SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

# An Introduction to Fractals and Iterated Function Systems

av

**Adriel Alander**

2022 - No K2

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

# An Introduction to Fractals and Iterated Function Systems

Adriel Alander

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Gregory Arone

2022

**Abstract**

Fractals are undeniably popular, but what exactly is it that they are, and how do we construct them? Without giving a formal definition of a fractal, we will see that fractals are metric spaces that are characterized by having a fractional Hausdorff dimension, and an infinitely intricate structure that repeats on all scales. To generate examples of fractals, we use a technique known as "Iterative Function Systems". It is based on the idea that the space of compact subsets of a complete metric space is itself a complete metric space. We will construct fractals as limits of Cauchy sequences in this "meta" space. By the means provided in M. F. Barnsley's *Fractals Everywhere* we use the random iteration algorithm to render images for the Sierpinski triangle, the Barnsley fern, a fractal tree, Koch snowflake, and the dragon curve. We find their fractal dimensions to be 1.585, 1.45, 1.407, 1.262, and 1.524 respectively, where for the Koch snowflake and the dragon curve we specifically looked at the dimension of their boundaries.

## Acknowledgements

Thank you to my supervisor, Gregory Arone, for all the help and patience—I could not have written this without you. Also a special thanks to my father, Tommy Alander, who helped me proofread and make my figures look prettier.

# Contents

# 1 Introduction

Fractals are fascinating mathematical objects that can be rendered in pretty pictures. Some also argue that fractals provide a good mathematical model for many phenomena in nature. Once you know what to look for you will notice that fractals, indeed, are everywhere. We come to see that fractal theory is used in many types of sciences and engineering, as well as artistry. Then, what exactly are fractals? How do we go about constructing them? Using M. F. Barnsley's book *Fractals Everywhere* we will enter the fascinating world of fractals and explore their properties!

## 1.1 What are Fractals?

With the shape of nature in mind, Benoit Mandelbrot formed a theory in geometry. His goal was to express and measure mathematically the roughness of objects in nature, such as mountains, trees, coastlines, etc. He introduced the word "fractal" in 1975 to describe these structures.

There is not a universally agreed upon definition of what a fractal is, though mathematicians in the field may feel that they recognize one when they see it. We will refrain from giving a formal definition of a fractal. Rather, we use it as a somewhat informal term describing spaces that possess one or more of the following closely related properties:

1. Fractional dimension. The ordinary topological dimension of a space is, by definition, an integer. But there is another notion of dimension, sometimes called the Hausdorff dimension. Hausdorff dimension agrees with topological dimension for "nice" spaces like manifolds, but in general it can take fractional values. Benoit Mandelbrot initially defined a fractal as "a set for which the Hausdorff dimension strictly exceeds the topological dimension." This definition encompasses a large proportion of spaces that people regard as fractals, but perhaps not all of them.

2. Self-similarity. A fractal is typically a set that is similar, via some standard geometric transformation, to a part of itself. In many cases a fractal can be written as a union of several scaled down copies of itself. By recursion, this often leads to the next property.

3. An infinitely intricate structure that repeats itself on all scales. This in return means that fractals are "nowhere differentiable", which implies that fractals cannot be measured in traditional ways, which was the problem that gave Mandelbrot the motivation to form his new theory [4].

We are going to be using something called "iterated function systems" to construct our fractals. This concept was first introduced by J. E. Hutchinson in 1981. It was later expanded upon by Michael F. Barnsley [3].

The main source we will be using is *Fractals Everywhere* by Michael F. Barnsley [2]. All of the following definitions, theorems, lemmas, and proofs can be found in the book unless otherwise stated.

## 1.2 Background

In this part of the text we will have a look at a few definitions and theorems that are not necessarily restricted to the theory of fractals, but rather may be categorized as standard topology. The purpose of this section is to work as a collection for the reader to refresh their knowledge as well as provide a dictionary with tools we will later use when forming our theory.

First we will include the definition of a metric space which we will be using moving forward. We are also covering Cauchy sequences and their significance to a complete metric space.

**Definition 1.1** (Metric Space)**.** *A metric space* $(\mathbf{X}, d)$ *is a space* $\mathbf{X}$ *together with a real-valued function* $d : \mathbf{X} \times \mathbf{X} \to \mathbb{R}$*, which measures the* distance *between pairs of points* $x$ *and* $y$ *in* $\mathbf{X}$*. We require that* $d$ *obeys the following axioms:*

1. $d(x, y) = d(y, x)\ \forall x, y \in \mathbf{X}$
2. $0 < d(x, y) < \infty\ \forall x, y \in \mathbf{X}, x \neq y$
3. $d(x, x) = 0\ \forall x \in \mathbf{X}$
4. $d(x, y) \leq d(x, z) + d(z, y)\ \forall x, y, z \in \mathbf{X}$.

*Such a function* $d$ *is called a* metric.

**Definition 1.2** (Cauchy Sequence)**.** *A sequence* $\{x_n\}_{n=1}^{\infty}$ *of points in a metric space* $(\mathbf{X}, d)$ *is called a* Cauchy sequence *if, for any given number* $\epsilon > 0$*, there is an integer* $N > 0$ *so that*

$$d(x_n, x_m) < \epsilon \quad \text{for all } n, m > N.$$

**Definition 1.3** (Complete Metric Space)**.** *A metric space* $(\mathbf{X}, d)$ *is* complete *if every Cauchy sequence* $\{x_n\}_{n=1}^{\infty}$ *in* $\mathbf{X}$ *has a limit* $x \in \mathbf{X}$.

Following three definitions cover specific properties amongst subsets of metric spaces.

**Definition 1.4** (Compact Subset)**.** *Let* $S \subset \mathbf{X}$ *be a subset of a metric space* $(\mathbf{X}, d)$*.* $S$ *is compact if every infinite sequence* $\{x_n\}_{n=1}^{\infty}$ *in* $S$ *contains a subsequence having a limit in* $S$.

**Definition 1.5** (Bounded)**.** *Let* $S \subset \mathbf{X}$ *be a subset of a metric space* $(\mathbf{X}, d)$*.* $S$ *is* bounded *if there is a point* $a \in \mathbf{X}$ *and a number* $R > 0$ *so that*

$$d(a, x) < R\ \forall x \in \mathbf{X}.$$

**Definition 1.6** (Totally Bounded)**.** *Let $S \subset \mathbf{X}$ be a subset of a metric space $(\mathbf{X}, d)$. $S$ is* totally bounded *if, for each $\epsilon > 0$, there is a finite set of points $\{y_1, y_2, \dots, y_n\} \subset S$ such that whenever $x \in \mathbf{X}$, $d(x, y_i) < \epsilon$ for some $y_i \in \{y_1, y_2, \dots, y_n\}$. This set of points $\{y_1, y_2, \dots, y_n\}$ is called an* $\epsilon$-net.

**Theorem 1.1.** *Let $(\mathbf{X}, d)$ be a complete metric space. Let $S \subset \mathbf{X}$. Then $S$ is compact if and only if it is closed and totally bounded.*

Proof to theorem above is not included in this paper, but may be found in Barnsley's *Fractals Everywhere* on page 20.

**Definition 1.7.** *Let $S$ denote a set of real numbers. Then the* infimum *of $S$ is equal to $-\infty$ if $S$ contains negative numbers of arbitrary large magnitude. Otherwise the infimum of $S = \max\{x \in \mathbb{R} : x \leq s \text{ for all } s \in S\}$. The infimum of $S$ always exists because of the nature of the real number system, and it is denoted by $\inf S$. The* supremum *of $S$ is similarly defined. It is equal to $+\infty$ if $S$ contains arbritary large numbers; otherwise it is the minimum of the set of numbers that are greater than or equal to all of the numbers in $S$. The supremum of $S$ always exists, and it is denoted by $\sup S$.*

Lastly we will introduce a notation for when we want to iterate transformations on a metric space.

