FRACTALS AS FIXED POINTS OF ITERATED FUNCTION SYSTEMS

CHRISTOPHER NATOLI

Abstract. This paper discusses one method of producing fractals, namely that of iterated function systems. We first establish the tools of Hausdorff measure and Hausdorff dimension to analyze fractals, as well as some concepts in the theory of metric spaces. The latter allows us to prove the existence and uniqueness of fractals as fixed points of iterated function systems. We discuss the connection between Hausdorff dimension and iterated function systems, and then study an application of fractals as unique fixed points in dynamical systems theory.

Contents

Introduction

Geometry is often concerned with smooth or regular sets and functions. Much can be said, however, about those sets that are not smooth or regular. Moreover, such objects can better represent many natural phenomena. We turn to fractal geometry as one means of illuminating non-smooth and irregular objects.

Fractals are sets that satisfy most if not all of the following properties:

- The set has a fine, intricate structure, with detail at arbitrarily small scales.
- It is too irregular to be described with classical geometry. Nonetheless, it has a simple, possibly recursive definition.
- It possesses some form of self-similarity, that is, it is composed of scaled copies of itself.
- Its dimension exceeds what we would intuitively think of as its dimension. Often its dimension is between integers. (We will explain this more rigorously later.)

Consider the Cantor set as an example of a fractal. Start with the segment $[0, 1]$ and remove the middle third, leaving you with segments $\left[0, \frac{1}{3}\right]$ and $\left[\frac{2}{3}, 1\right]$. Then remove the

Date: 26 August 2012.

middle third from each of those segment. Repeat this procedure indefinitely. The resulting object is called the Cantor set. See Figure 1 for a few iterations toward the Cantor set.

FIGURE 1. A few iterations toward the Cantor set, from [9].

Note how the Cantor set satisfies the conditions above to be a fractal. It is simply defined: it is constructed by just recursively removing the middle third of every segment. This definition gives rise to its self-similarity; a small section of the Cantor set is just a scaled version of the entire set. Indeed, this gives it fine structure at small scales. If you keep "zooming in" to the Cantor set, what you see is small versions of itself, with detail no matter how small the scale. Its dimension, as we will rigorously define in Section 1 and then compute, is $\frac{\log 2}{\log 3}$.

To better study fractals, we will need some tools from measure theory. In particular, we will use the Hausdorff measure, which begets the concept of Hausdorff dimension, to better understand the size of fractals. This will allow us to study fractals produced by one particular method, namely, iterated function systems. For every iterated function system, there exists a unique fixed point, which is a fractal. Proving this fact is a key part of this paper. To do so, we will need some tools in the theory of metric spaces, which we will establish before proving this theorem. We will then discuss how this theorem yields fractals and allows us to easily compute their dimension. We will also look into an application of this theorem and fractals in general to the theory of dynamical systems.

1. Hausdorff measure

Before we begin to analyze fractals, we must first introduce some concepts in measure theory. These tools will prove useful for studying the exotic nature of fractals. We will need to use a measure other than the usual Lebesgue measure. We recall the Lebesgue measure here:

Definition 1.1. Let $\{B_i\}$ be a countable collection of n-dimensional *boxes*, i.e., subsets of $\mathbb{R}^n$ such that $B_i = \{(x_1, \dots, x_n) \in \mathbb{R}^n : a_j \leq x_j \leq b_j \text{ for } a_j < b_j \text{ and } j = 1, \dots, n\}$. We will write the n-dimensional volume of B_i as $\mathrm{vol}(B_i) = (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n)$. Then the *Lebesgue measure* of a subset E of $\mathbb{R}^n$ is

$$\mathcal{L}^n(E) = \inf \left\{ \sum_i \mathrm{vol}(B_i) : E \subset \bigcup_i B_i \right\}.$$

Our intuitive understanding of length, area, and 3-dimensional volume is given by the Lebesgue measure, in particular, $\mathcal{L}^1$, $\mathcal{L}^2$, and $\mathcal{L}^3$, respectively. Note, however, that $\mathcal{L}^3$ cannot distinguish between a line and a plane: both have a 3-dimensional Lebesgue measure of 0. In general, Lebesgue measure in $\mathbb{R}^n$ provides no way to distinguish between subsets of $\mathbb{R}^n$ with dimension less than n. Indeed, we do not even *have* a rigorous definition of the dimension of a set (we have been relying on and will continue to rely on an intuitive understanding of dimension until we can construct a more precise one). A generalization of Lebesgue measure, known as Hausdorff measure, is needed to avoid both these limitations.

