2016

Fractal Shapes Generated by Iterated Function Systems

Mary Catherine McKinley
Louisiana State University and Agricultural and Mechanical College

FRACTAL SHAPES GENERATED BY ITERATED FUNCTION SYSTEMS

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Masters of Science

in

Mathematics

by
Mary Catherine McKinley
BS, Spring Hill College, 2015
December 2016

Acknowledgments

The author would like to thank Dr. Frank Neubrander for his support and guidance on this project; Dr. James Madden for his help with the Mathematica programs; Dr. Oliver Dasbach and Dr. James Oxley for their unending support and encouragement; James Frye for listening and believing in me; and most of all my family for being taught the beauty of fractals during this project and never complaining about it.

Table of Contents

List of Tables

List of Figures

Abstract

This thesis explores the construction of shapes and, in particular, fractal-type shapes as fixed points of contractive iterated function systems as discussed in Michael Barnsley's 1988 book "Fractals Everywhere." The purpose of the thesis is to serve as a resource for an undergraduate-level introduction to the beauty and core ideas of fractal geometry, especially with regard to visualizations of basic concepts and algorithms.

Chapter 1
Introduction

We begin by introducing some key ideas that will be important in explaining the space where shapes and fractal-type shapes live. According to Dictionary.com, a fractal is "a rough or fragmented geometric shape that can be subdivided in parts, each of which is (at least approximately) a smaller copy of the whole. Fractal shapes are generally self-similar (bits look like the whole) and independent of scale (they look similar, no matter how close you zoom in)." They also have the characteristic of having a non-integer dimension, a concept that will also be discussed below.

According to this definition, fractal shapes are two-dimensional shapes with special properties. Thus, fractals live in "the space of shapes;" that is, fractals are, in fact, a special type of shapes. What exactly is a shape? Let us recall that a subset X of $\mathbb{R}^2$ is *closed* if it contains all of its limit points; X is *bounded* if all of its points are contained in a circle of sufficiently large radius; X is *compact* if it is both closed and bounded. We denote by $H(\mathbb{R}^2)$ the set of compact subsets of $\mathbb{R}^2$. This is the space where fractal shapes live.

Definition 1.1 *A shape is a compact subset of* $\mathbb{R}^2$*. The set of all shapes is denoted by* $\mathcal{H}(\mathbb{R}^2)$*.*

A classical example of a fractal shape is the Sierpinski triangle, named after the Polish mathematician Waclav Sierpinski (1882-1969). One way to approximate the isosceles Sierpinski triangle is to continuously remove the open middle isosceles triangle, supposing that the original isosceles triangle is divided into four equivalent triangles as shown below:

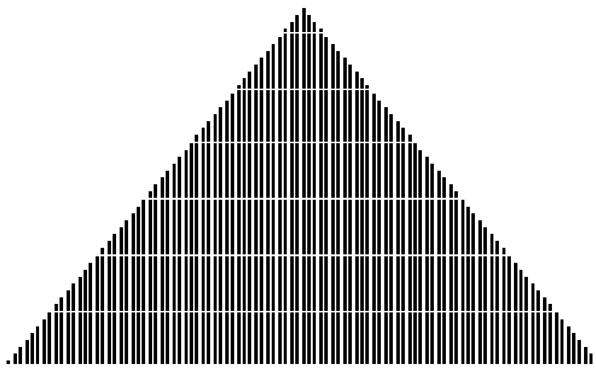

Figure 1.1: S_0 Original Triangle

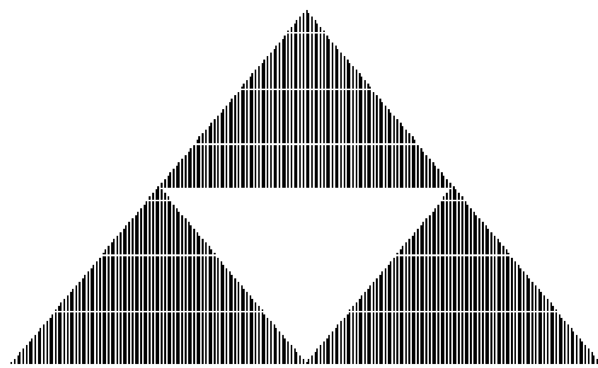

Figure 1.2: First Iteration

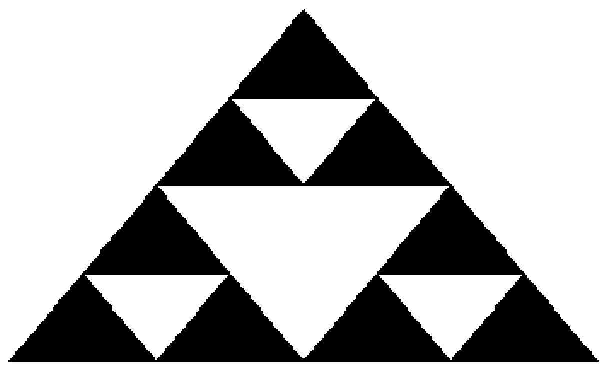

Figure 1.3: Second Iteration

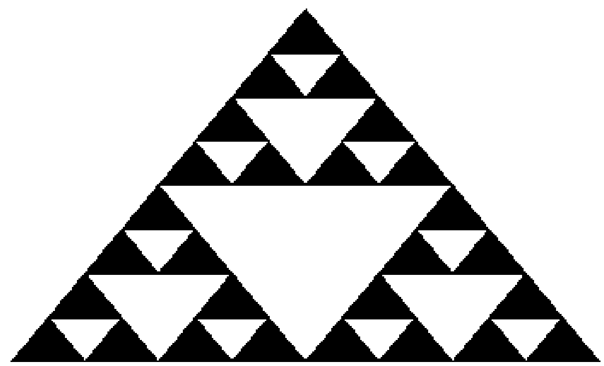

Figure 1.4: Third Iteration

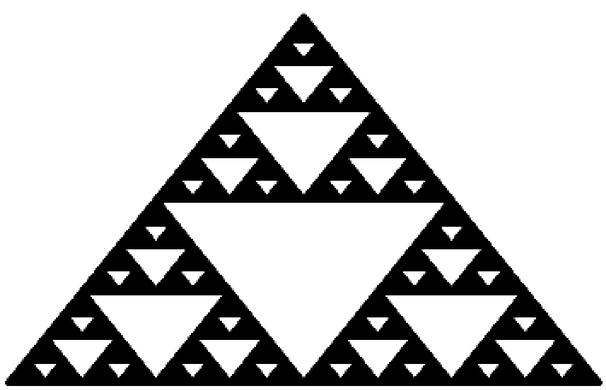

Figure 1.5: Fourth Iteration

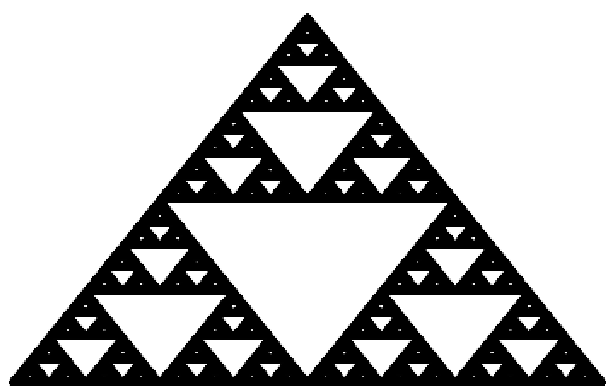

Figure 1.6: Fifth Iteration

It is important to note that this only works if our initiator is an isosceles triangle. Also it should be clear that the shapes $S_6, S_7, S_8, \ldots$ obtained after the sixth and higher iterations are visually the same sets due to the lack of sufficient resolution in the printed versions of the shapes. Also, by construction, the sets $S_1, S_2, S_3, \ldots$ will all be compact sets in $\mathcal{H}(\mathbb{R}^2)$ since we always remove the open middle triangle (that is, the middle triangle

without its boundary). What is not clear about the Siepinski triangles is that the "limit shape" $S_\infty :=$ "$\lim_{n\to\infty} S_n$" is still a compact set. To see this we use an alternative way to generate the limit set $S_\infty :=$ "$\lim_{n\to\infty} S_n$". This alternative way is known as the "iterated function system" approach, where we denote the original isosceles triangle by S_0 (Figure 1.1) with vertices at $(0,0)$, $(1,0)$, $(\frac{1}{2},1)$. Looking at the shape S_1 (Figure 1.2) we see that it consists of three copies of S_0, each scaled by a factor of $1/2$ and shifted in its appropriate position. Thus,

$$S_1 = T(S_0),$$

where the function $T : H(\mathbb{R}^2) \to H(\mathbb{R}^2)$ is defined as

$$T(S) := T_1(S) \cup T_2(S) \cup T_3(S) \tag{1.1}$$

with

$$T_1(S) := \frac{1}{2}S \quad \text{or } T_1(x,y) = (\frac{x}{2}, \frac{y}{2}),$$
$$T_2(S) := \frac{1}{2}S + (\frac{1}{2}, 0) \quad \text{or } T_2(x,y) = (\frac{x}{2} + \frac{1}{2}, \frac{y}{2}),$$
$$T_3(S) := \frac{1}{2}S + (\frac{1}{4}, \frac{1}{2}) \quad \text{or } T_3(x,y) = (\frac{x}{2} + \frac{1}{4}, \frac{y}{2} + \frac{1}{2}).$$

Now, analyzing the shape S_2 in Figure 1.3, one sees that it consists of three copies of the shape S_1, each reduced by a factor of $\frac{1}{2}$; that is,

$$S_2 = T(S_1) = T(T(S_0)) = T^2(S_0)$$

where T is defined as above. Continuing in this manner, one obtains the shapes: $S_0, S_1, \cdots, S_5$ above (see Figures 1.1 - 1.6).

The Sierpinski triangle S_∞ is the "limit-shape" that we obtain by continuing this process; that is,

$$S_\infty = \lim_{n\to\infty} T^n(S_0).$$

We can easily observe that S_n consists of 3^n triangles with

$$\text{perimeter } P_n = (\frac{3}{2})^n(1+\sqrt{5}) \text{ and area } A_n = \frac{3^n}{2^{2n+1}}.$$

Thus, S_0 has perimeter $1+\sqrt{5} \geq 2.3$ and area $\frac{1}{2}$, S_5 consists of $3^5 = 243$ triangles with total perimeter $(\frac{3}{2})^5(1+\sqrt{5}) \geq 24.5$ and total area $\frac{3^5}{2^4} \leq 0.12$, S_{10} consists of $3^{10} = 59,049$ triangles with total perimeter of $(\frac{3}{2})^{10}(1+\sqrt{5}) \geq 128.9$ and total area $\frac{3^{10}}{2^{21}} \leq 0.03$. Thus, the Sierpinski triangle S_∞ will have infinite perimeter and no area. In addition to examining the perimeter and area of the Sierpinski triangle, S_∞, we can look at its dimension.

One of the properties of fractal shapes is that they have non-integer dimensions. The way to calculate the dimension of fractal shapes is by using the (Hausdorff) Fractal Dimension formula developed by German Mathematician Felix Hausdorff [11] (1868-1942). It is a way of defining the complexity of self-similar shapes. The fractal dimension is defined as follows

$$D = \frac{\ln(N)}{\ln(r)}, \tag{1.2}$$

where N is the number of self-similar shapes and r is the scaling factor.

Let us first calculate the fractal dimension for some common Euclidean elements that exhibit self similarity characteristics. We will calculate the dimensions of a line and a square.

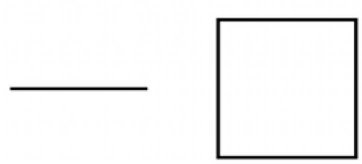

Figure 1.7: Euclidean Elements Exhibiting Self Similarity

Example 1.2 (Fractal Dimension of a Line). For a line L, whatever contractivity factor

we choose $N = r$. So, for example, if we scale the line by $\frac{1}{4}$ then it will take four shortened lines to be as long as our original L. Thus, $N = 4$ and $r = 4$. Therefore,

$$D = \frac{\ln(N)}{\ln(r)} = \frac{\ln(4)}{\ln(4)} = 1.$$

So, lines are one dimensional.

Example 1.3 (Fractal Dimension of a Square). Similarly, for a square S, if we choose $r = 2$ we are reducing the square by $\frac{1}{2}$. However, we will need four reduced squares to achieve the same size as our original S. Thus, $N = 4$. Therefore,

$$D = \frac{\ln(N)}{\ln(r)} = \frac{\ln(4)}{\ln(2)} = 2.$$

So, squares are two dimensional.

Example 1.4 (Fractal Dimension for Sierpinski triangle). We can calculate the fractal dimension of the Sierpinski triangle from this introduction. By our discussion, we know that there are three times as many triangles in each iteration of the Sierpinski triangle. Thus, $N = 3$. Also, we know that the contractivity factor for the Sierpinski triangle is $\frac{1}{2}$. Thus, the scaling factor will be $r = 2$. Therefore,

$$D = \frac{\ln(N)}{\ln(r)} = \frac{\ln(3)}{\ln(2)} \approx 1.58496$$

□

Why is S a shape though? It is clearly bounded, but why is it closed? To answer this question we have to provide in Chapter 2 some mathematical tools like complete metric spaces and the "Contraction Mapping Principle," also known as Banach's Fixed Point Theorem.

Before we begin the discussion of the "Contraction Mapping Principle", we briefly examine one of the oldest examples of a fractal object discussed in mathematical literature,

the coastline of Britain. Benoit Mandelbrot, a Polish-born, French and American mathematician [10], is credited with coining the word "fractal". Mandelbrot studied the coast of Britain and saw that the coastline was extremely rugged.

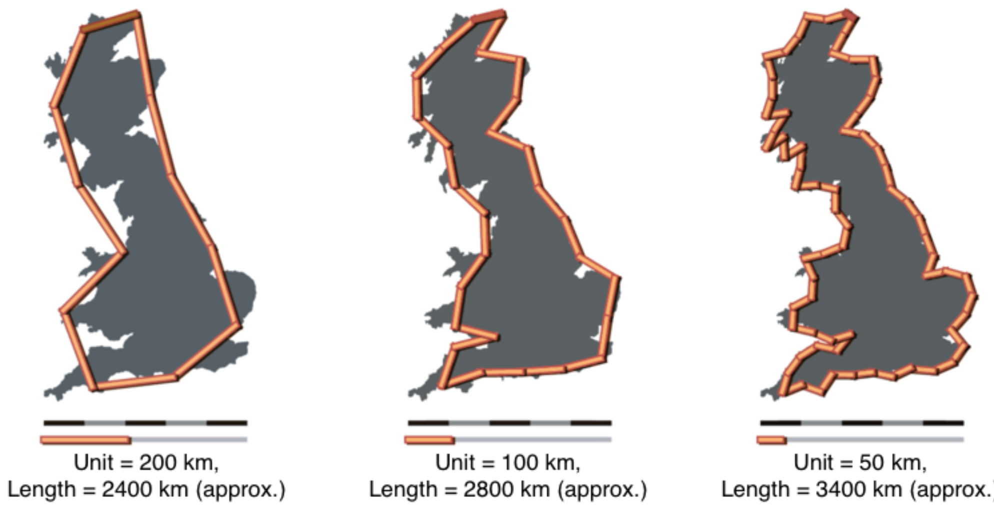

Figure 1.8: Measuring the British Coastline

One can see in the first image that a stretch of coast can be measured using a straight line from one section to another. However, this does not usually account for the irregularities and jagged parts of a country's coastline. Thus, the second image shows the process repeated, but using measuring units of half the length of the first. The approximate length of the coastline increases and the accuracy improves. Again, the third image uses measuring units half the distance of the second (or a quarter of the first) and estimates an even larger length.

Mandelbrot stated that the fractal dimension of the coastline of Britain is approximately 1.25 in [6], but he unfortunately did not specify what scaling factor he used and the number of self similar shapes that factor produced.

However, looking at the pictures above (taken from [2]), we see that when the scaling factor is $r = 2$, then we need about $N = \frac{28}{12} = \frac{7}{3}$ from the first image to the second and $N = \frac{68}{28} = \frac{17}{7}$ as many line segments to cover the length of the coast line from the second

image to the third. Thus, approximately, $D \approx \frac{\ln(N)}{\ln(r)} = \frac{\ln(\frac{7}{3})}{\ln(2)} = 1.22$ for our first N and $D \approx \frac{\ln(N)}{\ln(r)} = \frac{\ln(\frac{17}{7})}{\ln(2)} = 1.28$ for our second N. Both numbers that is pretty close to the one cited by Mandelbrot.

Using the number 1.25 as the dimension of the coastline, it follows that the number of N of measuring sticks needed when the unit measure is reduced by a factor of $\frac{1}{2}$ increases by $2^{1.25}$ (since $D\ln(r) = \ln(N)$ or $r^D = N$). This yields the following table.

Table 1.1: Calculating Fractal Dimension for the British Coastline

Unit	N	Length
100	28	2,800
50	$28 * 2^{1.25} \approx 66.6$	3,330
25	$28 * (2^{1.25})^2 \approx 158.4$	3,960
12.5	$28 * (2^{1.25})^3 \approx 376.7$	4,709
6.25	$28 * (2^{1.25})^4 = 896$	5,600
$100 * \frac{1}{2^n}$	$28 * 2^{1.25n}$	$28 * 2^{1.25n} * \frac{100}{2^n} = 2800 * 2^{\frac{n}{4}}$

Thus, according to Mandelbrot, if the unit measure is reduced by a factor of $(\frac{1}{2})^4 = \frac{1}{16}$, then the length of the coastline will be doubled to 5,600 km. Moreover, if the unit measure is reduced from 100 km by a factor of $(\frac{1}{2})^{23}$ to 1.1921 cm, then the length of the coast line is predicted to be 150,689 km, almost four times the circumference of the Earth.

Chapter 2
Complete Metric Spaces

We now begin the more formal discussion of fractal shapes with a series of definitions that will lead to the mathematical definition of $\mathcal{H}(\mathbb{R}^2)$.

Definition 2.1 *A metric space ($\boldsymbol{H}$, h) is a set $\boldsymbol{H}$ together with a positive, real-valued function $h : \boldsymbol{H} \times \boldsymbol{H} \to \mathbb{R}^+$ that measures the distance between elements A and B in $\boldsymbol{H}$; that is,*

(I) $h(A, B) = h(B, A)$ *for all* $A, B \in \boldsymbol{H}$.

(II) $h(A, B) = 0$ *if and only if* $A = B$.

(III) $h(A, A) = 0$ *for all* $A \in \boldsymbol{H}$.

(IV) $h(A, B) \leq h(A, C) + h(C, B)$ *for all* $A, B, C \in \boldsymbol{H}$.

A sequence $\{A_n\}(n \in \mathbb{N})$ of elements in a metric space ($\mathbf{H}$, h) is called a *Cauchy sequence* if, for any given number $\epsilon > 0$, there is an integer $N \in \mathbb{N}$ so that

$$h(A_n, A_m) < \epsilon \text{ for all } n, m \geq N.$$

A metric space, $(\mathbf{H}, d)$, is *complete* if every Cauchy sequence $\{A_n\}$ in $\mathbf{H}$ has a limit $A \in \mathbf{H}$. That is, $A = \lim_{n\to\infty} A_n$ if, for any given number $\epsilon > 0$, there is an integer $N \in \mathbb{N}$ such that

$$h(A_n, A) < \epsilon \text{ for all } n > N.$$

Example 2.2 The rational numbers $\mathbb{Q}$ and real numbers $\mathbb{R}$ are metric spaces with the canonical metric

$$h(x, y) = |x - y|$$

for $x, y \in \mathbb{Q}$ (or $\mathbb{R}$). To show that $\mathbb{Q}$ is not a complete metric space, we consider the function $f(x) = x^2 - 2$ and the sequence of rational numbers $\{x_n\}$ generated by Newton's Iterative Formula

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)} = \frac{x_n}{2} + \frac{1}{x_n} \text{ with } x_0 = 1.$$

Thus, we consider the sequence generated from the following Mathematica code.