**Definition 1.8.** *Let $f : \mathbf{X} \to \mathbf{X}$ be a transformation on a metric space. The* forward iterates *of $f$ are transformations $f^{\circ n} : \mathbf{X} \to \mathbf{X}$ defined by $f^{\circ 0}(x) = x$, $f^{\circ 1}(x) = f(x)$, $f^{\circ(n+1)}(x) = f \circ f^{(n)}(x) = f(f^{(n)}(x))$ for $n = 0, 1, 2, \dots$. If $f$ is invertible then the* backward iterates *of $f$ are transformations $f^{\circ(-m)}(x) : \mathbf{X} \to \mathbf{X}$ defined by $f^{\circ(-1)}(x) = f^{-1}(x)$, $f^{\circ(-m)}(x) = (f^{\circ m})^{-1}(x)$ for $m = 1, 2, 3, \dots$.*

Now we hopefully have a base that lets us understand the fractal theory we will be exploring moving forward.

# 2 The Metric Space of Fractals

Without giving an explicit definition of a fractal, we are interested in *a subset of a metric space* with certain properties. To find this, we must first establish a framework that provides tools that let us execute a search for such subsets.

## 2.1 Hausdorff Sets

Consider a metric space which points are subsets of a fixed metric space $(\mathbf{X}, d)$. We will construct fractals in $\mathbf{X}$ as limits of Cauchy sequences in this space of subsets of $\mathbf{X}$. Before we figure out exactly how to do so, let us give a proper definition.

**Definition 2.1.** *Let $(\mathbf{X}, d)$ be a complete metric space. Define $\mathcal{H}(\mathbf{X})$ to be the set of all compact subsets of $(\mathbf{X}, d)$ other than the empty set. We will call this set $\mathcal{H}(\mathbf{X})$ the* Hausdorff set *of* $\mathbf{X}$.

So we are looking at metric spaces that are *complete*, that is, metric spaces where all Cauchy sequences have a limit (see definition 1.3 for further details.) A common example of a complete metric space, which also happen to be the one we will be using, is the Euclidean plane ($\mathbb{R}^2$, Euclidean). Note that this property of completeness amongst the Euclidean spaces is not restricted to the plane, but actually for any $\mathbb{R}^n$ we get a complete metric space.

We are interested in Cauchy sequences in $\mathcal{H}(\mathbf{X})$, so we need to construct a complete metric on this set of sets. How exactly do we measure the distance between two sets of points? Let us define a way!

## 2.2 Distance Between Sets

In order to study convergent sequences we need to define a way to measure how "far away" two points are in the Hausdorff set. Since each point in the Hausdorff set $\mathcal{H}(\mathbf{X})$ is itself a subset of $\mathbf{X}$, we will define the distance between two sets. We begin this by looking at the distance from one point in $\mathbf{X}$ to an entire subset.

**Definition 2.2.** *Let* $(\mathbf{X}, d)$ *be a complete metric space. Consider* $x \in \mathbf{X}$ *and* $B \in \mathcal{H}(\mathbf{X})$*. Define*

$$d(x, B) = \min\{d(x, y) : y \in B\}$$

*to be the distance* from *the point* $x$ to *the set* $B$*.*

The assumption that $B$ is compact and non-empty guarantees that $d(x, B)$ is well-defined. We now have a value for the distance from a point $x \in \mathbf{X}$ to a set $B \in \mathcal{H}(\mathbf{X})$. We will use this definition in order to define the distance from a set $A \in \mathcal{H}(\mathbf{X})$ to another subset $B \in \mathcal{H}(\mathbf{X})$.

**Definition 2.3.** *Let* $(\mathbf{X}, d)$ *be a complete metric space. Let* $A, B \in \mathcal{H}(\mathbf{X})$*. Define the distance* from *the set* $A$ to *the set* $B$ *to be*

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

With this definition, we do not necessarily get that $d(A, B) = d(B, A)$. According to definition 1.1, this does not qualify as a metric due to its asymmetry. In order to obtain a valid metric we will need to adjust our definition.

**Definition 2.4** (Hausdorff Distance)**.** *Let* $(\mathbf{X}, d)$ *be a complete metric space. Consider* $A, B \in \mathcal{H}(\mathbf{X})$*. Define*

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

*to be the* Hausdorff distance *between two sets.*

By this we get a symmetric way to measure distances, but before we proceed we must indeed confirm that this forms a metric on $\mathcal{H}(\mathbf{X})$.

*Proof.* Considering our set $\mathcal{H}(\mathbf{X})$ and the function $h : \mathcal{H}(\mathbf{X}) \times \mathcal{H}(\mathbf{X}) \to \mathbb{R}$. For all $A,\, B,\, C \in \mathcal{H}(\mathbf{X})$, we have the following axioms:

1. $h(A,\, B) = h(B,\, A)$,
2. $0 < h(A,\, B) < \infty$,
3. $h(A,\, A) = 0$,
4. $h(A,\, B) \leq h(A,\, C) + h(C,\, B)$.

Axiom 1 we have already stated to be true. 3 is apparent since $h(A,\, A) = \max\{d(A,\, A),\, d(A,\, A)\} = d(A,\, A) = \max\{d(x,\, A) : x \in A\} = 0$. Since $A$ and $B$ are compact, $h(A,\, B) = d(a,\, b)$ for some $a \in A$ and $b \in B$, which tells us that $0 \leq h(A,\, B) < \infty$. Furthermore, if $A \neq B$ we can assume there is an $a \in A$ so that $a \notin B$. This follows that $h(A,\, B) \geq d(A,\, B) > 0$, which confirms axiom 2. Lastly, to show that axiom 4 is true we will first show that $d(A,\, B) \leq d(A,\, C) + d(C,\, B)$. For any $a \in A$ we have

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

Similarly,

$$\begin{aligned}
d(B,\, A) &\leq d(B,\, C) + d(C,\, A), \text{and therefore} \\
h(A,\, B) &= \max\{d(A,\, B),\, d(B,\, A)\} \\
&\leq \max\{d(B,\, C),\, d(C,\, B)\} + \max\{d(A,\, C),\, d(C,\, A)\} \\
&= h(B,\, C) + h(A,\, C),
\end{aligned}$$

as desired. □

This allows us to form the metric space $(\mathcal{H}(\mathbf{X}),\, h)$. This is the metric space that Barnsley expressed as "the space where fractals live", which will be the space in which we will work in moving forward.

## 2.3 The Complete Metric Space $(\mathcal{H}(\mathbf{X}),\, h)$

If the metric space $(\mathcal{H}(\mathbf{X}),\, h)$ is complete, we know that all Cauchy sequences have a limit. From this we can find a Cauchy sequence whose limit is the subset we are trying to find; our desired fractal. Before we prove that this metric space is complete, we need to introduce some tools that will help us in the proof. More specifically we are interested in the extension of a Cauchy sequence.

**Definition 2.5.** *Let $S \subset \mathbf{X}$ and let $\Gamma \geq 0$. Then $S + \Gamma = \{y \in \mathbf{X} : d(x, y) \leq \Gamma$ for some $x \in S\}$.*

$S + \Gamma$ is sometimes called, for example, in the theory of set morphology, the dilation of $S$ by a ball of radius $\Gamma$.