Definition 1.2. For any subset U of $\mathbb{R}^n$, the *diameter* of U is

$$\mathrm{diam}\, U = \sup \{|x - y| : x, y \in U\}.$$

Definition 1.3. Let E be a subset of $\mathbb{R}^n$ and let $\{U_i\}$ be a countable collection of sets that cover E such that $\operatorname{diam} U_i \leq \delta$ for some $\delta > 0$. For $s > 0$, the *s-dimensional Hausdorff measure* of E is

$$\mathcal{H}^s(E) = \lim_{\delta \to 0} \inf \left\{ \sum_{i=1}^{\infty} (\operatorname{diam} U_i)^s : E \subset \bigcup_{i=1}^{\infty} U_i \text{ and } \operatorname{diam} U_i \leq \delta \right\}.$$

This is similar to the Lebesgue measure. Both make use of a countable collection of sets that cover the set we wish to measure, and then calculate its measure from the "tightest" covering. Indeed, in $\mathbb{R}^n$, Lebesgue measure and n-dimensional Hausdorff measure agree (see Theorem 1.9 below). The Lebesgue measure, however, can only pick up n-dimensional volume, while the s-dimensional Hausdorff measure is sensitive to the volume of s-dimensional objects living in $\mathbb{R}^n$ (for $s < n$).

As δ decreases, potential coverings using large diameters are excluded from the set of possible coverings. Thus the infimum of this set is nondecreasing as δ decreases. This limit therefore exists, with $0 \leq \mathcal{H}^s(E) \leq \infty$.

To finish establishing the definition of Hausdorff measure, we will prove that it is an *outer* measure. (Note that it is a fact from measure theory that an outer measure can become a measure if restricted to measurable sets.)

Definition 1.4. A function μ is an *outer measure* if it is a nonnegative function defined on all subsets of $\mathbb{R}^n$ such that

- $\mu(\varnothing) = 0$,
- monotonicity: $\mu(A) \leq \mu(B)$ if $A \subseteq B$, and
- countable subadditivity: $\mu(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu(A_i)$ for a countable (or finite) sequence of sets $\{A_i\}$.

Theorem 1.5. *The Hausdorff measure is an outer measure.*

Proof. For the first condition, since the empty set is covered by any single set, the infimum of these sets has diameter 0. Hence the Hausdorff measure of the empty set is 0. For the second condition, note that every cover of B also covers A. Thus, the tightest cover of B *at least* covers A, i.e., $\mathcal{H}^s(B) \geq \mathcal{H}^s(A)$.

For the third condition, given A_i and δ, choose a set of coverings $\{U_{i,j}\}$ such that $A_i \subset \bigcup_{j=1}^{\infty} U_{i,j}$, $\operatorname{diam} U_{i,j} \leq \delta$, and

$$\sum_{j=1}^{\infty} (\operatorname{diam} U_{i,j})^s \leq \mathcal{H}^s(A_i) + \frac{\varepsilon}{2^i}$$

for $\varepsilon > 0$. Then $\bigcup_{i=1}^{\infty} A_i$ is covered by all $U_{i,j}$, i.e., $\bigcup_{i=1}^{\infty} A_i \subset \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} U_{i,j}$. Certainly,

$$\inf_{\substack{\bigcup_{i=1}^{\infty} A_i \subset \bigcup_{i=1}^{\infty} \bigcup_{j=1}^{\infty} U_{i,j} \\ \operatorname{diam} U_{i,j} \leq \delta}} \left\{ \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (\operatorname{diam} U_{i,j})^s \right\} \leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (\operatorname{diam} U_{i,j})^s \leq \sum_{i=1}^{\infty} \left(\mathcal{H}^s(A_i) + \frac{\varepsilon}{2^i} \right).$$

