we are dealing with sets that are not self-similar or when the scaling ratio is difficult to identify. If we can't use the similarity dimension, the most traditional method of calculating fractal dimension is the Hausdorff dimension. In order to develop Hausdorff dimension, we will first need to develop some heavy machinery. Let's start off with the definition of Hausdorff measure.

Definition 4.6 Let U be a set in a metric space (X, d). The *diameter* of U, denoted $|U|$, is defined to be $|U| = \sup\{|x - y| : x, y \in U\}$.

Let $F \subset \mathbb{R}^n$ and $s \in \mathbb{R}$ such that $s \geq 0$ Then for any $\delta > 0$, define

$$\mathcal{H}^s_\delta(F) = \inf \left\{ \sum_{i=1}^{\infty} |U_i|^s : \{U_i\} \text{ is a } \delta\text{-cover of F} \right\}.$$

So we are looking at the covers of F with diameter at most δ, but we want to make the sum as small as possible. Note that as δ decreases, the infimum $\mathcal{H}^s_\delta(F)$ increases. This is because as δ gets smaller, the number of possible covers of F decreases. The following figure illustrates this using the coastline of Britain for F. Note that we cover the coastline with a cover, then make the balls smaller. As the balls become smaller, we need more of them to cover the set.

When the number of elements in a set is reduced, some of the smaller elements may get thrown out, so the infimum could increase. However, if the smallest ones do not get thrown out, the infimum will stay the same. So as we decrease the value for δ, the sequence of elements of the Hausdorff measure is increasing. Thus the sequence of infima converges, so $\lim_{\delta \to 0} \mathcal{H}^s_\delta(F)$ exists. Unlike some disciplines of analysis, where limits at infinity are said not to exist, for the purposes of this thesis we will still say that the limit exists even if the limit is infinity. This gives us the definition of Hausdorff measure.

Definition 4.7 Let $F \subset \mathbb{R}^n$. Then the *s-dimensional Hausdorff measure* of F

is

$$\mathcal{H}^s(F) = \lim_{\delta \to 0} \mathcal{H}^s_\delta(F) = \lim_{\delta \to 0} \inf \left\{ \sum_i |U_i|^s : \{U_i\} \text{ is a } \delta\text{-cover of } F \right\}.$$

Theorem 4.8 Definition 4.7 defines a measure on R^n.

Proof:

1. Clearly $\mathcal{H}^s(\emptyset) = 0$ since any collection of sets covers the empty set.

2. Suppose $E \subseteq F$ and let $\{U_i\}$ be a δ-cover of F. Then $\{U_i\}$ is also a δ-cover of E. So there is a subcollection $\{V_j\}$ of $\{U_i\}$ such that $\{V_i\}$ is a cover of E. Then for each $V_j = U_i$ we have $|V_j| \leq |U_i| \leq \delta$ and there may be a fewer number of sets in $\{V_j\}$ than in $\{U_i\}$. So the sum for the $|V_j|$'s will possibly be smaller than the sum over the $|U_i|$'s. Thus $\mathcal{H}^s(E) \leq \mathcal{H}^s(F)$.

3. Our goal is to show that $\mathcal{H}^s\left(\bigcup_{k=1}^{\infty} F_k\right) \leq \sum_{k=1}^{\infty} \mathcal{H}^s(F_k)$ for any finite or countable sequence of sets.

   Let $F_1, F_2, ...$ be a finite or countable sequence of sets. If $\mathcal{H}^s(F_k) = \infty$ for some k, then the inequality holds trivially. So assume that $\mathcal{H}^s(F_k) < \infty$ for all k. Let $\epsilon > 0$. Then there is a δ-cover $\{U_i^{(k)}\}$ of F_k such that $\sum_i \left|U_i^{(k)}\right|^s < \mathcal{H}^s_\delta(F_k) + \frac{\epsilon}{2^k}$. Now take the sum over all k:

   $$\sum_{k=1}^{\infty} \sum_i \left|U_i^{(k)}\right|^s < \sum_{k=1}^{\infty} \mathcal{H}^s_\delta\left((F_k) + \frac{\epsilon}{2^k}\right) = \sum_{k=1}^{\infty} \mathcal{H}^s_\delta(F_k) + \epsilon \sum_{k=1}^{\infty} \frac{1}{2^k} = \sum_{k=1}^{\infty} \mathcal{H}^s_\delta(F_k) + \epsilon.$$

   Since $\{U_i^{(k)}\}$ is a cover of F_k, then $\left\{\bigcup_{k=1}^{\infty} U_i^{(k)}\right\}$ is a cover for $\bigcup_{k=1}^{\infty} F_k$. So

   $$\mathcal{H}^s_\delta\left(\bigcup_{k=1}^{\infty} F_k\right) \leq \sum_{k=1}^{\infty} \sum_i \left|U_i^{(k)}\right|^s < \sum_{k=1}^{\infty} \mathcal{H}^s_\delta(F_k) + \epsilon.$$

Since this holds for any $\epsilon > 0$, let $\epsilon \to 0$:

$$\mathcal{H}^s_\delta\left(\bigcup_{k=1}^{\infty} F_k\right) \leq \sum_{k=1}^{\infty} \mathcal{H}^s_\delta(F_k)$$

Now letting $\delta \to 0$ yields

$$\mathcal{H}^s\left(\bigcup_{k=1}^{\infty} F_k\right) \leq \sum_{k=1}^{\infty} \mathcal{H}^s(F_k).$$

∎

When dealing with fractals, it's hard to get anywhere without talking about some form of scaling. To define scaling formally, let $S : F \to F$ be a one-to-one function defined such that $|S(x) - S(y)| = c|x - y|$ for $x, y \in F$ and for some $c \in \mathbb{R}$ with $c > 0$. Then S is called a similarity transformation. Similarity transformations take a set and scales them up or down depending on c, we call c the scaling ratio. If $c > 1$, then the image will increase in size, if $c = 1$, then the image will be the same, and if $0 < c < 1$, then the image will decrease in size. Since we usually think of fractals as a set that decreases in size at a certain ratio, in most of our cases we will have $0 < c < 1$. There is a relationship between scaling and Hausdorff measure, described in the following theorem.

Theorem 4.9 Scaling property: Let S be a similarity transformation with scaling factor $\lambda > 0$ and let $F \subseteq \mathbb{R}^n$. Then the s-dimensional Hausdorff measure scales with a factor of λ^s. So $\mathcal{H}^s(S(F)) = \lambda^s \mathcal{H}^s(F)$, where $S(F)$ is the scaled version of F.

Proof:

Let $\{U_i\}$ be a δ-cover of F. Then since S(F) is a scaled down copy of F, there is a $\delta\lambda$-cover $\{V_i\}$ of S(F). We want to examine the sums of the covers for the scaled version S(F) and compare it to the covers of the original set F.

$$\sum_{i=1}^{\infty} |V_i|^s = \sum_{i=1}^{\infty} \lambda^s |U_i|^s = \lambda^s \sum_{i=1}^{\infty} |U_i|^s$$

Then any δ-cover of F will also cover $S(F)$ since the domain and codomain of S are the same. But we are not guaranteed that any cover of the scaled version S(F) will also cover F. So we have the following inequality.

$$\mathcal{H}^s_{\delta\lambda}(\mathrm{S}(F)) \leq \lambda^s \mathcal{H}^s_{\delta}(F)$$

So as $\delta \to 0$,

$$\mathcal{H}^s(\mathrm{S}(F)) \leq \lambda^s \mathcal{H}^s(F).$$

To get the opposite inequality, we will use the inverse function S^{-1} with scaling factor $\dfrac{1}{\lambda}$. Note that $S^{-1}(S(F))$ is a scaled up copy of $S(F)$ by a factor of $\dfrac{1}{\lambda}$. So any δ-cover of $S(F)$ will also cover $S^{-1}(S(F))$ when scaled. So if $\{V_i\}$ is a δ-cover of $S(F)$, there is a $\dfrac{\delta}{\lambda}$-cover $\{U_i\}$ of $S^{-1}(S(F)) = F$. Thus

$$\sum_{i=1}^{\infty} |U_i|^s = \left(\frac{1}{\lambda}\right)^s \sum_{i=1}^{\infty} |V_i|^s.$$

Thus we have the inequality

$$\mathcal{H}^s_{\frac{\delta}{\lambda}}(S^{-1}(S(F))) \leq \left(\frac{1}{\lambda}\right)^s \mathcal{H}^s_{\delta}(S(F))$$

So as $\delta \to 0$,

$$\mathcal{H}^s(F) \leq \left(\frac{1}{\lambda}\right)^s \mathcal{H}^s(S(F))$$

$$\lambda^s \mathcal{H}^s(F) \leq \mathcal{H}^s(S(F))$$

∎

There is yet another way of calculating the Hausdorff measure of a set. The n-dimensional Hausdorff measure is a constant multiple of the n-dimensional volume of the set. Let $\mathrm{vol}^n(F)$ denote the n-dimensional volume of F, computed using Lebesgue measure. If F is an open or closed set, then it can be shown that $\mathcal{H}^n(F) = c_n^{-1}\text{vol}^n(F)$, where c_n is the volume of an n-dimensional ball of diameter 1. If n is even, then $c_n = \dfrac{\pi^{\frac{n}{2}}}{2^n(\frac{n}{2})}$. If n is odd, then $c_n = \dfrac{\pi^{\frac{n-1}{2}}(\frac{n-1}{2})!}{n!}$. Let's briefly look at a few examples of this [5].

1. $\mathcal{H}^0(F)$ can be thought of as the number of points in F. Say we have five points. Then the measure would be $\sum_{k=1}^{5} |U_i|^0 = \sum_{k=1}^{5} 1 = 5$.

2. $\mathcal{H}^1(F)$ can be thought of as the arclength of a curve since $c_1 = 1$ and the $\text{vol}^1(F)$ is the length of a 1-dimensional line.

3. $\mathcal{H}^2(F) = c_n^{-1}\text{vol}^2(F) = \dfrac{4}{\pi}\,\text{area}(F)$ since here $n = 2$, which is even. Plugging 2 in for n in the appropriate formula above gives us $c_n = \dfrac{\pi}{4}$, so $c_n^{-1} = \dfrac{4}{\pi}$.

4. $\mathcal{H}^3(F) = c_n^{-1}\text{vol}^3(F) = \dfrac{6}{\pi}\text{vol}^3(F)$. This time $n = 3$, which is odd, so plug 3 into the appropriate formula gives us $c_n = \dfrac{\pi}{6}$, so $c_n^{-1} = \dfrac{6}{\pi}$.

Example 4.10 Let F be the unit circle in the xy-plane in $\mathbb{R}^3$. Then $\mathcal{H}^1(F) = \infty$ since F can be covered by infinitely many lines. The 2-dimensional measure is $\mathcal{H}^2(F) = \dfrac{4}{\pi} \cdot 2\pi = 4$. The 3-dimensional measure is $\mathcal{H}^3(F) = \dfrac{6}{\pi} \cdot 0 = 0$ since the circle has no volume in $\mathbb{R}^3$. So

$$\mathcal{H}^s(F) = \begin{cases} \infty & \text{if } s < 2 \\ 4 & \text{if } s = 2 \\ 0 & \text{if } s > 2 \end{cases}$$

$\square$

Using the definition of Hausdorff measure, $\mathcal{H}^s_\delta(F)$ is decreasing with s for $\delta < 1$. This is because the size of each U_i of the cover $\{U_i\}$ is less than one. So as s increases, $|U_i|^s$ decreases since $|U_i| \leq \delta \leq 1$, and raising small numbers to large powers makes the number smaller. Then as $\delta \to 0$, we have that $\mathcal{H}^s(F)$ is also decreasing. Thus for $t > s$,

$$\sum_i |U_i|^t = \sum_i |U_i|^{t-s}|U_i|^s \leq \delta^{t-s} \sum_i |U_i|^s = \delta^{t-s}\mathcal{H}^s(F).$$

Then if $\delta \to 0$, the right hand side is going to 0. So for all values t such that $t > s$, the t-dimensional Hausdorff measure is 0. So there is a critical value where $\mathcal{H}^s(F)$ "jumps" from ∞ to 0. This value is the Hausdorff dimension of F.

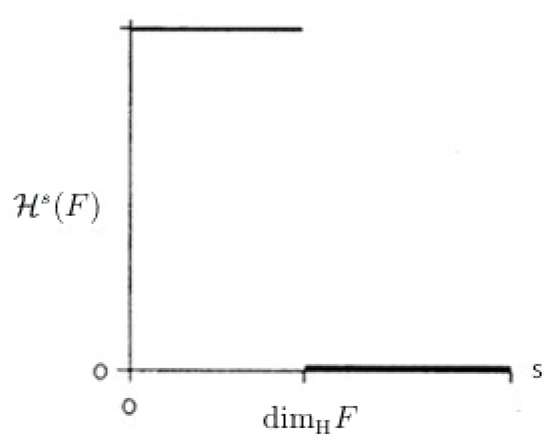
Definition 4.11 Let $F \subseteq \mathbb{R}^n$. Then the *Hausdorff dimension* of F is

$$\dim_{\mathrm{H}} F = \inf\{s \geq 0 : \mathcal{H}^s(F) = 0\} = \sup\{s : \mathcal{H}^s(F) = \infty\}.$$

So we have

$$\mathcal{H}^s(F) = \begin{cases} \infty & \text{if } 0 \leq s < \dim_{\mathrm{H}} F \\ 0 & \text{if } s > \dim_{\mathrm{H}} F \end{cases}.$$

If $s = \dim_{\mathrm{H}} F$, then $0 \leq \mathcal{H}^s(F) \leq \infty$, so $\mathcal{H}^s(F)$ could take on any value. See the figure.



One thing to be careful of when calculating the dimension is that the dimension is s. It is easy to conclude that the dimension is the value $\mathcal{H}^s(F)$, but that is the measure. Be sure to make the distinction between Hausdorff dimension and Hausdorff measure. Let us look at an example for calculating the Hausdorff dimension of a fractal.

Example 4.12 Let F be the Cantor set. We will now calculate the Hausdorff dimension of this set. Clearly, F can be split up into two halves: the left half and the right half. Call the left half $F_L = F \cap \left[0, \frac{1}{3}\right]$ and the right half $F_R = F \cap \left[\frac{2}{3}, 1\right]$. Then $F = F_L \cup F_R$ and this union is disjoint (so $F_L \cap F_R = \emptyset$). Thus $\mathcal{H}^s(F) = \mathcal{H}^s(F_L) + \mathcal{H}^s(F_R)$. Since the Cantor set is self-similar, each half is a scaled down copy of the whole thing, and it is scaled by a factor of $\lambda = \frac{1}{3}$. Thus by the scaling property, $\mathcal{H}^s(F_L) = \mathcal{H}^s(F_R) = \left(\frac{1}{3}\right)^s \mathcal{H}^s(F)$. So

$$\mathcal{H}^s(F) = \mathcal{H}^s(F_L) + \mathcal{H}^s(F_R) = \left(\frac{1}{3}\right)^s \mathcal{H}^s(F) + \left(\frac{1}{3}\right)^s \mathcal{H}^s(F) = \frac{2}{3^s}\mathcal{H}^s(F)$$

Thus $\mathcal{H}^s(F) = \frac{2}{3^s}\mathcal{H}^s(F)$. Note that $\mathcal{H}^s(F) \neq 0$ and $\mathcal{H}^s(F) \neq \infty$ [5]. Then if we assume that $s = \dim_{\mathrm{H}} F$, then $0 < \mathcal{H}^s(F) < \infty$ and we can divide both sides by $\mathcal{H}^s(F)$. This yields

$$\mathcal{H}^s(F) = \frac{2}{3^s}\mathcal{H}^s(F)$$

$$1 = \frac{2}{3^s}$$

$$3^s = 2$$

$$\ln 3^s = \ln 2$$

$$s \ln 3 = \ln 2$$

$$s = \frac{\ln 2}{\ln 3}$$

$\square$

This was a relatively simple example because we did it without actually calculating any of the Hausdorff measures. If those calculations need to be made, the process becomes much more difficult. In fact, working with the definition of Hausdorff dimension can be very daunting. It is usually easiest to find the Hausdorff dimension by finding upper and lower bounds or using some trick to find the dimension. The upper bound for the Hausdorff dimension is usually fairly easy to find, and luckily it is also usually the actual value of the dimension. To help us calculate the Hausdorff dimension of certain sets, we will need a few theorems. Theorem 4.13 gives us a way of finding upper bounds to the Hausdorff measure. This is then used to prove Theorem 4.15 and Theorem 4.16, which give a way of finding an upper bound for the Hausdorff dimension (remember Hausdorff measure is $\mathcal{H}^s(F)$ and Hausdorff dimension is the s value where the measure jumps from ∞ to 0. We usually care about the dimension more than the measure.)