**Lemma 2.1.** *Let $A$ and $B$ belong to $\mathcal{H}(\mathbf{X})$ where $(\mathbf{X}, d)$ is a metric space. Let $\epsilon > 0$. Then*

$$h(A, B) \leq \epsilon \iff A \subset B + \epsilon \text{ and } B \subset A + \epsilon.$$

*Proof.* Begin by showing that $d(A, B) \leq \epsilon \iff A \subset B+\epsilon$. Suppose $d(A, B) \leq \epsilon$. Then $\max\{d(a, B) : a \in A\} \leq \epsilon$ implies $d(a, B) \leq \epsilon$ for all $a \in A$. Hence for each $a \in A$ we have a $a \in B + \epsilon$, which means that $A \subset B + \epsilon$. Suppose $A \subset B + \epsilon$. Consider $d(A, B) = \max\{d(a, B) : a \in A\}$. Let $a \in A$. Since $A \subset B+\epsilon$, there is $a$, $b \in B$ so that $d(a, b) \leq \epsilon$ for all $a \in A$. Hence $d(a, B) \leq \epsilon$. This is true for each $a \in A$. So $d(A, B) \leq \epsilon$. Since, by definition, $h(A, B) = \max\{d(A, B), d(B, A)\}$, which means that $h(A, B) \leq \epsilon$ if and only if $d(A, B) \leq \epsilon$ and $d(B, A) \leq \epsilon$. We know that $d(A, B) \leq \epsilon \iff A \subset B + \epsilon$, and by symmetry we know that $d(B, A) \leq \epsilon \iff B \subset A+\epsilon$, so $h(A, B) \leq$ if and only if $A \subset B + \epsilon$ and $B \subset A + \epsilon$. □

The following lemma will show us a way to apply this theory on our Haussdorff sets.

**Lemma 2.2** (The Extension Lemma)**.** *Let $(\mathbf{X}, d)$ be a metric space. Let $\{A_n : n = 1, 2, \ldots, \infty\}$ be a Cauchy sequence of points in $(\mathcal{H}(\mathbf{X}), h)$. Let $\{n_j\}_{j=1}^{\infty}$ be an infinite sequence of integers*

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

*Suppose we have a Cauchy sequence $\{x_{n_j} \in A_{n_j} : j = 1, 2, 3, \ldots\}$ in $(\mathbf{X}, d)$. Then there is a Cauchy sequence $\{\tilde{x}_n \in A_n : n = 1, 2, \ldots\}$ such that $\tilde{x}_{n_j} = x_{n_j}$, for all $j = 1, 2, 3, \ldots$.*

*Proof.* We begin this proof by constructing the sequence $\{\tilde{x}_n \in A_n : n = 1, 2, \ldots\}$. For each $n \in \{1, 2, \ldots, n_1\}$, choose $\tilde{x}_n$ to be the closest point, or one of the closest points, in $A_n$ to $x_{n_1}$: $\tilde{x}_n \in \{x \in A_n : d(x, x_{n_1}) = d(x_{n_1}, A_n)\}$. Since $A_n$ is compact we know that such point exists. Similarly, for each $j \in \{2, 3, \ldots\}$ and each $n \in \{n_j + 1, \ldots, n_{j+1}\}$, choose $\tilde{x}_n \in \{x \in A_n : d(x, x_{n_{j+1}}) = d(x_{n_{j+1}}, A_n)\}$.

Now we show that $\{\tilde{x}_n\}$ is an extension of $\{x_{n_j}\}$ to $\{A_n\}$. By construction $\tilde{x}_{n_j} = x_{n_j}$ and $x_n \in A_n$. To show that it is a Cauchy sequence, let $\epsilon > 0$ be given. There is an $N_1$ so that $n_k$, $n_j \geq N_1$ implies $d(x_{n_k}, x_{n_j}) \leq \epsilon/3$. There is an $N_2$ so that $m$, $n \geq N_2$ implies

$$d(A_m, A_n) \leq \epsilon/3.$$

Let $N = \max\{N_1, N_2\}$ and note that, for $m, n \geq N$,

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

where $m \in \{n_{j-1}+1, n_{j-1}+2, \ldots, n_j\}$ and $n \in \{n_{k-1}+1, n_{k-1}+2, \ldots, n_k\}$. Since $h(A_m, A_{n_j}) < \epsilon/3$ there exists $y \in A_m \cap (\{x_{n_j}\} + \epsilon/3)$ so that $d(\tilde{x}_m, x_{n_j}) \leq \epsilon/3$. Similarly $d(x_{n_k}, \tilde{x}_n) \leq \epsilon/3$. Hence $d(\tilde{x}_m, \tilde{x}_n) \leq \epsilon$ for all $m, n > N$. □

We now have the tools necessary in order to prove that the metric space $(\mathcal{H}(\mathbf{X}), h)$ is complete. Before we do, we will properly formulate the theorem:

**Theorem 2.3** (The Completeness of the Space of Fractals)**.** *Let $(\mathbf{X}, d)$ be a complete metric space. Then $(\mathcal{H}(\mathbf{X}), h)$ is a complete metric space. Moreover, if $\{A_n \in \mathcal{H}(\mathbf{X})\}_{n=1}^{\infty}$ is a Cauchy Sequence, then*

$$A = \lim_{n\to\infty} A_n \in \mathcal{H}(\mathbf{X})$$

*can be characterized as follows:*

$$A = \{x \in \mathbf{X} : \text{there is a Cauchy sequence } \{x_n \in A_n\} \text{ that converges to } x\}.$$

This last expression is what we will be using as our definition for $A$ when proving this theorem.

*Proof.* In order to show that $A \neq \emptyset$ we will show that there exist a Cauchy sequence $\{a_i \in A_i\}$ in $\mathbf{X}$.
Find a sequence $N_1 < N_2 < \cdots < N_n < \cdots$ of positive integers so that

$$h(A_m, A_n) \leq \frac{1}{2^i} \text{ for } m, n \geq N_i.$$

Select an element $x_{N_1} \in A_{N_1}$. Since $h(A_{N_1}, A_{N_2}) \leq \frac{1}{2}$, we can pick $x_{N_2} \in A_{N_2}$ so that $d(x_{N_1}, x_{N_2}) \leq \frac{1}{2}$. Suppose we choose a finite sequence $x_{N_i} \in A_{N_i}$ where $d(x_{N_{i-1}}, x_{N_i}) \leq \frac{1}{2^{i-1}}$ where $i = 1, 2, \ldots, k$. Then, since $x_{N_k} \in A_{N_k}$, and $h(A_{N_k}, A_{N_{k+1}}) \leq \frac{1}{2^k}$, we can choose $x_{N_{k+1}} \in A_{N_{k+1}}$ such that $d(x_{N_k}, x_{N_{k+1}}) \leq \frac{1}{2^k}$. Assume we pick the $x_{N_{k+1}}$ that is the closest to $x_{N_k}$. By induction we are able to select an infinite sequence $\{x_{N_i} \in A_{N_i}\}$ where $d(x_{N_k}, x_{N_{k+1}}) \leq \frac{1}{2^k}$.