If we take the limit as δ approaches 0, we have the Hausdorff measure on the left side:

$$\mathcal{H}^s \left(\bigcup_{i=1}^{\infty} A_i \right) \leq \sum_{i=1}^{\infty} \left(\mathcal{H}^s(A_i) + \frac{\varepsilon}{2^i} \right) = \sum_{i=1}^{\infty} \mathcal{H}^s(A_i) + \varepsilon.$$

□

In fact, we can say even more about countable subadditivity. Since we will be working only with Borel sets (sets constructed by countable union, countable intersection, and complementation of open and closed sets), all our sets are *measurable* for Hausdorff measure. While the details of measurability do not concern us here, we will note that for disjoint Borel sets, we have equality in countable subadditivity, i.e., $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$.

With the definition of Hausdorff measure sufficiently established, we will proceed to study its properties and behavior. Indeed, it has many nice properties. For instance, Hausdorff measure scales as expected:

Theorem 1.6. *Let S be a similarity, i.e., a mapping $S : \mathbb{R}^n \to \mathbb{R}^n$ such that $|S(x) - S(y)| = \lambda|x - y|$ for all x, y in $\mathbb{R}^n$, with $\lambda > 0$. For a subset E of $\mathbb{R}^n$,*

$$\mathcal{H}^s(S(E)) = \lambda^s \mathcal{H}^s(E).$$

Proof. For any cover $\{U_i\}$ of E with $\operatorname{diam} U_i \leq \delta$, the set $S(U_i)$ covers $S(E)$, and $\operatorname{diam} S(U_i) \leq \lambda\delta$. Then

$$\inf_{\substack{S(E)\subset\bigcup_i S(U_i) \\ \operatorname{diam} S(U_i)\leq\lambda\delta}} \left\{ \sum_i (\operatorname{diam} S(U_i))^s \right\} \leq \sum_i (\operatorname{diam} S(U_i))^s = \lambda^s \sum_i (\operatorname{diam} U_i)^s$$

and so

$$\inf_{\substack{S(E)\subset\bigcup_i S(U_i) \\ \operatorname{diam} S(U_i)\leq\lambda\delta}} \left\{ \sum_i (\operatorname{diam} S(U_i))^s \right\} \leq \inf_{\substack{E\subset\bigcup_i U_i \\ \operatorname{diam} U_i\leq\delta}} \left\{ \lambda^s \sum_i (\operatorname{diam} U_i)^s \right\}.$$

Letting δ approach 0, we have $\mathcal{H}^s(S(E)) \leq \lambda^s \mathcal{H}^s(E)$. Repeating the proof after switching S with S^{-1}, λ with $\frac{1}{\lambda}$, and E with $S(E)$, we get $\mathcal{H}^s(E) \leq \frac{1}{\lambda^s}\mathcal{H}^s(S(E))$, or $\mathcal{H}^s(S(E)) \geq \lambda^s \mathcal{H}^s(E)$. ☐

Another nice property of Hausdorff measure is that for integral dimension, it agrees with Lebesgue measure. That is to say, for an integer n, the n-dimensional Hausdorff measure is equal to a scaling of the n-dimensional Lebesgue measure. Specifically, $\mathcal{L}^n(E) = 2^{-n}\alpha_n \mathcal{H}^n(E)$, where α_n is the volume of the unit n-ball.