Program: Newton's Method for $f(x) = x^2 - 2 = 0$

```
RecurrenceTable{{a[n+1] == a{n}/2 + 1/a[n], a[1] == 1}, a, {n, 1, 7}]
```

$$(1, \frac{3}{2}, \frac{17}{12}, \frac{577}{408}, \frac{665857}{470832}, \frac{886731088897}{627013566048}, \frac{1572584048032918633353217}{1111984844349868137938112} \ldots)$$

A straightforward induction shows that $1 \leq x_n \leq 2$ for all $n \in \mathbb{N}$. Let $\alpha := \sqrt{2}$. From basic calculus we know that:

$$f(x) = x^2 - 2,\ f'(x) = 2x,\ f''(x) = 2,\ f'''(x) = 0.$$

By Taylor's Theorem, see [7], and the fact that the n-th derivatives of f are zero for $n \geq 3$, it follows that

$$0 = f(\alpha) = f(x_n) + f'(x_n)(\alpha - x_n) + \frac{f''(x_n)}{2!}(\alpha - x_n)^2 + \frac{f'''(z)}{3!}(\alpha - x_n)^3,$$

where z is a number between α and x_n. Since $f'(x_n) \neq 0$ for all $n \in \mathbb{N}$, we obtain

$$\begin{aligned} 0 &= \frac{f(x_n)}{f'(x_n)} + \alpha - x_n + \frac{1}{2}\frac{f''(x_n)}{f'(x_n)}(\alpha - x_n)^2 \\ &= \alpha - x_{n+1} + \frac{1}{2}\frac{f''(x_n)}{f'(x_n)}(\alpha - x_n)^2 \\ &= \alpha - x_{n+1} + \frac{1}{2x_n}(\alpha - x_n)^2. \end{aligned}$$

The term x_{n+1} in the second line is obtained from Newton's Iterative Formula $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$. Also, the third line follows from $f'(x) = 2x$. This shows that $|\alpha - x_{n+1}| = \frac{1}{2x_n}(\alpha - x_n)^2 \leq \frac{1}{2}|\alpha - x_n|^2$, and therefore

$$|\alpha - x_n| \leq \frac{1}{2^n}|\alpha - x_0|^{2n} \leq \frac{1}{2^{3n}}$$

for all $n \in \mathbb{N}$. According to this estimate, the difference between

$$a_5 = \frac{665857}{470832} = 1.41421356237469\ldots$$

and $\sqrt{2}$ is at most $(\frac{1}{2})^2 \leq 3.06 \times 10^{-5}$. The true error is, in fact, less then 2×10^{-12} (since $\sqrt{2} = 1.41421356237310\ldots$).

Since every convergent sequence is a Cauchy sequence and since one knows that $\sqrt{2} \notin \mathbb{Q}$, it follows that $\mathbb{Q}$ is not complete since the Cauchy sequence $\{x_n\}$ does not converge in $\mathbb{Q}$ (only in $\mathbb{R}$). But clearly, $\mathbb{R}$ and $\mathbb{R}^2$ with the Euclidean metric

$$h(u, v) := \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2}$$

for $u = (u_1, u_2)$ and $v = (v_1, v_2) \in \mathbb{R}^2$ are complete metric spaces. $\square$

We have described above the set, $H(\mathbb{R}^2)$, of all compact subsets of $\mathbb{R}^2$, as the"space where fractal shapes live". We now explore the Hausdorff metric h that measures the distance between two sets $A, B \in H(\mathbb{R}^2)$ and show that $(H(\mathbb{R}^2), h)$ is a complete metric space. For $B \in H(\mathbb{R}^2)$ and $x \in \mathbb{R}^2$, define

$$d(x, B) := \min\{d(x, b),\, b \in B\} \text{ and } d(A, B) := \max\{d(a, B),\, a \in A\}.$$

The minimum value is attained since the set $B \in \mathcal{H}(\mathbb{R}^2)$ is compact and nonempty and since a metric is a continuous, real-valued function.

Theorem 2.3 *Let $A, B, \in H(\mathbb{R}^2)$. Define*

$$h(A, B) := \max\{d(A, B), d(B, A)\}.$$

Then $h(A, B)$ is a metric and $\mathcal{H} := (H(\mathbb{R}^2), h)$ is a complete metric space.

The metric $h(A, B)$ is called the *Hausdorff metric* and $\mathcal{H}$ is called the *Hausdorff space of shapes*. Before we prove Theorem 2.3, we will compute $h(A, B)$ for some $A, B \in \mathcal{H}$.

Example 2.4 In Figure 2.1, let A be the unit square with vertices at $p_0 = (0, 0), p_1 = (1, 0), p_2 = (0, 1)$, and $p_3 = (1, 1)$ and let $B_r := U(p_3, r)$ be a circle with radius, $r > 0$, and center $p_3 = (1, 1)$. Let $q_r = (1 - \frac{r\sqrt{2}}{2}, 1 - \frac{r\sqrt{2}}{2})$. Then one can calculate $d(A, B_r)$ and $d(B_r, A)$ as follows.

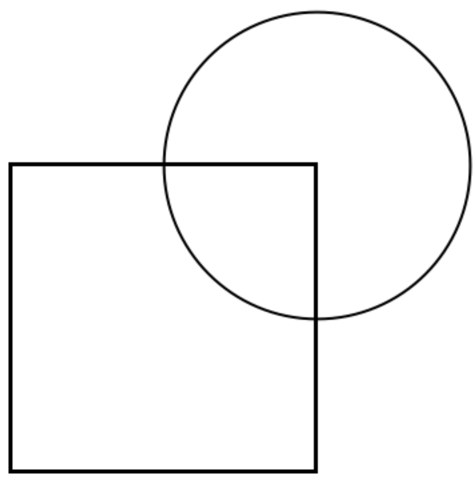

Figure 2.1: Hausdorff Distance Example

Since $d(A, B_r)$ is the largest distance from a point in the square A to the nearest point in the circle B_r, it follows that

$$d(A, B_r) = d(p_0, B_r) = d(p_0, q_r) = \sqrt{2} - r$$

if $0 < r \leq \sqrt{2}$ and $d(A, B_r) = 0$ if $r > \sqrt{2}$ (the circle covers the square). Similarly, since $d(B_r, A)$ is the largest distance from a point in the circle to the nearest point in the square,

one sees that

$$d(B_r, A) = d(u, A) = d(u, p_3) = r,$$

where $u = (u_1, u_2)$ is any point on the circle's boundary with $u_1 \geq 1$ and $u_2 \geq 1$. Therefore,

$$h(A, B_r) = \max\{\sqrt{2} - r, r\}.$$

□

To show that $h(A, B)$ is a complete metric, we need the following results.

Lemma 2.5 *Let $A, B \in \mathcal{H}$. Then there exist $\bar{a} \in A$ and $\bar{b} \in B$ such that $h(A, B) = d(\bar{a}, \bar{b})$.*

Proof: From the definition of a maximum, we obtain a sequence $\{a_n\}$ in A such that

$$d(A, B) = \max\{d(a, B),\, a \in A\} = \lim_{n\to\infty} d(a_n, B).$$

Similarly, there exists a point b_n in B such that $d(a_n, B) = d(a_n, b_n)$. By the compactness of A, there exists a subsequence $\{a_{n_k}\}$ of $\{a_n\}$ converging to $\bar{a}$, and by the compactness of B, there exists a subsequence $\{b_{n_k}\}$ of $\{b_n\}$ converging to $\bar{b}$. Therefore, by the continuity of the Euclidean metric, $d(a_{n_{k_j}}, b_{n_{k_j}})$ converges to $d(\bar{a}, \bar{b})$, which gives

$$d(A, B) = \lim_{j\to\infty} d(a_{n_{k_j}}, B) = \lim_{j\to\infty} d(a_{n_{k_j}}, b_{n_{k_j}}) = d(\bar{a}, \bar{b}).$$

Similarly, the same process can be used to show that $d(\bar{b}, \bar{a})$=d(B,A). Therefore,

$$h(A, B) = \max\{d(A, B), d(B, A)\} = d(\bar{a}, \bar{b})$$

from part (1) of Definition 2.1. □

Definition 2.6 *Let $A \in \mathcal{H}(\mathbb{R}^2)$ and $\epsilon > 0$. Then*

$$A + \epsilon := \{x \in \mathbb{R}^2 : d(x, a) \leq \epsilon \text{ for some } a \in A\} = \bigcup_{a \in A} U(a, \epsilon),$$

where $U(a, \epsilon) = \{x \in \mathbb{R}^2 : d(x, a) \leq \epsilon\}$ denotes the closed disk of radius ϵ around $a \in \mathbb{R}^2$.

Example 2.7 Below we see one way to implement $A + \epsilon$ in Mathematica using a square.

<u>Program:</u> $A + \epsilon$ (A Square)

```
dots = Flatten[Table[{i, j}, {i, -1, 1, .005}, {j, -1, 1, .005}], 1];
  *Makes a table of points from x = -1 to x = 1
Graphics[Point[dots]];
  *Plots the table of points to form a square
Graphics[{ Black, Map[Disk[#, 0.5] &, dots]}, Axes -> False,
 ImageSize -> 500]
  *Plots A + epsilon with epsilon = 0.5
```

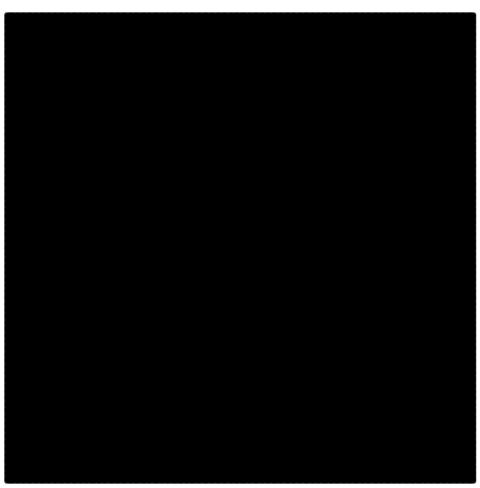

Figure 2.2: A, Square

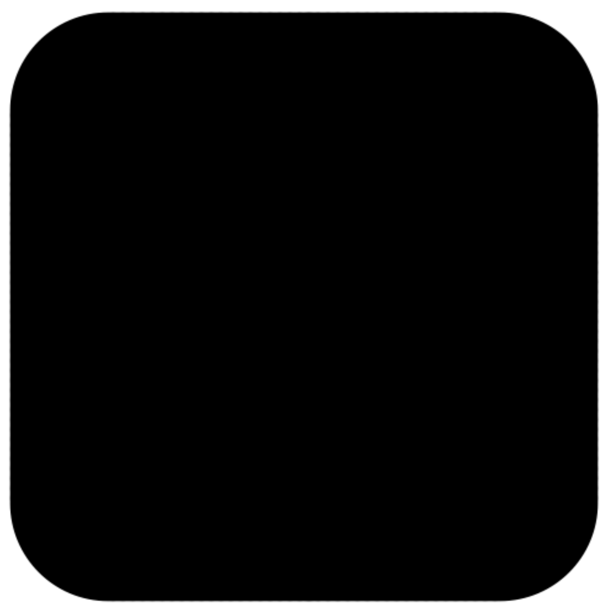

Figure 2.3: $A + \epsilon$ for Square

Program: $A + \epsilon$ (A Triangle)

```
dotstri = Select[dots, And[#[[2]] <= Sqrt[3] (#[[1]] + 1) - 1, #[[2]] <= -
    Sqrt[3] (#[[1]] - 1) - 1] &];
  *Makes an equilateral triangle by restricting which points from the square
     are plotted
Graphics[Point[dotstri]];
  *Plots the table of points to form an equilateral triangle
```

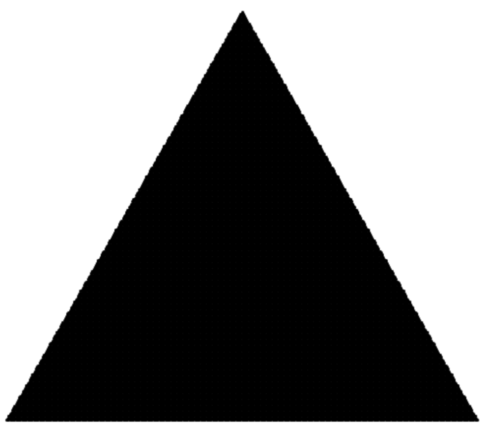

Figure 2.4: A, Triangle

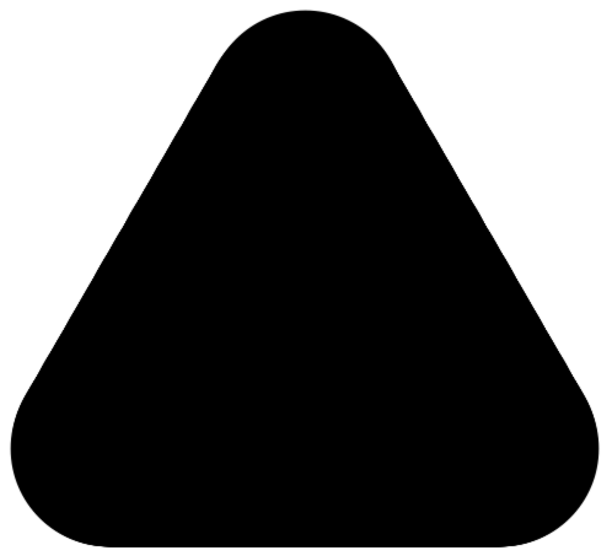

Figure 2.5: $A + \epsilon$ for Triangle

Program: $A + \epsilon$ (A Bowtie)

```
dotsv = Select[dots, -Abs[#[[1]]] <= #[[2]] <= Abs[#[[1]]] &];
  *Makes a bowtie by restricting which points from the square are plotted
Graphics[Point[dotsv]]
  *Plots the table of points to form a bowtie figure
```

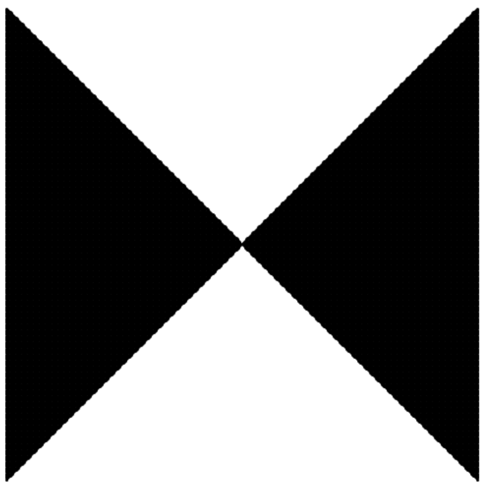

Figure 2.6: A, Bowtie

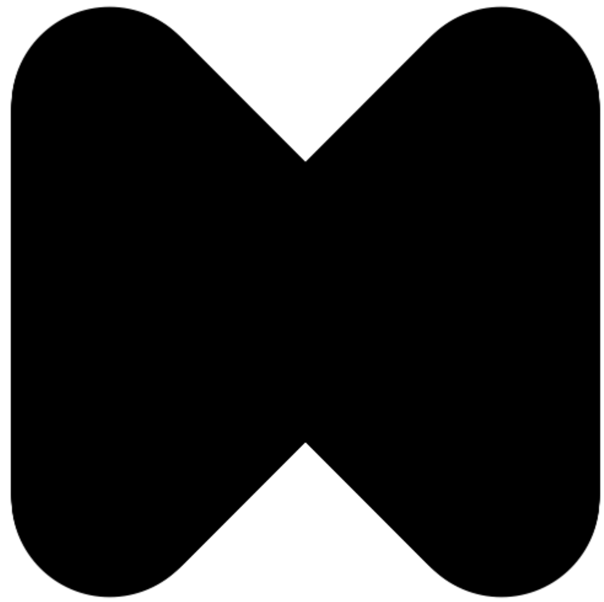

Figure 2.7: $A + \epsilon$ for Bowtie

These examples show that in general $A + \epsilon$ only smoothes corners of A with interior angles less then $180°$.

Lemma 2.8 *If $A \in \mathcal{H}(\mathbb{R}^2)$, then $A + \epsilon \in \mathcal{H}(\mathbb{R}^2)$.*

Proof: In order to show that $A + \epsilon \in \mathcal{H}(\mathbb{R}^2)$ we must show that $A + \epsilon$ is bounded and closed. First, we show that $A + \epsilon$ is bounded. If $A \in \mathcal{H}(\mathbb{R}^2)$, then $A \subset U(0, r)$ as defined in Definition 2.6 for some $r > 0$. Let $b \in A + \epsilon$. Then there exists $a \in A$ such that $b \in U(a, \epsilon)$. Therefore, $d(b, 0) \leq d(b, a) + d(a, 0) \leq r + \epsilon$. Thus $A + \epsilon \subseteq U(0, r + \epsilon)$. Now let $b_n \in A + \epsilon$ such that $b_n \to b$. For $A + \epsilon$ to be closed, we have to show that $b \in A + \epsilon$. For all n there exists $a_n \in A$ such that $b_n \in U(a_n, \epsilon)$. Since A is compact, there exists a subsequence $\{a_{n_i}\}$ that converges to some $a \in A$. Clearly $b_{n_i} \to b$. Thus

$$d(a, b) \leq d(a, a_{n_i}) + d(a_{n_i}, b_{n_i}) + d(b_{n_i}, b) \leq \epsilon$$

as $i \to \infty$. This shows that $b \in U(a, \epsilon) \subset A + \epsilon$. Therefore, $A + \epsilon \in \mathcal{H}(\mathbb{R}^2)$. $\square$

Lemma 2.9 *Let $A, B \in \mathcal{H}(\mathbb{R}^2)$ and $\epsilon > 0$. Then $h(A, B) \leq \epsilon$ if and only if $A \subseteq B + \epsilon$ and $B \subseteq A + \epsilon$. Moreover,*

$$h(A, B) = \min\{\epsilon \geq 0 : A \subseteq B + \epsilon \text{ and } B \subseteq A + \epsilon\}.$$

Proof: We show first that

$$d(A, B) = \max\{d(a, B), a \in A\} \leq \epsilon$$

if and only if $A \subseteq B + \epsilon$. Suppose $d(A, B) \leq \epsilon$. Then $d(a, B) \leq \epsilon$ for all $a \in A$. Since for all $a \in A$ there exists $b_0 \in B$ such that $d(a, B) = d(a, b_0)$, it follows that $a \in U(b_0, \epsilon)$. Thus, $A \subseteq B + \epsilon$. Now suppose

$$A \subseteq B + \epsilon = \bigcup_{b \in B} U(b, \epsilon).$$

Thus, for all $a \in A$ there exists $b \in B$ such that $d(a, b) \leq \epsilon$. Then $d(a, B) = \min\{d(a, b), b \in B\} \leq \epsilon$ for all $a \in A$. Therefore,

$$d(A, B) = \max\{d(a, B), a \in A\} \leq \epsilon.$$

This shows that $d(A, B) \leq \epsilon$ if and only if $A \subseteq B + \epsilon$. By interchanging A with B, we obtain that $d(B, A) \leq \epsilon$ if and only if $B \subseteq A + \epsilon$. Thus,

$$h(A, B) = \max\{d(A, B), d(B, A)\} \leq \epsilon$$

if and only if $A \subseteq B + \epsilon$ and $B \subseteq A + \epsilon$. Since $h(A, B) = d(\bar{b}, \bar{a})$ for some $\bar{a} \in A$, $\bar{b} \in B$, it follows from Lemma 2.5 that

$$h(A, B) = \min\{\epsilon \geq 0 : A \subseteq B + \epsilon \text{ and } B \subseteq A + \epsilon\}.$$

$\square$

The following "Extension Lemma" is crucial for the proof of the main result in this section, Theorem 2.3.

Lemma 2.10 *Let d be a a metric in the space $\mathbb{R}^2$. Let $\{A_n\}$ be a Cauchy sequence of sets in $(\mathcal{H}(\mathbb{R}^2), h)$. Let $\{n_j\}$ for $j \geq 1$ be an infinite sequence of integers*

$$0 < n_1 < n_2 < n_3 < \cdots .$$

Suppose that we have a Cauchy sequence $\{x_{n_j}\}$ with $x_{n_j} \in A_{n_j}$ in $(\mathbb{R}^2, d)$. Then there is a Cauchy sequence $\{\tilde{x}_n\} \in (\mathbb{R}^2, d)$ with $\tilde{x}_n \in A_n$ for $n \geq 1$ such that $\tilde{x}_{n_j} = x_{n_j}$ for all $j \geq 1$.

Proof: Let $n \in \{1, 2, \ldots, n_1\}$ and let $\tilde{x}_n$ be the closest point in A_n to x_{n_1} (such a point exists since A_n is compact). Clearly $\tilde{x}_{n_1} = x_{n_1}$. Similarly, let $j \geq 1$ and $n \in \{n_j + 1, n_j + 2, \ldots, n_{j+1}\}$ and let $\tilde{x}_n \in A_n$ be the closest point in A_n to $x_{n_{j+1}}$. Then $\tilde{x}_{n_{j+1}} = x_{n_{j+1}}$. Let $\epsilon > 0$. Since $\{x_{n_j}\}$ is a Cauchy sequence in $(\mathbb{R}^2, d)$, there exists N_1 such that $d(x_{n_k}, x_{n_j}) \leq \frac{\epsilon}{3}$ for all $n_k, n_j \geq N_1$. Since $\{A_n\}$ is a Cauchy sequence in $\mathcal{H}(\mathbb{R}^2)$, there exists N_2 such that $h(A_m, A_n) \leq \epsilon$ (and, therefore, $d(A_m, A_n) \leq \epsilon$ for all $m, n \geq N_2$. Let $N > \max\{N_1, N_2\}$ and observe that for $m, n > N$

$$\begin{aligned} d(\tilde{x}_m, \tilde{x}_n) &\leq d(\tilde{x}_m, x_{n_j}) + d(x_{n_j}, \tilde{x}_n) \\ &\leq d(\tilde{x}_m, x_{n_j}) + d(x_{n_j}, x_{n_k}) + d(x_{n_k}, \tilde{x}_n) \end{aligned}$$

where $m \in \{n_{j-1}+1, n_{j-1}+2, \ldots, n_j\}$ and $n \in \{n_{k-1}+1, n_{k-1}+2, \ldots, n_k\}$ and $n_j, n_k > N$. Then $d(\tilde{x}_m, x_{n_j})$ is the smallest distance from A_m to x_{n_j}. So $d(\tilde{x}_m, x_{n_j}) = d(x_{n_j}, A_m)$. Thus,

$$d(x_{n_j}, A_m) \leq d(A_{n_j}, A_m) \leq \frac{\epsilon}{3}.$$

Similarly, $d(x_{n_k}, \tilde{x}_n) \leq \frac{\epsilon}{3}$. Since $d(x_{n_j}, x_{n_k}) \leq \frac{\epsilon}{3}$ it follows that $d(\tilde{x}_m, \tilde{x}_n) \leq \epsilon$ for $m, n > N$.
$\square$

Proof of Theorem 2.3: It is easily seen that $h(A, A) = 0$ for all $A \in \mathcal{H}$ (item (III)). To show item (II), let $A \neq B$ for $A, B \in H(\mathbb{R}^2)$. Without loss of generality, we may assume there exists $a \in A$ such that $a \notin B$. Then, $h(A, B) \geq d(A, B) \geq d(a, B) > 0$. Therefore $h(A, B) = 0$ if and only if $A = B$.