Now we want to show that $\{x_{N_i}\}$ is a Cauchy sequence in $\mathbf{X}$. Let $\epsilon > 0$ and choose $N_\epsilon$ such that $\sum_{i=N_\epsilon}^{\infty} \frac{1}{2^i} < \epsilon$. For $m > n \geq N_\epsilon$ we get that

$$d(x_{N_m}, x_{N_n}) \leq d(x_{N_m}, x_{N_{m+1}}) + d(x_{N_{m+1}}, x_{N_{m+2}}) + \cdots + d(x_{N_{n-1}}, x_{N_n})$$
$$< \sum_{i=N_\epsilon}^{\infty} \frac{1}{2^i} < \epsilon.$$

By using the Extension lemma we find that there exists a convergent subsequence $\{a_i \in A_i\}$ where $a_{N_i} = x_{N_i}$. This tells us that $\lim a_i$ exists and by definition is in $A$, in other words, $A$ is *non-empty.*

Proceeding, we wish to show that $A$ is closed. Suppose that $\{a_i \in A_i\}$ is a sequence that converges to some point $a$. By showing that $a \in A$ we show that $A$ is indeed closed. For each positive integer $i$, there exists a sequence $\{x_{i,n} \in A_n\}$ such that $\lim_{n\to\infty} x_{i,n} = a_i$. We introduce two more sequences of positive numbers, one *increasing* $\{N_i\}_{i=1}^{\infty}$ such that $d(a_{N_i}, a) < \frac{1}{i}$, the other (not necessarily increasing) consists of integers $\{m_i\}$ such that $d(x_{N_i, m_i}, a_{N_i}) \leq \frac{1}{i}$. Thus $d(x_{N_i, m_i}, a) \leq \frac{2}{i}$. If we let $y_{m_i} = x_{N_i, m_i}$, we see that $y_{m_i} \in A_{m_i}$ and that $\lim_{i\to\infty} y_{m_i} = a$. Using the extension lemma again, we can extend $\{y_{m_i}\}$ to a convergent sequence $\{z_i \in A_i\}$, and so $a \in A$. This is what we wanted to show, therefore is $A$ closed. Since $\mathbf{X}$ is complete, we also know that $A$ is complete from this!

Before we prove that $A$ is totally bounded, we want to prove that for every $\epsilon > 0$, there is an $N$ such that for $n \geq N$, $A \subset A_n + \epsilon$. We will use this expression later in the proof!

Let $\epsilon > 0$. There exists an $N$ such that for some $m, n \geq N$, $h(A_m, A_n) \leq \epsilon$. Now let $n \geq N$, then for $m \geq n$, we get $A_m \subset A_n + \epsilon$. We want to show that $A \subset A_n + \epsilon$. Let $a \in A$. There is a sequence $\{a_i \in A_i\}$ that converges to $a$. We assume that $N$ is large enough so that for $m \geq N$, $d(a_m, a) < \epsilon$. Then we know that $a_m \in A_n + \epsilon$ since $A_m$ is a subset of $A_n + \epsilon$. Since $A_n$ is compact, it is possible to show that $A_n + \epsilon$ is closed. Because $a_m \in A_n + \epsilon$ for all $m \geq N$, $a$ must also be in $A_n + \epsilon$, which means that $A \subset A_n + \epsilon$ when $n$ is large enough.

We will prove that $A$ is totally bounded by contradiction. Suppose that $A$ was not totally bounded. Then for some $\epsilon > 0$ there would not exist a finite $\epsilon$-net. We could then find a sequence $\{x_i\}_{i=1}^{\infty}$ in $A$ such that $d(x_i, x_j) \geq \epsilon$ for $i \neq j$. We just proved that there exists an $n$ large enough so that $A \subset A_n + \frac{\epsilon}{3}$. For each $x_i$, there is a corresponding $y_i \in A_n$ for which $d(x_i, y_i) \leq \frac{\epsilon}{3}$. Since $A_n$ is compact, some subsequence $\{y_{n_i}\}$ of $\{y_i\}$ converges. From this converging sequence we can find points with a desired distance, in particular we select two points $y_{n_i}$ and $y_{n_j}$ such that $d(y_{n_i}, y_{n_j}) < \frac{\epsilon}{3}$. Using the triangle inequality we see

$$d(x_{n_i}, x_{n_j}) \leq d(x_{n_i}, y_{n_i}) + d(y_{n_i}, y_{n_j}) + d(y_{n_j}, x_{n_j}) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3},$$

which contradicts with the way $\{x_{n_i}\}$ was chosen. This would mean that $A$ is totally bounded. Since we already proved that $A$ is also complete, we also know that $A$ is *compact*.

The last piece we need to prove is that $\lim A_n = A$. We just showed that $A$ is compact, which would mean that $A \in \mathcal{H}(\mathbf{X})$. We have acquired tools that let us boil this down to showing that for $\epsilon > 0$, there is an $N$ such that, for $n \geq N$, $A_n \subset A + \epsilon$. We show this by letting $\epsilon > 0$ and finding $N$ so that for $m, n \geq N$, $h(A_m, A_n) \leq \frac{\epsilon}{2}$. Then for $m, n \geq N$, $A_m \subset A_n + \frac{\epsilon}{2}$. Let

$n \geq N$. We will show that $A_n \subset A + \epsilon$. Let $y \in A_n$. There exists an increasing sequence $\{N_i\}$ of integers such that $n < N_1 < N_2 < \cdots < N_k < \cdots$ and for $m, n \geq N_j$, $A_m \subset A_n + \frac{\epsilon}{2^{j+1}}$. Note that $A_n \subset A_{N_1} + \frac{\epsilon}{2}$. Since $y \in A_n$, there exists an $x_{N_1} \in A_{N_1}$ such that $d(y, x_{N_1}) \leq \frac{\epsilon}{2}$. Since $x_{N_1} \in A_{N_1}$, there is a point $x_{N_2} \in A_{N_2}$ such that $d(x_{N_1}, x_{N_2}) \leq \frac{\epsilon}{2^2}$. In a similar manner we can use induction to find a sequence $x_{N_1}, x_{N_2}, \ldots$, such that $x_{N_j} \in A_{N_j}$ and $d(x_{N_j}, x_{N_{j+1}}) < \frac{\epsilon}{2^{j+1}}$. Using the triangle inequality a number of times we can show that

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

and also that $\{x_{N_j}\}$ is a Cauchy sequence. From the way $n$ was chosen, each $A_{N_j} \subset A_n + \frac{\epsilon}{2}$. $\{x_{N_j}\}$ converges to a point $x$ since $A_n + \frac{\epsilon}{2}$ is closed, $x \in A_n + \frac{\epsilon}{2}$ also. Moreover, $d(y, x_{N_j}) \leq \epsilon$ implies that $d(y, x) \leq \epsilon$. We have thus shown that $A_n \subset A + \epsilon$ for $n \geq N$. This completes the proof that $\lim A_n = A$ and, since this was the last step, that $(\mathcal{H}(\mathbf{X}), h)$ is a complete metric space.

□

This is a rather long proof, but its result will be a powerful tool for us moving forward. We have now defined an environment where we wish to construct our fractals, but we have yet to figure out *how* to do so. In the next section we will have a closer look at different mappings on metric spaces and how we can use these combined with our metric space $(\mathcal{H}(\mathbf{X}), h)$ to create fractals.

# 3 IFS

There are many ways to go about when constructing a fractal. In this text our main focus will be on a method called *hyperbolic iterated function system*. Since this is a system of *functions*, we first need to have a closer look on what functions these are and what they look like on our metric space $(\mathcal{H}(\mathbf{X}), h)$.

## 3.1 Transformations

Barnsley mentions multiple types of transformations on different types of metric spaces, but since our main focus is on the plane $\mathbb{R}^2$, we will have a look at those relevant to that.

### 3.1.1 Affine Transformations

**Definition 3.1.** *A transformation $w : \mathbb{R}^2 \to \mathbb{R}^2$ of the form*

$$w(x_1, x_2) = (ax_1 + bx_2 + e, cx_1 + dx_2 + f),$$

*where $a$, $b$, $c$, $d$, $e$, and $f$ are real numbers, is called a (two-dimensional)* affine *transformation.*

A more convenient notation for this is to express this with matrices:

$$w(x) = w\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} = Ax + t,$$

where $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ and $t = \left(\begin{smallmatrix} e \\ f \end{smallmatrix}\right)$. $A$ can always be expressed with polar coordinates

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} r_1 \cos\theta_1 & -r_2 \sin\theta_2 \\ r_1 \sin\theta_1 & r_2 \cos\theta_2 \end{pmatrix},$$

which means that $(r_1,\, \theta_1)$ are the polar coordinates of the point $(a,\, c)$ and $(r_2,\, (\theta_2 + \pi/2))$ are those of point $(b,\, d)$. Like this it is easier to visualize what happens when you apply $A$ to points in space, since the vector $(1,\, 0)$ is rotated by angle $\theta_1$ and scaled by factor $r_1$, and $(0,\, 1)$ is rotated and scaled by $\theta_2$ and $r_2$. If we want to have some affine transformation that only allows translations, scaling, reflection, and rotation, we may set $r_1 = r_2$ and $\theta_1 = \theta_2$. We may give this sort of mapping a special name.

