The Collage Theorem and Its Applications

Faculty Advisor: Jane McDougall

Soyeon Kang

April 8, 2016

Abstract

In this expository paper, I introduce fractals with ways we can identify them such as self-similarity and non-integer dimensions. I build the concepts necessary for the proof of the Collage Theorem, including basic topological concepts on metric spaces, closed sets, and the convergence of sequences. With these definitions, we can establish what it means for a metric space to be complete. Once we have a complete metric space, we can prove the contraction mapping theorem and use a similar technique to prove The Collage Theorem. The Collage Theorem can then help us construct a method for programming a computer to generate the fractals. I discuss two types of algorithms to plot the points of an attractor that will resemble a fractal-like image. Finally, I implement these algorithms in Python to make the fractals on my own computer.

1 Introduction

How would you describe a fern using only triangles, polygonal shapes, circles, and points? It would be hard quite hard to do so with only the tools of Euclidean geometry. They do a good job of simplifying shapes in order for us conjecture about the properties of these smooth objects, but it doesn't translate to what we see around us everyday. That's why we turn to fractal geometry- it allows us to describe and classify more complex objects in a simple, but precise manner. However, through our studies, it's important to keep in mind that fractals are abstractions, a result of an infinitely iterative process. Mathematical fractals do not exist in the real world, but they help us capture the qualities of shapes and process of the numerous examples of natural fractals that do, although they might not exhibit the exact self-similarity. Thus the best we can do is find a method to approximate the fractal as closely as possible.

So what exactly makes a fractal? Remember, since fractals are ideals for an infinitely defined process, there is no formal definition. In most cases though, they share the following properties.

1. Self-similarity: the details of image is made up of small copies of the whole image
2. Structure at different scales: the details of the image is intricate regardless of magnification
3. Fractal dimension: a non-integer number in between two dimensions

We can usually inspect the image for the first two properties, but the fractal dimension is what quantifies those observations.

1.1 Fractal Dimension

Fractals have dimensions which are not found in the shapes that we encounter in Euclidean geometry. There are many kinds of dimensions but not all of them are useful, and others are more involved in measure theory thus is beyond the scope of this project, so I will be discussing two types of dimension that are relatively straightforward to calculate.

1.1.1 Self-similarity Dimension

For images where the whole image is composed of smaller copies of itself, we call that image *exactly self-similar.* This is case for classical fractals such as the Cantor set and Sierpinski's triangle. We define the dimension as the relation between the number of small copies in an object and its magnification factor. Thus the following formula is used to find the dimension

$$\text{number of small copies} = (\text{magnification factor})^{\text{dimension}}. \qquad [4]$$

It tells us how many small self-similar pieces of an object fit inside a large piece. Most of the time, fractals have self-similarity dimensions that are non-integer because they exhibit qualities of two different dimensions.

Example 1. If we look at the first iteration, we'll see that both of the lines are scaled down by a third, Thus, the Cantor set has magnification factor 3, with 2 small copies of itself:

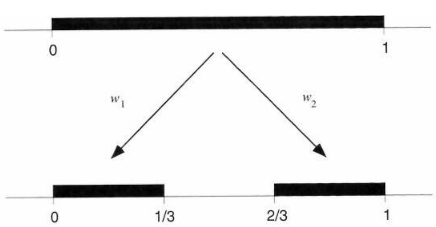
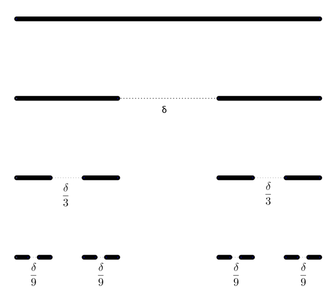


Figure from p.173 of Peitgen [5]



We can determine the dimension using the formula above:

$$\begin{aligned}
3^{\dim} &= 2 \\
\log(3^{\dim}) &= \log(2) \\
\dim \log(3) &= \log(2) \\
\dim &= \frac{\log(2)}{\log(3)} \approx 0.63
\end{aligned}$$

The Cantor set has a non-integer dimension in between 0 and 1. This makes sense because of the way the set is constructed. We start from a line (1 dimensional) on the closed interval $[0, 1]$ and break it down so that it eventually becomes a set of points (0 dimensional).

Example 2. We can find a dimension for Sierpinski's Triangle as well. It is constructed by taking the middle triangle out each iterations. Thus, each of our triangles will be $\frac{1}{2}$ of the whole, and we will obtain 3 small copies. Thus, $2^d = 3$ and $d = \frac{log(3)}{log(2)} \approx 1.585$.

Example 3. It is relatively easy to determine the dimension of shapes that are not fractals. For instance, take a cube and make "tiles" of it by splitting into 8 smaller cubes, we will have scaled down the width, length, and depth of the cube by a half. Thus, using the formula we have $8 = 2^d$, and our cube has a self-similarity dimension of 3 as expected.

1.1.2 Box-Counting Dimension

So far we've only talked about perfect or mathematical fractals, but we don't often encounter them in real life. The fractals found in nature, such as the fern are not exactly similar but are often times *statistically self-similar*. In those cases, we look at the box-counting diemnsion which looks at how many small copies of an ob ject are contained in a large copy by seeing how the volume or size of the overall shape changes as we change measurement scales. The process is this: lay a grid with boxes of side length ϵ over the fractal and count how many boxes are required to cover the set, this gives the value for $N(\epsilon)$. We want cover the fractal with as many little boxes as possible. We want the relationship between number of boxes of size s needed to cover the shape. As s gets smaller, N gets larger.

$$N(s) = k\left(\frac{1}{s}\right)^d$$

Taking the log of both sides, we can turn this expression into an equation for a line:

$$\begin{aligned}\log N(s) &= \log\left(k\left(\frac{1}{s}\right)^d\right)\\ \log N(s) &= \log k + d\log\left(\frac{1}{s}\right). \quad [4]\end{aligned}$$

Now we just have to find the slope d and y-intercept k by plotting $\log N(s)$ vs. $\log\frac{1}{s}$. Essentially the box-counting dimension is calculated by finding the limit with respect to an ϵ:

$$\dim_{box}(s) = \lim_{\epsilon\to 0}\frac{\log N(\epsilon)}{\log\left(\frac{1}{s}\right)}$$

2 Preliminaries

Now that we are able to identify the simple properties, we must build some terminology to describe them in a more precise way. We will turn to some topological concepts in metric spaces.