The commutativity of h (item (I)) follows because

$$\max\{d(A, B), d(B, A)\} = \max\{d(B, A), d(A, B)\} = h(B, A).$$

Finally to prove the triangle inequality (IV), we first prove the inequality for $d(A, B)$; that is, we show that

$$d(A, B) \leq d(A, C) + d(B, C)$$

for $A, B, C \in H(\mathbb{R}^2)$. Suppose $a \in A$. Then

$$\begin{aligned} d(a, B) &= \min_{b \in B} d(a, b) \leq \min_{b \in B}(d(a, c) + d(c, b)) \\ &= d(a, c) + \min_{b \in B} d(c, b) = d(a, c) + d(c, B) \end{aligned}$$

for all $c \in C$. Therefore,

$$\begin{aligned} d(a, B) &\leq \min_{c \in C} d(a, c) + d(c, B) = d(a, C) + d(c, B) \\ &\leq d(a, C) + \max_{c \in C} d(c, B) = d(a, C) + d(C, B) \end{aligned}$$

for all $a \in A$. This shows that,

$$d(A, B) \leq d(A, C) + d(C, B)$$

for all $A, B, C \in H(\mathbb{R}^2)$. Now,

$$h(A, B) = \max\{d(A, B), d(B, A)\}$$

$$\leq \max\{d(A,C)+d(C,B), d(B,C)+d(C,A)\}$$
$$\leq \max\{d(A,C), d(C,A)\}+\max\{d(C,B), d(B,C)\}$$
$$= h(A,C)+h(C,B)$$

for all $A, B, C \in H(\mathbb{R}^2)$. This shows that h is indeed a metric on $H(\mathbb{R}^2)$.

To show that $\mathcal{H} := (H(\mathbb{R}^2), h)$ is a complete metric space we let $\{A_n\} \in H(\mathbb{R}^2)$ be a Cauchy sequence and define

$$A := \{a \in \mathbb{R}^2 : \text{there exists a sequence } a_n \in A_n \text{ such that } a_n \to a\}.$$

We will show that $\mathcal{H}(\mathbb{R}^2)$ is complete; that is, $A \in \mathcal{H}(\mathbb{R}^2)$ and $A = \lim_{n\to\infty} A_n$.

We will divide this proof into the following steps:

(I) $A \neq \emptyset$;

(II) A is closed;

(III) For all $\epsilon > 0$ there is N such that $A \subset A_n + \epsilon$ for all $n \geq N$;

(IV) A is bounded;

(V) $\lim_{n\to\infty} A_n = A$.

(I) Let $\{A_n\}$ be a Cauchy sequence of sets in $\mathcal{H}(\mathbb{R}^2)$. Then for all $\epsilon \geq 0$ there exists $N \in \mathbb{N}$ such that

$$h(A_m, A_n) \leq \epsilon$$

for all $m, n \geq N$. By Lemma 2.9, there exists $\bar{a}_i \in A_i$ such that

$$h(A_m, A_n) = d(\bar{a}_n, \bar{a}_m).$$

Then $\{\bar{a}_n\}$ is a Cauchy sequence in $\mathbb{R}^2$. Therefore $a := \lim_{i\to\infty} \bar{a}_n$ exists and it follows that $A := \{a \in \mathbb{R}^2 : \text{there exists a sequence } a_n \in A_n \text{ such that } a_n \to a\}$ is non-empty.

(II) Let $a_i \in A$ and $a := \lim_{i\to\infty} a_i$. The set A is closed if $a \in A$. Now, for all i, there exist $a_{i,n} \in A_n$ such that $a_i = \lim_{n\to\infty} a_{i,n}$. It follows that there exists an increasing sequence $N_i \in \mathbb{N}$ such that

$$d(a_{N_i}, a) < \frac{1}{i}.$$

Also, there exists a sequence of integers m_i such that

$$d(a_{N_i}, a_{N_i,m_i}) < \frac{1}{i},$$

where $a_{N_i,m_i} \in A_{m_i}$. Thus $a_{N_i,m_i} \to a$ as $i \to \infty$ since.

$$d(a_{N_i,m_i}, a) \leq d(a_{N_i,m_i}, a_{N_i}) + d(a_{N_i}, a) \leq \frac{2}{i}.$$

By the Extension Lemma 2.10, there exists a sequence $\tilde{a}_i \in A_i$ such that $\tilde{a}_i \to a$. Thus, $a \in A$, and therefore A is closed.

(III) Let $\epsilon > 0$. There exists an N such that for $m, n \geq N$, $h(A_m, A_n) \leq \epsilon$. Let $n \geq N$. Then for $m \geq n$, $A_m \subset A_n + \epsilon$. We must show that $A \subset A_n + \epsilon$. To begin, let $a \in A$ and let $\{a_i\} \in A_i$ be a sequence that converges to a. Assume that N is large enough such that $m \geq N$, $d(a_m, a) < \epsilon$. Then $a_m \in A_n + \epsilon$ since $A_m \subset A_n + \epsilon$. Since A_n is compact, one can show that $A_n + \epsilon$ is compact from Lemma 2.8. So since $a_m \in A_n + \epsilon$ for all $m \geq N$, $a \in A_n + \epsilon$. Therefore, $A \subset A_n + \epsilon$ for n sufficiently large.

(IV) By Lemma 2.8 we know that $A_n + \epsilon$ is compact and, therefore, also bounded. From part (III) we know that $A \subset A_n + \epsilon$. Thus, A is bounded.

(V) By (IV), one knows that $A \in \mathcal{H}(\mathbb{R}^2)$. So, by (III) and Lemma 2.9, the proof that $\lim_{n\to\infty} A_n = A$ will be complete if we show that for all $\epsilon > 0$ there exists an N such that for $n \geq N$, $A_n \subset A + \epsilon$. To show that this is true, let $\epsilon > 0$ and choose N such

that $h(A_m, A_n) \leq \frac{\epsilon}{2}$ and $A_m \subset A_n + \frac{\epsilon}{2}$ for $m, n \geq N$. Let $n \geq N$. We show next that $A_n \subset A + \epsilon$. There exists an increasing sequence N_i of integers such that $n = N_o < N_1 < N_2 < N_3 < \cdots < N_k < \cdots$ and for $m, n \geq N_j$, $A_m \subset A_n + \frac{\epsilon}{2^{j+1}}$. Let $y \in A_n$. There is an $x_{N_1} \in A_{N_1}$ such that $d(y, x_{N_1}) \leq \frac{\epsilon}{2}$ and there is a point $x_{N_2} \in A_{N_2}$ such that $d(x_{N_1}, x_{N_2}) \leq \frac{\epsilon}{2^2}$. Using induction, one can find a sequence $x_{N_1}, x_{N_2}, x_{N_3}, \ldots,$ such that $x_{N_j} \in A_{N_j}$ and $d(x_{N_j}, x_{N_{j+1}}) \leq \frac{\epsilon}{2^{j+1}}$. Multiple uses of the triangle inequality show that

$$d(y, x_{N_j}) \leq \frac{\epsilon}{2} \text{ for all } j$$

and that $\{x_{N_j}\}$ is a Cauchy sequence. By the way n was chosen, $A_{N_j} \subset A_{N_0} + \frac{\epsilon}{2}$. Thus x_{N_j} converges to a point $x \in A$ by definition and since $A_n + \frac{\epsilon}{2}$ is closed, $x \in A_n + \frac{\epsilon}{2}$ as well. So, $d(y, x_{N_j}) \leq \epsilon$ implies that $d(y, x) \leq \epsilon$. Thus $A_n \subset A + \epsilon$ for $n \geq N$ and $\lim A_n = A$. Therefore $(\mathcal{H}(\mathbb{R}^2), h)$ is a complete metric space. $\square$

Chapter 3
The Contraction Mapping Principle

As we have seen in the introduction, certain fractal shapes like the Sierpinski triangle S_∞, are the result of an iterative construction. So,

$$T(S_0), T^2(S_0), T^3(S_0), \cdots \to S_\infty \text{ as } n \to \infty,$$

where T is a map from $\mathcal{H} := (H(\mathbb{R}^2), h)$ into itself with initial value $S_0 \in H(\mathbb{R}^2)$. As we will see below, a key assumption on the map T is that it is a contraction.

Definition 3.1 *A map $T : \mathcal{X} \to \mathcal{X}$ on a metric space $\mathcal{X} := (\boldsymbol{X}, h)$ is called a contraction if there is a constant $0 < s < 1$ such that $h(T(X), T(Y)) \leq s\, h(X, Y)$ for all $X, Y \in \mathcal{X}$. It is called non-expansive if $s = 1$.*

It is important to notice that non-expansive maps do not enlarge distances between two points, although the orbit of a point can go towards infinity. For example, if $T : \mathbb{R} \to \mathbb{R}$ is given by $T(x) = x + 1$, then T is distance-preserving (that is, $d(T(x), T(y)) = d(x, y)$), but $T^n(x) = x + n \to \infty$ as $n \to \infty$. As the following results shows, this cannot happen if T is a contraction.

Theorem 3.2 (Contraction Mapping Principle) *Let $T : \mathcal{X} \to \mathcal{X}$ be a contraction on a complete metric space $\mathcal{X} := (\boldsymbol{X}, h)$. Then T is continuous and possesses exactly one fixed point $F \in \mathcal{X}$. Moreover,*

$$\lim_{n\to\infty} T^n(S_0) = S_\infty$$

for any initial value $S_0 \in \mathcal{X}$, where $T^n(S_0) = T \circ T \circ \cdots \circ T(S_0)$.

Proof: Let $S_0 \in \mathcal{X}$. Let $0 < s < 1$ be a contractivity factor for T. Then

$$h(T^k(S_0), T^{n+k}(S_0)) \leq sh(T^{k-1}(S_0), T^{n+k-1}(S_0)) \leq \cdots \leq s^k h(S_0, T^n(S_0)) \tag{3.1}$$

for all $k, n \in \mathbb{N}_0$. Moreover,

$$\begin{aligned} h(S_0, T^n(S_0)) &\leq h(S_0, T(S_0)) + h(T(S_0), T^2(S_0)) + \cdots + h(T^{n-1}(S_0), T^n(S_0)) \\ &\leq (1 + s + s^2 + \cdots + s^{n-1}) h(S_0, T(S_0)) \\ &\leq \frac{1}{(1-s)} h(S_0, T(S_0)). \end{aligned}$$

Substituting this estimate into (3.1), one obtains

$$h(T^k(S_0), T^{n+k}(S_0)) \leq s^k \frac{1}{(1-s)} h(S_0, T(S_0)).$$

It follows that the sequence $T^n(S_0)$ $(n \in \mathbb{N}_0)$ is a Cauchy sequence. Since $\mathcal{X}$ is complete, this Cauchy sequence possesses a limit $S_\infty \in \mathcal{X}$; that is,

$$\lim_{n \to \infty} T^n(S_0) = S_\infty.$$

Now we show that S_∞ is a fixed point of T. Since T is contractive with contractivity factor $s > 0$, it is continuous because

$$d(T(x), T(y)) \leq s d(x, y) < \epsilon$$

whenever $d(x, y) < \delta$, where $\delta = \frac{\epsilon}{s}$. Thus,

$$T(S_\infty) = T(\lim_{n \to \infty} T^n(S_0)) = \lim_{n \to \infty} T^{n+1}(S_0) = S_\infty.$$

Finally, we want to show that the fixed point S_∞ is unique. So assume there are two fixed points, F and G, of T. Then $F = T(F)$ and $G = T(G)$. We have

$$h(F, G) = h(T(F), T(G)) \leq s h(F, G).$$

Since $0 < s < 1$ it follows that $h(F, G) = 0$ and therefore $F = G$. $\square$

Definition 3.3 *Let X be a complete metric space. An iterated function system X with generator T is a family of iterates $Id, T, T^2, T^3, \ldots$ generated by the function $T : X \to X$.*

In what follows, we take X to be the complete metric space $(\mathcal{H}, d)$ and maps $T : \mathcal{H} \to \mathcal{H}$. Often $T : \mathcal{H} \to \mathcal{H}$ is defined by

$$T(S) := T_1(S) \cup T_2(S) \cup \cdots \cup T_N(S),$$

where $T_i : \mathbb{R}^2 \to \mathbb{R}^2$ $(1 \leq i \leq N)$ are affine functions of the form

$$T_i(x, y) := (ax + by, cx + dy) + (e, f),$$

where a, b, c, d, e, and f are real numbers.

Lemma 3.4 *Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a continuous mapping on the metric space $(\mathbb{R}^2, d)$. Then T maps $\mathcal{H}(\mathbb{R}^2)$ into $\mathcal{H}(\mathbb{R}^2)$.*

Proof: Let S be a nonempty compact subset of $\mathbb{R}^2$. Then $T(S) = \{T(x) : x \in S\}$ is nonempty. We want to show that $T(S)$ is compact. Let $\{y_n\} = \{T(x_n)\}$ be an infinite sequence of points in S. Then $\{x_n\}$ is an infinite sequence of points S. Since S is compact, we know that there is a subsequence $\{x_{N_n}\}$ that converges to a point $\hat{x} \in S$. But then the continuity of T implies that $\{y_{N_n}\} = \{T(x_{N_n})\}$ is a subsequence of $\{y_n\}$ that converges to $\hat{y} = T(\hat{x}) \in T(S)$. $\square$

Lemma 3.5 *Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a contraction on the metric space $(\mathbb{R}^2, d)$ with contractivity factor s. Then $T : \mathcal{H}(\mathbb{R}^2) \to \mathcal{H}(\mathbb{R}^2)$ defined by*

$$T(X) := \{T(x) : x \in X\} \text{ for all } X \in \mathcal{H}(\mathbb{R}^2)$$

is a contraction mapping on $\mathcal{H}(\mathbb{R}^2, h)$ with contractivity factor s.

Proof: Since we know that if T is contractive then T is continuous, it follows that $T : \mathbb{R}^2 \to \mathbb{R}^2$ is continuous. So, by Lemma 3.4, T maps $\mathcal{H}(\mathbb{R}^2)$ into itself. Now let $X, Y \in \mathcal{H}(\mathbb{R}^2)$. Then,

$$\begin{aligned} d(T(X), T(Y)) &= \max\{\min\{d(T(x), T(y)) : y \in Y\} : x \in X\} \\ &\leq \max\{\min\{sd(x, y) : y \in Y\} : x \in X\} = sd(X, Y). \end{aligned}$$

Similarly,

$$d(T(Y), T(X)) \leq sd(Y, X).$$

Thus,

$$\begin{aligned} h(T(X), T(Y)) &= \max\{d(T(X), T(Y)), d(T(Y), T(X))\} \\ &\leq s \max\{d(X, Y), d(Y, X) \leq sd(X, Y). \end{aligned}$$

$\square$

Lemma 3.6 *Let* $(\mathbb{R}^2, d)$ *be a metric space and* $\mathcal{H}(\mathbb{R}^2)$ *denote the nonempty compact subsets of* $\mathbb{R}^2$. *Then*

$$h(A \cup B, C \cup D) \leq \max\{h(A, C), h(B, D)\}$$

for all $A, B, C, D \in \mathcal{H}(\mathbb{R}^2)$.

Proof: First we verify that

$$d(A, B \cup C) = \max\{d(B), d(C)\}$$

This is because

$d(A, B \cup C) = \max\{\min\{d(a, x) : a \in A\} : a \in B \cup C\}$

$$= \max\{\max\{\min\{d(a,b) : a \in A\} : b \in B\}, \max\{\min\{d(a,c) : a \in A\} : c \in C\}\}$$
$$= \max\{d(A,B), d(A,C)\}.$$

It follows that

$$d(A \cup B, C \cup D) = \max\{d(A \cup B, C), d(A \cup B, D)\}$$

Next we verify that

$$d(A \cup B, C) = \min\{d(A,C), d(B,C)\}$$

This is shown by

$$\begin{aligned} d(A \cup B, C) &= \max\{\min\{d(c,x) : x \in A \cup B\} : c \in C\} \\ &= \max\{\min\{\min\{d(c,a) : a \in A\}, \min\{d(c,b) : B \in B\}\}\} \\ &\leq \min\{\max\{\min\{d(c,a) : a \in A\} : c \in C\}, \max\{\min\{d(c,b) : b \in B\} : c \in C\}\} \\ &= \min\{d(A,C), d(B,C)\}. \end{aligned}$$

It follows that

$$d(A \cup B, C) \leq d(A,C) \text{ and } d(A \cup B, D) \leq d(A,D).$$

Therefore, we see that

$$d(A \cup B, C \cup D) \leq \max\{d(A,C), d(B,D)\}$$

$$\text{and } d(C \cup D, A \cup B) \leq \max\{d(C,A), d(D,B)\}$$

Thus,

$$\begin{aligned} h(A \cup B, C \cup D) &= \max\{d(A \cup B, C \cup D), d(C \cup D, A \cup B)\} \\ &\leq \max\{\max\{d(A,C), d(B,D)\}, \max\{d(C,A), d(D,B)\}\} \end{aligned}$$

$$\le \max\{d(A,C), d(B,D), d(C,A), d(D,B)\}$$
$$= \max\{\max\{d(A,C), d(C,A)\}, \max\{d(B,D), d(D,B)\}\}$$
$$= \max\{h(A,C), h(B,D)\}.$$

□

Lemma 3.7 *Let $(\mathbb{R}^2, d)$ be a complete metric space. Let $\{T_n : n = 1, 2, \ldots, N\}$ be contraction mappings on $(\mathcal{H}(\mathbb{R})^2, h)$. Let the contractivity factor for T_n by denoted by s_n for each n. Define $T : \mathcal{H} \to \mathcal{H}$ by*

$$\begin{aligned} T(X) &:= T_1(X) \cup T_2(X) \cup \cdots \cup T_N(X) \\ &= \bigcup_{n=1}^{n} T_n(X), \text{ for each } X \in \mathcal{H}(\mathbb{R}^2). \end{aligned}$$

Then T is a contraction mapping with contractivity factor $s = \max\{s_n : n = 1, 2, \ldots, N\}$.

Proof: Here we show the claim for $N = 2$, but inductively we see that the claim holds for all N. Let $X, Y \in \mathcal{H}$. Then we have

$$\begin{aligned} h(T(X), T(Y)) &= h(T_1(X) \cup T_2(X), T_1(Y) \cup T_2(X)) \\ &\le \max\{h(T_1(X), T_1(Y)), h(T_2(X), T_2(Y))\} \text{ from Lemma 3.6} \\ &\le \max\{s_1 h(X,Y), s_2 h(X,Y) \\ &\le s h(X,Y). \end{aligned}$$

□

Example 3.8 (The Sierpinski Triangle with Triangle Initiator)

Let $T : \mathcal{H} \to \mathcal{H}$ be defined as in (1.1). Then from the lemmas above we see that the functions T_i are contractions with contractivity factor $s_i = \frac{1}{2}$ $(1 \le i \le 3)$. Thus, by the

previous proposition, T is a contraction with contactivity factor $\frac{1}{2}$. By the Contraction Mapping Principle, for any initial value $S_0 \in \mathcal{H}$, the iterative function system

$$S_0, T(S_0), T^2(S_0), \cdots$$

converges to a unique fixed point $S_\infty \in \mathcal{H}$, called the Sierpinski Triangle.