Theorem 4.13 Let $F \subseteq \mathbb{R}^n$ and $f : F \to \mathbb{R}^m$ be a mapping where

$$|f(x) - f(y)| \leq c|x - y|^\alpha$$

for $x, y \in F$ and constants $c > 0$ and $\alpha > 0$. Then for any s,

$$\mathcal{H}^{s/\alpha}(f(F)) \leq c^{s/\alpha}\mathcal{H}^s(F).$$

Proof:

Let $\{U_i\}$ be a δ-cover of F. Consider $F \cap U_i$. By the hypothesis, $|f(F \cap U_i)| \leq c|F \cap U_i|^\alpha$, where the absolute values here denote the diameter of each set. Also, $c|F \cap U_i|^\alpha \leq c|U_i|^\alpha$ since $|U_i| \geq |F \cap U_i|$. Thus

$$|f(F \cap U_i)| \leq c|F \cap U_i|^\alpha \leq c|U_i|^\alpha \leq c\delta^\alpha.$$

Then the set $\{f(F \cap U_i)\}$ is an ϵ-cover of $f(F)$, where $\epsilon = c\delta^\alpha$. Then

$$|f(F \cap U_i)| \leq c|U_i|^\alpha$$

implies that

$$|f(F \cap U_i)|^{s/\alpha} \leq c^{s/\alpha}|U_i|^s$$

by raising both sides to the s/α power. On taking the infimum of both sides, we get

$$\mathcal{H}_\epsilon^{s/\alpha}(f(F)) \leq c^{s/\alpha}\mathcal{H}_\delta^s(F).$$

As $\delta \to 0$, $\epsilon \to 0$ also so

$$\mathcal{H}^{s/\alpha}(f(F)) \leq c^{s/\alpha}\mathcal{H}^s(F).$$

∎

A function f has the Hölder condition if $|f(x) - f(y)| \leq c|x - y|^\alpha$. If a function has this condition, then it is continuous. This is because if $\epsilon > 0$ and $|x_1 - x_2| < \delta$, then if we choose $\delta = \dfrac{\epsilon}{c}$, we have $|f(x_1) - f(x_2)| \leq c|x_1 - x_2| < c\delta = \epsilon$. So continuity follows by definition. If $\alpha = 1$ then f is called a *Lipschitz mapping* and $\mathcal{H}^s(f(F)) \leq c^s\mathcal{H}^s(F)$.

Example 4.14 Let $f : [0, 1] \to [0, 1]$ be defined such that $f(x) = \dfrac{1}{3}x$. Then for any $x, y \in [0, 1]$, $|f(x) - f(y)| = |\dfrac{1}{3}x - \dfrac{1}{3}y| = \dfrac{1}{3}|x - y|$, so f satisfies the Hölder condition. Furthermore, since $x - y \leq 0$ for all x, y, then $\dfrac{1}{3}|x - y| < \dfrac{1}{3}|x - y|^\alpha$ for any $\alpha > 1$. If we iterate f like we would a difference function, then f will produce the left half of the Cantor set.

□

Hölder and Lipschitz mappings are useful for finding upper bounds for the Hausdorff dimension of a set.

Theorem 4.15 Let $F \subseteq \mathbb{R}^n$ and suppose $f : F \to \mathbb{R}^m$ satisfies a Hölder condition: $|f(x) - f(y)| \leq c|x - y|^{\alpha}$, where $x, y \in F$. Then

$$\dim_{\mathrm{H}} f(F) \leq \frac{1}{\alpha} \dim_{\mathrm{H}} F.$$

Proof:

Since the Hölder inequality holds by the hypothesis, by Theorem 4.13 we have

$$\mathcal{H}^{s/\alpha}(f(F)) \leq c^{s/\alpha} \mathcal{H}^s(F)$$

for all s. Note that if $\dim_{\mathrm{H}} F < s$, then $\mathcal{H}^s(F) = 0$ by definition of Hausdorff dimension. But then $\mathcal{H}^{s/\alpha}(f(F)) = 0$ as well. Therefore, again by the definition of Haudorff dimension, $\dim_{\mathrm{H}} f(F) < \dfrac{s}{\alpha}$. Thus on taking the infimum of s we have

$$\dim_{\mathrm{H}} f(F) \leq \frac{\dim_{\mathrm{H}} F}{\alpha} < \frac{s}{\alpha}.$$

∎

An immediate corollary to this theorem is that if f is a Lipshitz transformation, that is if $\alpha = 1$, then we have $\dim_{\mathrm{H}} f(F) \leq \dim_{\mathrm{H}} F$. We will use this corollary for the following theorem, which gives a nice way of finding an upper bound for certain sets.

Theorem 4.16 If $\dim_{\mathrm{H}} F < 1$ for a set $F \subset \mathbb{R}^n$, then F is totally disconnected.

Proof:

Let $x, y \in F$ such that x and y are distinct. Let $f : \mathbb{R}^n \to [0, \infty)$ be defined such that $f(z) = |z - x|$. Then

$$|f(z) - f(w)| = ||z - x| - |w - x|| \leq |(z - x) - (w - x)| = |z - w|$$

by the reverse triangle inequality. This shows that the function f not only satisfies the Hölder condition, but is a Lipschitz transformation as well. Therefore by the corollary we have $\dim_{\mathrm{H}} f(F) \leq \dim_{\mathrm{H}} F < 1$. So if $s = 1$, then $f(F)$ has $\mathcal{H}^1$ measure zero by definition of Hausdorff measure ($\mathcal{H}^1$ measures the length of an interval, where length here is the normal way of describing geometric length). Let $r \in \mathbb{R}$ such that $r \notin f(F)$ and $0 < r < f(y)$. So there is no point $z \in F$ such that $f(z) = r$. Then we can break up F into two disjoint subsets, so

$$F = \{z \in F : |z - x| < r\} \cup \{z \in F : |z - x| > r\}.$$

But x is an element of the left subset of F since $f(x) = |x - x| = 0 < r$, and y is an element of the right subset since $f(y) = |y - x| > r$. Note that the closure of $\{z \in F : |z - x| < r\}$ is $\{z \in F : |z - x| = r\}$. So

$$\overline{\{z \in F : |z - x| < r\}} \cap \{z \in F : |z - x| > r\} = \emptyset$$

and

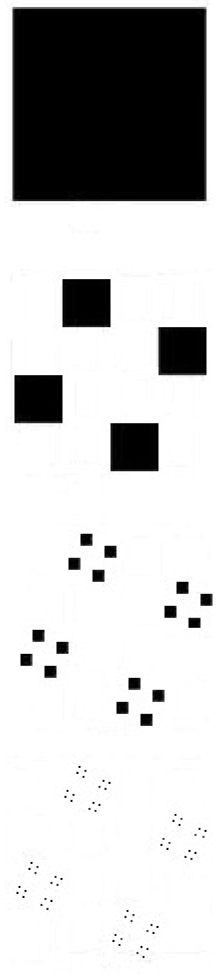
$$\overline{\{z \in F : |z - x| > r\}} \cap \{z \in F : |z - x| < r\} = \emptyset$$

. Thus F is totally disconnected by definition.

∎

The following example demonstrates how to calculate the Hausdorff dimension of a fractal when the Lipschitz condition holds. This is also a rather interesting fractal because it is a fractal that has an integer Hausdorff dimension, which very rarely occurs.

Example 4.17 Let F be the Cantor dust. The Cantor dust is constructed as follows. Take the unit square centered at $(\frac{1}{2}, \frac{1}{2})$ to be the initiator. For the generator, break apart the square into sixteen equal subsquares, but only keep four of them, and keep doing this. The squares that are kept are the ones illustrated in the following diagram.



Note that at the stage E_k, there are 4^k squares with side length 4^{-k}. So the diameter of each square is $4^{-k}\sqrt{2}$. Take $\delta = 4^{-k}\sqrt{2}$ so that all these squares are a δ-cover of F. In order to obtain the upper bound for the Hausdorff measure, we will need to appeal to definition 4.7. If $s = 1$, then since there are 4^k squares, we have

$$\mathcal{H}_\delta^1 \leq \sum_{i=1}^{\infty} |U_i| = \sum_{i=1}^{4^k} 4^{-k}\sqrt{2} = 4^k 4^{-k}\sqrt{2} = \sqrt{2}$$

. We have an inequality since we do not know the infimum of the set in

Definition 4.7, so there may be a smaller upper bound. Then taking the limit, we obtain $\lim_{\delta \to 0} \mathcal{H}^1_\delta(F) = \mathcal{H}^1(F) \leq \sqrt{2}$.

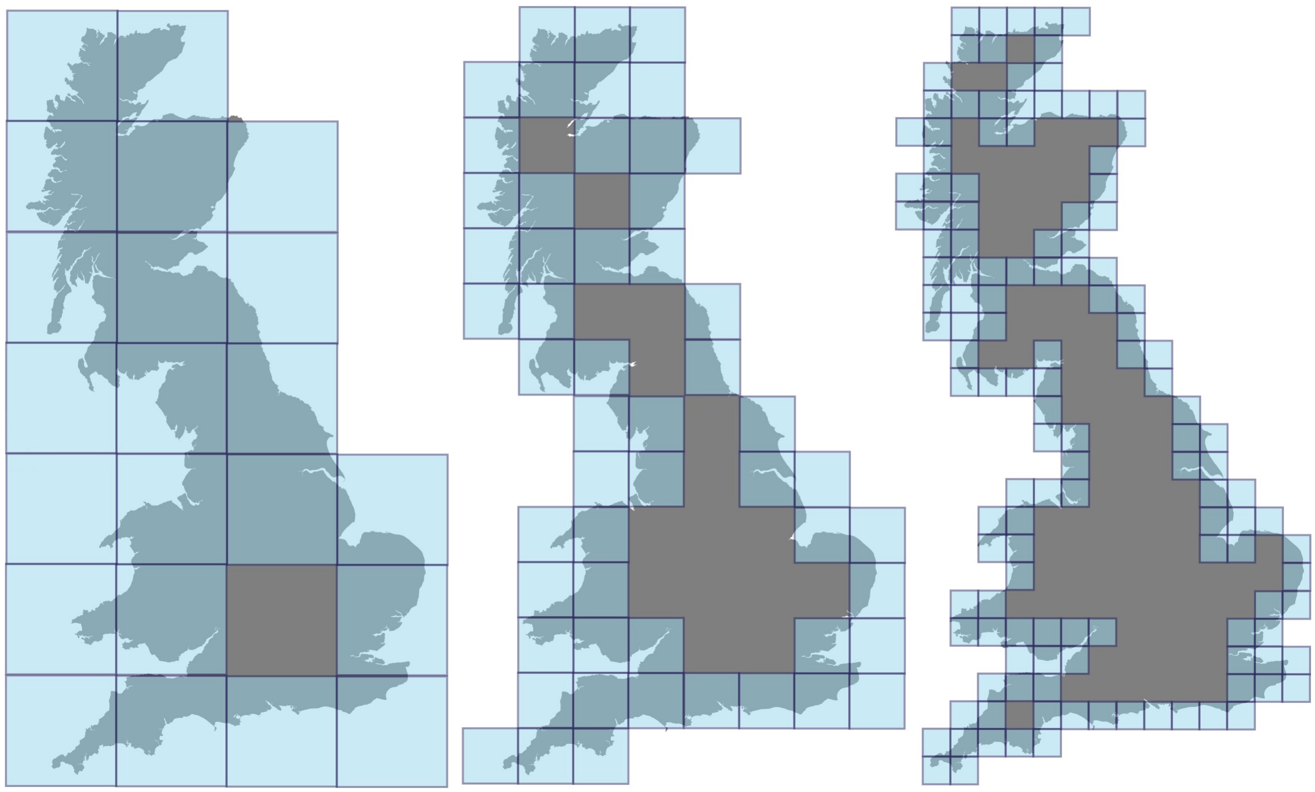
To find a lower bound, let $f(F)$ be the projection of F on the x-axis. The projection can be thought of as the shadow of this set on the x-axis. If a light was shone through F from above, no light would shine through the interval $[0, 1]$. Projections do not increase distances, so for any $x, y \in F$, $|f(x) - f(y)| \leq |x - y|$. Now since the length this interval is 1, we have

$$1 = \mathcal{H}^1([0, 1]) = \mathcal{H}^1(f(F)) \leq \mathcal{H}^1(F).$$

This shows that $1 \leq \mathcal{H}^1(F) \leq \sqrt{2}$, so $s = \dim_{\mathrm{H}} F = 1$ (again recall that the Hausdorff dimension is the power of the Hausdorff measure, so the 1 in $\mathcal{H}^1(F)$). So the Cantor dust is an unusual fractal in the sense that it has integer fractal dimension.

$\square$

Box-counting dimension is the most popular way of calculating fractal dimension since it is relatively easy to calculate. It closely resembles the similarity dimension in the way that logarithms are used, but the set need not be self-similar in order to use the box dimension. The idea behind using the box dimension is to put the set in question a grid and see how many boxes of a certain size are required to cover the set. Then start shrinking the boxes, so we get more boxes that are smaller, and see how the numbers change. The following picture illustrates this process if trying to calculate the coastline of Britain [12].



In order to find the box dimension, two calculations are needed: the lower box dimension and the upper box dimension. They are defined as follows.

Definition 4.18 Let $F \subseteq \mathbb{R}^n$ such that F is nonempty and bounded. Let $N_\delta(F)$ be the smallest number of sets with diameter δ needed to cover F. The *upper box-counting dimension* is defined as

$$\underline{\dim}_{\mathrm{B}} F = \limsup_{\delta \to 0} \left\{ \frac{\ln N_\delta F}{-\ln \delta} \right\}$$

and the *lower box-counting dimension* is defined as

$$\underline{\dim}_{\mathrm{B}} F = \liminf_{\delta \to 0} \left\{ \frac{\ln N_\delta F}{-\ln \delta} \right\}$$

If the two are equal, then we get the *box-counting dimension*:

$$\dim_{\mathrm{B}} F = \lim_{\delta \to 0} \frac{\ln N_\delta(F)}{-\ln \delta}$$

It must be noted that the value of $N_\delta(F)$ does not to a number of any specific kind of set. There are several different kinds of sets we could use, and the value of the box dimension will be the same. The most commonly used choices for $N_\delta(F)$ are:

1. The smallest number of closed balls with radius δ that cover F.

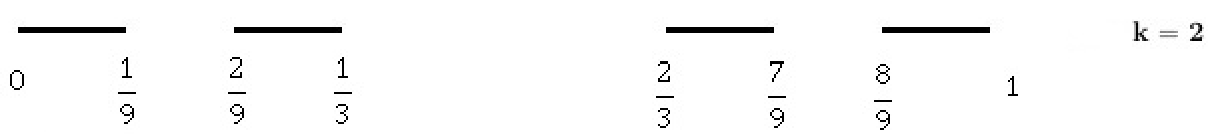
2. The smallest number of cubes with side length δ that cover F.

3. The smallest number of sets of diameter δ that cover F.

4. The largest number of disjoint balls with radius δ that cover F.

When using this definition, let $\delta > 0$ be small enough such that $-\ln\delta$ and $\ln N_\delta(F)$ are strictly positive. We only use this definition for nonempty and bounded sets to avoid $\ln 0$ and $\ln\infty$. Let us see an example using the box dimension.

Example 4.19 Let F be the middle third Cantor set. For any iteration E_k, the length of each piece is 3^{-k} and there are 2^k of them. Let $3^{-k} < \delta \leq 3^{-k+1}$. To understand what is going on, consider the case where $k = 2$. Then we would have 4 segments of length $\frac{1}{9}$ and $\frac{1}{9} < \delta \leq \frac{1}{3}$. Suppose that $\delta = \frac{1}{3}$.



The point $\frac{1}{6}$ is half way between $\frac{1}{9}$ and $\frac{2}{9}$. If we center a δ-ball around $\frac{1}{6}$, then we could cover the first two intervals except for the end points with that ball. In order to cover the endpoints, we would need two such δ-balls.

So to cover the whole set, we would need a total of 4 balls, which is the number of intervals in the set. If δ was just slightly greater than $\frac{1}{9}$, then we would still need 4 balls to cover the set. This indicates that if $3^{-k} < \delta \leq 3^{-k+1}$, then there would need to be 2^k balls to cover the whole set. Thus $N_\delta(F) = 2^k$. Note that as $\delta \to 0$, $k \to \infty$ since the 3 has a negative exponent. So we can calculate the upper box dimension using the formula and some basic calculus techniques:

$$\overline{\dim}_{\mathrm{B}} F = \limsup_{\delta \to 0} \frac{\ln N_\delta(F)}{-\ln \delta} = \limsup_{\delta \to 0} \frac{\ln 2^k}{-\ln 3^{-k+1}} = \limsup_{k \to \infty} \frac{\ln 2^k}{-\ln 3^{-k+1}} = \frac{\ln 2}{\ln 3}.$$

The lower box dimension is similar:

$$\underline{\dim}_{\mathrm{B}} F = \liminf_{\delta \to 0} \frac{\ln N_\delta(F)}{-\ln \delta} = \liminf_{\delta \to 0} \frac{\ln 2^k}{-\ln 3^{-k+1}} = \liminf_{k \to \infty} \frac{\ln 2^k}{-\ln 3^{-k+1}} = \frac{\ln 2}{\ln 3}.$$

So since the upper and lower box dimensions agree, the box dimension is

$$\dim_B F = \frac{\ln 2}{\ln 3}.$$

□

In this case, the box and Hausdorff dimensions happen to agree.However, this is not always the case, as we will see later.

Using the box dimension, there are ways for calculating the Hausdorff dimension without appealing to the definition. As stated earlier, this is nice because using the definition of Hausdorff dimension is usually very difficult to work with. To find upper bounds, we will usually use coverings of small sets, but since this is how we calculated the box dimension this will give us a relation between the two.