**Definition 3.2.** *A transformation $w : \mathbb{R}^2 \to \mathbb{R}^2$ is called a* similitude *if it is an affine transformation having one of the special forms*

$$w\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} r\cos\theta & -r\sin\theta \\ r\sin\theta & r\cos\theta \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}$$
$$w\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} r\cos\theta & r\sin\theta \\ r\sin\theta & -r\cos\theta \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}$$

*for some translation $(e,\, f) \in \mathbb{R}^2$, some real number $r \neq 0$, and some angle $\theta$, $0 \leq \theta < 2\pi$. $\theta$ is called the* rotation angle *while $r$ is called the* scale factor *or* scaling.

### 3.1.2 Contraction Mappings

**Definition 3.3.** *A transformation $f : \mathbf{X} \to \mathbf{X}$ on a metric space $(\mathbf{X},\, d)$ is called* contractive *or a* contraction mapping *if there is a constant $0 \leq s < 1$ such that*

$$d(f(x),\, f(y)) \leq s \cdot d(x,\, y) \forall x, \quad y \in \mathbf{X}.$$

*Any such number $s$ is called a* contractivity factor *for $f$.*

**Theorem 3.1** (The Contraction Mapping Theorem)**.** *Let $(\mathbf{X},\, d)$ be a complete metric space and let $f : \mathbf{X} \to \mathbf{X}$ be a contraction mapping. There is* exactly one *fixed point $x_f \in \mathbf{X}$ such that for any point $x \in \mathbf{X}$ the sequence $\{f^{\circ n}(x) : n = 1,\, 2,\, \ldots\}$ converges to $x_f$. That is,*

$$\lim_{n \to \infty} f^{\circ n}(x) = x_f, \quad \text{for each } x \in \mathbf{X}$$

*Proof.* Let $x \in \mathbf{X}$, and let $0 \leq s < 1$ be a contractivity factor for $f$. Then

$$d(f^{\circ n}(x),\, f^{\circ m}(x)) \leq s^{\min\{m,\, n\}} d(x,\, f^{\circ|n-m|})(x) \tag{1}$$

for all $m,\, n = 0,\, 1,\, 2, \ldots$, where we have fixed $x \in \mathbf{X}$. In particular, for $k = 0,\, 1,\, 2, \ldots$, we have

$$\begin{aligned} d(x,\, f^{\circ k}(x)) &\le d(x,\, f(x)) + (f(x),\, f^{\circ 2}(x)) + \ldots + d(f^{\circ(k-1)}(x),\, f^{\circ k}(x)) \\ &\le (1 + s + s^2 + \cdots + s^{k-1}) d(x,\, f(x)) \\ &\le (1-s)^{-1} d(x,\, f(x)), \end{aligned}$$

so substituting into equation 1 we obtain

$$d(f^{\circ n}(x),\, f^{\circ m}(x)) \le s^{\min\{m,\, n\}} \cdot (1-s)^{-1} \cdot (d(x,\, f(x)),$$

from which follows that $\{f^{\circ n}(x)\}_{n=0}^{\infty}$ is a Cauchy sequence. Since $\mathbf{X}$ is complete this Cauchy sequence possesses a limit $x_f \in \mathbf{X}$, and we have

$$\lim_{n\to\infty} f^{\circ n}(x) = x_f.$$

Now we need to show that $x_f$ is a fixed point of $f$. Since $f$ is contractive it is continuous, hence

$$f(x_f) = f(\lim_{n\to\infty} f^{\circ n}(x)) = \lim_{n\to\infty} f^{\circ(n+1)}(x) = x_f.$$

Lastly we need to check if it is possible to have more than one fixed point. Suppose there are. Let $x_f$ and $y_f$ be two fixed points of $f$. Then $x_f = f(x_f)$, $y_f = f(y_f)$, and

$$d(x_f,\, y_f) = d(f(x_f),\, f(y_f)) \le s d(x_f,\, y_f),$$

where $(1-s)d(x_f,\, y_f) \le 0$, which implies $d(x_f,\, y_f) = 0$ and hence $x_f = y_f$.

□

We are interested in these contraction mappings when applied to the metric space $(\mathcal{H}(\mathbf{X}),\, h)$. Now we want to know how a contraction mapping on the underlying metric $(\mathbf{X},\, d)$ translates to our Hausdorff space.

The first two things we are going to note are: A contraction mapping $w$ on the metric space $(\mathbf{X},\, d)$ is continuous, and $w$ maps $(\mathcal{H}(\mathbf{X}),\, h)$ into itself. So, by letting a contraction mapping act on a non-empty compact subset of $\mathbf{X}$ we in return get a compact subset of $\mathbf{X}$. The contraction mappings we speak of have only acted on specific points in the Hausdorff set, but we will now use these in order to create contraction mappings on $(\mathcal{H}(\mathbf{X}),\, h)$. Proofs to the following two lemmas can be found in Barnsley's book on pages 79–80.

**Lemma 3.2.** *Let $w : \mathbf{X} \to \mathbf{X}$ be a contraction mapping on the metric space $(\mathbf{X},\, d)$ with contractivity factor $s$. Then $w : \mathcal{H}(\mathbf{X}) \to \mathcal{H}(\mathbf{X})$ defined by*

$$w(B) = \{w(x) : x \in B\} \;\; \forall B \in \mathcal{H}(\mathbf{X})$$

*is a contraction mapping on $(\mathcal{H}(\mathbf{X}),\, h(d))$ with contractivity factor $s$.*

If we now want to combine multiple contraction mappings on $(\mathcal{H}(\mathbf{X}),\, h)$ to make new contraction mappings on $(\mathcal{H}(\mathbf{X}),\, h)$, we may use the following lemma:

**Lemma 3.3.** *Let* $(\mathbf{X},\, d)$ *be a metric space. Let* $\{w_n : n = 1,\, 2, \ldots,\, N\}$ *be contraction mappings on* $(\mathcal{H}(\mathbf{X}),\, h)$*, where* $s_n$ *is the contractivity factor for* $w_n$ *for each* $n$*. Define* $W : \mathcal{H}(\mathbf{X}) \to \mathcal{H}(\mathbf{X})$ *by*

$$W(B) = w_1(B) \cup w_2(B) \cup \ldots \cup w_n(B) = \cup_{n=1}^{N} w_n(B),$$

*for each* $B \in \mathcal{H}(\mathbf{X})$*. Then* $W$ *is a contraction mapping with contractivity factor* $s = \max\{s_n : n = 1,\, 2, \ldots,\, N\}$*.*

This in return gives us the final tool we need in order to define hyperbolic iterated function systems!

## 3.2 Hyperbolic Iterated Function Systems

We finally have the tools necessary in order to introduce the concept of *hyperbolic iterated function systems*. We start this off by defining what this is:

**Definition 3.4.** *A (hyperbolic)* iterated function system *(commonly referred to as "IFS") consists of a complete metric space* $(\mathbf{X},\, d)$ *together with a finite set of contraction mappings* $w_n : \mathbf{X} \to \mathbf{X}$*, with respective contractivity factors* $s_n$*, for* $n = 1,\, 2,\, \ldots,\, N$*. The notation for the IFS is* $\{\mathbf{X};\, w_n,\, n = 1,\, 2,\, \ldots,\, N\}$ *and its contractivity factor is* $s = \max\{s_n : n = 1,\, 2, \ldots,\, N\}$*.*

If we now combine all these mappings and let them act together in one transformation, we will find a unique fixed point similar to the one in the contraction mapping theorem. The following theorem is our main theoretical result. It combines Theorem 3.1, Lemma 3.2, and Lemma 3.3.