Program: Sierpinski Triangle (Triangle Initiator)

```
tri = Polygon[{{0, 0}, {1, 0}, {1/2, N[Sqrt[3]/2]}}];
  *Generates equilateral triangle
mydilate[gdata_, q_, s_] := Map[q + s (# - q) &, gdata, {-2}]
  *Graphics data with floating point numbers
  *q is a pair of numbers
  *s is contractivity factor
mapmydilate[gr_] := Map[mydilate[gr, #, 1/2] &, tri[[1]]]
  *Establishes contractivity factor of 1/2
  *Establishes the shape of our limit shape
Manipulate[Graphics[Nest[mapmydilate, tri, s]], {s, 0, 7, 1}]
  *Establishes the initiator as an equilateral Triangle
  *Generates 7 iterations of the Sierpinski triangle
```

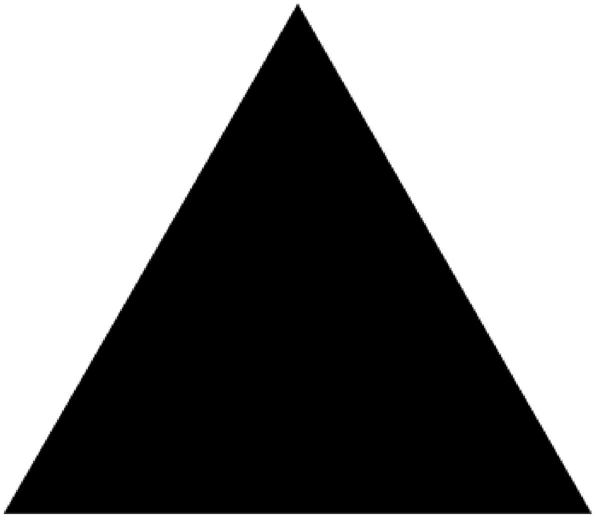

Figure 3.1: S_0 Original Triangle

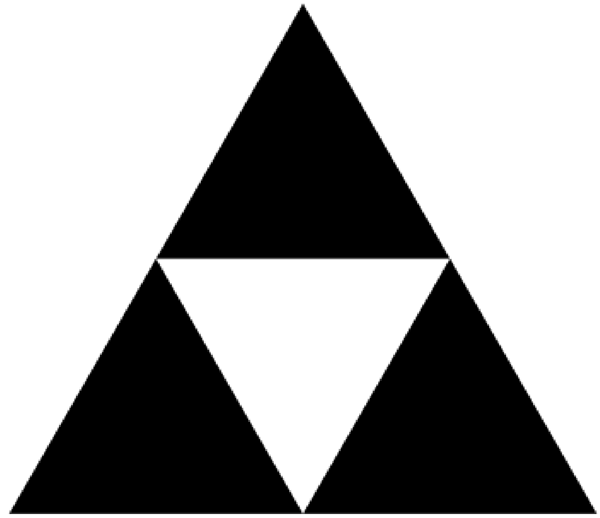

Figure 3.2: $T^1(S_0)$ after First Iteration

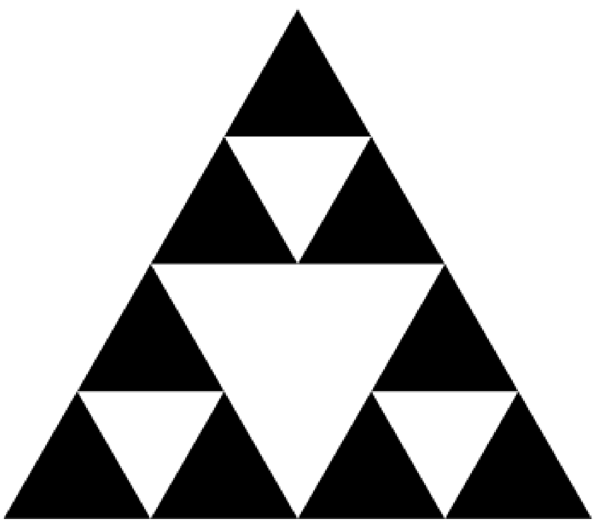

Figure 3.3: $T^2(S_0)$ after Second Iteration

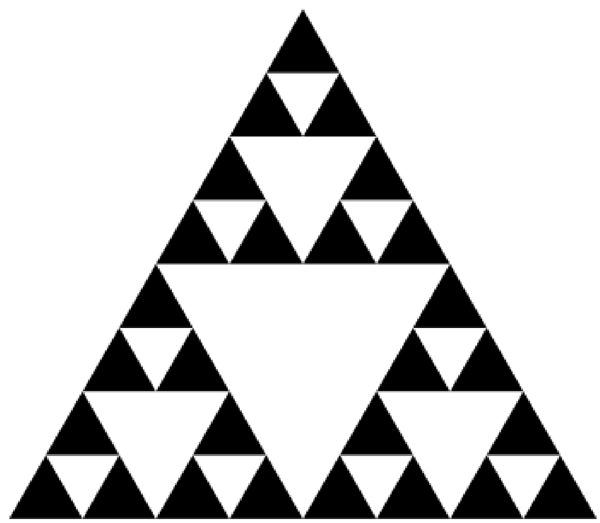

Figure 3.4: $T^3(S_0)$ after Third Iteration

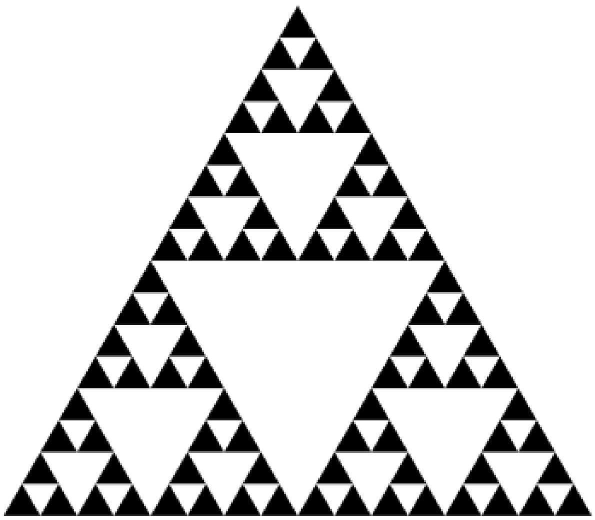

Figure 3.5: $T^4(S_0)$ after Fourth Iteration

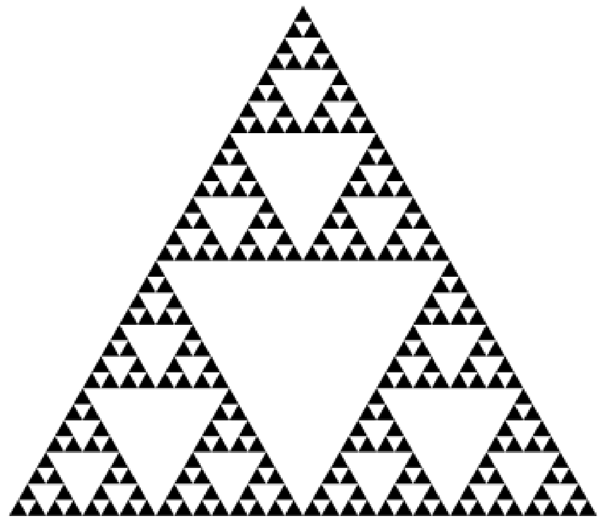

Figure 3.6: $T^5(S_0)$ after Fifth Iteration

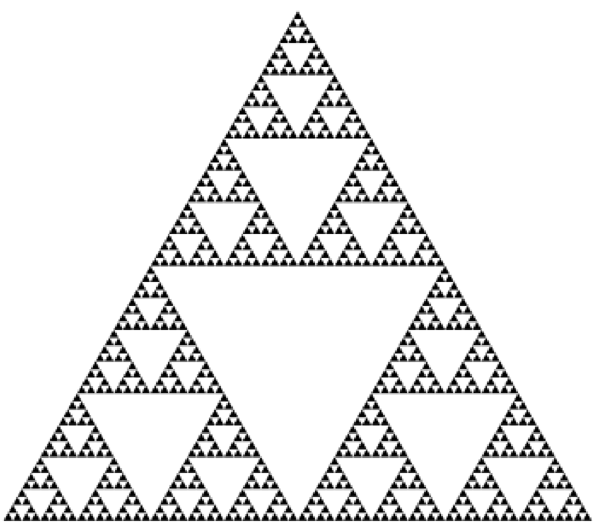

Figure 3.7: $T^6(S_0)$ after Sixth Iteration

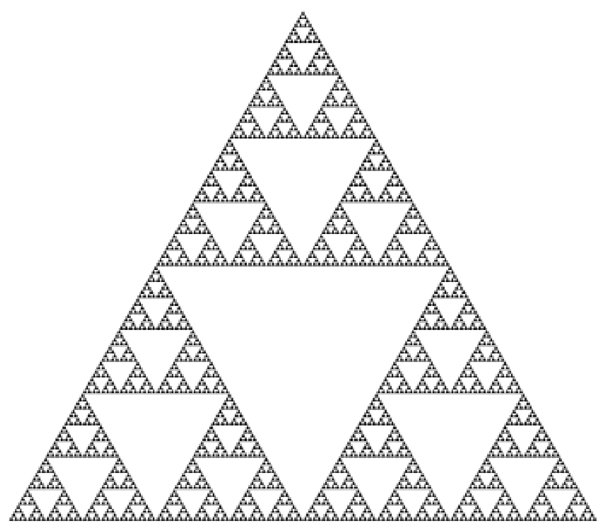

Figure 3.8: $T^7(S_0)$ after Seventh Iteration

□

With this approach, a fractal shape S_∞ is the unique fixed point of a contractive iterated function system and S_∞ can be obtained as the limit of recursively defined shapes where (at each level of magnification) after so many iterations of the contraction, there will not be any visible change in the graphical representation of S_∞. In other words, S_∞ is the limit of a sequence of shapes generates by an (IFS).

We will now look at two additional examples of the Sierpinski triangle, but the initiator will not be an equilateral triangle. This is to show that, given our set of iterated function systems, it does not matter what initiator we start with. Our limit shape will always be the Sierpinski Triangle.

Example 3.9 (The Sierpinski Triangle with Square Initiator)
Here we take our initiator to be the unit square

Program: Sierpinski Triangle (Square Initiator)

```
usq = Polygon[{{0, 0}, {0, 1}, {1, 1}, {1, 0}}];
  *Generates unit square
Manipulate[Graphics[Nest[mapmydilate, usq, s]], {s, 0, 7, 1}]
  *Establishes the initiator as the unit square
  *Generates 7 iterations of the Sierpinski triangle
```

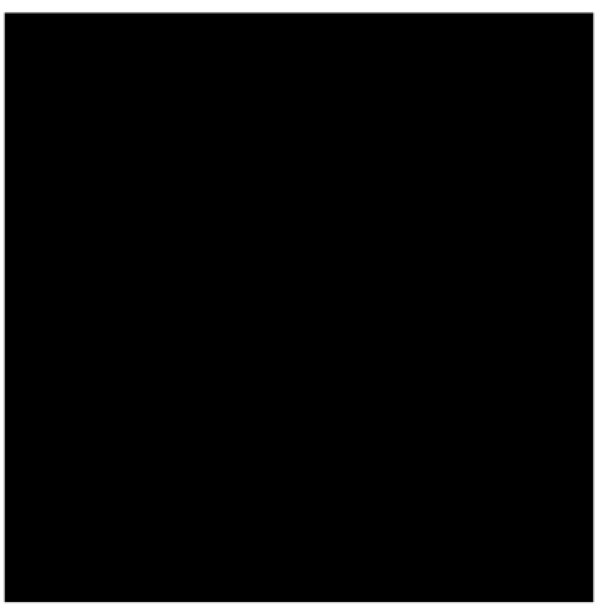

Figure 3.9: S_0 Original Square

Figure 3.10: $T^1(S_0)$ after First Iteration

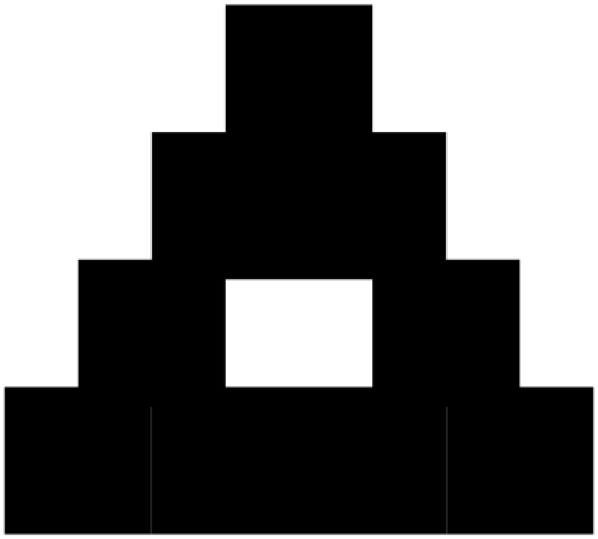

Figure 3.11: $T^2(S_0)$ after Second Iteration

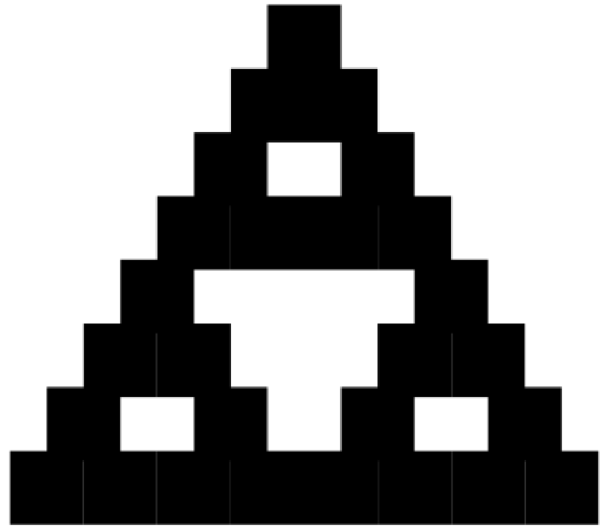

Figure 3.12: $T^3(S_0)$ after Third Iteration

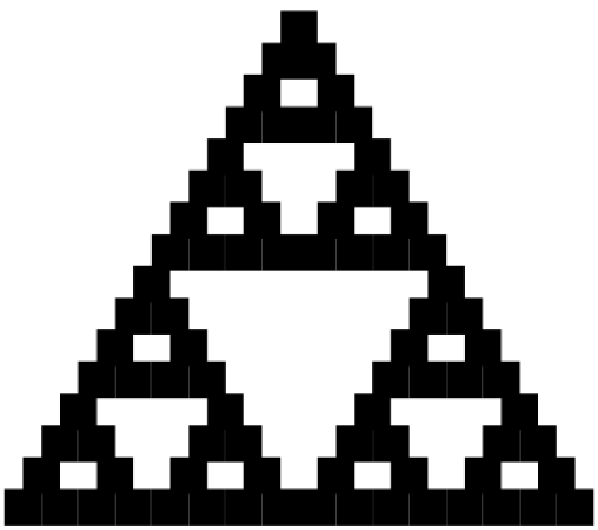

Figure 3.13: $T^4(S_0)$ after Fourth Iteration

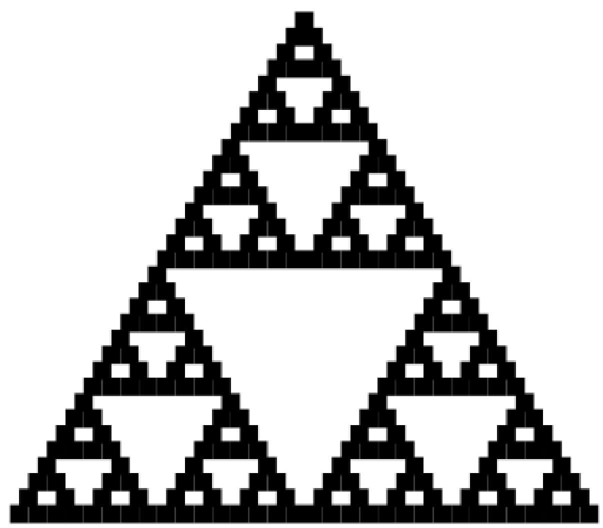

Figure 3.14: $T^5(S_0)$after Fifth Iteration

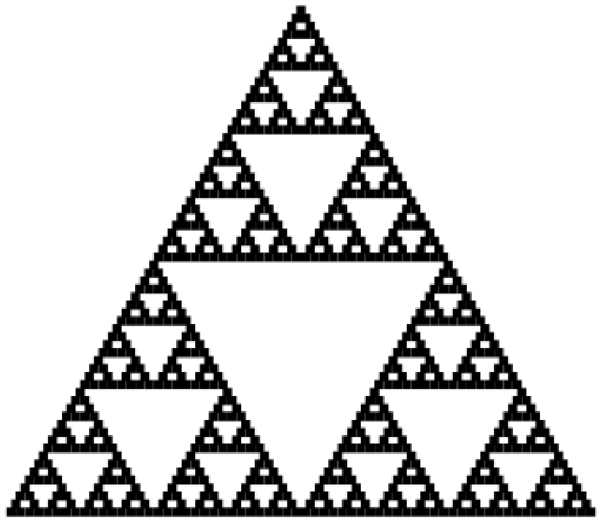

Figure 3.15: $T^6(S_0)$ after Sixth Iteration

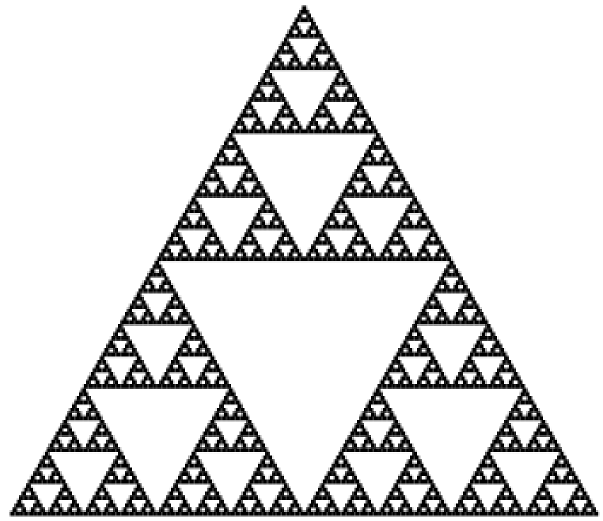

Figure 3.16: $T^7(S_0)$after Seventh Iteration

□

Example 3.10 (The Sierpinski Triangle with Circle Initiator)

We now take our initiator to be the unit circle.