2.1 Metric Spaces

A metric space uses the distance between sets to define some sort of structure on the space in which the sets exist. More specifically, we need a metric space so that we can conjecture about the limits and convergence of sequences in the space and then prove completeness of the space. These terms are then used to prove the contraction mapping theorem in the next section.

Definition 1. Let X be any non-empty set. A distance function $d : X \times X \to \mathbb{R}$ is called a *metric on* X if d satisfies the following:

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

Definition 2. A non-empty set X equipped with the metric d is called the *metric space* (X, d).

Example 4. The usual metric on a real line is called the Euclidean metric which is defined as

$$d(x, y) = |x - y|.$$

The same will hold for the distance on the Euclidean plane $\mathbb{R}^2$, where for $x = (x_1, x_2)$ and $y = (y_1, y_2)$, the distance function will be

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

which gives us the shortest distance between two points in the plane. Thus $(\mathbb{R}^2, d)$ would be our metric space.

Example 5. Another metric that is easy to prove is the dicrete metric:

$$d(x, y) = \begin{cases} 1, & \text{if } x \neq y \\ 0, & \text{if } x = y \end{cases}$$

Proof. We'll check each property in Definition 1:

1. $d(x, y) = 0$ or 1 so $d(x, y) \geq 0$
2. Suppose $d(x, y) = 0$, then by definition, $x = y$. Now say $x = y$, then $d(x, y) = d(x, x) = 0$
3. $d(x, y) = 1 = d(y, x)$
4. $d(x, z)$ can have two cases: (1) $x = z$, then $d(x, z) = 0$ and we are done. (2) $x \neq z$ then y can either be on x or z, or neither. Without loss of generality, let's say that $y = x$ then $y \neq z$ and, $d(x, y) + d(y, z) = 0 + 1$. If y is neither x or z, then $d(x, y) + d(y, z) = 1 + 1 \geq d(x, z)$.

$\square$

2.2 Properties of Closed Sets

Definition 3. Let S be a set of real numbers. The set S is *closed* if it contains all of its limit points.

Some properties of closed sets are:

1. The sets $\emptyset$ and $\mathbb{R}$ are closed.

2. Any union of a finite number of closed sets is closed.

3. Any intersection of an arbitrary collection of closed sets is closed.

4. The complement of a closed set is open. [1]

Definition 4. Let S be a set of real numbers that is bounded below and nonempty. If m is the greatest of all the lower bounds of A, then m is said to be the greatest lower bound of S or the infimum of A and we write $m = \inf A$. Similarily, if S is bounded above and l is the least of all the upper bounds of A, then $l = \sup S$. [1]

Completeness Axiom: A nonempty set of real numbers that is bounded below has a least upper bound (i.e. if A is nonempty and bounded above, then $\inf A$ exists and is a real number).

Definition 5. Let A be a set in the metric space (X, d). Then we write A_ϵ for the set of points that are no further than ϵ away from some point in A.

Proposition 1. *If A is closed and bounded, then A_ϵ is closed.*

Remark 1. This proof uses the **Bolzano-Weierstrass Theorem** which come in handy later on. It states that *if the sequence $\{x_n\}$ is bounded in X, it has a convergent subsequence $\{x_{n_m}\}$.*

Proof. Let p be an arbitrary limit point of A_ϵ. Then,

$$p = \lim_{n\to\infty} x_n$$

where $\{x_n\}_{n=1}^{\infty} \subset A_\epsilon$. Then for each $x_n \in \{x_n\}$ we can pick an element $a_n \in A$ such that

$$|x_n - a_n| \leq \epsilon$$

and construct a sequence $\{a_n\}_{n=1}^{\infty} \subset A$. By the Bolzano-Weierstrass Theorem, we know there exists a subsequence $\{a_{n_j}\}$ that converges to an $a \in A$ so that

$$|x_{n_j} - a_{n_j}| \leq \epsilon.$$

Since x_{n_j} converges to p and a_{n_j} converges to a:

$$|p - a| \leq \epsilon.$$

Thus $p \in A_\epsilon$ and since p was arbitrary, a_ϵ contains all of its limit points and is closed. □

Proposition 2. *If A is a closed set, and $r, s > 0$, $(A_r)_s \subseteq A_{r+s}$.*

Proof. Let $a \in (A_r)_s$ so that a is at most s away from some point in A_r. Then for some $\epsilon > 0$ and $a_1 \in A_r$ where a_1 is at most r away from some point $a_0 \in A$,

$$\begin{aligned}
|a - a_0| &\leq |a - a_1| + |a_1 - a_0| \\
&\leq (s - \epsilon) + (r - \epsilon) \\
&\leq r + s - 2\epsilon \\
&\leq r + s
\end{aligned}$$

Thus we have shown that an arbitrary point from the set $(A_r)_s$ belongs to the set A_{r+s}. □

2.3 The Hausdorff Metric

We want to define a distance function that captures the idea of the distance between two sets, and later in our applications two images We will first introduce one such distance that is not a metric, then make some adjustments to our distance function so that it satisfies the properties of a metric. A metric allows us to use the triangle inequality, which will be important in the contraction mapping theorem.

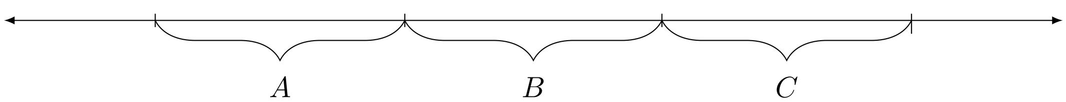
Example 6. Let $\mathcal{K}$ consist of the nonempty closed subsets of $[0,1] \times [0,1]$. For $A, B \in \mathcal{K}$, let's first define the distance between A and B:

$$dist(A,B) = \inf\{|a-b| : a \in A, b \in B\}.$$

This seems like an intuitive way to measure distance between sets but you will quickly see that this function does not satisfy the properties of a metric. For instance to assert the triangle inequality for this distance, we have to prove that

$$\inf\{|a-b|\} \leq \inf\{|a-c|\} + \inf\{|c-b|\}.$$

In the case shown below where the boundary of A is touching B and B touching C, the infimum of the distance between the points in A and B is 0, and same with B to C. However, the infimum of the distance between the points in A and C is greater than 0. Thus, it fails the triangle inequality and is not a metric on $\mathcal{K}$.