Theorem 4.20 If $\mathcal{H}^s(F) > 1$, then

$$\dim_{\mathrm{H}} F \leq \underline{\dim}_{\mathrm{B}} F \leq \overline{\dim}_{\mathrm{B}} F$$

Proof:

In order to prove this we will first establish that $\mathcal{H}^s_\delta(F) \leq N_\delta(F)\delta^s$, where $N_\delta(F)$ is the number of sets of diameter δ that can cover F. This inequality holds from the definition of Hausdorff measure,

$$\mathcal{H}^s_\delta(F) = \inf\left\{\sum_{i=1}^{\infty} |U_i|^s : \{U_i\} \text{ is a } \delta\text{-cover of } F\right\}.$$

$N_\delta(F)$ is the number of sets for one of the covers $\{U_i\}$. So $|U_i|^s \leq \delta^s$ and $\sum_{i=1}^{\infty} |U_i|^s \leq N_\delta(F)\delta^s$. Thus $N_\delta(F)\delta^s$ is an element of the set on the right side of the definition. So this gives an upper bound for the area that is covered by $\{U_i\}$ since the infimum of the set will be less than or equal to this value. Therefore $\mathcal{H}^s_\delta(F) \leq N_\delta(F)\delta^s$. Now if $\mathcal{H}^s(F) > 1$, then we have

$$\begin{aligned}
& 1 < \lim_{\delta\to 0} \mathcal{H}^s_\delta(F) \leq \lim_{\delta\to 0} N_\delta(F)\delta^s. \\
\Rightarrow \quad & 1 < \lim_{\delta\to 0} N_\delta(F)\delta^s \\
\Rightarrow \quad & 0 < \lim_{\delta\to 0} \ln(N_\delta(F)\delta^s) \\
\Rightarrow \quad & 0 < \lim_{\delta\to 0} (\ln N_\delta(F) + s\ln\delta) \\
\Rightarrow \quad & \lim_{\delta\to 0} -s\ln\delta < \lim_{\delta\to 0} \ln N_\delta(F) \\
\Rightarrow \quad & s < \lim_{\delta\to 0}\left(\frac{\ln N_\delta(F)}{-\ln\delta}\right) \\
\Rightarrow \quad & s \leq \liminf_{\delta\to 0}\left(\frac{\ln N_\delta(F)}{-\ln\delta}\right) = \underline{\dim}_{\mathrm{B}} F
\end{aligned}$$

So since $s = \dim_{\mathrm{H}} F$, $\dim_{\mathrm{H}} F \leq \underline{\dim}_{\mathrm{B}} F \leq \overline{\dim}_{\mathrm{B}} F$.

∎

This gives an upper bound for the Hausdorff dimension. Usually this will be a strict inequality, but for the self-similar fractals equality will hold.

To find lower bounds, we will be using mass distributions on the set, defined as follows.

Definition 4.21 A *mass distribution* μ on a set $F \subseteq \mathbb{R}^n$ is a measure such that $0 < \mu(F) < \infty$.

So a mass distribution is any measure that has a finite positive value on F. Mass distributions yield some powerful theorems that can again be used to find an upper bound for the Hausdorff dimension, even when the Hausdorff measure is less than 1. The following theorem is called the Mass Distribution Principle.

Theorem 4.22 Suppose μ is a mass distribution for a set F and suppose that for some $s \in \mathbb{R}$ there is a $c > 0$ and an $\epsilon > 0$ such that $\mu(U) \leq c|U|^s$ for any set U where $|U| < \epsilon$. Then $\mathcal{H}^s(F) \geq \dfrac{\mu(F)}{c}$ and $s \leq \dim_{\mathrm{H}} F \leq \underline{\dim}_{\mathrm{B}} F \leq \overline{\dim}_{\mathrm{B}} F$.

Proof:

Let $\{U_i\}$ be a cover of F. Then $0 < \mu(F)$ since μ is a mass distribution. Also, $\mu(F) \leq \mu\left(\bigcup_i U_i\right) \leq \sum_i \mu(U_i)$ by properties of measure. So if we assume that $|U_i| < \epsilon$,

$$
\begin{aligned}
& \sum_i \mu(U_i) \leq \sum_i c|U_i|^s \quad \text{by hypothesis} \\
& \sum_i c|U_i|^s = c\sum_i |U_i|^s \quad \text{by properties of sums} \\
\Rightarrow \quad & 0 < \mu(F) \leq \mu\left(\bigcup_i U_i\right) \leq \sum_i \mu(U_i) \leq c\sum_i |U_i|^s \\
\Rightarrow \quad & \mu(F) \leq c\sum_i |U_i|^s \\
\Rightarrow \quad & \frac{\mu(F)}{c} \leq \sum_i |U_i|^s
\end{aligned}
$$

So $\mathcal{H}^s_\delta(F) \geq \dfrac{\mu(F)}{c}$ for sufficiently small δ. Letting $\delta \to 0$ yields $\mathcal{H}^s(F) \geq \dfrac{\mu(F)}{c}$. Now $\mathcal{H}^s(F) > 0$ since $\mu(F) > 0$, thus $\mathcal{H}^s(F) = s$ or $\mathcal{H}^s(F) = \infty$. Therefore $\mathcal{H}^s(F) \geq s$, so $\dim_{\mathrm{H}} F \geq s$.

$\blacksquare$

Now let's look at an example using the mass distribution principle.

Example 4.23 Let F be the Cantor set. Recall that E_k denotes the kth iteration when constructing the Cantor set. Let μ be the mass distribution on F such that at each level, each interval of length 3^{-k} has mass 2^{-k}. Then $\mu(E_0) = 1$, $\mu(E_1) = \frac{1}{2} + \frac{1}{2} = 1$, $\mu(E_2) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1$, etc. Let U be a set with $|U| < 1$ and let k be a non-negative such that $3^{-(k+1)} \leq |U| < 3^{-k}$. Then U can intersect at most one interval of E_k, so

$$\mu(U) \leq 2^{-k} = (3^{\log_3 2})^{-k} = (3^{-k})^{\frac{\ln 2}{\ln 3}} \leq (3|U|)^{\frac{\ln 2}{\ln 3}}$$

So the dimension, s, is the power of the final piece of the above inequality. Note that

$$(3|U|)^{\frac{\ln 2}{\ln 3}} = 3^{\frac{\ln 2}{\ln 3}}|U|^{\frac{\ln 2}{\ln 3}} = 3^{\log_3 2}|U|^{\frac{\ln 3}{\ln 2}} = 2|U|^{\frac{\ln 2}{\ln 3}}.$$

So for the sake of the mass distribution principle, $c = 2$ and $s = \dfrac{\ln 2}{\ln 3}$. Thus by applying the mass distribution principle, $\mathcal{H}^{\frac{\ln 2}{\ln 3}}(F) \geq \dfrac{\mu(F)}{c} = \dfrac{1}{2}$ and $\dim_{\mathrm{H}} F = \dfrac{\ln 2}{\ln 3}$. $\square$

Now although box dimension has its benefits in that it is relatively easy to calculate, it has some major problems. At first, the following theorem may seem very useful, but there are some incredibly undesirable consequences as well.

Theorem 4.24 If $\overline{F}$ is the closure of F, then

$$\underline{\dim}_{\mathrm{B}}\overline{F} = \underline{\dim}_{\mathrm{B}}F \text{ and } \overline{\dim}_{\mathrm{B}}\overline{F} = \overline{\dim}_{\mathrm{B}}F.$$

Proof:

Let $B_1, B_2, ..., B_i$ be a finite collection of closed balls, each with radius δ. Then $\bigcup_{i=1}^{k} B_i$ is closed since it is a finite union of closed sets. So if $F \subseteq \bigcup_{i=1}^{k} B_i$, then $\overline{F} \subseteq \bigcup_{i=1}^{k} B_i$ also by definition of closure. So the smallest number of sets needed to cover F will also be the smallest number of sets needed to cover its closure $\overline{F}$. Then $\limsup_{\delta \to 0} \dfrac{\ln N_\delta(F)}{\ln \delta} = \limsup_{\delta \to 0} \dfrac{\ln N_\delta(\overline{F})}{\ln \delta}$ and the same is true for the lower limits. Thus $\overline{\dim}_{\mathrm{B}} F = \overline{\dim}_{\mathrm{B}} \overline{F}$ and $\underline{\dim}_{\mathrm{B}} B = \underline{\dim}_{\mathrm{B}} \overline{F}$.

∎

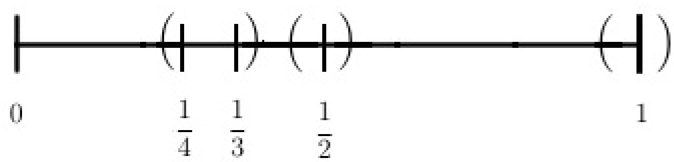
Let's investigate a devastating example of this theorem at work. Let $F = \mathbb{Q} \cap [0,1]$. Then $\overline{F} = [0,1]$. So since F is totally disconnected and is countable, F has Hausdorff dimension 0. But $\overline{F}$ has box dimension 1 since it is a line segment. So by the above theorem, F has box dimension 1 instead of box dimension 0. This means that any countable set could have non-zero box dimension. So a countable number of points in a set could cause the value of the box dimension to be much different than the value for the Hausdorff dimension for the set. In the following example, we have a subset of the rational numbers that has a different box dimension than the rationals. A subset of the rational numbers is still going to be countable, so in theory the dimension should be the same.

Example 4.25 Let $F = \{0, 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, ...\}$. To calculate the box dimension, let $0 < \delta < \dfrac{1}{2}$. Also let $k \in \mathbb{Z}$ such that $\dfrac{1}{k(k+1)} \leq \delta < \dfrac{1}{k(k-1)}$. Suppose $|U| \leq \delta$. Then U could only cover at most one point in $\{1, \frac{1}{2}, \frac{1}{3}, ..., \frac{1}{k}\}$, so we would need at least k sets with diameter δ to cover F. So we have $N_\delta(F) \geq k$ and calculating the lower box dimension gives us

$$\liminf_{\delta \to 0} \frac{\ln N_\delta(F)}{-\ln \delta} \geq \liminf_{k \to \infty} \frac{\ln k}{\ln(k(k-1))} = \liminf_{k \to \infty} \frac{\ln k}{\ln k + \ln(k-1)} = \frac{\frac{1}{k}}{\frac{1}{k} + \frac{1}{1-k}} = \frac{1}{2}.$$

Now let's find the upper box dimension. Again, suppose that $0 < \delta < \frac{1}{2}$ and

$k \in \mathbb{Z}$ such that $\dfrac{1}{k(k+1)} \leq \delta < \dfrac{1}{k(k-1)}$. Then k sets can cover the first $k+1$ points in the set. For an idea of what is happening, consider the case when $k = 3$. Then $\dfrac{1}{12} < \delta < \dfrac{1}{6}$. The following picture indicates that the first $3 = k$ sets of size δ can cover $4 = k+1$ elements. (Note that this picture is not completely drawn to scale. Each interval shown has length δ.)



The rest of the points can also be covered by k sets, keeping in mind that the set covering 0 will cover infinitely many points. So we have a total of $k + k = 2k$ sets for an upper bound. Thus $N_\delta(F) \leq 2k$. Therefore

$$\limsup_{\delta\to 0} \frac{\ln N_\delta(F)}{-\ln \delta} \leq \limsup_{k\to\infty} \frac{\ln 2k}{\ln(k(k-1))} = \limsup_{k\to\infty} \frac{\ln 2 + \ln k}{\ln k + \ln(k-1)} = \lim_{k\to\infty} \frac{\frac{1}{k}}{\frac{1}{k} + \frac{1}{k-1}} = \frac{1}{2}.$$

So since $\dfrac{1}{2} \leq \underline{\dim}_{\mathrm{B}} F \leq \dim_B F \leq \overline{\dim}_{\mathrm{B}} F \leq \dfrac{1}{2}$, we have that $\dim_B F = \dfrac{1}{2}$ by the Sandwich Theorem.

$\square$

So just a countable set of points can change the value of the box dimension. This is such a problem that we need to modify our definition of box dimension. The modified box dimension is hard to deal with because the difficulties that arose in calculating the Hausdorff dimension return when calculating the modified box dimension. Since modified box dimension is just as hard to work with as Hausdorff dimension, we won't discuss modified box dimension in this thesis.

Chapter 5

Iterated Function Systems

The ability to construct a fractal mathematically is incredibly useful and can be used as a way of simplifying the calculation of the dimension. Fractals can be defined using some form of chaotic function or some linear transformation on some starting set. These are called iterated function systems, which work by applying the linear transformation again and again to the original set. This is especially useful when dealing with self-similar fractals. Since a shape that is self-similar is an image with smaller and smaller copies of itself, it can be defined using a special type of function called a contraction.

Definition 5.1 Let $D \subseteq \mathbb{R}^n$. Then the mapping $S : D \to D$ is a *contraction* on D if $|S(x) - S(y)| \leq c|x - y|$ for all $x, y \in D$ and for some $c \in \mathbb{R}$ with $0 < c < 1$. All contractions are continuous. Note that all contractions also satisfy the Hölder condition.

A certain set of contractions can be used to generate a fractal. This set of contractions is defined as follows.

Definition 5.2 An *Iterated Function System* (IFS) is a finite collection of contractions $\{S_1, S_2, ..., S_m\}$, where $m \geq 2$.

Iterated functions systems give us a way of encoding a particular fractal. This makes it easy to communicate how a fractal can be generated mathematically. Also, some of our definitions from chaos theory can be defined in a new way using iterated function systems. Consider this new definition for an attractor.

Definition 5.3 If $F \subseteq D$ and F is nonempty and compact, then F is an *attractor* of an IFS if $F = \bigcup_{i=1}^{m} S_i(F)$.

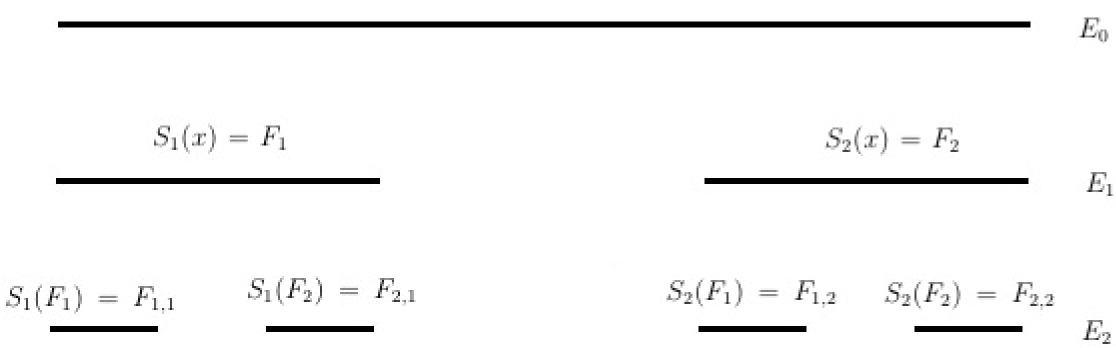
So an attractor is the actual fractal - the "last" stage of iteration. It is the image that the sequence of iterations approaches. The Cantor set is the limiting image of all of the iterations E_n, where E_n is the n^{th} stage of iteration. The opposite of an attractor is a *repellor*. The repellor is a set that as a function is iterated, some points (though perhaps not every point) near the repellor are iterated further and further away from the repellor. Attractors and repellors are either fixed or periodic points, so the actual attractor and repellor do not move under iteration.

Before we get to the theorems, let's look at a couple examples to get a feel for how these work. Let's construct the middle third Cantor set by using an IFS. The easiest way to figure out how to define the function is by looking at what happens to the end points, 0 and 1. Since we have two halves of the Cantor set, we will want the IFS to consist of two functions. Since fractals are defined by iteration, the domain of the function will change after each iteration. For this example, since the initiator, E_0, is the segment $[0, 1]$, this will be the domain for the first iteration.

Let $S_1, S_2 : E_n \to [0, 1]$ be defined such that

$$S_1(x) = \frac{1}{3}x \text{ and } S_2(x) = \frac{1}{3}x + \frac{2}{3},$$

where n is the nth iteration of the function. This way of defining the functions makes sense since because if we consider S_1, then the farthest point to the right, 1, is mapped to $\frac{1}{3}$, which is the farthest point to the right in the left half of the Cantor set. For S_2, 1 gets mapped to 1. Similarly, if we take the farthest point to the left, 0, then $S_1(0) = 0$ and $S_2(0) = \frac{2}{3}$. So for the first iteration, $S_1(E_0)$ is the line segment that starts at 0 and ends at $\frac{1}{3}$, and $S_2(E_0)$ is the line segment that starts at $\frac{2}{3}$ and ends at 1. Define $S_1(E_0) = F_1$ and $S_2(E_0) = F_2$. Now repeat the process, so take $S_1(E_1)$ and $S_2(E_1)$. Since there are two pieces of E_1, each function S_i will yield two pieces each. Define $S_1(F_1) = F_{1,1}$, $S_1(F_2) = F_{2,1}$, $S_2(F_1) = F_{1,2}$, and $S_2(F_2) = F_{2,2}$. See the diagram below to see how each segment is produced. Continuing this process infinitely many times, we obtain the attractor, F, for this IFS, which is the Cantor set.



Example 5.4 Let us now construct the Cantor dust using an iterated function system. The easiest way to construct the functions for this example is by looking at what happens at the corners. Since four squares get generated after the first iteration, we should have four contractions. Let us then define

$S_1, S_2, S_3, S_4 : E_n \to [0,1] \times [0,1]$ by

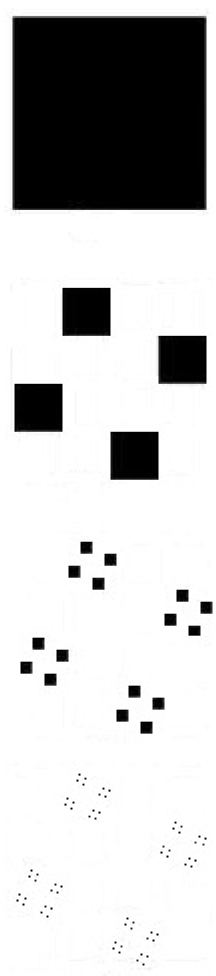
$$S_1(x,y) = \left(\frac{1}{4}x, \frac{1}{4}y + \frac{1}{2}\right)$$

$$S_2(x,y) = \left(\frac{1}{4}x + \frac{1}{4}, \frac{1}{4}y\right)$$

$$S_3(x,y) = \left(\frac{1}{4}x + \frac{1}{2}, \frac{1}{4}y + \frac{3}{4}\right)$$

$$S_4(x,y) = \left(\frac{1}{4}x + \frac{3}{4}, \frac{1}{4}y + \frac{1}{4}\right)$$

So for the first iteration, each function will map the unit square to one of the four squares in E_1. Continuing the process infinitely will lead us to the attractor, which is the Cantor dust.