To prove this, we will need the following two facts:

Lemma 1.7 (Isodiametric Inequality)**.** *Given a subset E of $\mathbb{R}^n$ with diameter d, the volume of E is at most the volume of a ball with diameter d, i.e., $\mathcal{L}^n(E) \leq \alpha_n \left(\frac{1}{2}d\right)^n$.*

Lemma 1.8 (Besicovitch Covering Lemma)**.** *Let μ be a measure on $\mathbb{R}^n$ and E be a subset of $\mathbb{R}^n$ such that $\mu(E) < \infty$. Also, let $\mathcal{C}$ be a collection of nontrivial closed balls with radius r and center x with* $\inf\{r :$ *ball in $\mathcal{C}$ centered at x*$\} = 0$ *for all $x \in E$. Then there exists a countable disjoint subcollection of $\mathcal{C}$ that covers μ almost all of E.*

We will leave these facts unproven and proceed to prove that for integral dimensions, the Lebesgue measure is just a rescaling of the Hausdorff measure.

Theorem 1.9. *For any set E in $\mathbb{R}^n$, we have $\mathcal{L}^n(E) = 2^{-n}\alpha_n \mathcal{H}^n(E)$.*

Proof. Choose a covering $\{U_i\}$ of E such that, for $\varepsilon > 0$, $\sum_i (\operatorname{diam} U_i)^n \leq \mathcal{H}^n(E) + \varepsilon$. By Lemma 1.7,

$$\mathcal{L}^n(U_i) \leq \alpha_n \left(\frac{1}{2}\operatorname{diam} U_i\right)^n$$
$$\mathcal{L}^n(E) \leq \sum_i \mathcal{L}^n(U_i) \leq \sum_i 2^{-n}\alpha_n (\operatorname{diam} U_i)^n \leq 2^{-n}\alpha_n \left(\mathcal{H}^n(E) + \varepsilon\right)$$
$$\mathcal{L}^n(E) \leq 2^{-n}\alpha_n \mathcal{H}^n(E).$$

We will split the converse into two cases. If $\mathcal{H}^n(E) = \infty$, then $\sum_i (\operatorname{diam} U_i)^n = \infty$ for all coverings $\{U_i\}$. In particular, for any covering by boxes U_i with $\operatorname{diam} U_i \leq \delta$, we have $\sum_i (\operatorname{diam} U_i)^n = \infty$. Note that $\operatorname{vol}(U_i)$ is just a nonzero scaling of $(\operatorname{diam} U_i)^n$. Therefore,

$$\mathcal{L}^n(E) = \inf_{E \subset \bigcup_i U_i} \left\{ \sum_i \operatorname{vol}(U_i) \right\} = \inf_{E \subset \bigcup_i U_i} \left\{ \sum_i k_i (\operatorname{diam} U_i)^n \right\} = \infty,$$

where $k_i \neq 0$ are constants.

In the final case, we have $\mathcal{H}^n(E) < \infty$. Given $\varepsilon > 0$, choose $\delta > 0$ such that

$$\mathcal{H}^n(E) \leq \inf \left\{ \sum_i (\operatorname{diam} U_i)^n : E \subset \bigcup_i U_i \text{ and } \operatorname{diam} U_i \leq \delta \right\} + \frac{2^n}{\alpha_n}\varepsilon. \tag{1.10}$$

Let $\mathcal{C}$ be a covering of E by closed balls contained in E with diameter at most δ. Apply Lemma 1.8, yielding a countable disjoint subcollection $\mathcal{D}$ of $\mathcal{C}$, which covers $\mathcal{H}^n$ almost all of E. Let F be the subet of E that is covered by $\mathcal{D}$, such that $\mathcal{H}^n(E \setminus F) = 0$. Also, let $\mathcal{D}'$ be a covering of $E \setminus F$ by balls with diameter at most δ, such that

$$\sum_{U \in \mathcal{D}'} \alpha_n \left(\frac{1}{2}\operatorname{diam} U\right)^n \leq \varepsilon.$$