Program: Sierpinski Triangle (Circle Initiator)

```
cir = Line[Table[N@{Cos[t], Sin[t]}, {t, 0, 2 Pi, Pi/30}]];
  *Generates unit circle
Manipulate[Graphics[Nest[mapmydilate, cir, s]], {s, 0, 7, 1}]
  *Establishes the initiator as the unit circle
  *Generates 7 iterations of the Sierpinski triangle
```

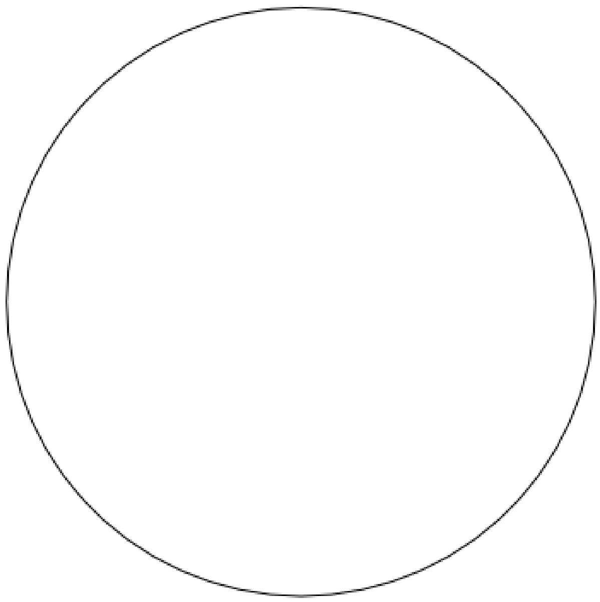

Figure 3.17: S_0, Original Circle

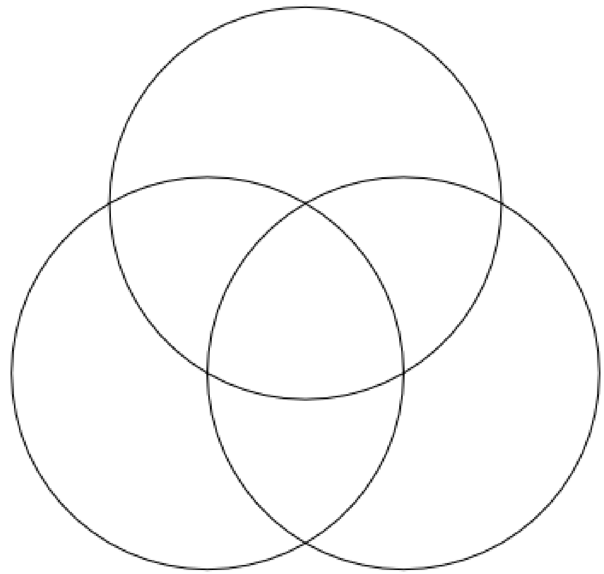

Figure 3.18: $T^1(S_0)$ after First Iteration

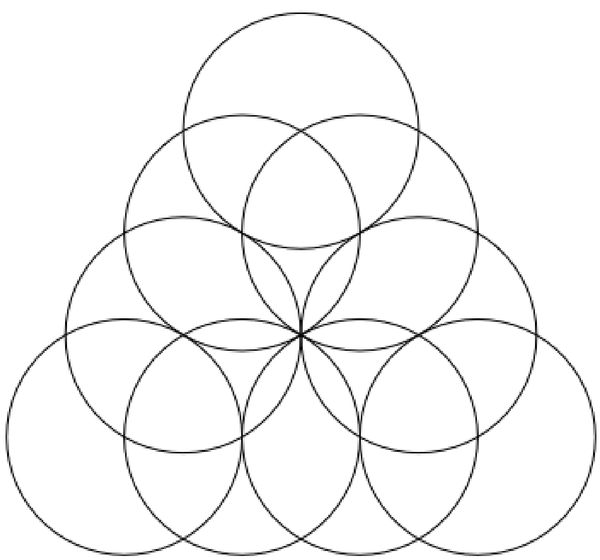

Figure 3.19: $T^2(S_0)$ after Second Iteration

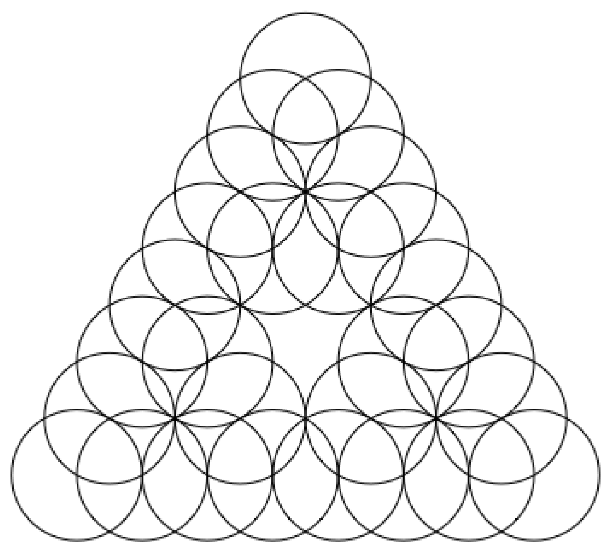

Figure 3.20: $T^3(S_0)$ after Third Iteration

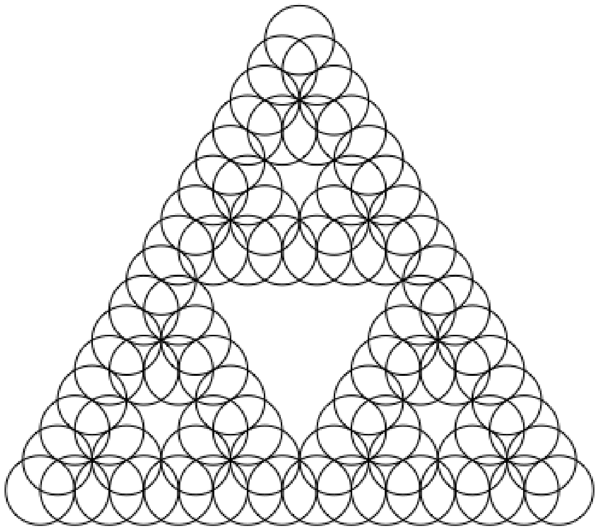

Figure 3.21: $T^4(S_0)$ after Fourth Iteration

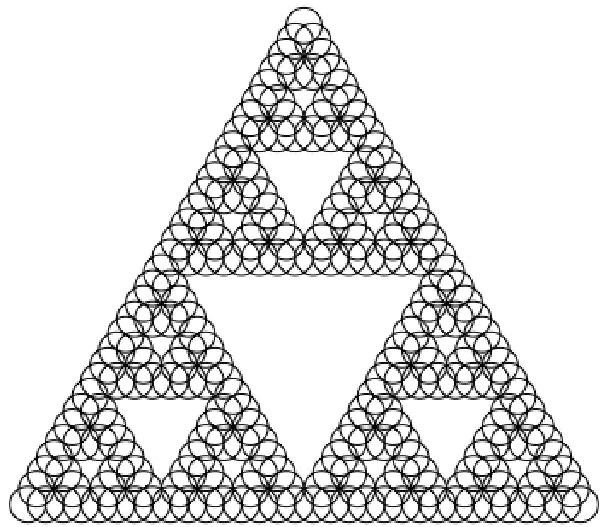

Figure 3.22: $T^5(S_0)$ after Fifth Iteration

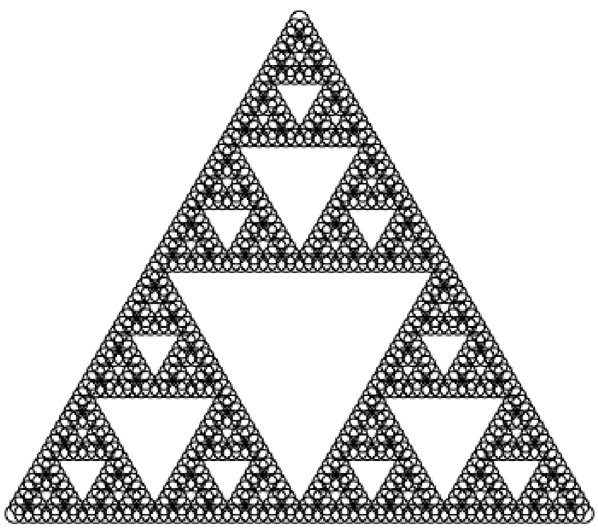

Figure 3.23: $T^6(S_0)$ after Sixth Iteration

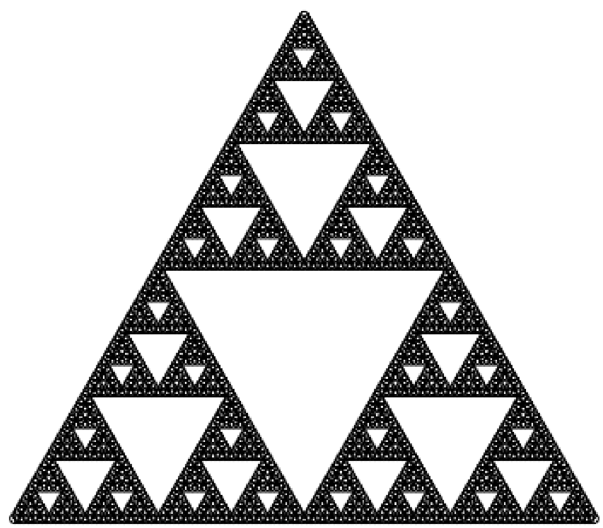

Figure 3.24: $T^7(S_0)$ after Seventh Iteration

□

Example 3.11 (Fern)

We consider the map $T : \mathcal{H}(\mathbb{R}^2) \to \mathcal{H}(\mathbb{R}^2)$ give by

$$T(S) := T_1(S) \cup T_2(S) \cup T_3(S) \cup T_4(S),$$

where

$$
\begin{aligned}
T_1(x,y) &= (0.2x - 0.26y, 0.23x + 0.22y + 1.6),\\
T_2(x,y) &= (0.85x + 0.04y, -0.04x + 0.85y + 1.6),\\
T_3(x,y) &= (-0.15x + 0.28y, 0.26x + 0.24y + .44),\\
T_4(x,y) &= (0, 0.16y).
\end{aligned}
$$

Then T_1, T_2, T_3, T_4 are contractions with contractivity factors

$$s_1 = \sqrt{0.1174}, s_2 = \sqrt{0.7241}, s_3 = \sqrt{0.1564}, s_4 = \sqrt{0.0256}.$$

Where s_1 is calculated below.

$$
\begin{aligned}
&d(T_1(x_1,y_1), T_1(x_2,y_2))\\
&\quad = d[(0.2x_1 - 0.26y_1, 0.23x_1 + 0.22y_1 + 1.6), (0.2x_2 - 0.26y_2, 0.23x_2 + 0.22y_2 + 1.6)]\\
&\quad = \sqrt{(0.2(x_1 - x_2) - 0.26(y_1 - y_2))^2 + (0.23(x_1 - x_2) + 0.22(y_1 - y_2))^2}\\
&\quad = \sqrt{\begin{array}{l}(0.2)^2(x_1 - x_2)^2 - (2)(0.2)(0.26)(x_1 - x_2)(y_1 - y_2) + (0.26)^2(y_1 - y_2)^2 +\\ (0.23)^2(x_1 - x_2)^2 + (2)(0.23)(0.22)(x_1 - x_2)(y_1 - y_2) + (0.22)^2(y_1 - y_2)^2\end{array}}\\
&\quad = \sqrt{\begin{array}{l}[(0.2)^2 + (0.23)^2](x_1 - x_2)^2 + [(0.26)^2 + (0.22)^2](y_1 - y_2)^2\\ + 2[(0.23)(0.22) - (0.2)(0.26)](x_1 - x_2)(y_1 - y_2)\end{array}}\\
&\quad = \sqrt{0.0929(x_1 - x_2)^2 + 0.116(y_1 - y_2)^2 - 0.0028(x_1 - x_2)(y_1 - y_2)}\\
&\quad = \sqrt{0.0929(x_1 - x_2)^2 + 0.116(y_1 - y_2)^2 - 2(0.03741\ldots)(x_1 - x_2)(0.03741\ldots)(y_1 - y_2)}\\
&\quad \leq \sqrt{0.0929(x_1 - x_2)^2 + 0.116(y_1 - y_2)^2 + ((0.03741\ldots)(x_1 - x_2))^2 + ((0.03741\ldots)(y_1 - y_2))^2}\\
&\quad \leq \sqrt{(0.0929 + 0.0014)(x_1 - x_2)^2 + (0.116 + 0.0014)(y_1 - y_2)}\\
&\quad \leq \sqrt{(0.1174)(x_1 - x_2)^2 + (0.1174)(y_1 - y_2)^2}
\end{aligned}
$$

by taking the maximum of $(0.0929 + 0.0014)$ and $(0.116 + 0.0014)$

$= \sqrt{0.1174}\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$

Therefore $s_1 = \sqrt{0.1174}$.

The inequality in our calculations comes from the fact that $-2ab \leq a^2 + b^2$ where $a = (0.03741\ldots)(x_1 - x_2)$ and $b = (0.03741\ldots)(y_1 - y_2)$. Thus, T is a contraction with contractivity factor $s = \max\{s_i\}$ for $1 \leq i \leq 4$. Now, we examine a program that takes a point initiator $S_0 = \{\frac{1}{2}, \frac{1}{2}\}$ and then through the iterated function systems $S_0, T(S_0), T^2(S_0), \ldots$ the image of the fixed point $S = \lim_{n\to\infty} T^n(S_0)$ is produced.

Program: Fern with Point Initiator

```
step[lsp_, lsf_] := DeleteDuplicates[Round[Join @@ Map[Map[#, lsp] &, lsf],
    .01]]
step[lsp_, funs_, n_] := Nest[step[#, funs] &, lsp, n]
gpoint[xxx_] := Graphics[{Black, PointSize[.002], Map[Point, xxx]},
    AspectRatio -> 1, Axes -> False, ImageSize -> 800]
fff[mm_, tt_][xx_] := mm.xx + tt
aa = fff[{{.2 , -.26}, {.23, .22}}, {0, 1.6}];
bb = fff[{{.85 , .04 }, {-.04 , .85 }}, {0, 1.6}];
cc = fff[{{-.15 , .28}, {.26, .24}}, {0, .44}];
dd = fff[{{0 , 0 }, {0 , .16 }}, {0, 0}];
ee = fff[{{1 , 0 }, {0 , 1 }}, {0, 0}];
lsf1 = {aa, bb, cc, dd};
gpoint[step[{{.5, .5}}, lsf1, 23]]
```

Figure 3.25: $T^5(S_0)$

Figure 3.26: $T^{10}(S_0)$

Figure 3.27: $T^{15}(S_0)$

Figure 3.28: $T^{20}(S_0)$

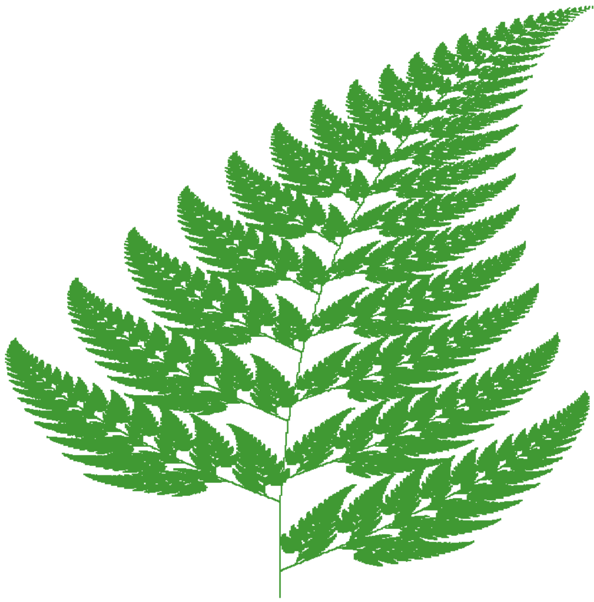

Figure 3.29: $T^{25}(S_0)$

Figure 3.30: $T^{30}(S_0)$

As more iterations are completed, the fern gets fuller. The fine details start to merge together as more points are placed within the limit shape. The are "evenly" dispersed over the image. In our next chapter, we will explore another method of generating limit shapes S_∞ for contractive iterated function systems that provides also great detail while understanding the need for computational speed. $\square$

Chapter 4
The Barnsley Method

As we saw in the previous section, the Sierpinski triangle and the Barnsley fern are fixed points S_∞ of contractive maps T on $\mathcal{H}(\mathbb{R}^2)$ that are attracting; that is, $S_\infty = \lim_{n\to\infty} T^n(S_0)$ for any initial set $S_0 \in \mathcal{H}(\mathbb{R}^2)$. The problem with approximating the shapes S_∞ is the large amount of memory needed to store the coordinates of the point of the n^{th} iterate $T^n(S_0)$, even for small values of n. For example, if $T = T_1 \cup T_2 \cup T_3 \cup T_4$ is the Barnsley fern map and S_0 consists of one point, then $T^n(S_0)$ consists of $4^n = 2^{2n}$ points. Clearly, when using the biggest laptop money can buy, going past $n = 15$ will take an eternity (since $2^{30} > 10^9$). However, as Figure 3.27 shows, the resolution of $T^{15}(S_0)$ is still quite unrefined. The way we handle this problem in Figures 3.28 - 3.30 is by

1. rounding the coordinates of each point to 0.01, and then
2. removing duplicates of points in the last evaluation of $T^n(S_0)$.

The fact that "rounding" does not destroy the quality of the graphical representation of the limit set S_∞ (the Barnsley fern) lies in the fact that S_∞ is an "attractor;" that is points close to S_∞ get even closer in subsequent iterations. The Barnsley method that will be discussed just in this section is even more radical in "removing" points from the list by creating the following "random walk".

Definition 4.1 (The Barnsley Method) *Let $T : \mathcal{H}(\mathbb{R}^2) \to \mathcal{H}(\mathbb{R}^2)$ be given as $T(S) := T_1(S) \cup T_2(S) \cup \cdots \cup T_N(S)$, where $T_i : \mathbb{R}^2 \to \mathbb{R}^2$ are given functions with transition probability distribution $p = (p_1, \ldots, p_n)$. Let $S_0 = \{S_0\}$ be an initial point in $\mathcal{H}(\mathbb{R}^2)$ and let $S_n := T_{i(k)}(T_{i(n-1)} \circ T_{i(n-2)} \circ \cdots \circ T_{i(1)})$ where $T_{i(k)} \in \{T_1, \ldots, T_N\}$ is randomly chosen with probability $p_{i(k)} \in \{p_1, \ldots, p_n\}$. Then the set of points $\{S_0, S_1, S_2 \ldots S_m\}$ is called a Barnsley random walk of length m.*

Clearly, if T is a contraction and if the initial point $S_0 = \{S_0\}$ is taken from the limit set (fixed point) $S_\infty := \lim_{n\to\infty} T^n(\Box)$ (for some initial set $\Box \in \mathcal{H}(\mathbb{R}^2)$, then the random walk $\{S_0, S_1, S_2 \ldots S_m\}$ will be a subset of S_∞ for all $m \in \mathbb{N}$

In fact, if the transition probability $(p_1, \ldots, p_N)$ are chosen appropriately, then the random walk will "fill out" the limit set S_∞ quite effectively for moderate-sized values of m (like $m = 50,000$).

Applying this method to the Sierpinski function $T = T_1 \cup T_2 \cup T_3$ (with uniform probability of $\frac{1}{3}$) or the Barnsley fern function $T = T_1 \cup T_2 \cup T_3 \cup T_4$ with probability distribution $(p_1, p_2, p_3, p_4) = (0.07, 0.85, 0.07, 0.01)$ reduces the number of stored points significantly since at each iteration only one point is added to the list of points that will approximate the limit shape S_∞.

Program: Barnsley Random Walk for Sierpinski Triangle, $m = 2000$, $p = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$

```
T1[{x_, y_}] := {.5 x, .5 y};
T2[{x_, y_}] := {.5 x + .5, .5 y};
T3[{x_, y_}] := {.5 x + .25, .5 y + .5};
x0 = 1/2; y0 = 1/2;
list = {{x0, y0}};
For[n = 0, n < 2000, n++,
p = RandomInteger[{1, 99}];
If[p <= 33, q = T1[list[[n]]]];
If[p > 33 && p <= 66, q = T2[list[[n]]]];
If[p > 66, q = T3[list[[n]]]];
AppendTo[list, q];];
ListPlot[list, Axes -> False, PlotStyle -> Black]
```

Figure 4.1: 20 Iterations

Figure 4.2: 200 Iterations

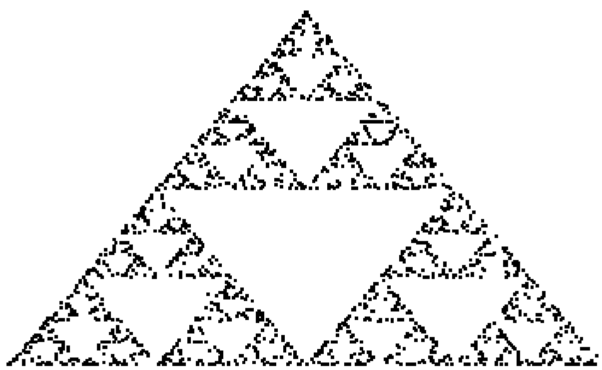

Figure 4.3: 2,000 Iterations

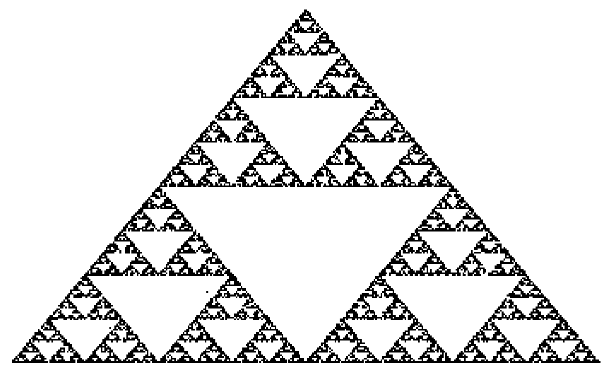

Figure 4.4: 20,000 Iterations

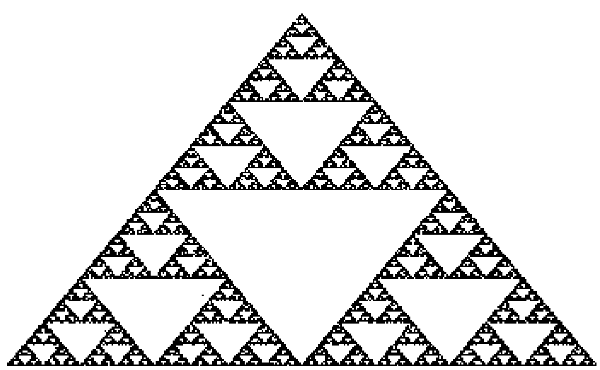

Figure 4.5: 30,000 Iterations

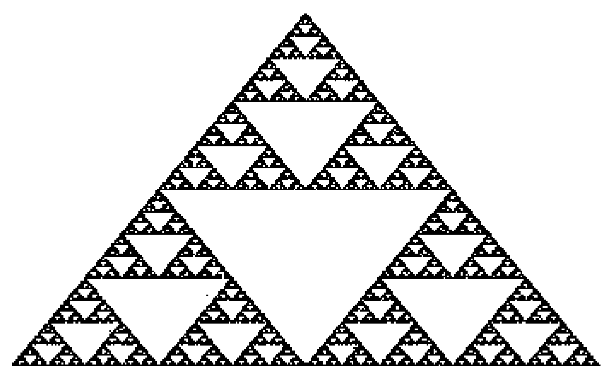

Figure 4.6: 50,000 Iterations

In Figure 4.1, we see how the first 20 iterations of this random walk produces a random walk of 20 points in the Sierpinski triangle and Figure 4.6 consists of a random walk of 50,000 stepping points within the Sierpinski triangle, filling the shape nicely. However, if one were to change the probability distribution from the uniform distribution of $(p_1, p_2, p_3) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ to a skewed distribution like $(0.8, 0.1, 0.1)$, then -clearly- the random

walk will spend more time in the lower left third's of the Sierpinski triangle instead of filling out the shape.