So we are going to turn to the Hausdorff metric, which captures the idea of the "greatest distance that a point in A can be from the set B or a point in B from the set A".

Definition 6. Define $\mathcal{K}$ to be the subspace that consists of the non-empty closed subsets of $[0,1] \times [0,1]$. For $A \in \mathcal{K}$ and $\delta \geq 0$, let A_δ be the union of all closed disks of radius δ centered at points of A. Thus a point $a \in A_\delta$ is no at most δ away from some point in A. Define

$$\rho(A,B) = \inf\{\delta \geq 0 : A \subseteq B_\delta, B \subseteq A_\delta\}.$$

We can also build from the Euclidean to get the distance of the a point to a set $d(x,A) = \inf\{|x-y| : y \in A\}$. Then to define the distance between two sets A and B define $d(X,Y) = \sup\{d(x,Y) : x \in X\}$. Finally, we will get the Hausdorff distance between the sets by taking the greater of the two:

$$\rho(A,B) = \max\{d(X,Y), d(Y,X)\}$$

Before we can prove that the Hausdorff Metric is actually a metric, we must look at some more properties of the closed set A.

2.4 Properties

Proposition 3. *If $\epsilon < \epsilon'$, $A_\epsilon \subset A_{\epsilon'}$*

Proof. Pick an arbitrary point $x_\epsilon \in A_\epsilon$. By definition, this means that x_ϵ is within ϵ of some point in A, call it x. Then $|x_\epsilon - x| < \epsilon$. So if

$$\begin{aligned} \epsilon &< \epsilon' \\ |x_\epsilon - x| &< \epsilon', \end{aligned}$$

it means that x_ϵ is within ϵ of some point $x \in A$. Thus $x_\epsilon \in A_{\epsilon'}$ and $A_\epsilon \subset A_{\epsilon'}$. □

Proposition 4. *For $\epsilon > 0$, $\rho(A, B) \leq \epsilon$ if and only if $A \subseteq B_\epsilon$ and $B \subseteq A_\epsilon$.*

Proof. Suppose $\rho(A, B) \leq \epsilon$. Define the set

$$\mathcal{S} := \{\delta > 0 : A \subset B_\delta, B \subset A_\delta\}.$$

For the first case, let $\epsilon > \inf S$. Then by definition of an infimum, ϵ cannot be a lower bound, so there must exist an $s \in \mathcal{S}$ where $s < \epsilon$ and $A \subset B_s, B \subset A_s$. By Proposition 3 we know that if $s < \epsilon$, $A_s \subset A_\epsilon$ and $B_s \subset B_\epsilon$. Thus, $A \subset B_\epsilon$ and $B \subset A_\epsilon$. For the second case, let $\epsilon = \inf \mathcal{S}$. Then by compactness properties of $\mathcal{S}$, the infimum of the set ϵ is actually the minimum so $\epsilon \in \mathcal{S}$. Thus $A \subseteq B_\epsilon$ and $B \subseteq A_\epsilon$. Conversely, Suppose $A \subseteq B_\epsilon$ and $B \subseteq A_\epsilon$. But then $\epsilon \in \mathcal{S}$ by definition of A_ϵ and B_ϵ. Then, the infimum of $\mathcal{S}$ is either ϵ or less than ϵ. Thus, $\inf \mathcal{S} = \rho(A, B) \leq \epsilon$. □

Now we want to prove that the Hausdorff distance is actually a metric.

Theorem 1. *The Hausdorff distance ρ defined above is a metric on X.*

We just need to prove that the Hausdorff distance satisfies the properties in Definition 1.

Proof. Let $A, B, C \in \mathcal{K}$.

1. By definition, our Hausdorff distance $\rho(A, B) = \inf\{\delta \geq 0\}$.
2. Suppose $\rho(A, B) = 0$, then $A \subseteq B$ and $B \subseteq A$ but that means $A = B$. Conversely, say $A = B$ so that A is radius 0 from any point in B and vice versa. Then $\delta = 0 = \rho(A, B)$.
3. $\rho(A, B) = \inf\{\delta \geq 0 : A \subseteq B_\delta, B \subseteq A_\delta\} = \inf\{\delta \geq 0 : B \subseteq A_\delta, A \subseteq B_\delta\} = \rho(B, A)$
4. Triangle Inequality: Let $r = d(A, B)$, so that $A \subseteq B_r, B \subseteq A_r$, and $s = d(B, C)$ so that $B \subseteq C_s, C \subseteq B_s$. By Proposition 2, $A_{r+s} \supseteq (A_r)_s \supseteq B_s \supseteq C$ and $C_{r+s} \supseteq (C_s)_r \supseteq B_r \supseteq A$. Then we have that

   $$C \subseteq A_{r+s} \text{ and } A \subseteq C_{r+s}.$$

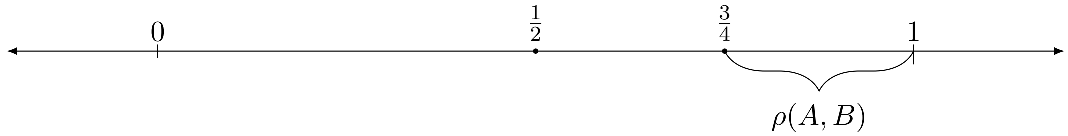
   Thus $r + s$ belongs to the set $T = \{\delta > 0 : \delta > 0 : A \subseteq C_\delta, C \subseteq A_\delta\}$. If we take the infimum of T, we have the lower bound of the set, i.e. $\inf T \leq \delta$ for every $\delta \in T$. Then we can say

   $$\begin{aligned} d(A, C) &= \inf T \\ &\leq r + s \\ &\leq d(A, B) + d(B, C). \end{aligned}$$