Then $\mathcal{D} \cup \mathcal{D}'$ covers E, with $\operatorname{diam} U \leq \delta$ for all $U \in \mathcal{D} \cup \mathcal{D}'$. From equation (1.10),

$$\begin{aligned}
\mathcal{H}^n(E) &\leq \sum_{U \in \mathcal{D} \cup \mathcal{D}'} (\operatorname{diam} U)^n + \frac{2^n}{\alpha_n}\varepsilon \\
2^{-n}\alpha_n \mathcal{H}^n(E) &\leq \sum_{U \in \mathcal{D} \cup \mathcal{D}'} \alpha_n \left(\frac{1}{2}\operatorname{diam} U\right)^n + \varepsilon \\
&= \sum_{U \in \mathcal{D}} \mathcal{L}^n(U) + \sum_{U \in \mathcal{D}'} \alpha_n \left(\frac{1}{2}\operatorname{diam} U\right)^n + \varepsilon \\
&\leq \mathcal{L}^n(E) + \varepsilon + \varepsilon.
\end{aligned}$$

(Note that the equality above follows from that fact that each U is a ball, and consequently $\alpha_n \left(\frac{1}{2}\operatorname{diam} U\right)^n = \mathcal{L}^n(U)$.) □

This theorem gives us some sense of what the Hausdorff measure may evaluate to, at least for integral dimension. Like the Lebesgue measure, it may evaluate to 0 for sets whose dimension is too low (e.g., the 2-dimensional measure of a line is 0) or infinity for sets whose dimension is too high (e.g., the 2-dimensional measure of a cube is infinity). For sets of just the right dimension, the measure becomes meaningful.

We can demonstrate this more rigorously. Let $\{U_i\}$ be a covering of a subset E of $\mathbb{R}^n$ with $\operatorname{diam} U_i \le \delta$, and let $t > s$. Then

$$\sum_i (\operatorname{diam} U_i)^t = \sum_i (\operatorname{diam} U_i)^{t-s} (\operatorname{diam} U_i)^s \le \sum_i \delta^{t-s} (\operatorname{diam} U_i)^s.$$

Taking infima and letting δ approach 0, we find that $\mathcal{H}^t(E) = 0$ if $\mathcal{H}^s(E) < \infty$. Thus there is a critical value s such that $\mathcal{H}^t(E) = 0$ if $t > s$ and $\mathcal{H}^t(E) = \infty$ if $t < s$. We call this value the dimension of E:

Definition 1.11. The *Hausdorff dimension* of a subset E of $\mathbb{R}^n$ is

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

Note that if $s = \dim_{\mathrm{H}} E$, then $\mathcal{H}^s(E)$ may be zero, infinite, or some positive real number in-between.

As an example, we will calculate the Hausdorff dimension of the Cantor set C. Its self-similarity gives us a quick way to find its dimension, although rigor will have to be suspended until we find this value. Let C_L be the left part of C, where $C_L = C \cap [0, \frac{1}{3}]$, and let $C_R = C \cap [\frac{2}{3}, 1]$ be the right part. Note that C_L and C_R are just scalings of C by the ratio $\frac{1}{3}$, and that C is a disjoint union of C_L and C_R. Therefore,

$$\mathcal{H}^s(C) = \mathcal{H}^s(C_L) + \mathcal{H}^s(C_R) = \left(\frac{1}{3}\right)^s \mathcal{H}^s(C) + \left(\frac{1}{3}\right)^s \mathcal{H}^s(C),$$

by Theorem 1.6. Assuming $0 < \mathcal{H}^s(C) < \infty$ (which will be proven shortly), we have $1 = 2\left(\frac{1}{3}\right)^s$ and so $s = \frac{\log 2}{\log 3}$.