Program: Barnsley Random Walk for Sierpinski Triangle, $m = 2000$, $p = (0.8, 0.1, 0.1)$

```
T1[{x_, y_}] := {.5 x, .5 y};
T2[{x_, y_}] := {.5 x + .5, .5 y};
T3[{x_, y_}] := {.5 x + .25, .5 y + .5};
x0 = 1/2; y0 = 1/2;
list = {{x0, y0}};
For[n = 0, n < 2000, n++,
p = RandomInteger[{1, 100}];
If[p <= 80, q = T1[list[[n]]]];
If[p > 80 && p <= 90, q = T2[list[[n]]]];
If[p > 90, q = T3[list[[n]]]];
AppendTo[list, q];];
ListPlot[list, Axes -> False]
```

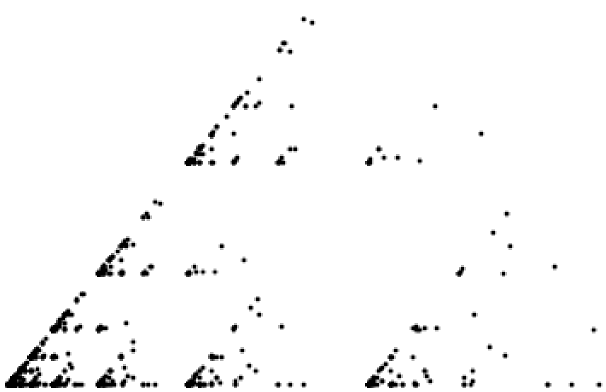

Figure 4.7: 2,000 Iterations

Figure 4.8: 50,000 Iterations

It is somehow fascinating (at least at first) that a random walk after 50,000 footprints (stepping periods) leads always to the same overall footprints although the paths taken are entirely different. Thus, if the maps T_i defining a random walk are contractions, randomness

is tamed in that the walk cannot be all over the place, but only within the limit structure (shape) S_∞.

If we apply the random walk method to the Barnsley fern function $T = T_1 \cup T_2 \cup T_3 \cup T_4$ it turns out that the probability distribution $(0.07, 0.85, 0.07, 0.01)$ produces a random walk that fills out the limit shape S_∞ (Barnsley fern) optimally. That is, all other probability distributions yield an image of S_∞ (fern) with far less overall detail.

Program: Barnsley Random Walk for Barnsley Fern, $m = 5000$, $p = (0.07, 0.85, 0.07, 0.01)$

```
T1[{x_, y_}] := {.2 x − .26 y, .23 x + .22 y + 1.6};
T2[{x_, y_}] := {.85 x + .04 y, −.04 x + .85 y + 1.6};
T3[{x_, y_}] := {−.15 x + .28 y, .26 x + .24 y + .44};
T4[{x_, y_}] := {0, .16 y};
x0 = 1/2; y0 = 1/2;
list = {{x0, y0}};
For[n = 0, n < 5000, n++,
 p = RandomInteger[{1, 100}];
 If[p <= 85, q = T2[list[[n]]]];
 If[p > 85 && p <= 92, q = T1[list[[n]]]];
 If[p > 92 && p <= 99, q = T3[list[[n]]]];
 If[p > 99, q = T4[list[[n]]]];
 AppendTo[list, q];
 ]; ListPlot[list, PlotRange -> All, AspectRatio -> 1,
 PlotStyle -> RGBColor[.25, .60, .2], Axes -> False]
```

Figure 4.9: Fern

Now that we have established two numerical effective methods to visualize the limit sets or fixed points S_∞ of contractive iterative function systems $T : \mathcal{H}(\mathbb{R}^2) \to \mathcal{H}(\mathbb{R}^2)$, we now turn to a question that we could not answer

Question:

1. Characterize those shapes $S \in \mathcal{H}(\mathbb{R}^2)$ that are limit sets for some contractive iterated function system $T : \mathcal{H}(\mathbb{R}^2) \to \mathcal{H}(\mathbb{R}^2)$.

2. Let $H_0(\mathbb{R}^2) := \{S \in \mathcal{H}(\mathbb{R}^2) : S = T(S)$ for some contraction $T : \mathcal{H}(\mathbb{R}^2) \to \mathcal{H}(\mathbb{R}^2)\}$. Is $H_0(\mathbb{R}^2)$ dense in $\mathcal{H}(\mathbb{R}^2)$?

An answer to these questions would be interesting from a data-comparison point of view. That is, given $\epsilon > 0$, for which two dimensional shapes $S \subset \mathcal{H}(\mathbb{R}^2)$ can one find an elementary function $T_s : \mathcal{H}(\mathbb{R}^2) \to \mathcal{H}(\mathbb{R}^2)$ and that

$$h(S, S_\infty) \leq \epsilon,$$

where $S_\infty = \lim_{n\to\infty} {T_s}^n(S_0)$ for some (all) initial sets $S_0 \in \mathcal{H}(\mathbb{R}^2)$?

For example, one of the most widely known fractal shapes is the "Mandelbrot Set"

$$M := \{c \in \mathbb{R}^2 : T_c^n(0) \text{ stays bounded}\},$$

where $T_c : \mathbb{R}^2 \to \mathbb{R}^2$ is given by $T_c(z) = z^2 + c$ or

$$T_c(x, y) = (x^2 - y^2 + c_1, 2xy + c_2)$$

for $z = (x, y) = x + iy$ and $c = (c_1, c_2) = c_1 + ic_2$.

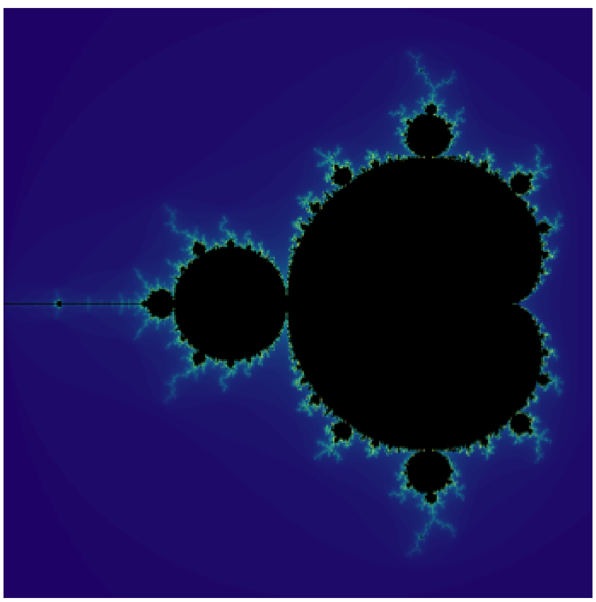

Figure 4.10: Mandelbrot Set

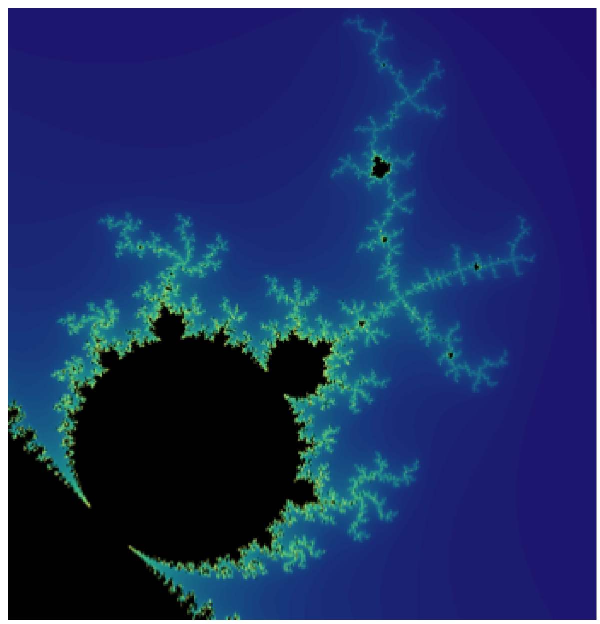

Figure 4.11: Zoomed Mandelbrot Set

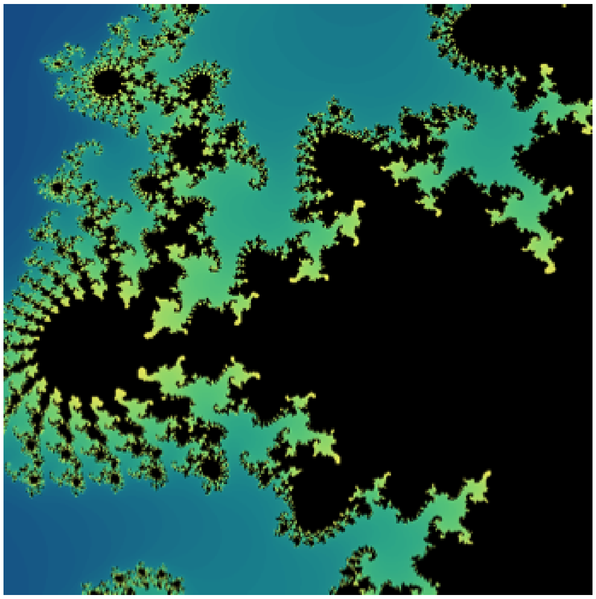

Figure 4.12: Zoomed Mandelbrot Set

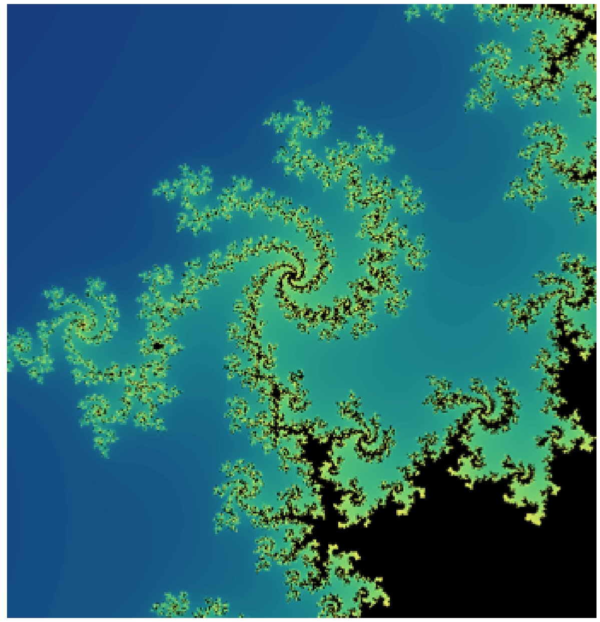

Figure 4.13: Zoomed Mandelbrot Set

It is well known that $M \in \mathcal{H}(\mathbb{R}^2)$ and that M is a fractal whose boundary δM has Hausdorff dimension 2. What is not known is if there is a contraction $T : \mathcal{H}(\mathbb{R}^2) \to \mathcal{H}(\mathbb{R}^2)$ such that $T^n(S_0) \to M$ for all $S_0 \in \mathcal{H}(\mathbb{R}^2)$, or that for each $\epsilon > 0$ there is a contraction $T_\epsilon : \mathcal{H}(\mathbb{R}^2) \to \mathcal{H}(\mathbb{R}^2)$ such that $h(M, T_\epsilon^n(S_0)) \leq \epsilon$ for all $n \geq n_0$ and $S_o \in \mathcal{H}(\mathbb{R}^2)$.

Chapter 5
Looking Forward

So far, we looked at shapes and fractals generated by contractive iterated function systems. As demonstrated, this situation is governed by the Contraction Mapping Principle. But what happens if the map $T : \mathcal{H}(\mathbb{R}^2) \to \mathcal{H}(\mathbb{R}^2)$ is not contractive but "only" non-expansive; that is,

$$h(T(X), T(Y)) \leq h(X, Y)$$

for all $X, Y \in \mathcal{H}(\mathbb{R}^2)$? This situation seems to be far more challenging. As mentioned above, the map $T : \mathbb{R} \to \mathbb{R}, T(x) := x + 1$, is non-expansive , that is, $(d(T(x), T(y)) = d(x, y)$ for all $x, y \in \mathbb{R}$), but $T^n(x) = x + n \to \infty$. The fact that iterated function systems $I, T, T^2, \ldots$ for non-expansive maps T may not converge to a fixed point leads to a much more difficult (and interesting) dynamical behavior.

Without going into much detail, we study one particular example of an iterated function system $I, T, T^2, \ldots$ defined by a non-expansive map $T : \mathcal{H}(\mathbb{R}^2) \to \mathcal{H}(\mathbb{R}^2)$. Let $R : \mathbb{R} \to \mathbb{R}$ be a rotation by $60°$; that is,

$$R = \begin{bmatrix} \frac{1}{2} & \frac{-\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix},$$

where $\frac{1}{2} = \cos(\frac{\pi}{3})$ and $\frac{\sqrt{3}}{2} = \sin(\frac{\pi}{3})$. Then R is distance-preserving, that is, $d(R(u), R(v)) = d(u, v)$ for all $u, v \in \mathbb{R}^2$). For $S \in \mathcal{H}(\mathbb{R}^2)$ define

$$T(S) = T_1(S) \cup T_2(S) \cup T_3(S),$$

where $T_1(u) = R(u), T_2(u) = R(u) + (\frac{1}{2}, \frac{1}{2})$, and $T_3(u) = R(u) - (\frac{1}{2}, \frac{1}{2})$ for $u \in \mathbb{R}^2$. Then T is distance preserving since T_1, T_2, T_3 are distance preserving and

$$|T^n(S_0)| := \max\{|T^n(u)| : u \in S_0\} \to \infty$$

as $n \to \infty$. To see this, we look first at the map $T_2(u) = R(u) + p, u \in \mathbb{R}^2, p = (\frac{1}{2}, \frac{1}{2})$.

Then $T_2^2(u) = T_2(R(u) + p) = R^2(u) + R(p) + p$ and

$$\begin{aligned} T_2^n(u) &= R^n(u) + \sum_{i=0}^{n-1} R^i(p) \\ &= R^n(u) + (I - R^n)(I - R)^{-1}(p) \end{aligned}$$

since $(I - R) \sum_{i=0}^{n-1} R^i = I - R^n$ or $\sum_{i=0}^{n-1} R^i = (I - R^n)(I - R)^{-1}$.

Now, $(I - R)^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{-\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{-\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} = R$, and therefore

$$\begin{aligned} T_2^n(u) &= R^n(u) + (I - R^n)R(p) \\ &= R^n(u) + R(p) - R^{n+1}(p). \end{aligned}$$

It follows that

$$\begin{aligned} T_2^6(u) &= R^6(u) + R(p) - R^7(p) \\ &= u + R(p) - R(p) = u. \end{aligned}$$

Similarly, $T_3^6(u) = T_1^6(u) = u$. However, one can show (by elementary matrix multiplication) that

$$(T_1T_3T_3T_1T_2T_2)(u) = u + \sqrt{3}(1, -1)$$

and therefore,

$$(T_1T_3T_3T_1T_2T_2)^n(u) = u + n\sqrt{3}(1, -1).$$

This shows that

$$|T^n(S_0)| \to \infty$$

for all $S_0 \in \mathcal{H}(\mathbb{R}^2)$. In particular, $T = T_1 \cup T_2 \cup T_3$ is not periodic although T_1, T_2, T_3 are periodic with period 6. The following program computes the orbit $S_0 \cup T(S_0) \cup T^2(S_0) \cup \cdots \cup T^n(S_0)$ for $S_0 = \{(0.4, 0.2)\}$.

Program: Rotations by 60° and shifts

```
step[lsp_, lsf_] := DeleteDuplicates[Round[Join @@ Map[Map[#, lsp] &, lsf],
    .005]]
step[lsp_, funs_, n_] := Nest[step[#, funs] &, lsp, n]
gpoint[xxx_] := Graphics[{Black, PointSize[.001], Map[Point, xxx]},
    AspectRatio -> 1, Axes -> True, ImageSize -> 500]
fff[mm_, tt_][xx_] := mm.xx + tt
a = Cos[Pi/3]; (*Spinning Wheels*)
b = Sin[Pi/3];
c = -Sin[Pi/3];
d = Cos[Pi/3];
e1 = .5;
e2 = .5;
f1 = -.5;
f2 = -.5;
aa = fff[{{a , b}, {c, d}}, {0, 0}];
bb = fff[{{a , b}, {c, d}}, {e1, e2}];
cc = fff[{{a , b}, {c, d}}, {f1, f2}];
ee = fff[{{1, 0}, {0, 1}}, {0, 0}];
lsf1 = {aa, bb, cc, ee};
gpoint[step[{{.4, .2}}, lsf1, 18]]
```

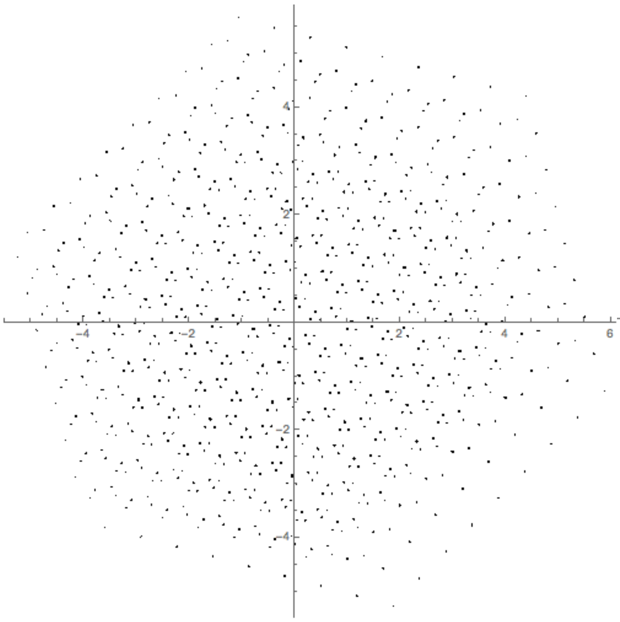

Figure 5.1: Twelve Iterations

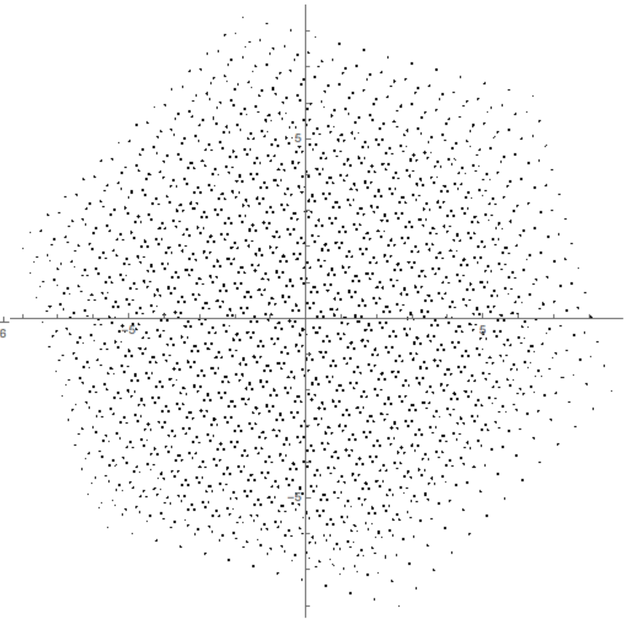

Figure 5.2: Eighteen Iterations

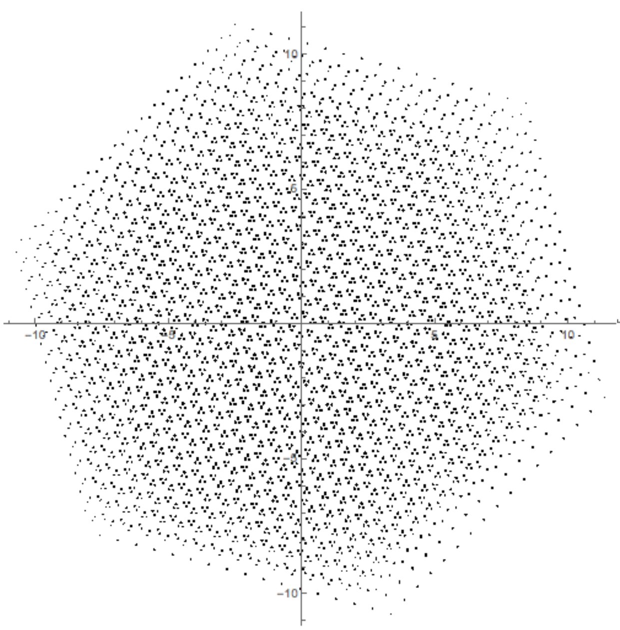

Figure 5.3: Twenty-Four Iterations

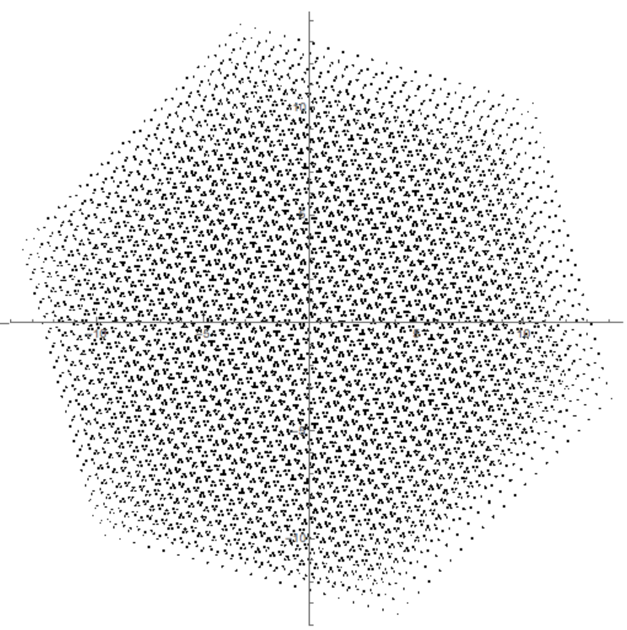

Figure 5.4: Thirty Iterations

The resulting orbits O_{24} and O_{30} suggest that they are always self-similar subsets of the plane. The following zooms illustrate it is not at exactly clear at the moment what O_∞ might look like (but, it cannot be all of $\mathbb{R}^2$ since it will be a countable union of countable subsets of $\mathbb{R}^2$, and therefore countable), but that O_∞ will still be a self-similar shape since the zoomed illustrations for each iteration produce the exact same image.