□

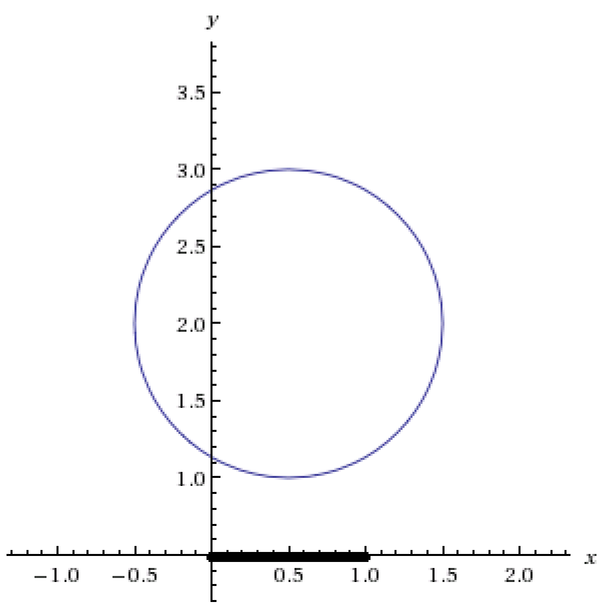
Every IFS has a unique attractor, but in order to prove this, we will need a way of measuring the distance between two sets. Since distance is measured using a metric space, we will use a metric called the Hausdorff metric.

Definition 5.5 Let $D \subseteq \mathbb{R}^n$ and let $\mathcal{S}$ be the set of all non-empty compact subsets of D. Let $A, B \subseteq \mathcal{S}$. Then the *Hausdorff metric* on $\mathcal{S}$ is defined as

$$d(A, B) = \inf\{\delta : A \subseteq N_\delta(B) \text{ and } B \subseteq N_\delta(A)\}.$$

Recall that $N_\delta(A)$, the δ-neighborhood of A, is the set of points within a distance of δ from A.

Example 5.6 Let A be the circle defined by $\left(x - \frac{1}{2}\right)^2 + (y-2)^2 = 1$ and let B be the interval [0,1]. To find the Hausdorff distance between A and B, we must look at the neighborhoods of each set. In order for $A \subseteq N_\delta(B)$, we would need $\delta > 3$. The distance between the bottom point on the circle ($\frac{1}{2}$,1) and (0,0) is $\frac{\sqrt{5}}{2}$. So in order for $B \subseteq N_\delta(A)$, we would need $\delta > \frac{\sqrt{5}}{2}$. So the Hausdorff distance between A and B is $\frac{\sqrt{5}}{2}$ on taking the infimum.



□

Let's briefly verify the metric axioms to verify that the Hausdorff metric is a metric.

1. If $A \neq B$, then $d(A, B) > 0$ since $\delta > 0$.

2. $d(A, B) = 0$ if and only if $\delta = 0$ for both $N_\delta(A)$ and $N_\delta(B)$ if and only if $A = B$.

3. Clearly, $d(A, B) = d(B, A)$ by how d is defined.

4. The last axiom is the most difficult to prove. Let $a \in A$. Then there is some $b \in B$ such that $d(a, b) \leq d(A, B)$. Likewise, for this same b, there is some $c \in C$ such that $d(b, c) \leq d(B, C)$. Therefore

$$d(a, b) + d(b, c) \leq d(A, B) + d(B, C)$$

by adding the inequalities. But since $d(a, b)$ and $d(b, c)$ are real numbers, by the triangle inequality we have

$$d(a, c) \leq d(a, b) + d(b, c) \leq d(A, B) + d(B, C).$$

So there is a $c \in C$ for each $a \in A$ such that $A \subseteq N_{d(a,c)}(C)$ and vice versa. Hence $d(A, C) \leq d(A, B) + d(B, C)$.

Given an IFS, are we guaranteed that there is an attractor? And if there is one, could there be multiple attractors? The following theorem addresses this issue, and proves that given any IFS, there will always be one and only one attractor.

Theorem 5.7 Let $\{S_1, S_2, ..., S_n\}$ be a set of contractions that defines an IFS on some domain $D \subseteq \mathbb{R}^n$. So $|S_i(x) - S_i(y)| \leq c_i|x - y|$ for all $1 \leq i \leq n$.

Then there exists a unique attractor $F = \bigcup_{i=1}^{n} S_i(F)$. Define a transformation S on the class $\mathcal{S}$ of nonempty compact sets by $S(E) = \bigcup_{i=1}^{m} S_i(E)$ for $E \in \mathcal{S}$ and let S^k denote the k^{th} iterate of S. Then $F = \bigcap_{k=1}^{\infty} S^k(E)$, for every $E \in \mathcal{S}$ such that $S_i(E) \subset E$ for all i.

Proof:

Let $\mathcal{S}$ denote the set of nonempty compact subsets D. Let $E \in \mathcal{S}$ such that $S_i(E) \subseteq E$ for all i. Then $S_i^k(E) \subseteq S_i^{k-1}(E)$ since S_i is a contraction. So we have the decreasing sequence $\left\{S_i^k(E)\right\}_{i=1}^{\infty}$. Since each $S^k(E)$ is nonempty and compact and since $S_i^k(E) \subseteq S_i^{k-1} \subseteq ... \subseteq S_i(E) \subseteq E$, the intersection $F_E = \bigcap_{k=1}^{\infty} S_i^k(E)$ is also nonempty and compact. Then since $F_E = \bigcap_{k=1}^{\infty} S^k(E)$, we have that $S_i(F_E) = S_i\left(\bigcap_{k=1}^{\infty} S^k(E)\right) = F_E$ for each i. Thus $\bigcup_{i=1}^{n} S_i(F_E) = F_E$, hence F_E is an attractor and E generates an attractor.

To prove uniqueness, let $A, B \in \mathcal{S}$ such that A and B are both attractors of the IFS. By definition of an attractor, we have $A = \bigcup_{i=1}^{n} S_i(A)$ and $B = \bigcup_{i=1}^{n} S_i(B)$. Since $A, B \in \mathcal{S}$, we have

$$d(S(A), S(B)) = d\left(\bigcup_{i=1}^{m} S_i(A), \bigcup_{i=1}^{m} S_i(B)\right) \leq \max_{1 \leq i \leq m} d(S_i(A), S_i(B))$$

Then since each S_i is a contraction, we have

$$d(A, B) = d(S(A), S(B)) \leq \max_{1 \leq i \leq m} c_i \cdot d(A, B)$$

. But since each $c_i < 1$, then $d(A, B) \leq d(A, B)$. Since distance is non-negative, this can only happen if $d(A, B) = 0$, therefore $A = B$.

∎

There are two main purposes for an iterated function system. One thing we can do with an IFS is encode the information for generating a certain attractor. This information can then be decoded by displaying the image of the attractor. In order to encode an IFS, we will need to use linear transformations.

Definition 5.8 A *linear transformation* is a mapping $T : \mathbb{R}^n \to \mathbb{R}^n$ with

1. $T(x + y) = T(x) + T(y)$ and

2. $T(\alpha x) = \alpha T(x)$ for all $\alpha \in \mathbb{R}$

In terms of matrices, if T is a linear transformation in $\mathbb{R}^n$, then

$$T(\mathbf{x}) = A\mathbf{x}$$

where $\mathbf{x}$ is a vector in $\mathbb{R}^n$ and A is some matrix. For example, in $\mathbb{R}^2$ we would have $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}$$

Definition 5.9 A mapping $S : \mathbb{R}^n \to \mathbb{R}^n$ is an *affine transformation* if

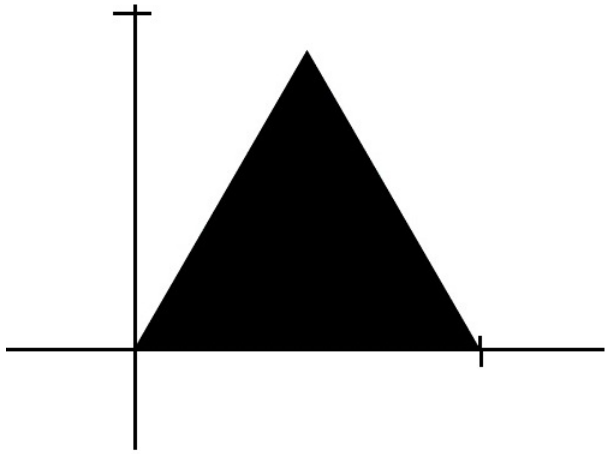
$$S(x) = T(x) + \mathbf{b},$$

where T is a linear transformation in $\mathbb{R}^n$ and $\mathbf{b}$ is a vector in $\mathbb{R}^n$.

In general, $S(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$. In $\mathbb{R}^2$, affine transformations will have the form

$$S\begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} e \\ f \end{bmatrix}.$$

Example 5.10 Let $A = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$. Let F_0 be the solid equilateral triangle with base 1 and height $\dfrac{\sqrt{3}}{2}$ situated as follows:



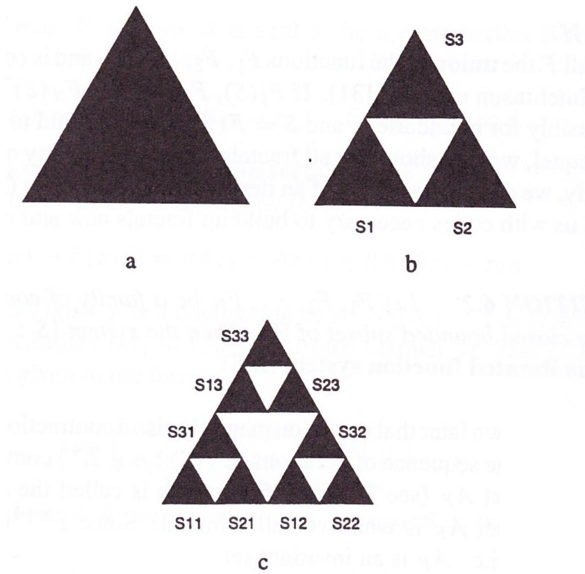
This triangle is acted on by the following contractions.

$$S_1 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$S_2 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix}$$

$$S_3 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \frac{1}{4} \\ \frac{\sqrt{3}}{4} \end{bmatrix}$$

Then $\{S_1, S_2, S_3\}$ is an IFS. The first few iterations will look like the following picture.



So the sequence of iterations converges to the Sierpinski triangle. Thus the Sierpinski triangle is the attractor, F, for this IFS.

Since each S_i is an affine transformation, they each have the form

$$S_i = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} e \\ f \end{bmatrix}.$$

We can encode an IFS by simply labeling what $a, b, c, d, e, \text{and} f$ are for each S_i. For the case of the Sierpinski triangle, the code is

	a	b	c	d	e	f
S_1	$\frac{1}{2}$	0	0	$\frac{1}{2}$	0	0
S_2	$\frac{1}{2}$	0	0	$\frac{1}{2}$	$\frac{1}{2}$	0
S_3	$\frac{1}{2}$	0	0	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{\sqrt{3}}{4}$

$\square$

One useful type of linear transformation is a rotation matrix. A rotation matrix takes a point in $\mathbb{R}^n$ and rotates it by a certain angle. Since we are working with contractions, we will use the rotation matrix to rotate a point, then we will also scale the point by a certain value. In $\mathbb{R}^2$, if a point is rotated about the origin and scaled, the rotation and scaling can be represented by the matrix

$$A = \begin{bmatrix} r\cos\theta & -s\sin\theta \\ r\sin\theta & s\cos\theta \end{bmatrix}$$

The factor r scales the x value by some amount and the s value scales the y value. If $0 < r < 1$ and $0 < s < 1$, then the transformation will be a contraction (the point will be nearer to the origin), and if $r > 1$ and $s > 1$, then the transformation will be a dilation (the point will move further from the origin). If $r = s = 1$, then the point will only rotate about the origin but keep the same distance from the origin. So we can rotate a point (x, y) around the origin by a certain angle by solving the following equation for (x', y'):

$$\begin{bmatrix} r\cos\theta & -s\sin\phi \\ r\sin\theta & s\cos\phi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix}$$

Example 5.11 Let's find the IFS for the Koch curve. Note that the generator of the Koch curve has four pieces, so there will be 4 transformations on the line segment $[0, 1]$. When creating this IFS, some of the segments will need to be rotated. A line segment can be rotated by using some simple trigonometry. The first segment is the easiest since we can just scale down the segment [0,1] by $\frac{1}{3}$. To obtain the next piece, we will need to scale down by a factor of $\frac{1}{3}$ and rotate the segment by 60 degrees or $\dfrac{\pi}{3}$ radians and translate. For the third piece we will need to scale down by a factor of $\frac{1}{3}$ and rotate the segment by 120 degrees or $\dfrac{2\pi}{3}$ and translate. The fourth piece is like the first piece, just translated by $\dfrac{2}{3}$ to the left. These are explicitly represented

$$S_1\begin{pmatrix}x\\y\end{pmatrix} = \begin{bmatrix}\frac{1}{3} & 0\\ 0 & \frac{1}{3}\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}$$

$$\begin{aligned}
S_2\begin{pmatrix}x\\y\end{pmatrix} &= \frac{1}{3}\begin{bmatrix}\cos\frac{\pi}{3} & -\sin\frac{\pi}{3}\\ \sin\frac{\pi}{3} & \cos\frac{\pi}{3}\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} + \begin{bmatrix}\frac{1}{3}\\0\end{bmatrix}\\
&= \frac{1}{3}\begin{bmatrix}\frac{1}{2} & -\frac{\sqrt{3}}{2}\\ \frac{\sqrt{3}}{2} & \frac{1}{2}\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} + \begin{bmatrix}\frac{1}{3}\\0\end{bmatrix}\\
&= \begin{bmatrix}\frac{1}{6} & -\frac{\sqrt{3}}{6}\\ \frac{\sqrt{3}}{6} & \frac{1}{6}\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} + \begin{bmatrix}\frac{1}{3}\\0\end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
S_3\begin{pmatrix}x\\y\end{pmatrix} &= \frac{1}{3}\begin{bmatrix}\cos\frac{2\pi}{3} & -\sin\frac{2\pi}{3}\\ \sin\frac{2\pi}{3} & \cos\frac{2\pi}{3}\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} + \begin{bmatrix}\frac{2}{3}\\0\end{bmatrix}\\
&= \frac{1}{3}\begin{bmatrix}-\frac{1}{2} & -\frac{\sqrt{3}}{2}\\ \frac{\sqrt{3}}{2} & -\frac{1}{2}\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} + \begin{bmatrix}\frac{2}{3}\\0\end{bmatrix}\\
&= \begin{bmatrix}-\frac{1}{6} & -\frac{\sqrt{3}}{6}\\ \frac{\sqrt{3}}{6} & -\frac{1}{6}\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} + \begin{bmatrix}\frac{2}{3}\\0\end{bmatrix}
\end{aligned}$$

$$S_4\begin{pmatrix}x\\y\end{pmatrix} = \begin{bmatrix}\frac{1}{3} & 0\\ 0 & \frac{1}{3}\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} + \begin{bmatrix}\frac{2}{3}\\0\end{bmatrix}$$

This can be represented by the following code for this IFS. Observe that it is easier just to state the code rather than give all the functions.

$$
\begin{array}{ccccccc}
 & a & b & c & d & e & f \\
S_1 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 \\
S_2 & \frac{1}{6} & -\frac{\sqrt{3}}{6} & \frac{\sqrt{3}}{6} & \frac{1}{6} & \frac{1}{3} & 0 \\
S_3 & -\frac{1}{6} & -\frac{\sqrt{3}}{6} & \frac{\sqrt{3}}{6} & \frac{1}{6} & \frac{2}{3} & 0 \\
S_4 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{2}{3} & 0
\end{array}
$$

$\square$

There are objects in nature, such as trees and ferns, that can be constructed using an IFS. We can use the following theorem to construct such shapes. The theorem was first proved by Michael Barnsley and Lyman Hurd in 1993. The fractal construction of the fern is attributed to Barnsley and is also named after him.

Theorem 5.12 Collage Theorem: Let $\{S_1, S_2, ..., S_n\}$ be an IFS, so

$$|S_i(x) - S_i(y)| \leq c|x - y|$$

for all $x, y \in D \subseteq \mathbb{R}^n$, for all $i \in \{1, ..., n\}$, and for some $0 < c < 1$. Let $E \subseteq D$ be a compact set. Then

$$d(E, F) \leq d\left(E, \bigcup_{i=1}^{m} S_i(E)\right) \frac{1}{1-c}$$

where F is the attractor of the IFS and d is the Hausdorff metric.

Proof:

$$\begin{aligned}
d(E,F) &\le d\left(E, \bigcup_{i=1}^{m} S_i(E)\right) + d\left(\bigcup_{i=1}^{m} S_i(E), F\right) && \text{by the triangle inequality} \\
&= d\left(E, \bigcup_{i=1}^{m} S_i(E)\right) + d\left(\bigcup_{i=1}^{m} S_i(E), \bigcup_{i=1}^{m} S_i(F)\right) && \text{by definition of an attractor} \\
&\le d\left(E, \bigcup_{i=1}^{m} S_i(E)\right) + cd(E,F) && \text{since each } S_i \text{ is a contraction}
\end{aligned}$$

Then solving the inequality for $d(E,F)$, we have

$$\begin{aligned}
d(E,F) &\le d\left(E, \bigcup_{i=1}^{m} S_i(E)\right) + cd(E,F) \\
d(E,F) - cd(E,F) &\le d\left(E, \bigcup_{i=1}^{m} S_i(E)\right) \\
d(E,F)(1-c) &\le d\left(E, \bigcup_{i=1}^{m} S_i(E)\right) \\
d(E,F) &\le \frac{1}{1-c} \quad d\left(E, \bigcup_{i=1}^{m} S_i(E)\right).
\end{aligned}$$

∎

The collage theorem gives a way of going backwards in terms of iterated function systems. We have seen how we can produce an image from a given IFS, but how can we create an IFS given a certain image, such as a fern or a tree? If we have guessed an IFS that seems to be working, how do we know that this IFS will still hold for the image at infinity?