We must now show that $\mathcal{H}^s(C)$ is indeed a positive real number for $s = \frac{\log 2}{\log 3}$. Define E_k as the kth iteration toward the Cantor set. Note that it contains 2^k closed intervals of length 3^{-k}. Then E_k covers C and

$$\mathcal{H}^s(C) \le 2^k 3^{-ks} = 2^k 3^{-k\frac{\log 2}{\log 3}} = 1.$$

To get a lower bound, let $\{U_i\}$ be any covering by intervals. Expand each U_i slightly so that it makes an open cover of C; by the compactness of C, there exists a finite subcover of C. Thus we can assume that $\{U_i\}$ is a finite collection of intervals. Define an integer k for each U_i such that $3^{-(k+1)} \le \operatorname{diam} U_i < 3^{-k}$ Since each interval of E_k is separated by at least 3^{-k}, U_i can intersect only one interval of E_k. For $j \ge k$, U_i can intersect at most

$$2^{j-k} = 2^j 3^{-ks} = 2^j 3^s 3^{-s(k+1)} \le 2^j 3^s (\operatorname{diam} U_i)^{-s}$$

intervals of E_j. Choose j large enough so that $3^{-(j+1)} \le \operatorname{diam} U_i$ for all U_i. Since $\{U_i\}$ intersects all 2^j intervals, and since each U_i intersects at most $2^j 3^s (\operatorname{diam} U_i)^{-s}$ intervals, we have

$$\sum_i 2^j 3^s (\operatorname{diam} U_i)^{-s} \ge 2^j$$

$$\sum_i (\operatorname{diam} U_i)^{-s} \ge 3^{-s}$$

$$\mathcal{H}^s(C) \ge \frac{1}{2}.$$

Hence, our assumption earlier that $0 < \mathcal{H}^s(C) < \infty$ was safe, and the dimension of the Cantor set is indeed $\dim_{\mathrm{H}} C = s = \frac{\log 2}{\log 3}$. Note that we have also given some rough bounds for the s-dimensional Hausdorff measure of C, namely, $\frac{1}{2} \leq \mathcal{H}^s(C) \leq 1$. It can be shown that $\mathcal{H}^s(C) = 1$ (a proof of a generalization of this fact can be found in [6]).

2. Metric spaces

We must make another detour before analyzing fractals. To guarantee the existence and uniqueness of fractals as produced by contracting similarities, we will need some tools from the theory of metric spaces. This will allow us to more rigorously define contractions, and it will give us the necessary fixed point theorem.

Definition 2.1. A pair (X, d), where $d : X \times X \to \mathbb{R}$, is a *metric space* if d satisfies

- $d(x, y) \geq 0$, with equality only when $x = y$,
- $d(x, y) = d(y, x)$, and
- $d(x, y) \leq d(x, z) + d(z, y)$ for all z in X.

Then d is called a *metric*. If every Cauchy sequence in X converges, X is a *complete* metric space.

Let D be a closed subset of $\mathbb{R}^n$. For a subset E of D, write $N_\delta(E)$ for the δ-neighborhood of E, i.e.,

$$N_\delta(E) = \{x \in D : |x - a| < \delta \text{ for some } a \in E\}.$$

Then let Ω be the set of all nonempty compact subsets of D. We associate the following metric to Ω:

Definition 2.2. The *Hausdorff metric* on Ω for any $A, B \in \Omega$ is

$$d_{\mathrm{H}}(A, B) = \inf \{\delta : A \subset N_\delta(B) \text{ and } B \subset N_\delta(A)\}.$$

Theorem 2.3. *The set Ω associated with the Hausdorff metric d_{H} is a complete metric space.*

Proof. The nonnegativity of the Hausdorff metric follows from its definition. Also, let ℓ be the least distance between any point in A and any point in B, and note that $\operatorname{diam} A, \operatorname{diam} B < \infty$ since A and B are bounded. Then certainly $A \subset N_{\operatorname{diam} A + \ell + \operatorname{diam} B}(B)$ and $B \subset N_{\operatorname{diam} A + \ell + \operatorname{diam} B}(A)$, so $d_{\mathrm{H}}(A, B) \leq \operatorname{diam} A + \ell + \operatorname{diam} B < \infty$.