□

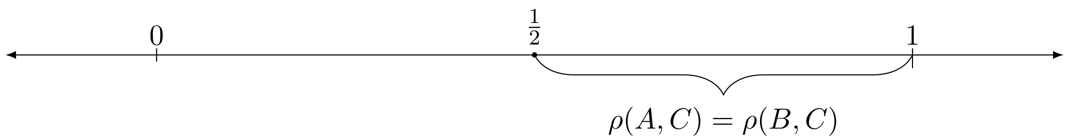
Example 7. Let $A = \left\{\frac{1}{n} : n \in \mathbb{N}\right\} \cup \{0\}$, $B = [0, 1]$, and $C = \left\{\frac{1}{2}\right\}$. Calculate the Hausdorff distances $\rho(A, B)$, $\rho(B, C)$, and $\rho(A, C)$. [1]

1. $\rho(A, B) = \inf\{\delta > 0 : A \subset B_\delta \text{ and } B \subset A_\delta\}$. Every point in A is already a point in B. But for the point $\frac{3}{4} \in B$, it has to travel at least $\frac{1}{4}$ across the number line in either direction to get to a point in A. Thus $\rho(A, B) = \max\{0, \frac{1}{4}\} = \frac{1}{4}$



2. $\rho(B, C) = \inf\{\delta > 0 : B \subset C_\delta \text{ and } C \subset B_\delta\}$. Here, C is just the one point at $\frac{1}{2}$ so that is already in B. But then for B_δ to reach C, δ has to be at least $\frac{1}{2}$.

3. $\rho(A, C) = \inf\{\delta > 0 : A \subset C_\delta \text{ and } C \subset A_\delta\}$. Similarly, δ has to be at $\frac{1}{2}$ in order for any point



in A_δ to reach $\frac{1}{2}$.

3 Completeness of a Set in $\mathbb{R}$

Definition 7. Let $\{x_n\}$ be a sequence of real numbers. Then if $\epsilon > 0$, there exists an integer $N \in \mathbb{N}$ such that for every $n \geq N$, $|x_n - L| < \epsilon$, the sequence $\{x_n\}$ converges to the limit L as $n \to \infty$. This is also notated $\{x_n\} \to L$. [1]

Definition 8. Let (X, d) be a metric space and $\{x_n\}$ be a sequence of real numbers in X. $\{x_n\}$ is called a *Cauchy sequence* if for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that for every $n, m \geq N$, $d(x_n - x_m) < \epsilon$. [1]

Definition 9. A metric space (X, d) is said to be *complete* if every Cauchy sequence $\{x_n\} \in X$ converges to a point $x \in X$.

Example 8. The set of real numbers and the plane $\mathbb{R}^2$ with the Euclidean metric is complete, but $\mathbb{Q}$ is not (consider the sequence $\{a_n\} = \frac{x_n}{x_{n+1}}$).

Lemma 1. *Let $\{F_n\}$ be a decreasing sequence of closed, bounded, non-empty sets in $\mathbb{R}^k$- i.e. $F_1 \supseteq F_2 \supseteq \ldots$. Then $F = \bigcap_{n=1}^{\infty} F_n$ is also closed, bounded, and non-empty.* [6]

Proof. For every n, select an element $x_n \in F_n$ to form the sequence $\{x_n\}$. Then by the Bolzano-Weisterass Theorem, there exists a subsequence $\{x_{n_m}\}$ that converges to x_0, an element of $\mathbb{R}^k$.
Now we want to show that $x_0 \in F$. If we choose an arbitrary $n_o \in n$, then it will suffice to show that $x_0 \in F_{n_0}$, since F is the intersection of all F_{n_0} and our choice of n_0 was arbitrary. Then, if we choose an $m \geq n_0$, then $n_m \geq n_o$ because values of n_m just form a subset of n. Now we have that $\{x_{n_m}\} \in F_{n_m} \subset F_{n_0}$. Then, $\{x_{n_m}\}$ consists of points in F_{n_0} and converges to x_0. Since F_{n_0} is closed, it contains all of its limits points, thus $x_0 \in F_{n_0}$. $\square$

Lemma 2. *Let F_n and F be defined as in Lemma 1. Then for $\delta \geq 0$, $\rho(F, F_n) < \delta$ and equivalently, F_n converges to F.*

Proof. Suppose that F_n does not converge to F. Then, there must be some δ such that $\rho(F_n, F) \geq \delta$. This means that either $F \subset (F_n)_\delta$ or $F_n \subset F_\delta$ fails. But we know that F is always contained in F_n so the second statement must be false. By Lemma 1, we have found a sequence $\{x_{n_m}\} \in F_n$ that converges to the limit point $x_o \in F$. Thus there does nota $\delta \geq 0$ such that $\rho(F_n, F) \geq \delta$ and we arrive at a contradiction. □

Theorem 2. *The subspace $\mathcal{K} \subset \mathbb{R}^2$ with the Hausdorff metric, (X, ρ) is complete.*

Proof. Let $\{x_n\}$ be an arbitrary Cauchy sequence in $\mathcal{K}$. Then for $\epsilon > 0$, there exists an N such that for $n, m \geq N$, $\rho(x_n, x_m) < \epsilon$. Also for $n, m \geq N$ $(X_n)_\epsilon \supset X_m$ where X_n is the set of points in the sequence $\{x_n\}$. Thus if we set $m = k$, we can say

$$(X_n)_\epsilon \supset \bigcup_{k=n}^{\infty} x_k.$$

Now we define the sets

$$H_n = \mathbf{cl}\left(\bigcup_{k=n}^{\infty} x_k\right) \text{ and } H = \bigcap_{n=1}^{\infty} H_n.$$

Since $(X_n)_\epsilon$ is closed,

$$\begin{aligned} (X_n)_\epsilon &\supseteq \mathbf{cl}\left(\bigcup_{k=n}^{\infty} x_k\right) \\ &\supseteq H_n \\ &\supseteq H. \end{aligned}$$