Here is what the collage theorem is doing. We start with some target image in the real world that we want to create using an IFS. Call this target image A. Then our initiator, or starting set, will be acted on by the functions S_i. Then if the Hausdorff distance between the target image and the Nth iteration of our initiator is "close", that is, if $d\left(A, \bigcup_{i=1}^{N} S_i(A)\right) < \epsilon$, then this gives a good approximation for the attractor, F, so $d(A,F) < \dfrac{\epsilon}{1-c}$.

Example 5.13 Construction of a fern [2]

The Barnsley fern is constructed from four affine transformations. There are many different varieties of the fern, depending on the value of the entries for the affine transformation matrices. Though technically these could be constructed by hand, the amount of iterations necessary are so many, that a computer is needed to construct the image in a timely fashion.

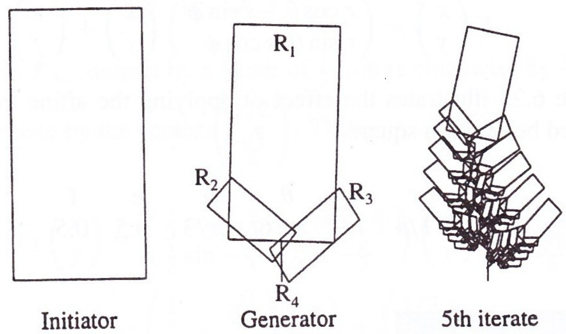
We can construct the Barnsley fern using an IFS. Each function in the IFS will have the form

$$S_i = \begin{bmatrix} r\cos\theta & -s\sin\phi \\ r\sin\theta & s\cos\phi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} e \\ f \end{bmatrix}$$

The initiator of the IFS is a rectangle. Now depending on the desired appearance of the fern, the values for the constants of each S_i may vary. One such IFS code acting on the rectangle (where the angles are in degrees) is

	r	s	θ	ϕ	e	f
S_1	0.85	0.85	-2.5	-2.5	0	1.6
S_2	0.3	0.34	49	49	0	1.6
S_3	0.3	0.37	120	-50	0	0.44
S_4	0	0.16	0	0	0	0

The following figure shows how each of the four functions is working. Notice how each iteration generates four new pieces.



Act on each of the pieces with the same functions over and over again, and we will start to get close to the desired fern. This is where the collage theorem comes in to play. When have we gone far enough in our iterations? We have gone far enough when the difference between the target image and the produced image is less than some value ϵ. The way this works is that we choose the ϵ and target image at the beginning. After iterating, we know we have gone far enough once the iterations and the target image differ by the chosen ϵ.

$\square$

Using iterated function systems and the collage theorem is how some computer graphics are created. An artist has a particular landscape in mind, perhaps from reality, then to create it in the digital world, they need to apply some set of functions to create the landscape. Many popular movies have fractal computerized graphics. For example, one such movie is Star Wars episode III Revenge of the Sith. A particular example is towards the end of the movie on the lava planet, as illustrated in the following picture. The lava flows were computer generated using several layers of fractal images.

Using the IFS, the dimension of fractals satisfying a certain condition can be easily calculated. Recall, though, that an IFS can only be defined for a fractal that is self-similar. The condition required for the calculation of these fractals is called the *open set condition*.

Definition 5.14 Each S_i for the IFS has the *open set condition* if there is a non-empty, bounded open set V where $\bigcup_{i=1}^{m} S_i(V) \subset V$ and the union is disjoint.

If a fractal has this condition, then the dimension of a fractal F is the solution for s given by the equation

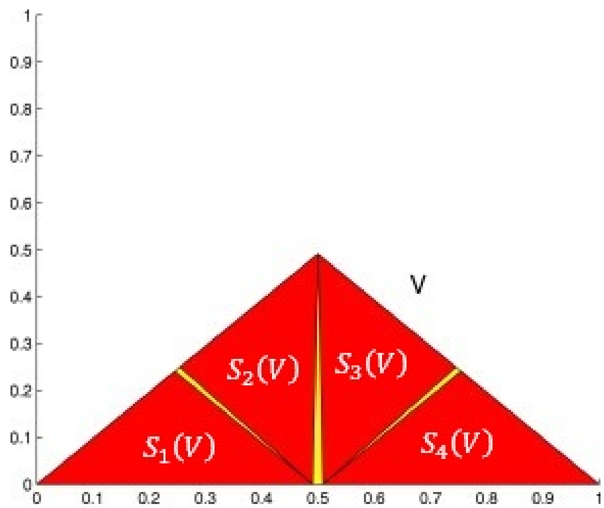
$$\sum_{i=1}^{m} c_i^s = 1$$

where c_i is the contracting factor for the IFS. This value for s is $s = \dim_{\mathrm{H}}(F)$, the value where the Hausdorff measure jumps from infinity to 0. This condition ensures that the values of the contractions S_i do not overlap too much. Let's look at a few examples to illustrate the open set condition.

Example 5.15 If F is the Cantor set, let $V = (0, 1)$. Then $\bigcup_{i=1}^{2} S_i(V)$ is a subset of $(0, 1)$ since both functions scale down V and $S_1(V)$ and $S_2(V)$ are disjoint by how each S_i is defined. Thus this satisfies the open set condition.

□

Example 5.16 Let F be the Koch curve. Recall that the curve is constructed by removing the middle third of the unit interval [0,1], and replacing it with an equilateral triangle with length $\frac{1}{3}$ and removing the base. Note that the highest point of the first iteration of the Koch curve is the point $(\frac{1}{2}, \frac{\sqrt{3}}{2})$. So let V be the interior of an isosceles triangle with no boundary and with base 1 and height $\frac{\sqrt{3}}{2}$. Since the triangle has no boundary, it is an open set in $\mathbb{R}^2$. Using the IFS described for the Koch curve above, apply each S_i to V. Each $S_i(V)$ will be a solid triangle with no boundary that has been scaled down and rotated. So $S_1(V)$ will take the triangle and scale it down by $\frac{1}{3}$. The resulting image of each $S_i(V)$ is labeled in the figure below.



Note that since we are not including the boundary of the triangle, none of the images of $S_i(V)$ intersect. Therefore, $\bigcup_{i=1}^{4} S_i(V)$ is disjoint. This union is also contained in the original set V. Thus the open set condition is satisfied for the Koch curve. $\square$

Proposition 5.17 Suppose that the open set condition holds for contractions $\{S_1, S_2, ..., S_m\}$ with ratios $0 < c_i < 1$. If F is the attractor for the IFS, then $\dim_{\mathrm{H}}(F) = \dim_B(F) = s$, where s is the solution to the equation

$$\sum_{i=1}^{m} c_i^s = 1$$

and for this value of s, we have $0 < \mathcal{H}^s(F) < \infty$.

The proof of this proposition can be found in [5].

Example 5.18 Recall that the IFS for the Cantor Set is $S_1(x) = \dfrac{1}{3}x$ and $S_2(x) = \dfrac{1}{3}x + \dfrac{2}{3}$. We have already shown that the Cantor set satisfies the open set condition, which means we can calculate the dimension using the equation

$\displaystyle\sum_{i=1}^{m} c_i^s = 1$ for s. Each stage of the Cantor set is scaled by $c = \frac{1}{3}$, so

$$\sum_{i=1}^{2} \left(\frac{1}{3}\right)^s = 1$$
$$2 \cdot \left(\frac{1}{3}\right)^s = 1$$
$$\left(\frac{1}{3}\right)^s = \frac{1}{2}$$
$$s \ln \frac{1}{3} = \ln \frac{1}{2}$$
$$s = \frac{\ln \frac{1}{2}}{\ln \frac{1}{3}}$$
$$s = \frac{\ln 2}{\ln 3}$$

□

Example 5.19 Since the Koch curve satisfies the open set condition, as stated above, we can calculate its dimension. Each c_i for the Koch curve is $\frac{1}{3}$. So

$$\sum_{i=1}^{4} \left(\frac{1}{3}\right)^s = 1$$
$$4 \cdot \left(\frac{1}{3}\right)^s = 1$$
$$\left(\frac{1}{3}\right)^s = \frac{1}{4}$$
$$s \ln \frac{1}{3} = \ln \frac{1}{4}$$
$$s = \frac{\ln \frac{1}{4}}{\ln \frac{1}{3}}$$
$$s = \frac{\ln 4}{\ln 3}$$

□

So we have yet another method for calculating the dimension of a fractal and a method for constructing self-similar fractals mathematically. But what if the fractal is extremely complex? Some fractals cannot be constructed by using an iterated function system. Some of these fractals are called Julia sets, which are incredibly beautiful and intricate. We will finally be able to discuss these sets in the following chapter.

Chapter 6

Julia Sets

Julia sets were first discovered and studied by Gaston Julia around the time of World War I. These are incredibly intricate shapes, but Julia did not have the luxury of studying them with the help of a computer. They were not well developed until the 1970s, when Mandelbrot was able to study them in more detail using a computer to generate images. It was then that he discovered the famous Mandelbrot set, which is a Julia set. The most interesting Juila sets are fractals and are usually constructed using a complex function. To construct Julia sets we will need to use difference equations for our functions. Before we actually start studying Julia sets, let's review a bit of complex numbers and complex functions.

The *modulus* of a complex number $z = x + iy = re^{i\theta}$ is $|z| = \sqrt{x^2 + y^2}$. This is the distance that z is from the origin. The argument, denoted $\arg(z)$ is the angle at which z lies, moving counterclockwise from the positive x-axis. In terms of polar coordinates, we can define $r = |z|$ and $\theta = \tan^{-1}\left(\frac{y}{x}\right)$, so θ is the argument of z. Let's look at a quick example of this.

Example 6.1 Let $z = 3+4i$. Then z is the point $(3, 4)$ on the complex plane.

The modulus of z is $|z| = \sqrt{3^2 + 4^2} = 5$, so z is 5 units from the origin. Also, $\arg(z) = \tan^{-1} \frac{5}{4}$ is the angle that z is lifted from the real axis. $\square$

Definition 6.2 A function f is called *analytic* if its derivative exists at all points in its domain. Just like with real numbers, the complex derivative is defined as $\lim_{\zeta \to 0} \frac{f(z+\zeta) - f(z)}{\zeta}$, where $z \in \mathbb{C}$ and $\zeta \in \mathbb{C}$. For the purposes of this thesis, we will only be working with functions whose domain is the complex numbers.

Definition 6.3 A *critical point* z for a function f is a point where $f'(z) = 0$.

Most of the theorems of chaos theory that we used for real numbers still hold for complex numbers, with perhaps some slight alterations. Just like with real numbers, fixed points to a complex difference equation are solutions to the equation $f(z) = z$ and k-periodic points are solutions to the equation $f^k(z) = z$. The proofs for each of these mimics the proof of the corresponding proposition for the real numbers.

Proposition 6.4 Let z^* be a fixed point of an analytic function f. Then

1. If $|f'(z^*)| < 1$, then z^* is asymptotically stable. (This means that z^* will be an element of the attractor, in terms of an associated iterated function system.)

2. If $|f'(z^*)| > 1$, then z^* is unstable. (This means that z^* is an element of the repellor, in terms of iterated function systems.)

Proposition 6.5 Let z be a k-periodic point of an analytic function f. Then

1. If $|f'(z)f'(f(z))f'(f(f(z)))...f'(f^{k-1}(z))| < 1$, then z is asymptotically stable.

2. If $|f'(z)f'(f(z))f'(f(f(z)))...f'(f^{k-1}(z))| > 1$, then z is unstable.

The proof of these propositions can be found in [4]. If z is an attracting point, then there is a $\delta > 0$ such that any point in the δ-neighborhood centered at z will converge to z under iteration of f. That is, $f^k(w) \to z$ as $k \to \infty$. If z is a repelling point, then for any $\delta > 0$, there exists a point w in the δ-neighborhood around z such that $f^k(w) \nrightarrow z$ as $k \to \infty$.

Example 6.6 Let's find the fixed points and 2-periodic points of $f(z) = z^3$ and determine their stability. The fixed points are the solutions to $f(z) = z$, so

$$z^3 = z$$

$$z^3 - z = 0$$

$$z(z^2 - 1) = 0$$

$$z(z - 1)(z + 1) = 0.$$

So the fixed points are $z_1 = 0$, $z_2 = 1$, and $z_3 = -1$. Note that $f'(z) = 3z^2$, so $|f'(z_1)| = |f'(0)| = 0 < 1$, so z_1 is stable. Likewise, $|f'(z_2)| = |f'(1)| = 3 > 1$ and $|f'(z_3)| = |f'(-1)| = 3 > 1$, so z_2 and z_3 are both unstable.

To find the 2-periodic points, solve $f^2(z) = z$, so

$$(z^3)^3 = z$$

$$z^9 - z = 0$$

$$z(z^8 - 1) = 0.$$

The solutions to this equation are $z_1 = 0$ and the eighth roots of unity. Recall that roots of unity are the n solutions to the polynomial $z^n - 1 = 0$. These solutions all lie on the unit circle and are an angle of θ apart, where θ

is given by $\dfrac{2\pi}{n}$. So for our situation, the eighth roots of unity are 1, w, w^2, w^3, w^4, w^5, w^6, and w^7, where w is given by

$$w = \cos\left(\frac{2\pi}{8}\right) + i\sin\left(\frac{2\pi}{8}\right) = \cos\left(\frac{\pi}{4}\right) + i\sin\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}$$

Now we must exclude 1 and $w^4 = -1$ since these are the fixed points that we found in the first part of this example. For the others, we need to find $|f'(w^i)f'(f(w^i))|$ in order to determine stability. For w, we have

$$|f'(w)f'(f(w))| = |(3w^2)(3w^6)| = |9w^8| = 9 > 1$$

so w is unstable. In the same way, for any w^i that we plug in, we will always get $|f'(w^i)f'(f(w^i))| = |9w^{8i}| = 9 > 1$. So each eighth root of unity, except for 1 and -1 is unstable.

Now let's determine what the cycles are. Consider $f(w)$ for a moment. We know that $f(w) = w^3$ and $f^2(w) = f(w^3) = w^9 = w$. So $\{w, w^3\}$ is a 2-cycle. For w^2, we have $f(w^2) = w^6$ and $f(w^6) = w^{18} = w^2$, so $\{w^2, w^6\}$ is a 2-cycle. The only two we have left are w^5 and w^7, so $\{w^5, w^7\}$ is a 2-cycle as well. $\square$

As stated earlier, Julia sets are fractals generated by iterations of a complex function. We are now ready for the definition of a Julia set.

Definition 6.7 Let $f : \mathbb{C} \to \mathbb{C}$ be a complex function. Then the *filled in Julia set*, $K(f)$, is defined as

$$K(f) = \{z \in \mathbb{C} : f^k(z) \text{ is bounded}\}.$$

The *Julia set*, $J(f)$, is the boundary of $K(f)$. The boundary of $K(f)$ is denoted as $\partial K(f)$.

Let $z \in \mathbb{C}$. Then $z \in J(f)$ if for any neighborhood around z, there are points a and b such that $f^k(a) \to \infty$ and $f^k(b) \nrightarrow \infty$ since z lies on the boundary of $K(f)$.

Example 6.8 Let $f(z) = z^2$. Then

$$\begin{aligned}
&\text{if } |z| > 1, f^k \to \infty \\
&\text{if } |z| = 1, f^k \to 1 \\
&\text{if } |z| < 1, f^k \to 0
\end{aligned}$$

So $K(f)$ is the filled in unit disc $|z| \leq 1$. The Julia set, $J(f)$ is the unit circle $|z| = 1$ since it is the boundary of K. In this extremely simple case, $J(f)$ is not a fractal, in the sense that a circle is not infinitely detailed. □

In our cases for Julia sets, we will mainly be looking at complex functions of the form $f_c(z) = z^2 + c$, where $c \in \mathbb{C}$ is a constant. Let $w \in \mathbb{C}$ be a periodic point of f with period p. A "quick" way of determining whether w is attracting or repelling is by looking at the complex derivative, $(f^p(w))' = \lambda$, where $f^p(w)$ means that f is composed p times. If $0 \leq |\lambda| < 1$, then w is attractive, and if $|\lambda| \geq 1$, then w is repelling. By attractive, we mean that points near w will be attracted to the orbit of w, and repelling means that points nearby will move away from the orbit.

When dealing with a given function of the form $f_c(z) = z^2 + c$, it is useful to examine the function's conjugacy class. Conjugacy is an idea usually studied in group theory, but we will need to use it a little here. In the group theory sense, conjugacy is defined as follows.

Definition 6.9 Let G be a group and let $g_1, g_2 \in G$. Then g_1 and g_2 are *conjugates* if there is an element $h \in G$ such that $hg_1h^{-1} = g_2$. We say that g_1 and g_2 are in the same *conjugacy class*.

A group is partitioned by the conjugacy classes, meaning that given an element in G, that element will belong to one conjugacy class. In our case, we will use this knowledge to aid us in classifying quadratic functions.