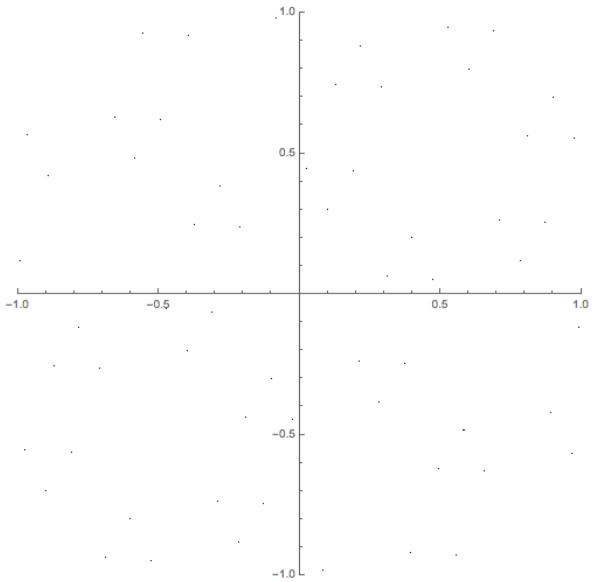

Figure 5.5: Twelve Iterations

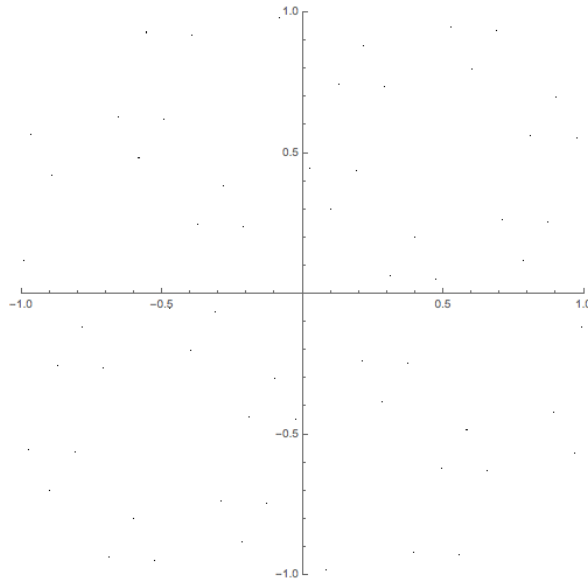

Figure 5.6: Eighteen Iterations

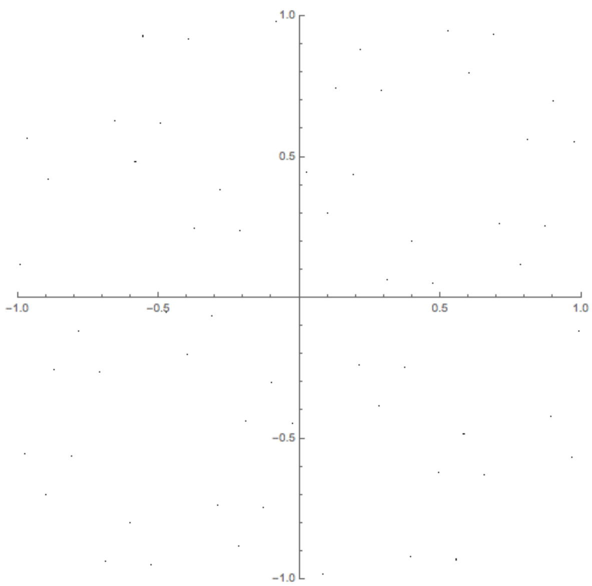

Figure 5.7: Twenty-Four Iterations

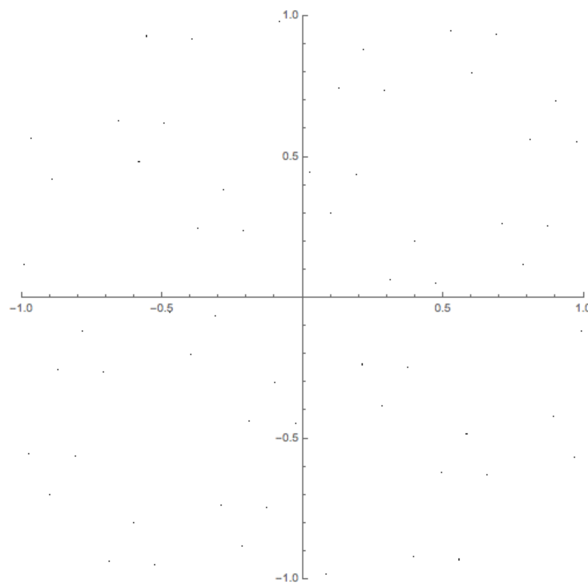

Figure 5.8: Thirty Iterations

It would be interesting to have a precise mathematical description of the sets $T^n(S_0)$, a task that I leave happily for others.

The orbits $O_n := S_0 \cup T(S_0) \cup T^2(S_0) \cup \cdots \cup T^n(S_0)$ seem to converge to a limit set $O_\infty \subset \mathbb{R}^2$ that is invariant under T. Because of computational limitations we can only compute O_n for $n \leq 30$. The resulting orbit O_{30} suggests that the limit orbit O_∞ is a true subset of the plane, invariant under T, and self-similar.

It is also interesting to look at random walks defined by $T = T_1 \cup T_2 \cup T_3$ with uniform transition probabilities $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. Whereas in the contractive case the random walk is confined to the attractor $S_\infty = \lim_{n\to\infty} T^n(S_0)$, the situation is more interesting for the non-expansive map T as defined above. Here, the random walk is confined to the set O_∞, but is far from "filling out" O_∞.

Program: Random Wheels

```
a = Cos[Pi/3];
b = Sin[Pi/3];
c = -Sin[Pi/3];
d = Cos[Pi/3];
e1 = 1/2; e2 = 1/2;
f1 = -1/2; f2 = -1/2;
xx[x_, y_, 1] := a*x + b*y;
yy[x_, y_, 1] := c*x + d*y;
xx[x_, y_, 2] := a*x + b*y + e1;
yy[x_, y_, 2] := c*x + d*y + e2;
xx[x_, y_, 3] := a*x + b*y + f1;
yy[x_, y_, 3] := c*x + d*y + f2;
u = .4; v = .2;
l := {{u, v}};
For[i = 1, i < 30000, i++,
  n = Random[Integer, {1, 3}];
  uu = xx[u, v, n];
  vv = yy[u, v, n];
  u = uu;
  v = vv;
  AppendTo[l, {u, v}]];
ListPlot[l, AspectRatio -> 1]
```

The following 12 images represent 12 consecutive runs of the random walks defined by T. Starting at (0.4, 0.2) with 15,000 total steps taken.

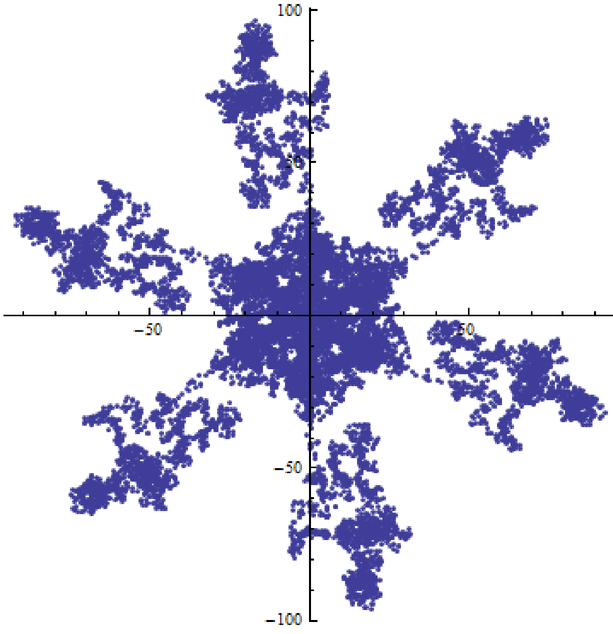

Figure 5.9: Run 1

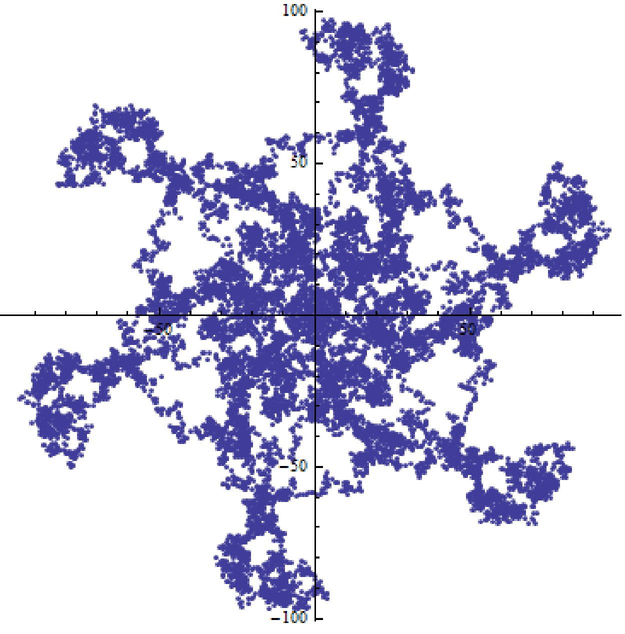

Figure 5.10: Run 2

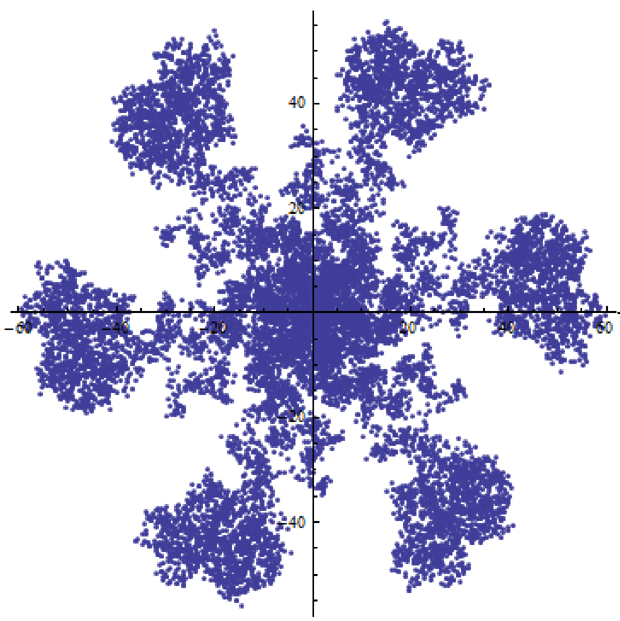

Figure 5.11: Run 3

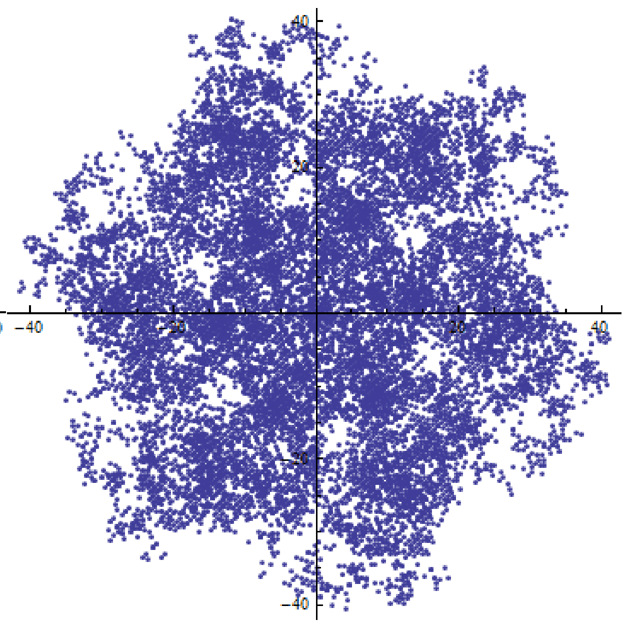

Figure 5.12: Run 4

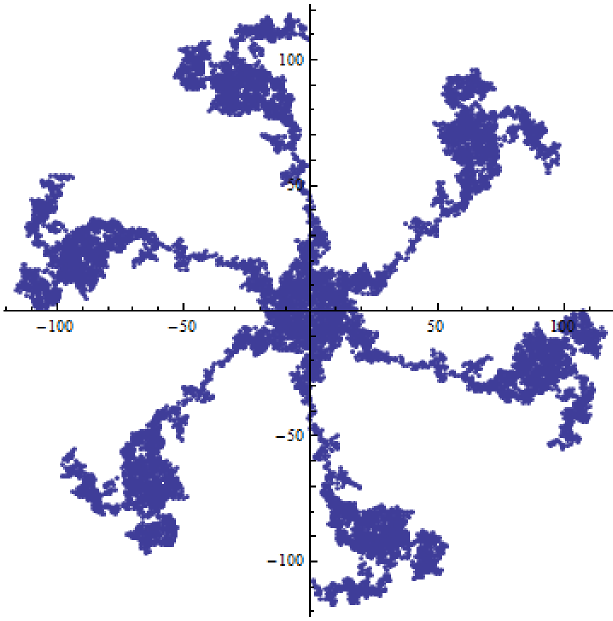

Figure 5.13: Run 5

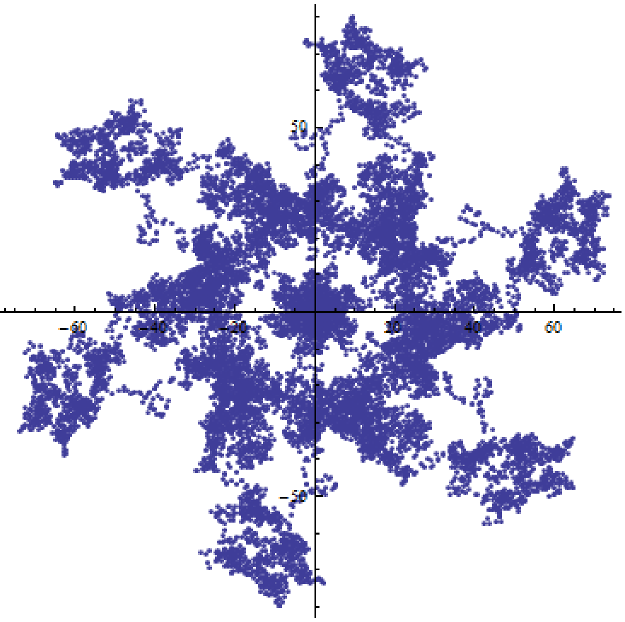

Figure 5.14: Run 6

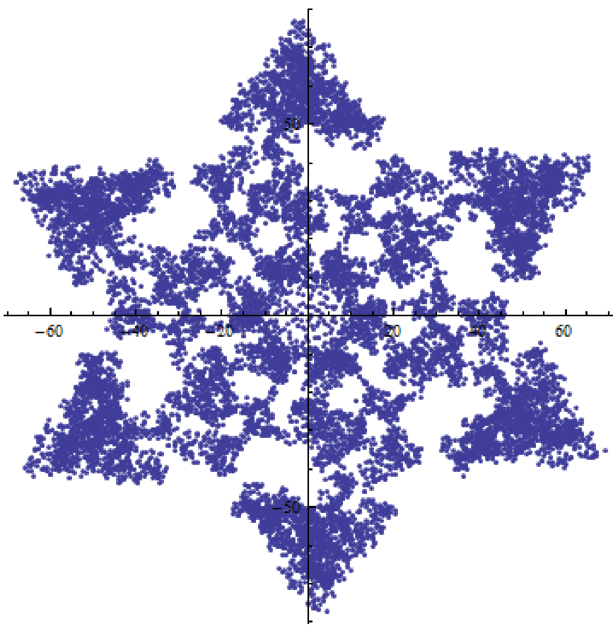

Figure 5.15: Run 7

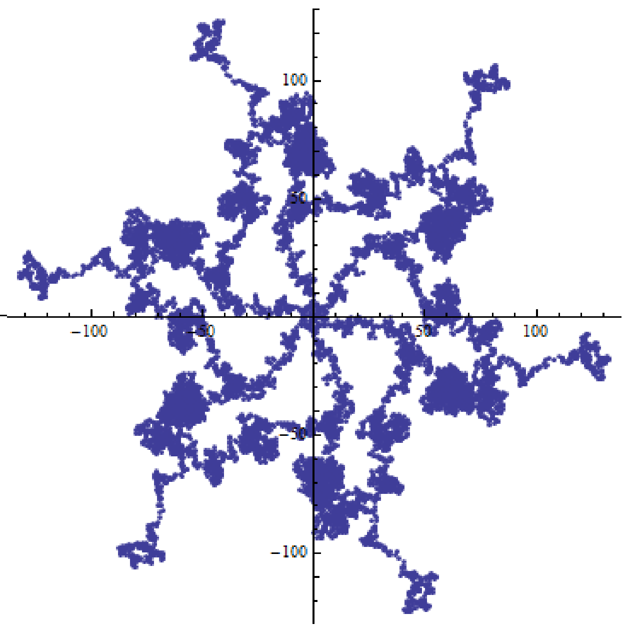

Figure 5.16: Run 8

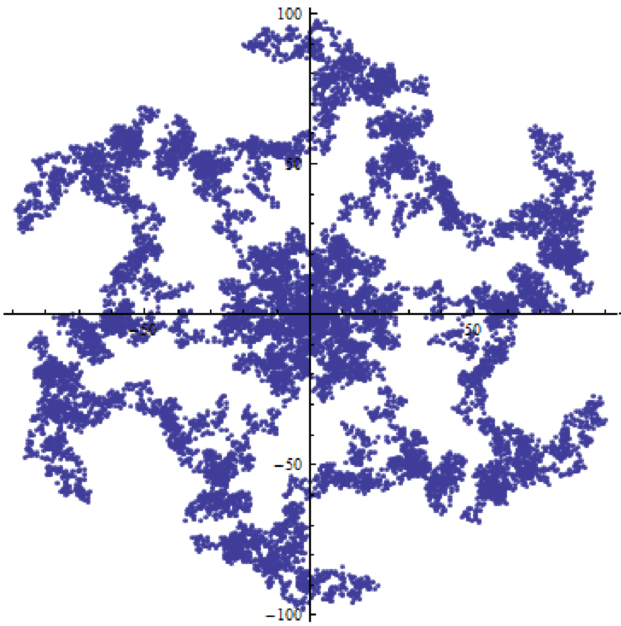

Figure 5.17: Run 9

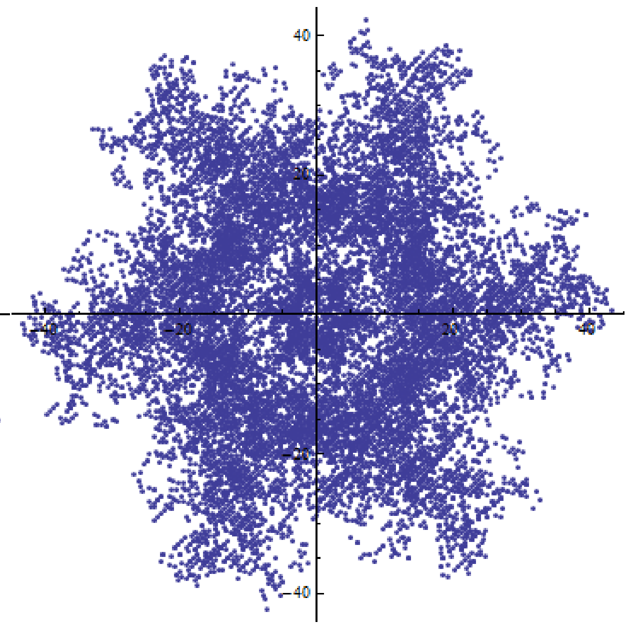

Figure 5.18: Run 10

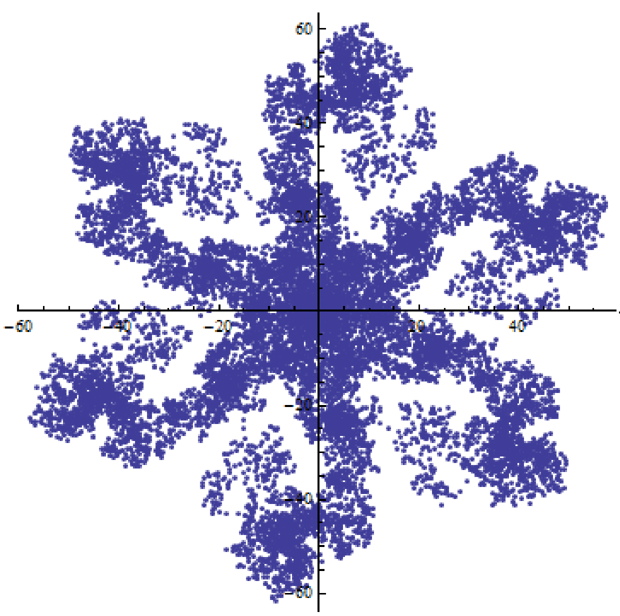

Figure 5.19: Run 11

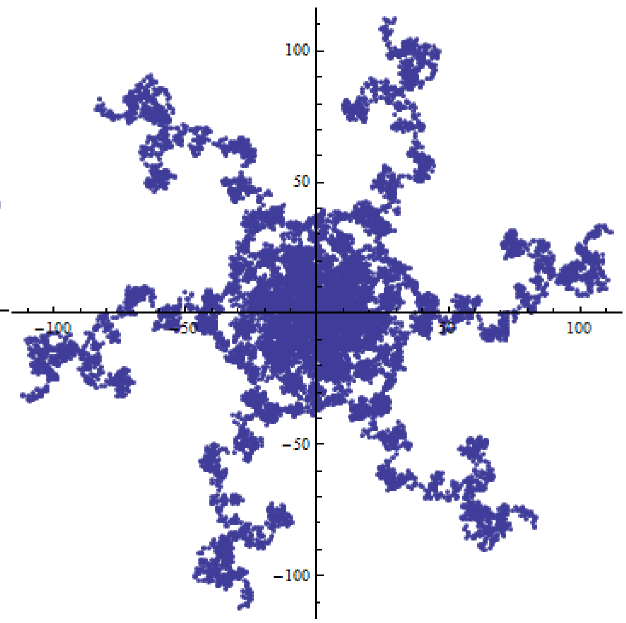

Figure 5.20: Run 12

The following 12 pictures represent a random walk consisting of 30,000 steps. Again, the footprints of the walks are self-similar and the images themselves clearly "related". How they are related exactly remains to be explored.