We also know that $H_n \to H$ so there exists an $M \in \mathbb{N}$ such that $H_n \subset H_\epsilon$ for $n \geq M$. Then for $n \geq M$,

$$x_n \in H_n \subset H_\epsilon.$$

Thus we have shown that $X_n \subset H_\epsilon$ and $H \subset (X_n)_\epsilon$ which are the conditions we need for convergence. Thus, $\rho(X_n, H) < \epsilon$ and our chosen Cauchy sequence $\{x_n\} \in X_n$ converges to H. □

4 Contraction Mappings

Definition 10. Let (X, ρ) be a metric space and let $A : X \to X$ be a mapping of X onto itself. Then A is called a *contraction mapping* if there exists a number $a \in [0, 1)$ such that

$$\rho\big(A(x), A(y)\big) \leq \alpha\, \rho(x, y) \text{ for all } x, y \in X.$$

Definition 11. Let (X, ρ) be a metric space and let $A : X \to X$. If a point $x \in X$ satisfies the equation $A(x) = x$, it is called a *fixed point.*

Theorem 3 (Contraction Mapping Theorem)**.** *Let (X, d) be a complete metric space. Then a contraction map A defined on (X, d) has a unique fixed point.*

Proof. Let $x_0 \in X$ and contruct a sequence $\{x_n\}$ so that $x_n = A(x_{n-1}) = A^n(x_0)$ for $n \in \mathbb{N}$. Then, $x_1 = A(x_0)$, $x_2 = A(x_1) = A^2(x_0), \ldots$, and in general $x_{n+1} = A(x_n) = A^{n+1}(x_0)$.

We want to show that $\{x_n\}$ is a Cauchy sequence: Let $n \leq m$ and A be a contraction mapping:

$$\rho(x_n, x_m) = \rho(A^n x_0, A^m x_0) \leq \alpha^n \rho(x_0, x_{m-n}).$$

Then using triangle inequality:

$$\begin{aligned}
\alpha^n \rho(x_0, x_{m-n}) &\leq \alpha^n \rho(x_0, x_1) + \rho(x_1, x_2), + \ldots + \rho(x_{m-n-1}, x_{m-n}) \\
&\leq \alpha^n \rho(x_0, x_1)[1 + \alpha + \ldots \alpha^{m-n-1}] && \text{from factoring and the definition of } x_n \\
&\leq \alpha^n \rho(x_0, x_1) \left(\frac{1}{1-n}\right) && \text{by substituting in the geometric series.}
\end{aligned}$$

The geometric series converges to 0 as $n \to \infty$, thus our sequence $\{x_n\}$ is Cauchy. Then by definition of completeness, there exists a point $x \in X$ such that x_n converges to x as $n \to \infty$. It seems like this point would be a good candidate for our fixed point. So let's see if our point x solves the equation $A(x) = x$. We know from the continuity of A that we can do the following rearrangements:

$$A(x) = A \lim_{n\to\infty} x_n = \lim_{n\to\infty} Ax_n = \lim_{n\to\infty} x_{n+1} = x$$

which satisfies our definition of a fixed point.

Thus we have shown that our point limit point $x \in X$ is a fixed point and all that's left is to show uniqueness: Suppose there are distinct $x, y \in X$ such that $Ax = x$ and $Ay = y$ and $\rho(Ax, Ay) = \rho(x, y)$. But we know $\rho(Ax, Ay) \leq \alpha\rho(x, y)$. Since $\alpha < 1$, in order for

$$\rho(x, y) \leq \alpha\rho(x, y) \text{ to be true, we must have that } \rho(x, y) = 0$$

Then because ρ is a metric, $\rho(x, y) = 0$ means $x = y$. Thus we have arrived at a contradiction and the fixed point x must be unique. [2]

□

4.1 Proof of the Collage Theorem

Proposition 5. *For every $B, C, D, E \in \mathcal{K}$, $\rho(B \cup C, D \cup E) \leq \max\big\{\rho(B, D), \rho(C, E)\big\}$.*

Proof. Without loss of generalization, let's say that $\rho(C, E) \leq \rho(B, D)$ so that $\rho(B, D) = \max\big\{\rho(B, D), \rho(C, E)\big\}$. Now, we want to show that $\rho(B \cup C, D \cup E) \leq \rho(B, D)$. Let

$$\rho(B, D) = \inf\{\delta > 0 : B \subset D_\delta \text{ and } D \subset B_\delta\} = \delta_0.$$

We are gauranteed an infimum δ_0 for this set because it is bounded below (Completeness Axiom). Then $\rho(C, E) \leq \delta_0$ so we know $C \subset E_{\delta_0}$ and $E \subset C_{\delta_0}$. Now we have that

$$B \cup C \subset D_{\delta_0} \cup E_{\delta_0} \subset (D \cup E)_{\delta_0}.$$

A similar argument can be made for $D \cup E \subset (B \cup C)_{\delta_0}$ and then we will have shown that

$$\rho(B \cup C, D \cup E) \leq \delta_0 = \max\{\rho(B, D), \rho(C, E)\}.$$

□

Definition 12. An *iterated function system* (**IFS**) is a finite set of contraction mappings on a complete metric space. It is notated as the set $\{w_1, w_2, \dots .w_n\}$ for $n \in \mathbb{N}$ with respective contraction factors $\alpha_1, \alpha_2, \dots, \alpha_n$.

Definition 13. Let A be any compact subset of $\mathbb{R}^2$. Then if $W(A) = \cup_{i=1}^{\infty} w_i$, where $w_i : i = 1, \dots, n$ are contractions with factor $\alpha = \max\left\{a_1, a_2, \dots, a_n\right\}$, we call W a *Hutchinson operator.*

Definition 14. Let $W(A)$ be a Hutchinson operator for the IFS $\{w_1, w_2, \dots, w_n\}$. Then the set that satisfies the equation $W^n(A) = A$ is called the *attractor* of the IFS.