Let $f_c : \mathbb{C} \to \mathbb{C}$ be the complex function $f_c(z) = z^2 + c$. If $h(z) = \alpha z + \beta$ with $\alpha, \beta \in \mathbb{C}$ and $\alpha \neq 0$, then $h^{-1}(z) = \dfrac{z - \beta}{\alpha}$. Using h, we will see that f_c is conjugate to some other quadratic function f. In particular, we have

$$\begin{aligned} h^{-1}(f_c(h(z))) &= h^{-1}(f_c(\alpha z + \beta)) \\ &= h^{-1}((\alpha z + \beta)^2 + c) \\ &= h^{-1}(\alpha^2 z^2 + 2\alpha\beta z + \beta^2 + c) \\ &= \frac{\alpha^2 z^2 + 2\alpha\beta z + \beta^2 + c - \beta}{\alpha} \\ &= f(z) \end{aligned}$$

Thus f and f_c are in the same conjugacy class. Elements in the same conjugacy class behave in the same way, so if we examine the Julia set for f_c for a certain value of c, we are examining the Julia set for any function f in the conjugacy class of f_c since f_c is conjugate to f. Thus the Julia set for any quadratic polynomial is geometrically similar to f_c for some $c \in \mathbb{C}$. So we have partitioned the class of quadratic functions using conjugacy. This makes it easier to examine these functions since by studying one function, we are really studying an entire equivalence class of functions. In this way, we can potentially study all Julia sets of quadratic functions if we can find all of the conjugacy classes.

Example 6.10 Consider the function $f_0(z) = z^2$. Let $\alpha = 1$ and $\beta = 1$. Then $h^{-1}(f_0(h(z))) = f(z) = z^2 + 2z$. By solving $f(z) = z$ we find that the fixed points of f are 0 and -1. Note that $f'(z) = 2z + 2$. Since $f'(0) = 2 > 1$, 0 is a repelling fixed point, and since $f'(-1) = 0 < 1$ -1 is an attracting fixed point. The boundary of the basin of attraction for -1 is the circle $|z - 1| = 1$.

So the Julia set of f_0 is shifted left 1 to the Julia set of f. This illustrates that Julia sets are mapped to Julia sets under $h^{-1}f_c h$. $\square$

The quintessential example of a Julia set is the boundary of the Mandelbrot set. The Mandelbrot set, mentioned earlier in chapter 2 is extremely detailed.

Definition 6.11 The *Mandelbrot set* is defined as

$$M = \{c \in \mathbb{C} : J(f_c) \text{ is connected}\}$$

.

This definition of the Mandelbrot set is very difficult to work with computationally. So we will derive an equivalent definition that is easier to work with, but in order to do this we will need several theorems to prove their equivalence. These theorems will also help us achieve a better understanding for Julia sets in general.

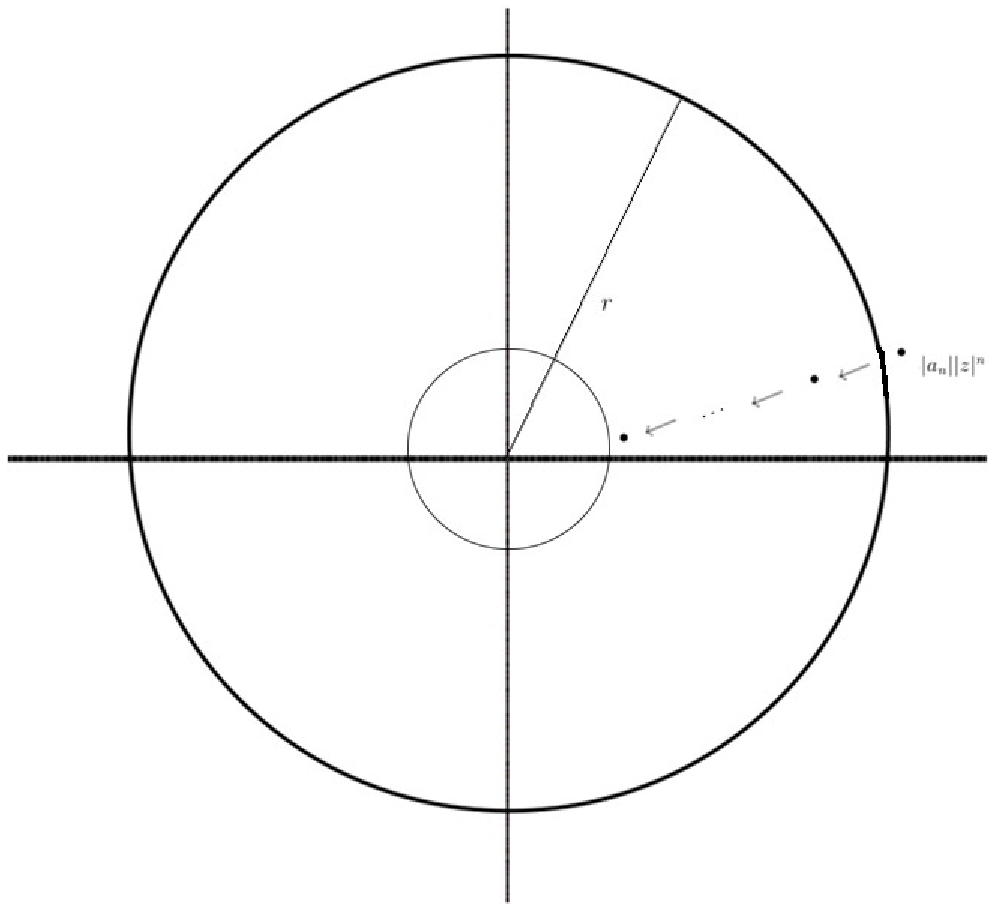
Theorem 6.12 Let f be a polynomial, $f(z) = a_n z^n + a_{n-1} z^{n-1} + ... + a_1 z_1 + a_0$, where $a_n \neq 0$. Then there is a number $r \in \mathbb{R}$ such that if $|z| > r$, then $|f(z)| > 2|z|$. If $|f^m(z)| \geq r$, then $f^k(z) \to \infty$ as $k \to \infty$. So either $f^k(z) \to \infty$ or $\{f^k(z) : k \in \mathbb{N}\}$ is bounded.

Proof:

Choose r to be $\sqrt[n-1]{\dfrac{4}{|a_n|}}$. Then if $|z| \geq r$ we have

$$
\begin{aligned}
& |z| \geq \sqrt[n-1]{\frac{4}{|a_n|}} \\
\Rightarrow \quad & |z|^{n-1} \geq \frac{4}{|a_n|} \\
\Rightarrow \quad & \frac{1}{2}|z|^{n-1} \geq \frac{2}{|a_n|} \\
\Rightarrow \quad & \frac{1}{2}|a_n||z|^{n-1} \geq 2 \\
\Rightarrow \quad & \frac{1}{2}|a_n||z|^{n} \geq 2|z|
\end{aligned}
$$

Also if r is large enough and $|z| > r$, then $|z|^n$ will be much larger than $|z|^{n-1}$ and all the other smaller powers of $|z|$. Consider $a_n z^n + a_{n-1} z^{n-1}$. At worst, this sum will have a distance of $|a_n||z|^n - |a_{n-1}||z|^{n-1}$ from the origin. The same will hold for each of the subsequent terms. This is described in the picture below.



So if $|z| > r$ for a large enough r, we have

$$\begin{aligned}
|a_n z^n + ... + a_1 z + a_0| &\geq |a_n||z|^n - (|a_{n-1}||z|^{n-1} + ... + |a_1||z| + |a_0|) \\
&\geq |a_n||z|^n \\
&\geq \frac{1}{2}|a_n||z|^n
\end{aligned}$$

So

$$|a_n||z|^n - (|a_{n-1}||z|^{n-1} + ... + |a_1||z| + |a_0|) \geq \frac{1}{2}|a_n||z|^n$$

implies that

$$\frac{1}{2}|a_n||z|^n \geq |a_{n-1}||z|^{n-1} + ... + |a_1||z| + |a_0|.$$

Thus

$$\begin{aligned}|f(z)| &= |a_n z^n + a_{n-1}z^{n-1} + ... + a_0| \\ &= |a_n z^n - (-a_{n-1}z^{n-1} - ... - a_1 z - a_0)| \\ &\geq |a_n||z|^n - (|a_{n-1}||z|^{n-1} + ... + |a_1||z| + |a_0|) \quad \text{by the reverse triangle inequality}\end{aligned}$$

Since $|a_{n-1}z^{n-1} + ... + a_0| \leq \frac{1}{2}|a_n||z^n|$, then subtracting $|a_{n-1}||z|^{n-1} + ... + |a_0|$ from $|a_n||z|^n$ leaves more than half of $|a_n||z|^n$, so

$$\begin{aligned}|f(z)| &\geq \tfrac{1}{2}|a_n||z^n| \\ &\geq 2|z|\end{aligned}$$

Now assume that $|f^m(z)| \geq r$ for some $m \in \mathbb{N}$. Then

$$|f^{m+1}(z)| = |f\left(f^m(z)\right)| \geq 2|f^m(z)|$$

by what we proved above. So

$$\begin{aligned}|f^{m+k}(z)| &= |f\left(f^{m+k-1}(z)\right)| \\ &\geq 2|f^{m+k-1}(z)| \\ &= 2|f(f^{m+k-2}(z))| \\ &\geq 2^2|f^{m+k-2}(z)| \\ &\vdots \\ &\geq 2^m|f^k(z)|\end{aligned}$$

Taking the limit, we have that $|f^k(z)| \to \infty$ as $k \to \infty$. So the k^{th} iterates of the function will eventually go to ∞ once $|z|$ gets large enough. For the simple example of the circle above, the r value is any number greater than 1 since for all $|z| \geq \sqrt[n-1]{1} = r$, the iterates $f^k(z) \to \infty$. ∎

Theorem 6.13 $J(f) = f(J) = f^{-1}(J)$. That is, J gets mapped onto itself under both f and f^{-1} (we say that J is invariant under f and f^{-1}).

Proof: Let $z \in J(f)$ and $\epsilon > 0$. Consider the ϵ-ball around $f(z)$. There is a δ such that $f(B_\delta(z)) \subseteq B_\epsilon(f(z))$ since f is continuous. Since z is in $J(f)$, there exists $w \in B_\delta(z))$ with $f^k(w) \to \infty$, so $f^k(f(w)) \to \infty$, which means $f(w) \notin K(f)$. There also exists $v \in B_\delta(z)$ such with $f^k(v) \nrightarrow \infty$, so $f^k(f(v)) \nrightarrow \infty$, thus $v \in K(f)$. So $B_\epsilon(f(z)) \cap K(f) \neq \emptyset$ and $B_\epsilon(f(z)) \cap \mathbb{C} \backslash K(f) \neq \emptyset$. Thus $f(z) \in J(f)$, so $z \in f^{-1}(J)$, hence $J(f) \subseteq f(J) \subseteq f^{-1}(J)$.

Now suppose $z \in f^{-1}(J)$, so $f(z) \in J(f)$. Pick $\epsilon > 0$ and a ball of radius ϵ around z. Note that polynomials are open maps, so $f(B_\epsilon(z))$ contains a ball of radius δ around $f(z)$. There exists $w', v' \in B_\delta(f(z)) \subseteq f(B_\epsilon(z))$ where $f^k(w') \to \infty$ and $f^k(v') \nrightarrow \infty$. But since they are in $f(B_\epsilon(z))$ as well, $w' = f(w)$ for some $w \in B_\epsilon(z)$ and $v' = f(v)$ for some $v \in B_\epsilon(z)$. Therefore $z \in J(f)$. Thus $f^{-1}(J) \subseteq f(J) \subseteq J(f)$, hence $J(f) = f(J) = f^{-1}(J)$.

∎

Definition 6.14 A family of functions $\{g_k\}$ is called *normal* on an open subset U of $\mathbb{C}$ if every sequence of functions from $\{g_k\}$ has a subsequence that converges uniformly on every compact subset of U either to ∞ or to a bounded analytic function. The family is *normal at a point* in $w \in U$ if there is some open subset V of U containing w such that $\{g_k\}$ is a normal family on V.

This definition is used in Montel's theorem, which is an ingenious result that we will use for the construction of our knowledge of Julia sets.

Proposition 6.15 Montel's Theorem: Let $\{g_k\}$ be a family of complex analytic functions on an open domain U. If $\{g_k\}$ is not a normal family, then for all $z \in \mathbb{C}$ with at most one exception, we have $g_k(w) = z$ for some $w \in U$ and some k.

The proof of this theorem requires too much knowledge of complex analysis for our purposes. The proof of this theorem may be found in [14].

Theorem 6.16 For a function f, the Julia set of f is

$$J(f) = \{z \in \mathbb{C} : \text{the family } \{f^k\} \text{ is not normal at } z\ \}$$

Proof:

Case 1: Suppose $z \in J(f)$. Then for $\delta > 0$, $U = N_\delta(z)$ is an open set and since $z \in J(f)$, there is a point $w \in U$ such that $f^k(w) \to \infty$. But $f^k(z) \nrightarrow \infty$. Thus $\{f^k\}$ does not converge to ∞ uniformly, so $\{f^k\}$ is not normal at z.

Note that if $z \notin J(f)$, then either z is in the interior of $K(f)$, denoted $\text{int}(K)$, or $z \in \mathbb{C} \backslash K(f)$.

Case 2: Suppose $z \notin J(f)$ such that $z \in \text{int}(K)$. Note that $U = N_\delta(z)$ is open for all $\delta > 0$. Suppose δ is small enough such that $U \subset K(f)$. Then $f^k(w) \nrightarrow \infty$ for all $w \in U$. So there is at least one $\zeta \in \mathbb{C}$ such that $g^k(w) \neq \zeta$ for any $w \in U$ or for any k since $f^k(w)$ is bounded. So by the contrapositive of Montel's Theorem, $\{f^k\}$ in normal on U.

Case 3: Suppose $z \notin J(f)$ such that $z \in \mathbb{C} \backslash K(f)$ and let $r > 0$ be defined as in Theorem 6.12. Since $z \in \mathbb{C} \backslash K(f)$, we have that $|f^k(z)| > r$ for some k. Then for $U = N_\delta(z)$ and for δ small enough, for all $w \in U$ we have $|f^k(w)| > r$. So $|f^k(w)| \to \infty$ uniformly on U by Theorem 6.12. Thus $\{f^k(w)\}$ is normal for all $w \in U$. So the only case when $\{f^k(z)\}$ is not normal is when $z \in J(f)$.

∎

Theorem 6.17 Let f be a polynomial, $w \in J(f)$ and let U be any neighborhood of w. Then for each $j \in \{1, 2, ...\}$ the set $W \equiv \bigcup_{k=j}^{\infty} f^k(U) = \mathbb{C}$, except possibly for a single point. Any such exceptional point is not in $J(f)$, and does not depend on w or U.

Proof:

Since $w \in J(f)$, then for $j \in \{1, 2, ...\}$ we have $\{f^k(w)\}_{k=j}^{\infty}$ is not normal by Theorem 6.16. Then by Montel's Theorem, for any $z \in \mathbb{C}$ there is a function g_k such that $g_k(z) = w$ with at most one exception.

Now suppose there is a point $\zeta \notin W$. If there is some point $z_0 \in \mathbb{C}$ with $f(z_0) = \zeta$, then $z_0 \notin W$ since $f(W) \subseteq W$ by how W is defined. From Montel's Theorem, $\mathbb{C} \backslash W$ can be at most one point, so $z_0 = \zeta$. So the first part of the theorem is satisfied.

Now we need to show that $\zeta \notin J(f)$ and is independent of w and U. Note that since $z_0 = \zeta$, we know that this is the only solution to the equation $f(z) - \zeta = 0$. So f is a polynomial of degree n such that $f(z) - \zeta = c(z - \zeta)^n$ for some constant $c \in \mathbb{C}$. Now suppose that z and ζ are sufficiently close so that $|z - \zeta| < \delta$ for δ sufficiently small. Then $f^k(z) - \zeta \to 0$ as $k \to \infty$ since f is uniformly convergent on $\{z : |z - \zeta| < \delta\}$. So $\{f^k\}$ is normal at ζ by definition, so $\zeta \notin J(f)$. Note that the point ζ is only dependent on the function f. ∎

Lemma 6.0.1 If $w \in J(f)$ is a fixed point, then w is not attracting.

Proof: Let U be a neighborhood of w. Then since $w \in J(f)$ there are points in U that iterate to ∞ and there are other points in U that do not iterate to ∞. So w does not attract any neighborhood of points.

∎

Definition 6.18 For an attracting fixed point w, define the set

$$A(w) = \{z \in \mathbb{C} : f^k(z) \to w \text{ as } k \to \infty\}$$

to be the *basin of attraction* for the point w. So A is the set of points that are attracted to w.

Theorem 6.19 Let w be a fixed point of a polynomial f. Then the basin of attraction $A(w)$ is an open set.

Proof:

Consider a δ-neighborhood U around w. Since w is attracting, we can find a δ such that all the points in U are attracted to w. So $U \subseteq A(w)$. Let $z \in A(w)$. Then $f^k(z) \in U$ for some k, so $z \in f^{-k}(U)$. Now since f is a polynomial, it is continuous. Therefore, since U is open and f is continuous, $f^{-1}(U)$ is open by the topological definition of a continuous function. We can do this for any z we like, so on taking the union $\bigcup_{k=1}^{\infty} f^{-k}(U) = A(w)$. Since the union of a collection of open sets is open, $A(w)$ is open.

∎

Theorem 6.20 Let w be an attractive fixed point of f. Then $\partial A(w) = J(f)$.

Proof:

Let $z \in J(f)$. Then $f^k(z) \in J(f)$ by Theorem 6.13. Note that $w \notin J(f)$ by Lemma 6.0.1. Then $f^k(z) \nrightarrow w$ since $f^k(z) \in J(f)$. So $z \notin A(w)$ by definition. But by Theorem 6.17, if $U = N_\delta(z)$ is a neighborhood around z, then there are points in U whose iterates converge to w. So since we can make δ as small as we like, $z \in \partial A(w)$. So $J(f) \subseteq \partial A(w)$.