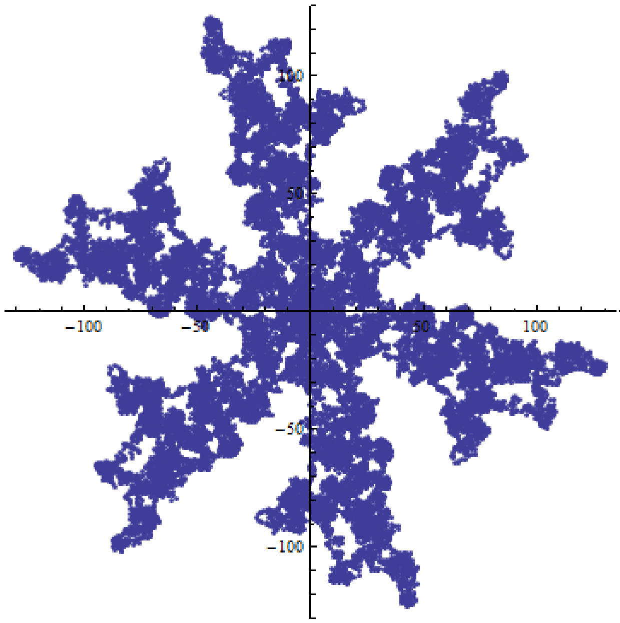

Figure 5.21: Run 1

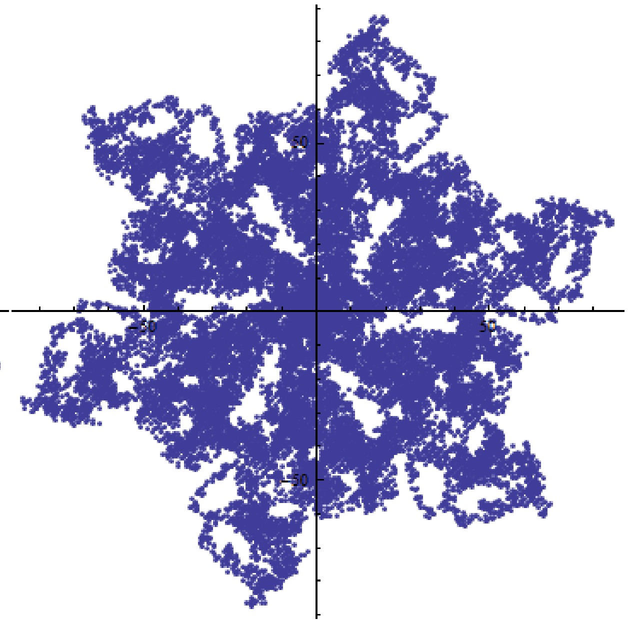

Figure 5.22: Run 2

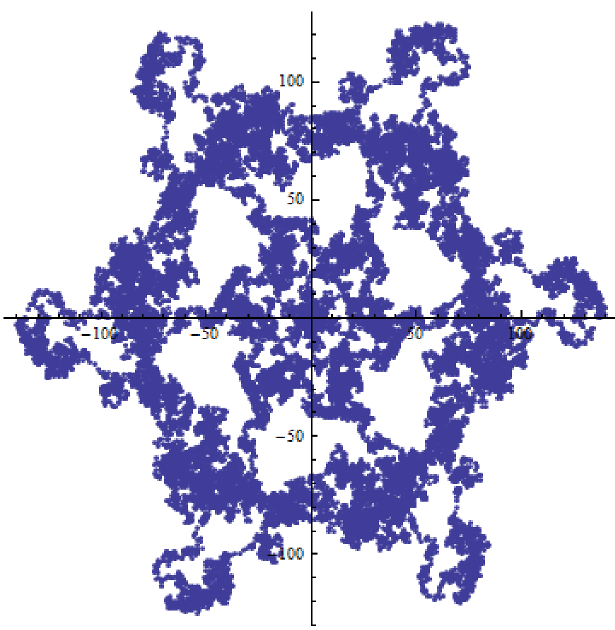

Figure 5.23: Run 3

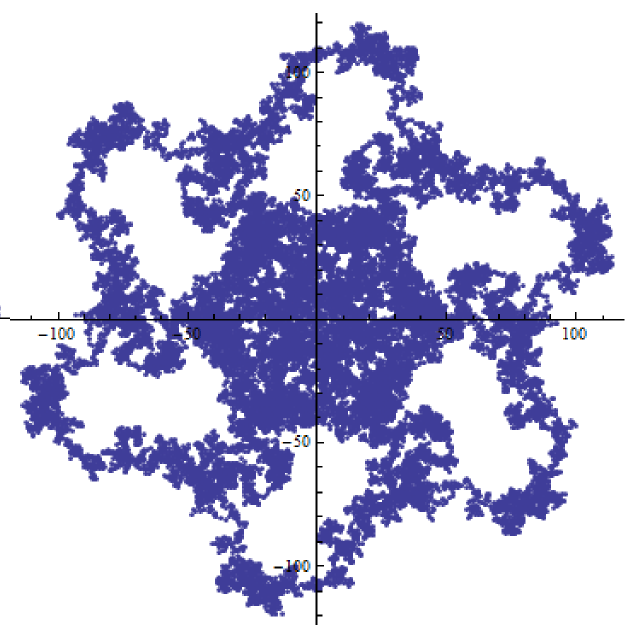

Figure 5.24: Run 4

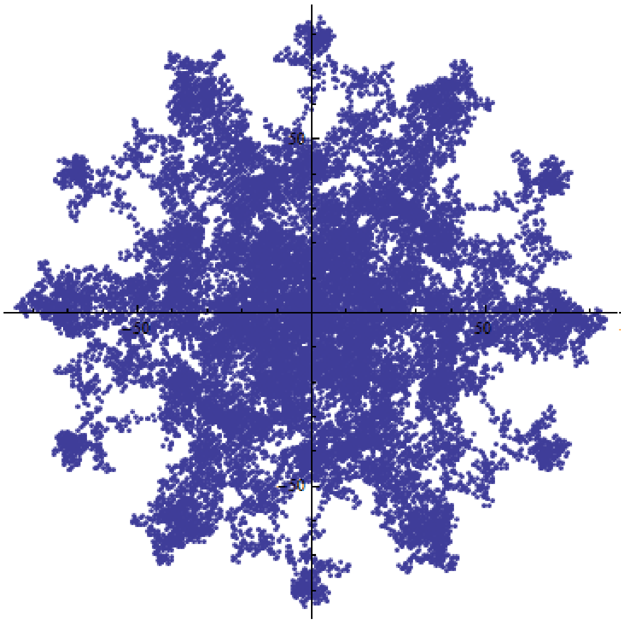

Figure 5.25: Run 5

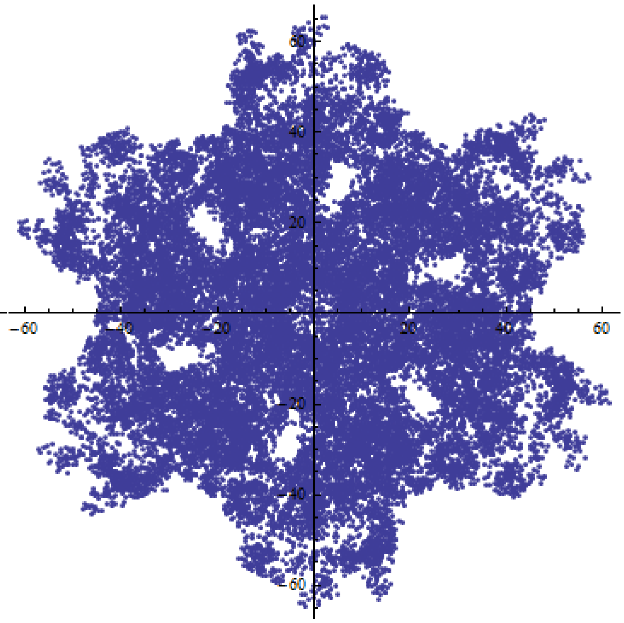

Figure 5.26: Run 6

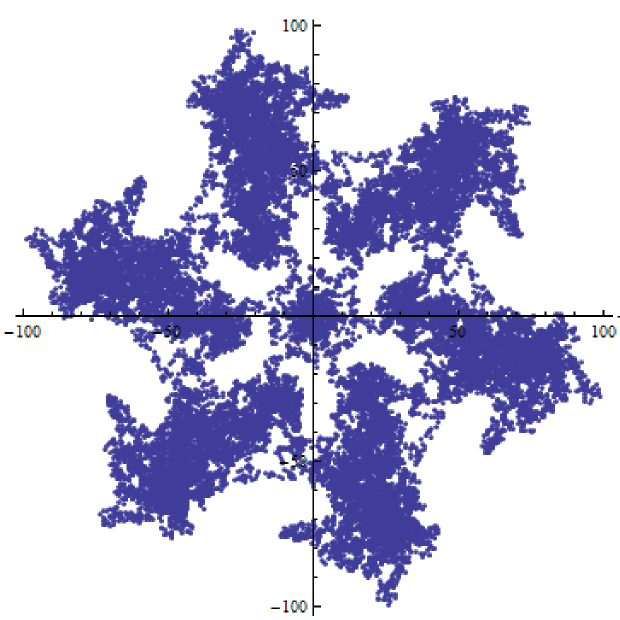

Figure 5.27: Run 7

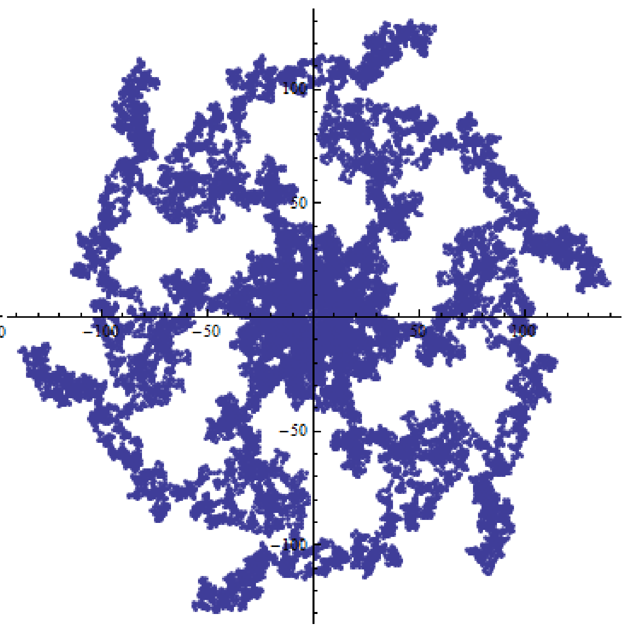

Figure 5.28: Run 8

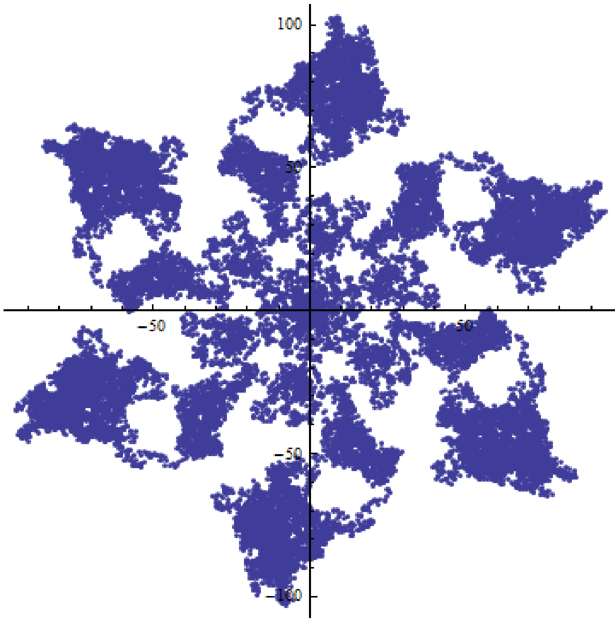

Figure 5.29: Run 9

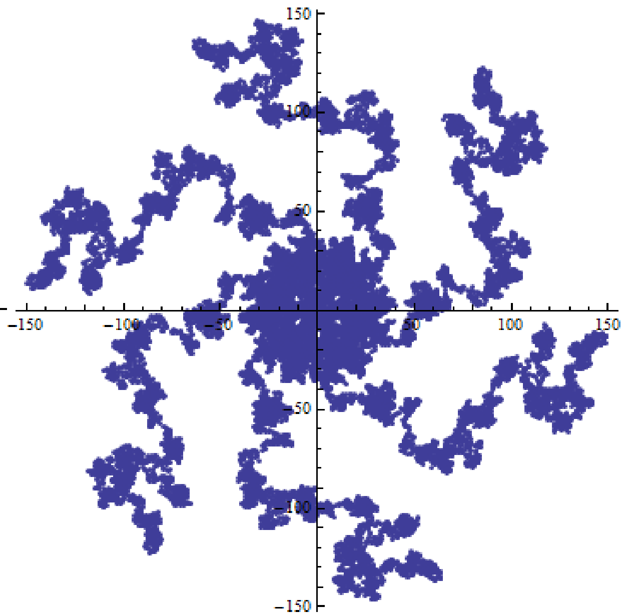

Figure 5.30: Run 10

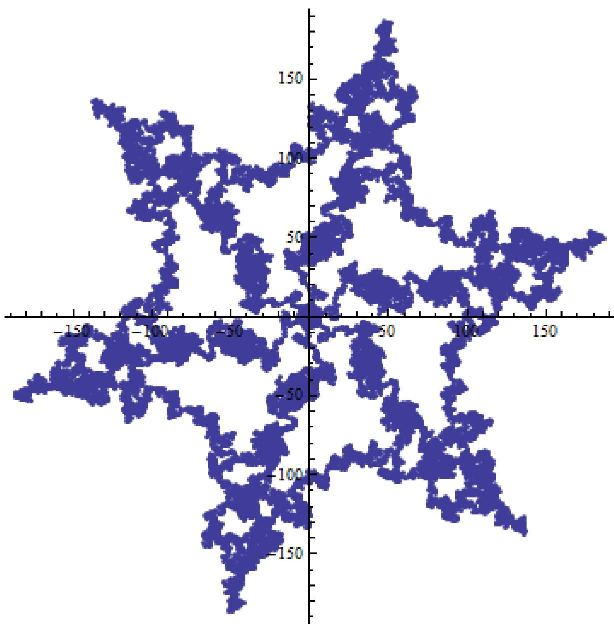

Figure 5.31: Run 11

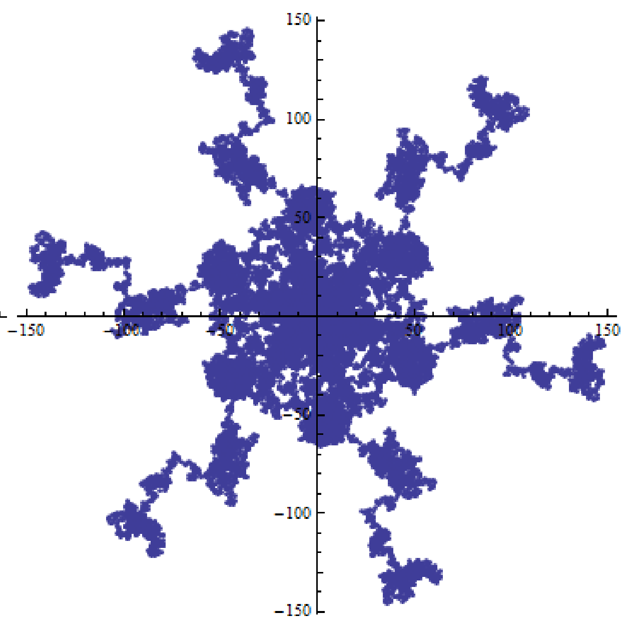

Figure 5.32: Run 12

References

[1] Barnsley, Michael Fielding. Fractals Everywhere. 2nd ed. Boston, MA: Academic Professional, 1993. Print.

[2] British Coastline. https://upload.wikimedia.org/wikipedia/commons/2/20/Britain-fractal-coastline-combined.jpg.

[3] "Chapter 4: Calculating Fractal Dimensions." Web. 21 Oct. 2016.

[4] Losa, G.A., Ristanovic, D., Ristanovic, D., Zaletel, I. and Beltraminelli, S. (2016) From Fractal Geometry to Fractal Analysis. Applied Mathematics, 7, 346-354. http://dx.doi.org/10.4236/am.2016.74032

[5] Mandelbrot, Benoit B. The Fractal Geometry of Nature. New York: W.H. Freeman, 1983. Print.

[6] Mandelbrot, Benoit B. "How Long is the Coast of Britain?" (1967). Web. 1 Aug 2016.

[7] Stewart, James. Calculus: Early Transcendentals. Boston: Cengage Learning, 2016. Print.

[8] Wolfram Mathematica. "Plot the Mandelbrot Set: New in Mathematica 10." Web. 21 Oct. 2016.

[9] "Michael Barnsley." Wikipedia: The Free Encyclopedia. Wikimedia Foundation, Inc. 22 July 2004. Web. 10 Aug. 2016.

[10] "Benoit Mandelbrot." Wikipedia: The Free Encyclopedia. Wikimedia Foundation, Inc. 22 July 2004. Web. 10 Aug. 2016.

[11] "Felix Hausdorff". Wikipedia: The Free Encyclopedia. Wikimedia Foundation, Inc. 22 July 2004. Web. 10 Aug. 2016.

Vita

Mary Catherine McKinley, a native of Huntsville, Alabama, received her Bachelor's of Science from Spring Hill College in Mobile, Alabama double majoring in Mathematics with a Concentration in Computer Analysis and Hispanic Studies in 2015. She was accepted into the Louisiana State University Graduate School to continue her education in Mathematics beginning the fall of 2015. She has taught as a Teaching Assistant for LSU during her time in graduate school. She anticipates graduating with her Master's of Science in Mathematics in December 2016. She plans to continue pursing her dream of being a mathematics educator and eventually earn her Ph.D.