For A, B elements of Ω, suppose $d_{\mathrm{H}}(A, B) = 0$ but $A \neq B$, i.e., there exists an x in $\mathbb{R}^n$ such that x is in A but not B, or x is in B but not A. Without loss of generality, choose the former. If x is a distance $\varepsilon > 0$ away from B, then an $\frac{\varepsilon}{2}$-neighborhood of B does not contain x and therefore does not contain A. So $d_{\mathrm{H}}(A, B) > \frac{\varepsilon}{2}$. Contradiction. If, on the other hand, x is not a positive distance away from B, then there exists b_n in B such that $|x - b_n| < \frac{1}{n}$. As n tends to infinity, b_n approaches x, and thus x is in the closure of B. Since B is closed, x is in B. Contradiction.

The symmetry of d_{H} follows immediately from the symmetry in its definition.

Given A, B, C elements of Ω, let a be an element of A. Then there exists some b in B such that $|a - b| < d_{\mathrm{H}}(A, B)$. Similarly, there exists some c in C such that, for the b given above, $|b - c| < d_{\mathrm{H}}(B, C)$. Adding the two and applying the triangle equality for Euclidean distance, we find that for all a, there exists some c such that

$$|a - c| < d_{\mathrm{H}}(A, B) + d_{\mathrm{H}}(B, C).$$

Thus, $A \subset N_{d_{\mathrm{H}}(A,B) + d_{\mathrm{H}}(B,C)}(C)$ by definition. Following the same argument but reversing the order, we have $C \subset N_{d_{\mathrm{H}}(A,B) + d_{\mathrm{H}}(B,C)}(A)$. Therefore, $d_{\mathrm{H}}(A, C) \leq d_{\mathrm{H}}(A, B) + d_{\mathrm{H}}(B, C)$, and so Ω is a metric space.

To prove that it is complete, let a sequence $\{A_n\}$ in Ω be a Cauchy sequence with respect to the Hausdorff metric. Select a subsequence $\{A_k\}$ of $\{A_n\}$ such that $d_{\mathrm{H}}(A_k, A_{k+1}) \leq 2^{-k}$. Then there exists a sequence $\{x_k\}$ in D, where x_k is in A_k, such that

$$|x_k - x_{k+1}| < 2^{-k}.$$

Note that $\{x_k\}$ is a Cauchy sequence, and since every Cauchy sequence in $\mathbb{R}^n$ converges, $\{x_k\}$ converges to some point x in D (since D is closed). By the triangle inequality, we have

$$|x_k - x| \leq |x_k - x_{k+1}| + |x_{k+1} - x_{k+2}| + \cdots < 2^{-k} + 2^{-k-1} + \cdots = 2^{-k+1} \tag{2.4}$$

Let A be the set of all x as defined above, i.e., all limits of sequences $\{x_k\}$ such that x_k is in A_k and $|x_k - x_{k+1}| < 2^{-k}$. Note that A is nonempty. Also, for any x in A, it follows from equation (2.4) that there is some x_k in A_k such that $|x_k - x| < 2^{-k+1}$. Thus, $A \subset N_{2^{-k+1}}(A_k)$, and so A is bounded. If we let $\overline{A}$ be the closure of A, then $\overline{A}$ is nonempty and compact, and so $\overline{A}$ is an element of Ω. Certainly, we also have that $\overline{A} \subset N_{2^{-k+1}}(A_k)$. To achieve the converse, note that it also follows from equation (2.4) that for any x_k in A_k, there exists an x in A (and thus in $\overline{A}$, too) such that $|x_k - x| < 2^{-k+1}$. Then we have $A_k \subset N_{2^{-k+1}}(\overline{A})$, and so $d_{\mathrm{H}}(A_k, \overline{A}) < 2^{-k+1}$. Therefore $\{A_k\}$ converges to $\overline{A}$.

Since the subsequence $\{A_k\}$ of the Cauchy sequence $\{A_n\}$ converges to a point, it follows that the whole sequence $\{A_n\}$ also converges to the same point (proving this fact in general is a standard exercise in the theory of metric spaces, so we will leave it unproven here). □

We have now established the metric space Ω that we intend to work with. To better study and make use of Ω, we need one more tool in the theory of metric spaces, namely a theorem guaranteeing us fixed points under a contraction.