**Theorem 3.4.** *Let* $\{\mathbf{X} : w_n,\, n = 1,\, 2,\, \ldots,\, N\}$ *be a hyperbolic iterated function system with contractivity factor* $s$*. The transformation* $W : \mathcal{H}(\mathbf{X}) \to \mathcal{H}(\mathbf{X})$ *defined by*

$$W(B) = \cup_{n=1}^{N} w_n(B), \quad \text{for all } B \in \mathcal{H}(\mathbf{X}),$$

*is a contraction mapping on the complete metric space* $(\mathcal{H}(\mathbf{X}),\, h(d))$ *with contractivity factor* $s$*. In other words*

$$h(W(B),\, W(C)) \leq s \cdot h(B,\, C), \quad \text{for all } B,\, C \in \mathcal{H}(\mathbf{X}).$$

*There is a unique fixed point* $A \in \mathcal{H}(\mathbf{X})$ *such that*

$$A = W(A) = \cup_{n=1}^{N} w_n(A)$$

*and is given by* $A = \lim_{n\to\infty} W^{\circ n}(B)$ *for any* $B \in \mathcal{H}(\mathbf{X})$*.*

The attractor of a hyperbolic iterated function system is, in many cases, a fractal. It follows from Theorem 3.4 that the attractor equals to a union of scaled down copies of itself, which is one of our defining characteristics of a fractal. We will see later that in many cases the attractor has a fractional Hausdorff dimension. The figures we later generate depict the attractor of some iterated function systems, and in these we can tell that these have the infinitely intricate, nowhere differentiable structure that we expect of fractals.

In order to find the attractor of a given IFS, we need to start with a non-empty compact subset $A \in \mathbf{X}$, apply the functions $w_1, w_2, \ldots, w_n$ to $A$, take the union $w_1(A) \cup w_2(A) \cup \ldots \cup w_n(A)$, and then repeat the procedure many times. By theorem 3.4, the process converges to the attractor, so after enough iterations we will be close to it. In practice, rather than take the union of all $w_i(A)$, we will only apply one randomly selected function $w_i$ at each step. This is the Random Iteration Algorithm, which we will explore further in the next section.

## 3.3 Algorithms

Despite us often thinking of fractals as "cool pictures", we have yet to actually generate pictures of fractals. With the theory we have discussed, we can render pictures of attractors of different IFS using algorithms and computers. For simplicity's sake we will focus on fractals with the underlying space to be in $\mathbb{R}^2$.

If we consider hyperbolic IFS $\{\mathbb{R}^2; w_i : i = 1, 2, \ldots, N\}$, where each of its mapping is an affine transformation. Using the same notation as Barnsley we can express such map $w_i(x)$ as

$$w_i(x) = w_i \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} e_i \\ f_i \end{pmatrix} = A_i x + t_i.$$

Now if we use this we can express our IFS in tables. Tables 1–5 show examples on some IFS we will later generate pictures of. These also include a probability–column, which we will return to in section 3.3.1.

| $w$ | $a$ | $b$ | $c$ | $d$ | $e$ | $f$ | $p$ |
|---|---|---|---|---|---|---|---|
| 1 | 1/2 | 0 | 0 | 1/2 | 1 | 1 | 0.33 |
| 2 | 1/2 | 0 | 0 | 1/2 | 1 | 50 | 0.33 |
| 3 | 1/2 | 0 | 0 | 1/2 | 50 | 50 | 0.34 |

Table 1: IFS for the Sierpinski triangle.

| $w$ | $a$ | $b$ | $c$ | $d$ | $e$ | $f$ | $p$ |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0.16 | 0 | 0 | 0.01 |
| 2 | 0.85 | 0.04 | $-0.04$ | 0.85 | 0 | 1.6 | 0.85 |
| 3 | 0.2 | $-0.26$ | 0.23 | 0.22 | 0 | 1.6 | 0.07 |
| 4 | $-0.15$ | 0.28 | 0.26 | 0.24 | 0 | 0.44 | 0.07 |

Table 2: IFS for the Barnsley fern.

| $w$ | $a$ | $b$ | $c$ | $d$ | $e$ | $f$ | $p$ |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0.5 | 0 | 0 | 0.05 |
| 2 | 0.42 | $-0.42$ | 0.42 | 0.42 | 0 | 0.2 | 0.4 |
| 3 | 0.42 | 0.42 | $-0.42$ | 0.42 | 0 | 0.2 | 0.4 |
| 4 | 0.1 | 0 | 0 | 0.1 | 0 | 0.2 | 0.15 |

Table 3: IFS for a fractal tree.

| $w$ | $a$ | $b$ | $c$ | $d$ | $e$ | $f$ | $p$ |
|---|---|---|---|---|---|---|---|
| 1 | $1/2$ | $-\sqrt{3}/6$ | $\sqrt{3}/6$ | $1/2$ | 0 | 0 | 0.142 |
| 2 | $1/3$ | 0 | 0 | $1/3$ | $1/\sqrt{3}$ | $1/3$ | 0.143 |
| 3 | $1/3$ | 0 | 0 | $1/3$ | $1/\sqrt{3}$ | $-1/3$ | 0.143 |
| 4 | $1/3$ | 0 | 0 | $1/3$ | $-1/\sqrt{3}$ | $1/3$ | 0.143 |
| 5 | $1/3$ | 0 | 0 | $1/3$ | $-1/\sqrt{3}$ | $-1/3$ | 0.143 |
| 6 | $1/3$ | 0 | 0 | $1/3$ | 0 | $2/3$ | 0.143 |
| 7 | $1/3$ | 0 | 0 | $1/3$ | 0 | $2/3$ | 0.143 |

Table 4: IFS for the Koch snowflake.

| $w$ | $a$ | $b$ | $c$ | $d$ | $e$ | $f$ | $p$ |
|---|---|---|---|---|---|---|---|
| 1 | $1/2$ | $-1/2$ | $1/2$ | $1/2$ | 0 | 0 | 0.5 |
| 2 | $-1/2$ | $-1/2$ | $1/2$ | $-1/2$ | 1 | 0 | 0.5 |

Table 5: IFS for the dragon curve.

#### 3.3.1 The Random Iteration Algorithm

Let $\{\mathbf{X}; w_1, w_2, \ldots, w_N\}$ be a hyperbolic IFS, where for each $w_i$ $(i = 1, 2, \ldots, N)$ we have a probability $p_i > 0$ with $\sum_{i=1}^{n} p_i = 1$. Choose $x_0 \in \mathbf{X}$ and then recursively select

$$x_n \in \{w_1(x_{n-1}), w_2(x_{n-1}), \ldots, w_N(x_{n-1})\} \quad \text{for } n = 1, 2, 3, \ldots,$$

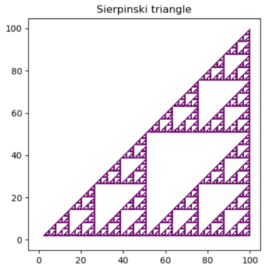
where each event $x_n = w_i(x_{n-1})$ has the probability $p_i$. In return we end up with the sequence $\{x_n : n = 0, 1, 2, \ldots\} \subset \mathbf{X}$.