Theorem 4 (The Collage Theorem)**.** *Let (X, d) be a complete metric space, H be a subset of the space $\mathcal{K}$ as defined above. Choose an IFS $\{w_1, w_2, \dots, w_n\}$ with contractivity factor $\alpha \in [0, 1)$ and define the Hutchinson operator W as the union of the transformations in the IFS.*

Then for some $\epsilon \geq 0$,

$$\rho\big(H, W(H)\big) \leq \epsilon \quad \textit{implies} \quad \rho(H, A) \leq \frac{\epsilon}{1-\alpha},$$

where A is the attractor of the IFS. Equivalently,

$$\rho(H, A) \leq \frac{\rho\big(H, W(H)\big)}{1-\alpha}.$$

[3]

Proof. First, we must show that $W : \mathcal{K} \to \mathcal{K}$ is a contraction mapping.
We will prove this by induction starting with $n = 2$. Let $H_1, H_2 \subseteq X$. Then by Proposition 5,

$$\begin{aligned}\rho\big(W(H_1), W(H_2)\big) &= \rho\big(W_1(H_1) \cup W_2(H_2), W_1(H_2) \cup W_2(H_2)\big)\\ &\leq \max\left\{\rho\big(W_1(H_1), W_1(H_2)\big), \rho\big(W_2(H_1), W_2(H_2)\big)\right\}\\ &= \max\left\{\, \alpha_1\, \rho(H_1, H_2),\ \alpha_2\, \rho(H_1, H_2)\right\}\\ &= \alpha\, \rho(H_1, H_2)\end{aligned}$$

where $\alpha = \max\{\alpha_1, \alpha_2\}$ So that in general:

$$\begin{aligned}\rho\big(W(H_1), W(H_2)\big) &= \rho\big(W_1(H_1) \cup W_2(H_2) \cup \ldots \cup W_n(H_1), W_1(H_2) \cup W_2(H_2) \cup \ldots \cup W_n(H_2)\big)\\ &\leq \max\left\{\rho\big(W_1(H_1), W_1(H_2)\big), \rho\big(W_2(H_1), W_2(H_2)\big), \dots, \rho\big(W_n(H_1), W_n(H_2)\big)\right\}\\ &\leq \max\left\{\alpha_1\rho(H_1, H_2), \alpha_2\rho(H_1, H_2), \dots, \alpha_n\rho(H_1, H_2)\right\}\\ &\leq \alpha\, \rho(H_1, H_2)\end{aligned}$$

where $\alpha = \max\{\alpha_1, \alpha_2, \dots, \alpha_n\}$. Therefore, W is a contraction mapping with contractivity factor α. We can now use the Contraction Mapping Theorem which tells us that $W(H)$ has a unique fixed point/attractor A where $W(A) = A$. It also gives us a method to find that fixed point. We choose any $x_0 \in H$ and iterate successively to get our next points: $x_1 = W(x_0)$ and in general $x_{n+1} = Wx_n = W^n(x_0)$. We use the notation W^n denote the the composition of W with itself n times so that $W^{n+1}(x) = W(W^n(x_0))$. Then our sequence $\{W^n(x_0\}$ will converge to a unique fixed point, called the attractor A. Since our choice of x_0 was arbritrary, we can say that $\{W^n(H)\}$

converges to A so that $\lim_{n\to\infty} W^n(H) = A$. We also know that W is a continous function, so $\rho(H, A) = \rho\big(H, \lim_{n\to\infty} W^n(H)\big) = \lim_{n\to\infty} \rho\big(H, W^n(H)\big)$. By the triangle inequality, we can say

$$\begin{aligned}
\rho(H, A) &\leq \lim_{n\to\infty} \rho\big(H, W(H)\big) + \rho\big(W(H), W^2(H)\big) + \cdots + \rho\big(W^{n-1}(H), W^n(H)\big) \\
&\leq \lim_{n\to\infty} (1 + \alpha + \cdots + \alpha^{n-1})\rho\big(H, W(H)\big) \\
&\leq \frac{\rho\big(H, W(H)\big)}{1-\alpha}
\end{aligned}$$

Thus if $\rho\big(H, W(H)\big) \leq \epsilon$,

$$\rho(H, A) \leq \frac{\rho\big(H, W(H)\big)}{1-\alpha} \leq \frac{\epsilon}{1-\alpha}.$$

□

What this theorem tells us is that if we can make a collage of the image whose Hausdorff distance to the idealized fractal is small, we can then say the distance between the attractor of our IFS and the fractal will be small. Although it doesn't immediately solve the inverse problem of finding an IFS for a certain fractal, it does guide us on how we should approach the problem- by covering the fractal with transformations of the fractal as closely as possible. It also gaurantees that an attractor of the IFS exists and is unique. So regardless of what our initial point is, we will converge to the same attractor.

5 Algorithmic Generation of Fractals

Now that we know how to generate the attractor of our IFS, or at least what we should aim for when designing our IFS, we can implement our methods using the algorithms discussed in the following section. Both use the ideas that result from the Collage Theorem, which tells us that if the distance between our desired fractal and the union of the fractal under contractive mappings is small, then our IFS will converge to a unique attractor, whose distance to the fractal will be small. Thus, if we choose our IFS wisely, it will produce an attractor which will resemble our desired fractal. Luckily, we already have some parameters from the examples in Barnsley [3] and Peitgen et al [5] and we will use those for our examples.

5.1 The Deterministic Algorithm

This algorithm is an example of using the Hutchinson operator from the Collage Theorem. We draw a rough outline of the object and then cover it as closely as possible by a number of smaller similar or affine copies to form a sort of "collage". Then we iterate with our resulting collage until we start to see our desired fractal-like image emerge.

Using the terms we discussed above: we take any initial set H such as a unit square and iterate the operator W from our IFS $\{w_1, w_2, \ldots w_n\}$ k times to get the k^{th} iterate/approximation $W^k(H)$to H for a suitable value k. Then for k large enough, we get the important identity $W(H) = H$ where the k^{th} iterate of the square starts to resemble the fractal and becomes the fixed point of our IFS.