Now suppose $z \in \partial A(w)$ and suppose for sake of contradiction that $z \notin J(f)$. Then by Theorem 6.16, the family $\{f_k\}$ is normal at z, so for a neighborhood V around z, the sequence $\{f_k\}$ has a subsequence that converges uniformly on a compact subset of V to a bounded analytic function or to ∞. Note that $V \cap A(w)$ is non-empty since $z \in \partial A(w)$ by the hypothesis, and it is open since V and $A(w)$ are both open. Clearly, on $V \cap A(w)$ the subsequence of $\{f_k\}$ converges to w. So the subsequence converges to a constant function. If a function converges to a constant on an open and connected subset of an open set, then the function converges to a constant on the whole open set [5]. So since the subsequence converges to a constant function on $V \cap A(w)$, it converges to w on all of V. Therefore, every point in V converges to w under iteration, which means $V \subseteq A(w)$. But now $z \notin \partial A(w)$ since all points around z map to w, which contradicts the hypothesis. Thus $z \in J(f)$, so $\partial A(w) \subseteq J(f)$. Therefore, $J(f) = \partial A(w)$.

∎

Example 6.21 Let's see an example of this theorem, considering a simple Julia set. Let $c = 0$, so $f_0(z) = z^2$. The fixed points of f are $z = 0$, $z = 1$. Note that $f'(z) = 2z$, so $f'(0) = 0 < 1$ and $f'(1) = 2 > 1$, implying that 0 is attracting and 1 is repelling. The Julia set for this function is the unit circle, since any complex number z with $|z| > 1$ will iterate to infinity, and any z with $|z| \leq 1$ will stay bounded. The boundary of the bounded points is the unit circle. The fixed point $z = 1$ is on the Julia set, which means that it is not attracting (which we already found). The basin of attraction for $z = 0$ is the set of points inside the unit circle, but not the boundary. So the boundary of the unit circle is the boundary of the basin of attraction, and the boundary of the unit circle is $J(f_0)$. □

Definition 6.22 A *loop*, C, is a smooth (differentiable), closed, simple (non-self-intersecting) curve in the complex plane.

Lemma 6.0.2 Let C be a loop in the complex plane. Then

(a) If c is in the interior of C, then $f_c^{-1}(C)$ is a loop, and the inverse image of the interior of C is the interior of $f_c^{-1}(C)$.

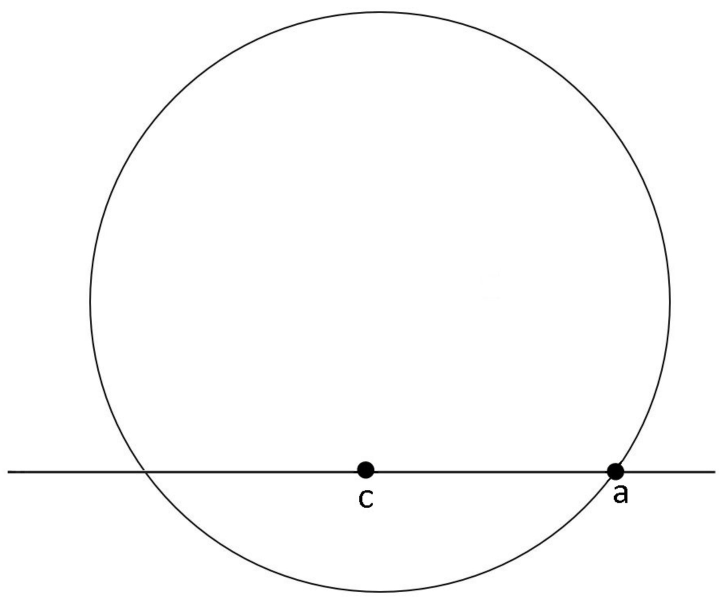
(b) If c lies on C, then $f_c^{-1}(C)$ is a figure eight with self-intersection at 0, such that the inverse image of the interior of C is the interior of the two loops.

(c) If c is outside C, then $f_c^{-1}(C)$ comprises two disjoint loops, with the inverse image of the interior of C the interior of the loops.

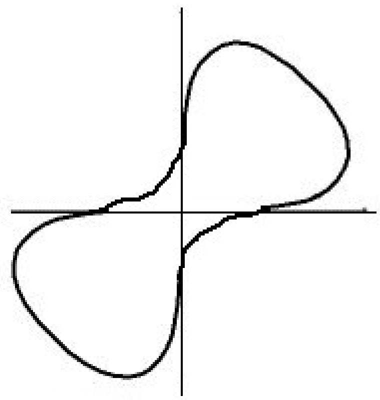
Proof:

First of all, if we prove that this theorem holds for any circle in the complex plane, then by continuously transforming the circle, we can transform the circle into any closed loop that we want. So essentially, if we prove the lemma for the circle, we prove it for any closed loop in the complex plane.

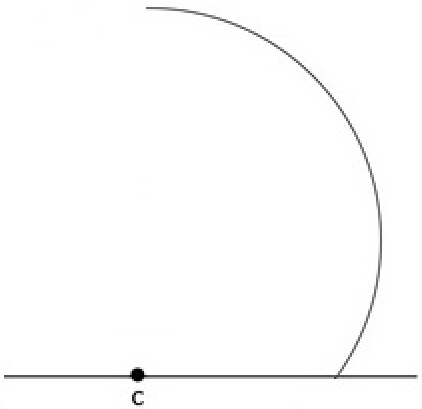
(a) Suppose c is inside of C and suppose C is a circle. We can translate the circle such that c is the origin. Suppose circle looks like the following picture, where c is the origin.

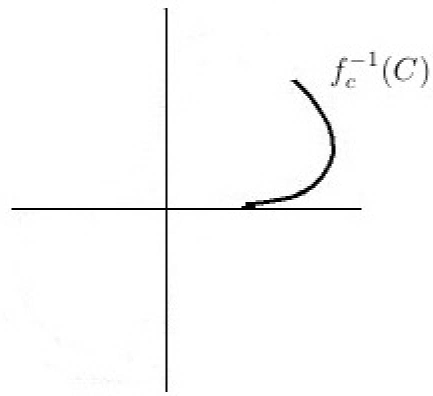


The inverse image will look something like

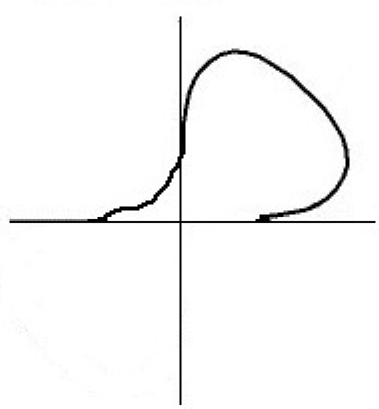


Since $c = 0$, $f_c(z) = z^2$, so $f^{-1}(c) = \pm\sqrt{z}$. Let w be a point on C and think of w using polar coordinates, so $w = re^{i\theta}$. Then on taking the square root, we get $f_c^{-1}(w) = \pm\sqrt{r}e^{i\frac{\theta}{2}}$. Start with w as the point a on the x-axis and consider only the positive square root for the inverse function. As w travels around the circle, $f^{-1}(w)$ traces another loop in the complex plane. The point furthest from c is when w reaches an angle of $\frac{\pi}{2}$. So the inverse function will be furthest from the origin at an angle of $\frac{\pi}{4}$. So tracing out the first $\frac{\pi}{2}$ radians will give us:

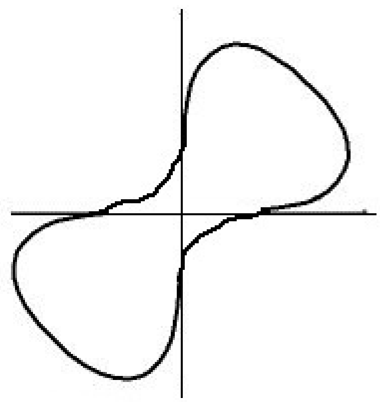




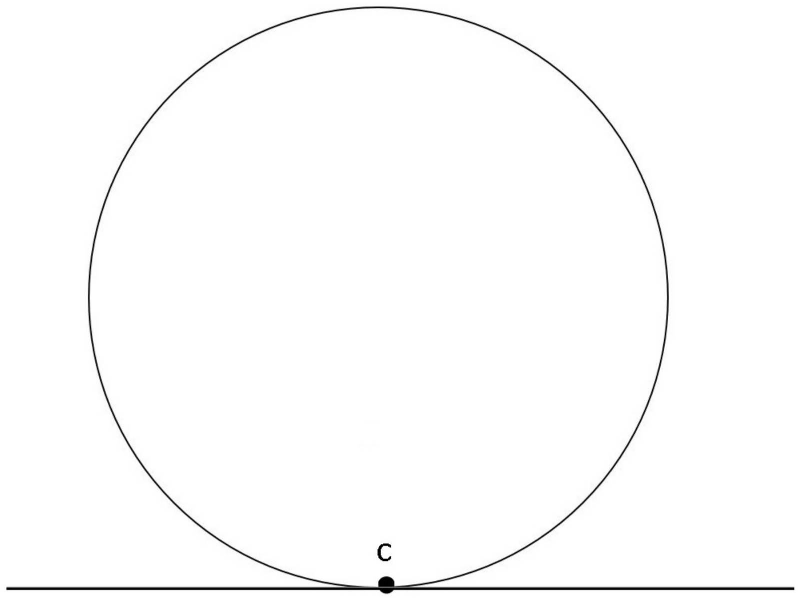
Likewise, the point closest to c is when w is at the angle $\frac{3\pi}{2}$, so the inverse function is closest to the origin at an angle of $\frac{3\pi}{4}$. Upon doing a full 2π rotation around the circle, the inverse function maps out a rotation of π around the new loop. So after a full 2π rotation we are the same distance away from c as when we started. So the inverse function at an angle of π will be at $-\sqrt{w}$.



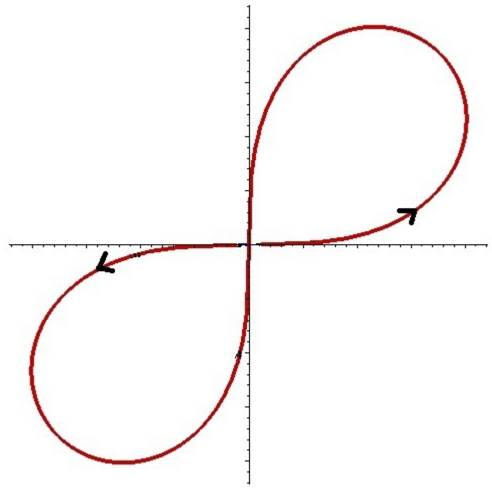
However, as we go around the loops a second loop is drawn at the same time on taking the negative square root. Taking the negative inverse function will map out a loop exactly as the positive inverse function, only reflected through the origin. Putting all these pieces together, we should get a loop that looks like the following picture.



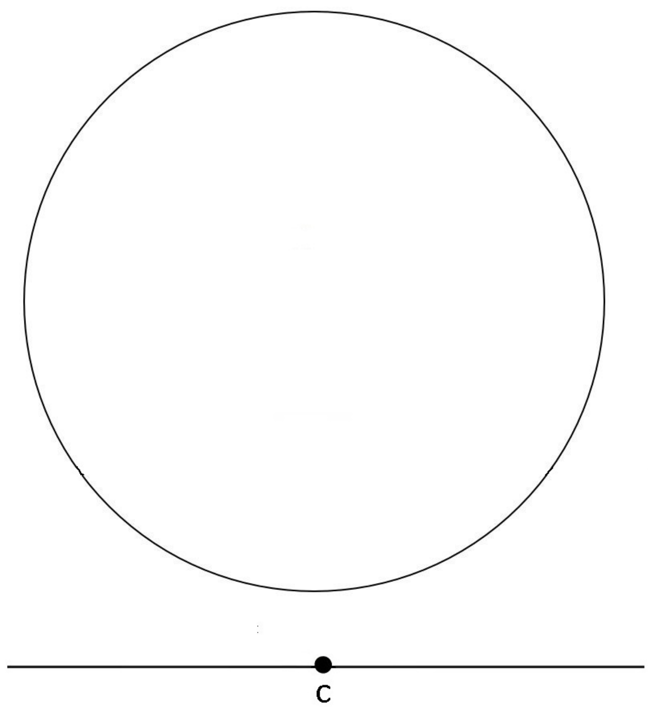
(b) Suppose c is the origin and C is a circle situated as follows.



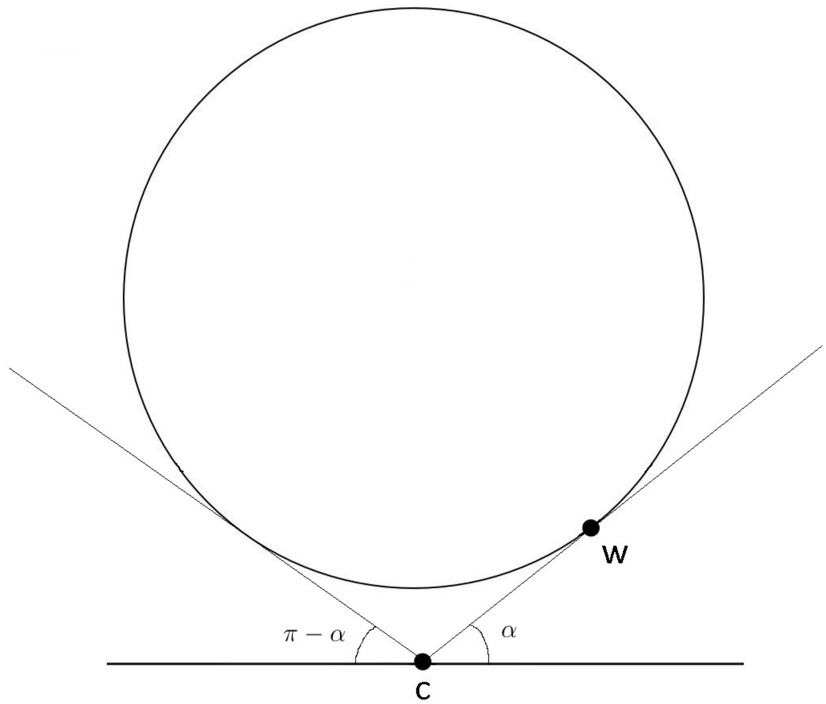
Choose w to start at c and move around the circle and for now consider only the positive square roots for the inverse function. The furthest point on the circle from c is the point at the angle $\frac{\pi}{2}$, so the point furthest from c for the inverse function happens at $\frac{\pi}{4}$. When w is traced around the circle at an angle of π, it has made a full rotation and returns to c. So for the inverse function, at the angle $\frac{\pi}{2}$ the loop returns to the origin. However, as the negative function values are taken, another loop is formed the mirrors the first loop. The two loops are connected at the origin. So $f_c^{-1}(C)$ should look something like this:



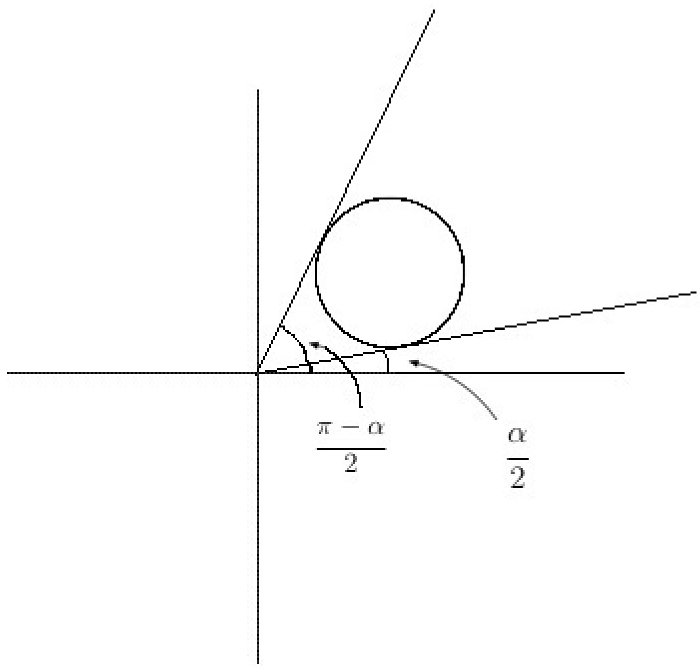
(c) Again suppose that c is the origin, and now C lies above the real line:



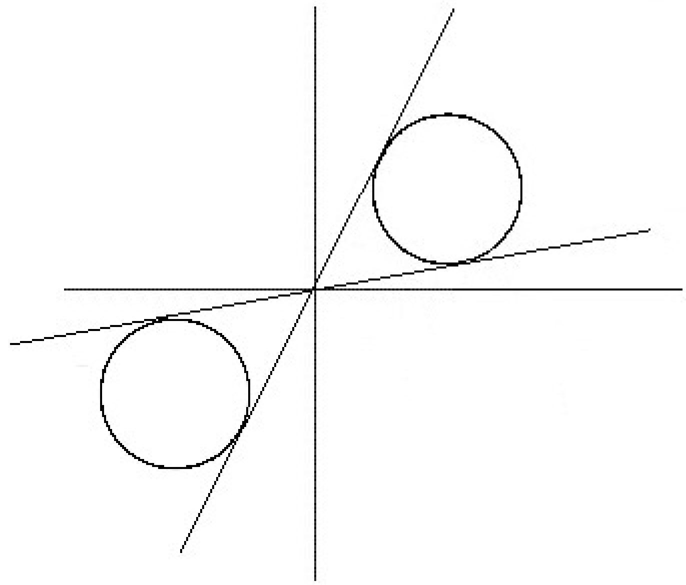
Note that when $\theta = 0$, there is no point on the circle, so there will be no point on f^{-1} for $\theta = 0$ either. Let α be the smallest angle for which a point on the circle is hit, and let this first point be the initial point w.