When using the Random Iteration Algorithm we need to associate probabilities $p_i > 0$ to each $w_i$. Typically one chooses the probabilities as follows:

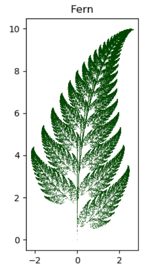
$$p_i \approx \frac{|\det A_i|}{\sum_{i=1}^{N} |A_i|} \qquad \text{for } i = 1,\, 2, \ldots,\, N.$$

In cases where $\det A_i = 0$, $p_i$ should be assigned a small positive number, such as 0.001. By giving the functions these probabilities we get a roughly even distribution of points over the attractor. Each function's image signify part of the attractor. Since the determinant of $A_i$ in some sense is related to the area of the image, we can calculate approximately the fraction of the attractor that is contributed from the image of $w_i$, hence distribute the points accordingly.

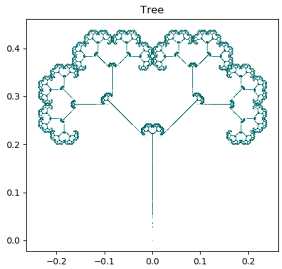
Together with the IFS shown in tables 1–5 we can now use this method in order to visualize these fractals. The results are shown in figure 1.



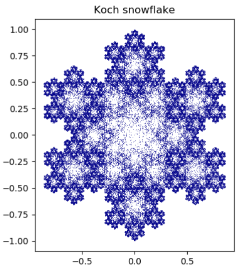

(a) Sierpinski Triangle (Table 1)



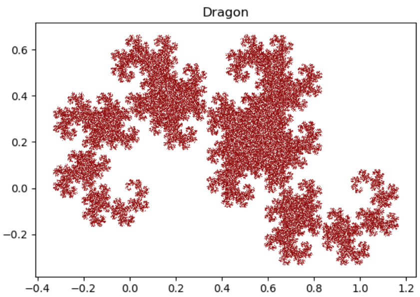

(b) Barnsley Fern (Table 2)




(c) Fractal Tree (Table 3)




(d) Koch Snowflake (Table 4)




(e) Dragon Curve (Table 5)

Figure 1: Generated images using the random iteration algorithm. Each figure consists of 100, 000 points of data. The IFS used can be found in tables 1–5.

# 4 Fractal Dimensions

## 4.1 New Definition of Dimensions

When talking about dimensions in an everyday-setting it is common to refer to them as different directions in space, or just as how many coordinates you need in order to give a point in space. When discussing fractals and their "roughness", we might want to introduce a new definition for dimensions, namely the *fractal dimension.*

**Definition 4.1.** *Let $A \in \mathcal{H}(\mathbf{X})$ where $(\mathbf{X}, d)$ is a metric space. For each $\epsilon > 0$ let $\mathcal{N}(A, \epsilon)$ denote the smallest number of closed balls of radius $\epsilon > 0$ needed to cover $A$. If*

$$D = \lim_{\epsilon \to 0} \left\{ \frac{\ln(\mathcal{N}(A, \epsilon))}{\ln(1/\epsilon)} \right\}$$

*exist, then $D$ is called the* fractal dimension *of $A$. We will also use the notation $D = D(A)$ and will say "$A$ has a fractal dimension $D$."*

We will visualize this definition better using the following example:

**Example 4.1.** Consider the square with the corners $(1, 1)$, $(1, -1)$, $(-1, -1)$, and $(-1, 1)$. We can cover this square using one ball with radius $\sqrt{2}$ with its center at $(0, 0)$. If we divide the radius by 2, we see that we can cover our square with four balls with the radius $\sqrt{2}/2$, their respective centers at $(1/2, 1/2)$, $(1/2, -1/2)$, $(-1/2, -1/2)$, and $(-1/2, 1/2)$. Doing this again, we get 16 balls with the radius $\sqrt{2}/4$. Continuing this we can see that, using closed balls with the radius $\frac{\sqrt{2}}{2^n}$, where $n = 0, 1, \ldots$, we need $4^n$ balls to cover the square. From the definition we get

$$D = \lim_{n \to \infty} \left\{ \frac{\ln\left(4^n\right)}{\ln\left(2^n/\sqrt{2}\right)} \right\} = 2.$$

In other words we see that the fractal dimension of a square is 2. This example indicates that the fractal dimension agrees with classical dimension for "nice" spaces.

This definition provide a somewhat tricky method of calculating the fractal dimension. Therefore we may want to consider alternatives to make this process easier.

## 4.2 Determining the Fractal Dimension

Depending on what type of fractals we are working with, we can use different methods when determining their fractal dimension. We will see that when it comes to fractal dimensions in nature, we will have to numerically estimate the answer, while some of our computer generated fractals we are able to theoretically determine their fractal dimension.

**Theorem 4.1.** *Let $\{\mathbb{R}^m; w_1, w_2, \ldots, w_N\}$ be a hyperbolic IFS, and let $A$ denote its attractor. Suppose $w_n$ is a similitude of scaling factor $r_n$ for each $n \in \{1, 2, 3, \ldots, N\}$. If the IFS is totally disconnected or just-touching then the attractor has fractal dimension $D(A)$, which is given by the unique solution of*

$$\sum_{n=1}^{N} |r_n|^{D(A)} = 1, \qquad D(A) \in [0,\, m].$$

*If the IFS is overlapping, then* $\overline{D} \geq D(A)$*, where* $\overline{D}$ *is the solution of*

$$\sum_{n=1}^{N} |r_n|^{\overline{D}} = 1, \qquad \overline{D} \in [0,\, \infty).$$

Moving forward, we will use the notation $D$ for $D(A)$. To see a sketch of proof for this theorem, please check Barnsley's book on page 183.

This is a method that only works when our IFS consists of similitudes. Out of our five previously generated fractals, three of these fit this criteria. Using this theorem we can calculate their fractal dimension. Doing this, we will quickly see that for two of these—the dragon curve and Koch snowflake—we get our dimension to be 2. How come we get non-fractional dimensions when these are supposedly fractals? The IFS in table 5 and table 4 are expressed for the *area* of the fractals, while these figure's *boundary* have a fractional dimension.

Now, let us have a closer look at our fractals' (or their boundaries') fractal dimensions!

### 4.2.1 Sierpinski Triangle

As for the Sierpinski triangle we can immediately conclude the scaling factor for all our functions to be 1/2, which—by using theorem 4.1—we get the following:

$$\sum_{n=1}^{3} r^D = 1$$

$$D = \frac{\ln(1/3)}{\ln(1/2)} \approx 1.585$$

### 4.2.2 Barnsley Fern

It is not possible to calculate the fractal dimension of the Barnsley fern using this method. With lack of a better source, according to the wiki page on *Math Images* it is only possible to estimate the fractal dimension, and when this has been estimated as about $D = 1.45$ [1].

### 4.2.3 Fractal Tree

For the first function $w_1$ in table 3 we will not be able to determine any scaling factor as this is not a similitude. However, this function simply works as a way to fill out the "tree-trunk." By removing it, we instead get the IFS for the branch tips as shown in figure 2. By ignoring $w_1$, we are able to determine the fractal dimension.

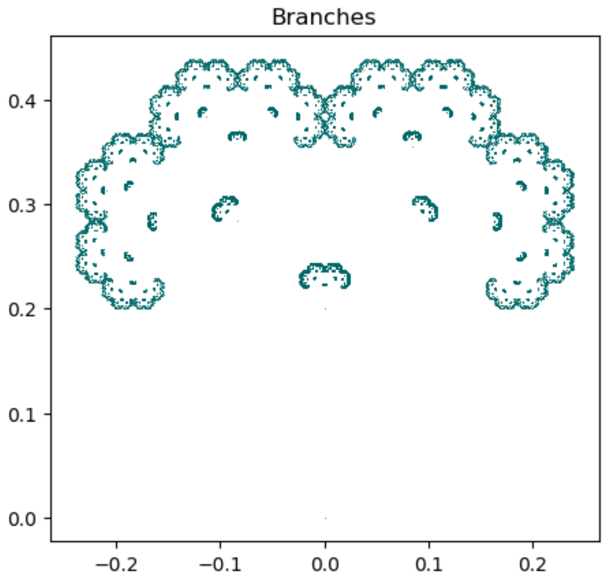



Figure 2: Generated image of the tips of the branches in a fractal tree using the random iteration algorithm. Figure consists of $100,000$ points of data. The IFS used can be found in table 3.