The deterministic algorithm can also be described with an analogy of the Multiple Reduced Copy Machine (MRCM) with multiple lenses. As input, it will take an image and set of parameters for each lens. The parameters specify values for our transformations on our given image. The kinds of transformations that MRCM takes are called affine linear transformations, which include scaling, shearing, reflection, rotation and translation. They take the parameters r, s, θ, ψ, e, f where

$$\begin{bmatrix} x_{new} \\ y_{new} \end{bmatrix} = \begin{bmatrix} r\cos\theta & -s\sin\psi \\ r\sin\theta & s\sin\psi \end{bmatrix} \begin{bmatrix} x_{old} \\ y_{old} \end{bmatrix} + \begin{bmatrix} e \\ f \end{bmatrix}$$

which we can simplify with the coefficients a, b, c, d, e, f and the matrix

$$\begin{bmatrix} a & b & e \\ c & d & f \end{bmatrix} \text{ where } \begin{bmatrix} x_{new} \\ y_{new} \end{bmatrix} = \begin{bmatrix} ax + by + e \\ cx + dy + f \end{bmatrix}$$

As output, it will assemble all of the transformed images ionto one image. The resulting image is then fed back into the machine as the input for iterative process.

Example 9. To create a "blueprint" for the Barnsley fern, we have the following picture. One of the transformations make most of the fern, two transformations make the smaller leaves, and the last transformation form the stem.

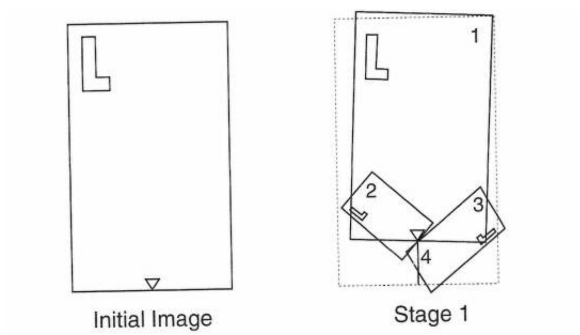


Figure 1: Collage blueprint for the fern [5]

Then if we feed the blueprint back into the machine and then iterate successively with the obtained collages, we will eventually produce a relatively detailed fern. The fern on the left of the following figure is after 5 iterations, and on the right after 10 iterations.

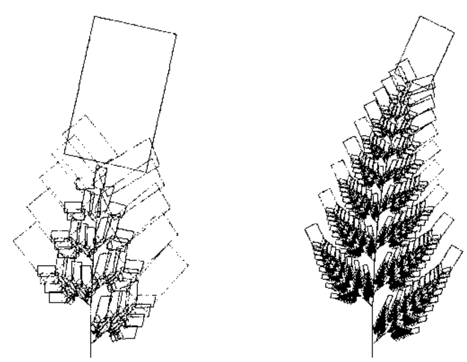


Figure 2: After 5 and 10 iterations [5]

However after a few more iterates, this process will take a considerably more time. Take as our initial set a rectangle whose length takes up 1000 pixels. Then if we want to reproduce our fern so that each rectangle becomes a point or 1 pixel, we use this formula to calculate how many iterations N we needto do. Since our fern contracts 85% in our blueprint, N is given by

$$1 = 1000 \times .85^N.$$

Thus it will take about iterations in order to get our fern to be somewhat detailed. Then we would have to calculate 42 iterates, which amounts to drawing $1 + 4^1 + 4^2 + \cdots + 4^N = \frac{4^{N+1}-1}{3}$ rectangles. For $N = 42$, this is a lot of rectangles - on the order of 2×10^{25}. Even with a very fast computer, this would take many many years. [5]

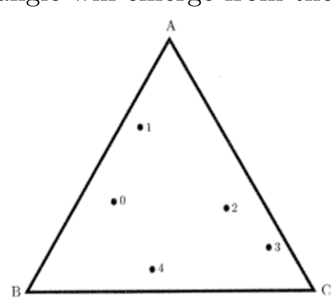
5.2 The Random Iteration Algorithm

This algorithm uses the idea of the chaos game. The chaos game was formulated by Michael Barnsley as method to generate classical fractals. The game goes as follows:

1. Pick a polygon and any initial point inside that polygon.
2. Choose a vertex of the polygon at random.
3. Move a fraction of the length towards the chosen vertex.
4. Repeat steps (2-3) with the new point.

By successively iterating in this way, we will produce our fractal shape. This method will plot points in random order all over the attractor.

Example 10. The chaos game for the Sierpinski triangle goes as follows: Start somewhere inside a triangle, then move halfway to a corner A, B, or C and then take that point and repeat 1-2. After a while the Sierpinski triangle will emerge from the image that the points trace out.



We can formalize the process of the chaos by using the concepts we discussed earlier such contraction mappings and convergent sequences. Let's say we have a fractal image H which we would like to approximate with the attractor of an IFS $\{w_1, w_2, \ldots, w_n\}$. We start with any initial point $x_0 \in \mathbb{R}$ and select a contraction $w_i : i = 1, 2, \ldots, n$ at random. Let $x_1 = w_{i_1}(x_0)$ and continue this way, choosing w_{i_n} at random and letting $x_n = w_{i_n}(x_{n-1})$ for $k \in \mathbb{N}$. We can construct a sequence $\{x_n\}$ with the points x_n for $n \in \mathbb{N}$. Then for n large enough, the sequence $\{x_n\}$ will converge to the attractor of the IFS and $\{x_k\}$ will be randomly distributed across F.