The inverse function maps this point to a point with angle $\dfrac{\alpha}{2}$ from the origin. The last point hit on the C is at an angle of $\pi - \alpha$. Call this point w'. This point is mapped to a point on f^{-1} with angle $\dfrac{\pi - \alpha}{2}$. Starting at w, the point will move around the circle until it hits w'. The inverse function at this stage has mapped out part of a loop. Then continuing from w', the inverse function will start to head back to its starting point. Note that the furthest and closest points to c on C are both at an angle of $\dfrac{\pi}{2}$. So on $f_c^{-1}(C)$ the closest and furthest points to the origin are at an angle of $\dfrac{\pi}{4}$.



Now consider the negative square root. The initial point with angle α will start on f^{-1} at an angle of α below the *negative* x-axis. The inverse will trace out a second circle here. The inverse function in this case will look like the following picture:



Since the Lemma holds for all three cases when C is a circle and c is the origin, the lemma will hold for any loop C and any c we wish.

∎

Recall that the Mandelbrot set is defined as

$$M = \{c \in \mathbb{C} : J(f_c) \text{ is connected}\}.$$

The following theorem gives us an equivalent definition.

Theorem 6.23 The Fundamental Theorem of the Mandelbrot Set. The Mandelbrot set can be defined as

$$M = \{c \in \mathbb{C} : \{f_c^k(0)\} \text{ is bounded}\} = \{c \in \mathbb{C} : f_c^k(0) \nrightarrow \infty \text{ as } k \to \infty\}$$

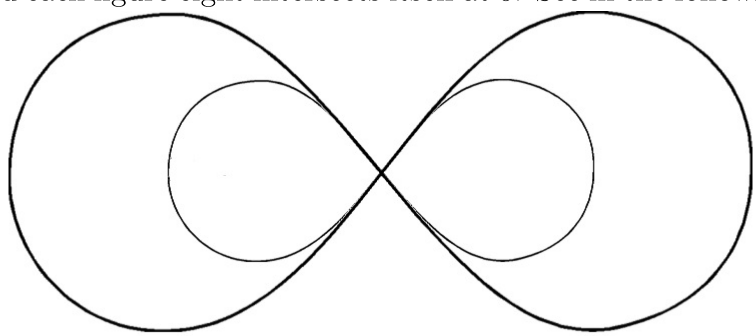
.

Proof:

The second two sets are equivalent by Theorem 6.12, since $\{f_c^k(0)\}$ is bounded if and only if $f_c^k(0) \nrightarrow \infty$. We will prove the first equality by proving that $\{f_c^k(0)\}$ is bounded if and only if it is connected.

Suppose that $\{f_c^k(0)\}$ is bounded. Let C be a large circle in $\mathbb{C}$ so that C contains all the points of $\{f_c^k(0)\}$. Note that $f_c(0) = c$, so c is in the interior of C. So by Lemma 6.0.2 (a), $f_c^{-1}(C)$ is a loop where the interior of C is mapped to the interior of $f_c^{-1}(C)$. Then $f_c^2(0) = f_c(c)$ is in the interior of C and $f_c^{-2}(C)$ is in the interior of $f_c^{-1}(C)$ by applying Lemma 6.0.2 (a) again. Continuing on in this fashion, we have that $\{f_c^{-k}(0)\}$ will be a loop inside a loop inside a loop and so on. Define K to be the set of points on or in the interior of each loop $f_c^{-k}(C)$ for all k. So if $z \in \mathbb{C} \backslash K$, then $f_c^k(z) \to \infty$. Thus the basin of attraction of the fixed point at ∞ is the boundary of K by Theorem 6.20. So K is the filled in Julia set $K(f_c)$ and $\partial K(f_c) = \partial A(\infty) = J(f_c)$. Note that K is the intersection of infinitely many closed, simply connected sets. So K is also closed and simply connected by the theorems of topology. Thus ∂K is simply connected, so the Julia set $J(f_c)$ is simply connected.

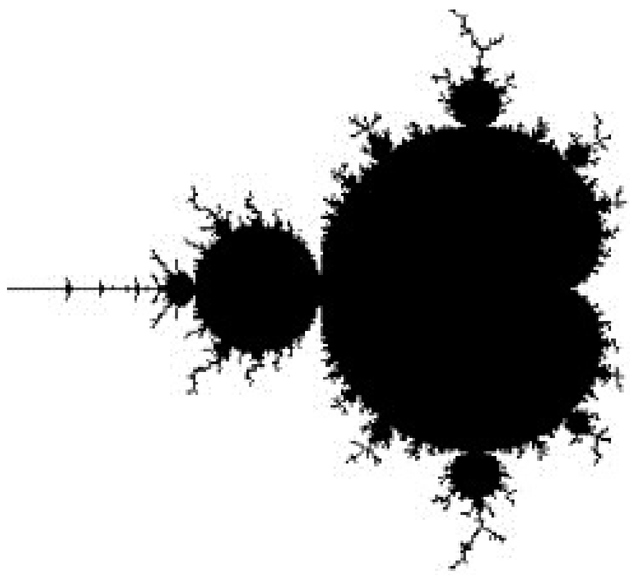
Now to prove that $J(f_c)$ connected implies $\{f_c^k(0)\}$ is bounded we will prove the contrapositive. So suppose that $\{f_c^k(0)\}$ is not bounded, and we will prove that $J(f_c)$ is disconnected. As before, let C be a large circle in $\mathbb{C}$ such that all points outside C go to ∞. Note, however, that this time 0 iterates to ∞ as well by the hypothesis, even though 0 is in the interior of C. Suppose there is some number n such that the point $f_c^n(0)$ lies on C. So if $k < n$, then $f_c^k(0)$ is inside C and if $k > n$, then $f_c^k(0)$ is outside C. As before, create a sequence of loops where each new loop is inside the previous one. Note that since $f_c^n(0) \in C$, we have $f_c^{n-1}(c) \in C$ so $c \in f_c^{1-n}(C)$. (A rough proof of this is that if we have $f(a) \in B$, then $a \in f^{-1}(B)$.) So since $c \in f_c^{1-n}(C)$, we can apply Lemma 6.0.2 (b). We will do this in a unique way. By the lemma, we know that $f_c^{-1}(C)$ is a figure eight, where the interior of C is mapped inversely to the interior of the figure eight and the figure eight intersects itself at 0. Do this again so that $f_c^{-2}(C)$ is a figure eight inside the figure eight created by $f_c^{-1}(C)$, and each figure eight intersects itself at 0. See in the following picture.



Keep doing this n times, and define $E = f_c^{-n}(C)$. Now the entirety of the Julia set $J(f_c)$ must lie on the interior of E. This is because the Julia set is invariant under f_c by Theorem 6.13, so $f_c(J) = J$. So as we take more and more iterations of f_c, the Julia set must remain the same, thus it must be contained in the smallest of all the figure eights, which is E. Note that $J(f_c)$ must lie on both sides of the figure eight since $f_c(E)$ maps onto the next figure eight. If the Julia set was only in, say, the right half of E, then parts of E would still have to map to the left side. But then these points would not iterate to infinity, and so would be in the Julia set. Recall that $0 \to \infty$. Then $0 \notin J(f_c)$. Thus 0 disconnects the Julia set inside E, since $J(f_c)$ lies on both halves of E. Hence the Julia set is disconnected. ∎

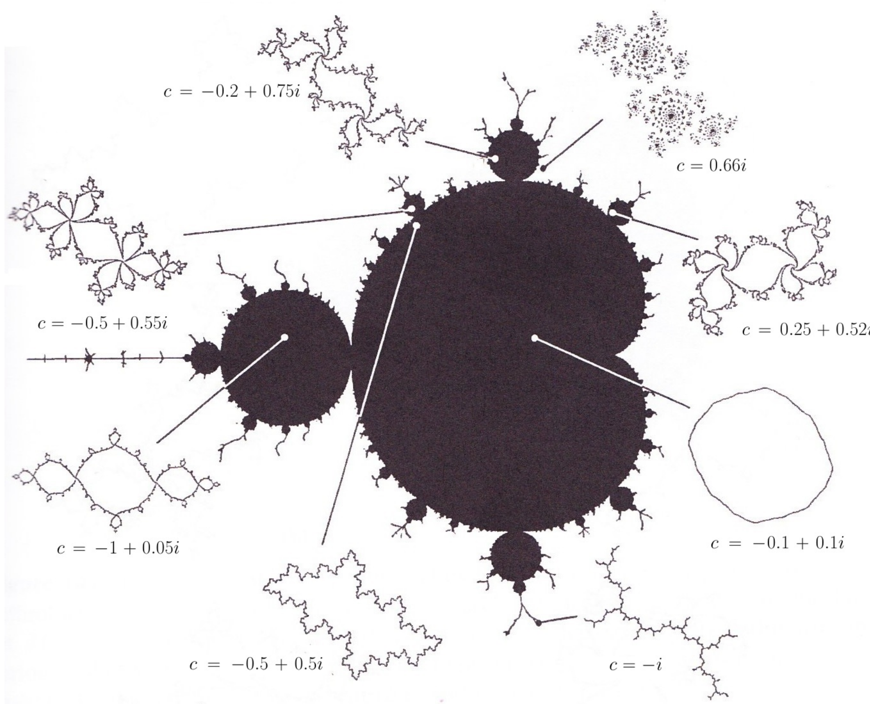
So the Mandelbrot set is the set of points c that cause the function $f(z) = z^2 + c$ to remain bounded when 0 is plugged in for z and iterated infinitely many times. This is how computer images of the Mandelbrot set are generated. A computer will iterate a value of c and see whether its modulus becomes greater than some number r, usually 2. If it does, then $c \notin M$, and if it does not then $c \in M$.

Many images of the Mandelbrot set contain different colors. The computer will color a point $c \notin M$ according to how fast or slow the points iterates to infinity. The points that remain bounded are all colored the same, usually black.



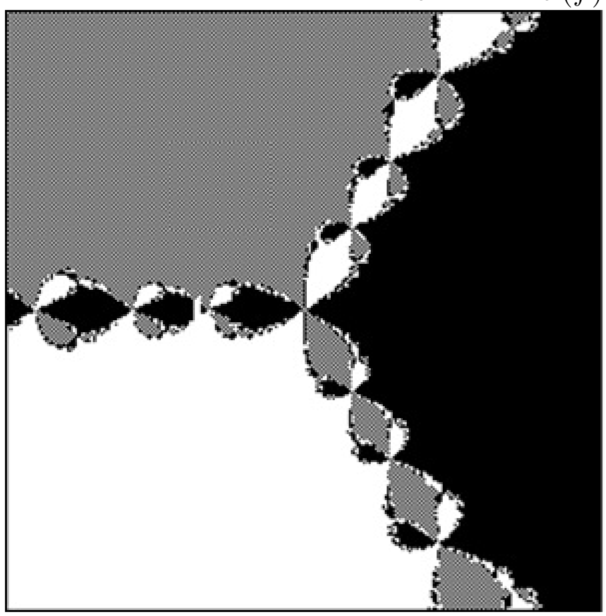
The Mandelbrot set itself has a cardioid-like shape with "buds" attached. These buds have buds of their own, and the buds of the buds have their own buds, and so on. Also there are strings that seem to randomly branch out from the buds and the cardioid. Each of the strings contain copies of the Mandelbrot set, but in a smaller scale. So the Mandelbrot set is an incredibly intricate set. Consider the Hausdorff dimension of the boundary for Mandelbrot set. Because of its jaggedness, it seems as though it should have dimension less than 2 since it does not seem quite 2-dimensional. However, because its boundary is so intricate and detailed the boundary has Hausdorff dimension 2. Note that the original definition of fractal would then say that this set is not a fractal since the dimension is an integer.

Let $c \in M$ such that c is not a fixed or periodic point. For this specific value of c, find all the points in the orbit of $f_c^k(0)$. This set of points will be a Julia set contained inside the Mandelbrot set. The following picture shows the different Julia sets obtained by choosing different values of c.



Example 6.24 Recall Newton's method from calculus. Let $p(x)$ be a function with continuous derivative. Newton's Method uses the difference function $f(x) = x - \dfrac{p(x)}{p'(x)}$ and $f^k(x)$ will converge to one of the zeros of p, assuming that $p'(x)$ is not zero and an appropriate initial value was chosen. For the purposes of Julia sets, we will examine Newton's method in the complex plane so we will be using the notation $f(z) = z - \dfrac{p(z)}{p'(z)}$. Notice that a point z is a zero of p if and only if z is a fixed point of f since then we would have $f(z) = z$.

Let $p(z) = z^3 - 1$. Then $f(z) = z - \dfrac{z^3 - 1}{3z^2} = \dfrac{2z^3 - 1}{3z^2}$. Note that the zeros of p are the cube roots of unity $z_1 = 1, z_2 = e^{\frac{2\pi i}{3}}$, and $z_3 = e^{\frac{4\pi i}{3}}$, so these are also the attracting fixed points of f. Using a computer to graph this situation on the complex plane, we get the image in the figure. The basin of attraction for z_1 is colored black, the basin of attraction for z_2 is colored grey, and the basin of attraction for z_3 is colored white. By Lemma 6.20, the boundary of the basins of attraction form the Julia set $J(f)$.



The infinite complexity of this image implies that the image is a fractal, but Newton's method doesn't end here. This was only one example of a fractal produced from Newton's method. We could change the original function to something else and get a completely different image.

Another thing to note is the boundary for the basin of attraction for any of the fixed points. Let $w \in J(F)$. Then for any neighborhood $N_\delta(w)$, there will be points of all three colors inside $N_\delta(w)$ because of how detailed this image is. Thus $w \in \partial A(z_i)$ for any $i \in \{1, 2, 3\}$. This is just to illustrate Theorem 6.20 working. $\square$

We have talked about how chaos theory effects fractals, fractal dimension, and Julia sets. This thesis could be expanded by exploring more applications of fractals in science and technology. Another expansion for this thesis is to look into fractal topology and how different fractals behave topologically depending on their dimensions. Applications in technology could arise from a fractal's topological structure. Further into applications, depending on the material used, a fractal-shaped object may be unstable if created using a certain material. If the object breaks because it is too weak, then either a different material will need to be used, or a fractal with different dimension will need to be constructed. Also in regards to topology, one could research how to take a product of two fractals and what the effects of this would be.

Of course, there is no such thing as an *exact* fractal in nature because fractals are infinite. But we can use our theories of exact fractals to come up with ideas for the approximate fractals we see in nature. Scientists use approximations all the time when dealing with ideas such as gravity and friction. Hopefully as these objects are studied further, they will give rise to even more powerful applications. Only time will tell.

$\mathcal{FIN}$

Bibliography

[1] "A Better 3-D Printer." *MIT Technology Review*. N.p., n.d. Web. 23 May 2013. <http://www.technologyreview.com/view/427273/infinity-computer-calculates-area-of-sierpinski-carpet-exactly/>.

[2] "Barnsley Fern." *Wikipedia*. Wikimedia Foundation, 05 June 2013. Web. 04 June 2013. <http://en.wikipedia.org/wiki/Barnsley_fern>.

[3] "CatSynth." *CatSynth RSS*. N.p., n.d. Web. 31 May 2013. <http://www.ptank.com/blog/2006/09/the-logistic-function-and-its-discontents/>.

[4] Elaydi, Saber. *Discrete Chaos.* Boca Raton: Chapman & Hall/CRC, 2000. Print.

[5] Falconer, K. J. *Fractal Geometry: Mathematical Foundations and Applications.* Chichester, England: Wiley, 2003. Print.

[6] Gibson, Carrie Rose. Chaos Theory. Thesis. Eastern Washington University, 2007. Print.

[7] Gleick, James. *Chaos: Making a New Science.* [S.l]: Scientific American Modern Classic, 2004. Print.

[8] "Hunting the Hidden Dimension." *Nova.* Narr. Neil Ross. Public Broadcasting System. WGBH, 24 Aug. 2008. Television.

[9] Li, T. Y., and J. A. Yorke. "Period Three Implies Chaos." *American Mathematical Monthly* 82.10 (1975): 985-92. Print.

[10] "List of Fractals by Hausdorff Dimension." *Wikipedia.* Wikimedia Foundation, 16 May 2013. Web. 30 May 2013. `<http://en.wikipedia.org/wiki/List_of_fractals_by_hausdorff_dimension>`.

[11] "Mandelbrot Set." *Wikipedia.* Wikimedia Foundation, 28 May 2013. Web. 31 May 2013. `<http://en.wikipedia.org/wiki/Mandelbrot_set>`.

[12] "MinkowskiBouligand Dimension." *Wikipedia.* Wikimedia Foundation, 24 May 2013. Web. 02 June 2013. `<http://en.wikipedia.org/wiki/MinkowskiBouligand_dimension>`.

[13] Munkres, James Raymond. *Topology, a First Course.* Englewood Cliffs Prentice-Hall, 1975. Print.

[14] Narasimhan, Raghavan, and Yves Nievergelt. *Complex Analysis in One Variable.* Boston: Birkhäuser, 2001. Print.

[15] "Math Research, Tips and Tricks." : *Pascal's Triangle and Prime Numbers.* N.p., n.d. Web. 23 May 2013. `<http://malsmath.blogspot.com/2006/12/pascals-triangle-and-prime-numbers.html>`.

[16] Ribeiro, Marcelo B., and Alexandre Y. Miguelote. “Fractals and the Distribution of Galaxies.” *Brazilian Journal of Physics* 28 (1998): Web. 30 May 2013. <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-97331998000200007>.

[17] *Star Wars, Episode III, Revenge of the Sith.* By George Lucas. Dir. George Lucas. Prod. George Lucas and Rick McCallum. Perf. Ewan McGregor, Natalie Portman, and Hayden Christensen. 20th Century Fox, 2005.

Vita

Author: Robert DeLorto

Place of Birth: Santa Clara, California

Undergraduate Schools Attended: North Idaho College
University of Idaho

Degrees Awarded: Bachelor of Science in mathematics, 2011, University of Idaho

Professional Experience: Graduate Instructor, Eastern Washington University, Cheney, Washington, 2011-2013