We can see that $w_2$ and $w_3$ have the same scaling factor, $r_2 = r_3 = 0.59$. As for $w_4$ we see that $r_4 = 0.1$. Solving $2 \cdot 0.59^D + 0.1^D = 1$ numerically we get that $D = 1.407$ [5].

#### 4.2.4 Koch Curve

For the Koch snowflake in figure 1d, we can divide the boundary into three smaller segments that are all (rotated) copies of themselves (see figure 3.) All of these segments have the same scaling/rotation matrices $A_i$, but differs in placement. Since we only need $A_i$ when calculating the fractal dimension with our methods, we know that the fractal dimension of the boundary of the whole Koch snowflake will be the same as the one of the segment. Using table 6, we get the IFS for the segment shown in figure 3.

| $w$ | $a$ | $b$ | $c$ | $d$ | $e$ | $f$ | $p$ |
|---|---|---|---|---|---|---|---|
| 1 | $1/3$ | 0 | 0 | $1/3$ | 0 | 0 | 0.25 |
| 2 | $1/6$ | $-\sqrt{3}/6$ | $\sqrt{3}/6$ | $1/6$ | $1/3$ | 0 | 0.25 |
| 3 | $1/6$ | $\sqrt{3}/6$ | $-\sqrt{3}/6$ | $1/6$ | $1/2$ | $\sqrt{3}/6$ | 0.25 |
| 4 | $1/3$ | 0 | 0 | $1/3$ | $2/3$ | 0 | 0.25 |

Table 6: IFS for the Koch curve segment.

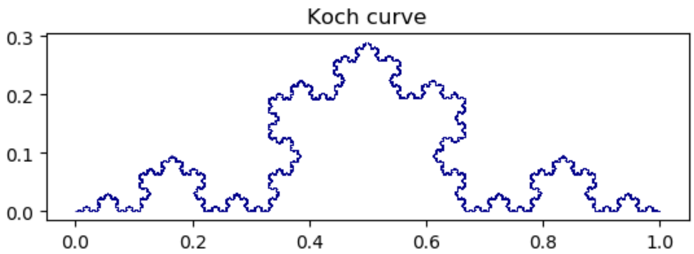



Figure 3: Generated image of part of the Koch snowflake's boundary using the random iteration algorithm. Figure consists of 100,000 points of data. The IFS used can be found in table 6.

From the IFS stated in table 6 applied to theorem 4.1 we get

$$\sum_{n=1}^{4}\left(\frac{1}{3}\right)^{D} = 4\left(\frac{1}{3}\right)^{D} = 1$$

Which gives us the dimension $D = \frac{\ln(4)}{\ln(3)} \approx 1.262$ [6].

#### 4.2.5 Dragon Curve

We have already concluded that based on the IFS stated in table 5 we get the fractal dimension equal to 2. If we want to consider the IFS for the boundary instead, we need to first break it down into four smaller pieces. The IFS from these are directly taken from Larry Riddle's website (see reference [7].) From these we can, similarly as we did with the Koch snowflake, consider only one of these four segments.

| $w$ | $a$ | $b$ | $c$ | $d$ | $e$ | $f$ | $p$ |
|---|---|---|---|---|---|---|---|
| 1 | $1/2$ | $-1/2$ | $1/2$ | $1/2$ | 0 | 0 | 0.5 |
| 2 | $-1/4$ | $-1/4$ | $1/4$ | $-1/4$ | $1/2$ | $1/2$ | 0.25 |
| 3 | $1/4$ | $1/4$ | $-1/4$ | $1/4$ | $1/2$ | $1/2$ | 0.25 |

Table 7: IFS for dragon curve's boundary segment.

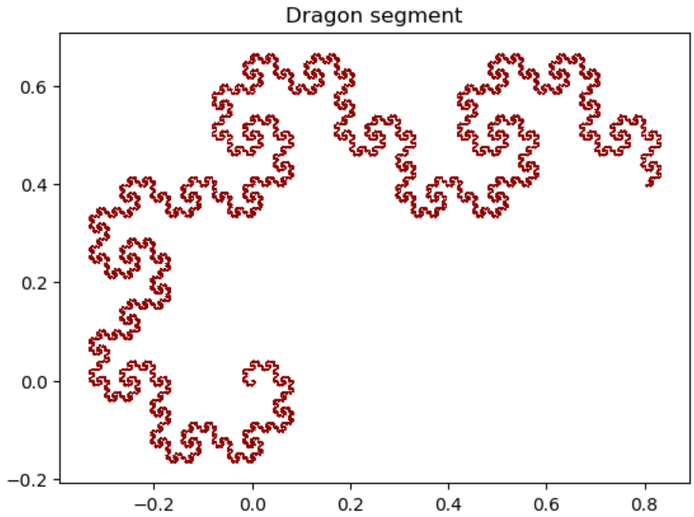



Figure 4: Generated image of part of the dragon curve's boundary using the random iteration algorithm. Figure consists of 100,000 points of data. The IFS used can be found in table 7.

Using the values in table 7 we find that the scaling factor for $w_1$ is $r_1 = \frac{1}{\sqrt{2}}$, while $r_2 = r_3 = \frac{1}{2\sqrt{2}}$. We get the expression

$$2\left(\frac{1}{2\sqrt{2}}\right)^D + \left(\frac{1}{\sqrt{2}}\right)^D = 1$$

Riddle solves this equation, but we can also find $D$ numerically to be $D = 1.524$.

# 5 Conclusion

We have introduced the concept of fractals and noted that by using a new definition of dimensions, as a mean to measure "roughness", we can measure fractals to have non-integer *fractal* dimensions.

We defined the complete metric space $(\mathcal{H}(\mathbf{X}), h)$, and in this used *hyperbolic iterated function systems* to construct fractals. We visualized the Sierpinski triangle, the Barnsley fern, a fractal tree, Koch snowflake, and the dragon curve, using the *random iteration algorithm*. We also determined the fractal dimensions of these to be 1.585, 1.45, 1.407, 1.262, and 1.524 respectively, where for the Koch snowflake and the dragon curve we specifically looked at the dimension of their boundaries. Though, it is unconfirmed whether the Barnsley fern's fractal dimension of 1.45 is a correct estimate.

## References


[1] Barnsley fern. `https://mathimages.swarthmore.edu/index.php/Barnsley_Fern`.

[2] M.F. Barnsley. *Fractals Everywhere*. Dover Books on Mathematics. Dover Publications, 2012.

[3] Claire David. On fractal properties of weierstrass-type functions. *Proceedings of the International Geometry Center*, 12, 10 2019.

[4] B.B. Mandelbrot, B. Mandelbrot, W. H. Freeman, and Company. *The Fractal Geometry of Nature*. Einaudi paperbacks. Henry Holt and Company, 1983.

[5] Larry Riddle. `https://larryriddle.agnesscott.org/ifs/pythagorean/symbinarytree.htm`.

[6] Larry Riddle. `https://larryriddle.agnesscott.org/ifs/kcurve/kcurve.htm`.

[7] Larry Riddle. `https://larryriddle.agnesscott.org/ifs/heighway/heighwayBoundary.htm`.