Example 11. With simple images where all the contraction factors are the same, we can assign equal probabilities to all of the transformations which is what I have done here with Sierpinski's triangle.

a	b	c	d	e	f
.5	0	0	.5	0	0
-.5	0	0	.5	.5	0
.5	0	0	-.5	0	.5

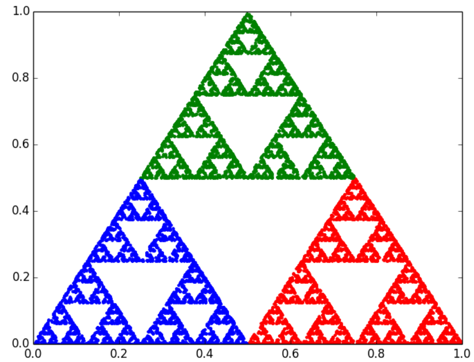
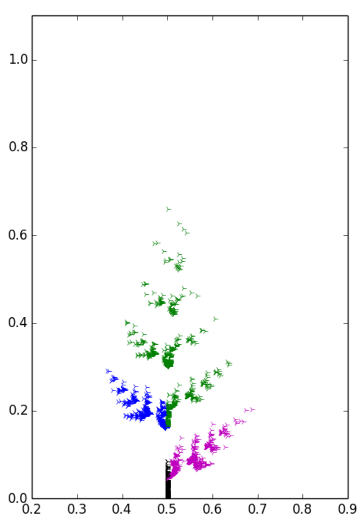


Figure 3: Sierpinski Triangle

However with a more natural fractal that doesn't necessarily have strictly self-similar properties, we will be fairly disappointed with our results when we assign equal probabilities, even with 10000 iterations.



i	a	b	c	d	e	f
1	.849	.037	-.037	.849	.075	.1830
2	.197	-.266	.226	.197	.4	.0470
3	-.15	.283	.26	.237	.575	-.084
4	0	0	0	.160	.5	0

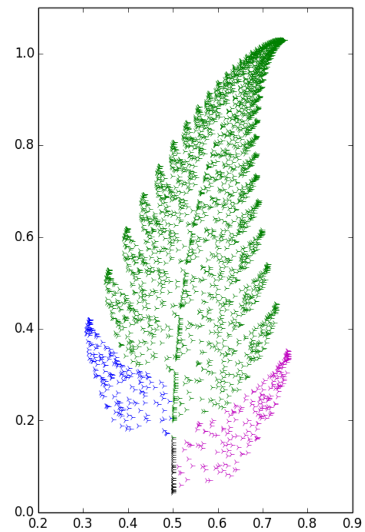
Figure 4: IFS table for the Barnsley Fern

To improve on this picture, we can do two things: throw out the first few thousand iterations and assign probabilities to each transformation. Since our sequence is converging to the attractor, we don't need to plot the points that we obtained in the beggining of the sequence. The probabilities are mass distributions given by

$$p_i = \frac{\det w_i}{\sum_{i=1}^{n} \det w_i}$$

where the sum of all the probabilities is 1. The theory behind this involves measures and is beyond the scope of this project, but we can use this to produce much more satisfying result with the same number of iterations.

i	1	2	3	4
p	.01	.85	.07	.07



Our refined algorithm, taking in account of mass distributions is as follows:

1. Take any initial point $x_0 \in H$.
2. Choose a transformation w_i chosen based on probability p_i.
3. Apply the chosen transformation onto our point.
4. Repeat steps (2-3) with the new point.

This method will converge to our attractor much quicker than with equal probabilities.

6 Future Work

The main inspiration for studying fractals for my capstone project was to find a way to implement them into a more abstract process, mainly in musical composition. For example, one could use the Cantor set as a basis for the rhythmic structure of a piece or as a model of how to scale and move a melody along in time. As for the implementation of 2-dimensional transformations, I have yet to figure out what each scaling, rotation, and translation would do to a set of pitches and durations. In the future, I would like to work with these fractal-like structures to sort of hear what fractals sound like. I would also like to learn more about the measure theory in order to get a better understanding of how to determine the Hausdorff measure and dimension of a fractal.

7 Acknowledgements

I would like to thank Jane McDougall for her tremendous patience and generosity in helping me with this project. Also, thanks to everyone else in the Colorado College Math and Computer Science Department.

8 Appendix

Python Script for Generating a Fern

```
import numpy as np
import random
import pylab as pl
import matplotlib

# define some matrices that are transformations
w1 = [np.matrix([[0,0],[0,.16]]), np.matrix([[0.5],[0]])]
w2 = [np.matrix([[.849, .037],[-.037, .849]]), np.matrix([[0.075],[0.1830]])]
w3 = [np.matrix([[.197,-.226],[.226,.197]]), np.matrix([[0.4],[0.049]])]
w4 = [np.matrix([[-.15,.283],[.26,.237]]), np.matrix([[0.575],[-0.084]])]

# set probabilities of each transformation
p1 = .01
p2 = .85
p3 = .07
p4 = .07

# initialize
x_0 = 0.5
y_0 = 0
n = 10000 # number of iterations

point = np.matrix([[x_0],[y_0]])

# pick function depending on probs
for c in range(n):
    r = random.random()
    markerstyle='2'
    if c > 7000:
        if r <= p1:
            point = w1[0] * point + w1[1]
            pl.plot(point.flat[0], point.flat[1], color='k',  marker=markerstyle)
        elif r <= p1 + p2:
            point = w2[0] * point + w2[1]
            pl.plot(point.flat[0], point.flat[1], color='g',  marker=markerstyle)
            # print point
        elif r <= p1 + p2 + p3:
            point = w3[0] * point + w3[1]
            pl.plot(point.flat[0], point.flat[1], color='b',  marker=markerstyle)
        else:
            point = w4[0] * point + w4[1]
            pl.plot(point.flat[0], point.flat[1],color='m', marker=markerstyle)

pl.xlim(.2,.9)
pl.ylim(0,1.1)

pl.show()
```

References

[1] B.S. Thompson A.M. Bruckner, J.B. Bruckner. *Elementary Real Analysis*. ClassicalRealAnalysis.com, second edition, 2008.

[2] B.S. Thompson A.M. Bruckner, J.B. Bruckner. *Real Analysis*. ClassicalRealAnalysis.com, second edition, 2008.

[3] Michael Barnsley. *Fractals Everywhere*. Dover, second edition, 2012. (Original work published 1993).

[4] David P. Feldman. *Chaos and Fractals: An Elementary Introduction*. Oxford, first edition, 2012.

[5] Dietmar Saupe Heinz-Otto Peitgen, Hartmut Jürgens. *Chaos and Fractals: New Frontiers of Science*. Undergraduate Texts in Mathematics. Springer-Verlag, second edition, 2004.

[6] Kenneth A. Ross. *Elementary Analysis: The Theory of Calculus*. Undergraduate Texts in Mathematics. Springer, second edition, 2013.