Definition 2.5. Given a metric space X with metric d, a mapping $S : X \to X$ is a *contraction* if

$$d(S(x), S(y)) \leq c\, d(x, y)$$

for some $c < 1$ and for all $x, y \in X$. If equality holds everywhere, we call S a *contracting similarity.*

Theorem 2.6 (Banach's Contraction Mapping Theorem)**.** *If X is a complete metric space, and if S is a contraction of X into X, then there exists one and only one x in X such that $S(x) = (x)$.*

Proof. Let x_0 be an arbitrary point in X. Recursively define the sequence $\{x_n\}$ such that $x_{n+1} = S(x_n)$. Then

$$d(x_{n+1}, x_n) = d(S(x_n), S(x_{n-1})) \leq c\, d(x_n, x_{n-1}) \leq c^n d(x_1, x_0).$$

For $n < m$, it follows from the triangle inequality that

$$\begin{aligned} d(x_n, x_m) &\leq \sum_{i=n+1}^{m} d(x_i, x_{i-1}) \\ &\leq \sum_{i=n+1}^{m} c^{i-1} d(x_1, x_0) \\ &\leq \frac{c^n}{1-c} d(x_1, x_0). \end{aligned}$$

Since $c < 1$, $d(x_n, x_m)$ approaches 0 for large n, and so $\{x_n\}$ is a Cauchy sequence. Since X is complete, $\{x_n\}$ converges to some $x = \lim_{n\to\infty} x_n$ in X.

Note that S is continuous since it is a contraction. Therefore,

$$S(x) = S\left(\lim_{n\to\infty} x_n\right) = \lim_{n\to\infty} S(x_n) = \lim_{n\to\infty} x_{n+1} = x,$$

and so x is indeed a fixed point.

To prove the uniqueness of x, suppose that there exists a second fixed point y. Note that $d(S(x), S(y)) \le c\, d(x, y)$. Since $S(x) = x$ and $S(y) = y$, we have $d(x, y) \le c\, d(x, y)$. This implies that $d(x, y) = 0$, or $x = y$. □

3. Iterated function systems

The fractals we are interested in studying are all self-similar, i.e., they are made up of several contracted copies of themselves. Clearly, fractals are fixed points under these contractions. We will use the concept of iterated function systems, under which fixed points are called attractors, to formalize this method of constructing fractals.

Definition 3.1. An *iterated function system* (abbreviated "IFS") is a collection of contractions $\{S_1, S_2, \ldots, S_m\}$, with $m \ge 2$, on a closed subset D of $\mathbb{R}^n$. A nonempty compact subset F of D is an *attractor* of the IFS if

$$F = \bigcup_{i=1}^{m} S_i(F).$$

Let us consider the Cantor set C as an example of a fractal that can be described through an IFS. Let D be the interval $[0, 1]$ and define the IFS as the contractions

$$S_1(x) = \frac{1}{3}x \qquad \text{and} \qquad S_2(x) = 1 - \frac{1}{3}x. \tag{3.2}$$

Note that $S_1(C)$ produces the left half of the Cantor set and $S_2(C)$ produces the right half. Thus their union is the Cantor set and, indeed, the Cantor set is the attractor of this IFS.

An important property of IFSs is that each IFS has a unique attractor. To prove this, we will apply Theorem 2.6 to the contraction $S(\cdot) = \bigcup_{i=1}^{m} S_i(\cdot)$ on the complete metric space Ω.

Theorem 3.3. *Given an IFS defined by contractions $\{S_1, \ldots, S_m\}$ on a closed subset D of $\mathbb{R}^n$, let Ω be the set of all nonempty compact subsets of D, and for an element E of Ω, let $S(E) = \bigcup_{i=1}^{m} S_i(E)$. Then there exists a unique attractor F of the IFS, i.e., a unique set F such that $S(F) = F$. Furthermore, for every set E in Ω such that $S(E) \subset E$, we have*

$$F = \bigcap_{k=0}^{\infty} S^k(E),$$

where $S^k(E) = \underbrace{S(S(\cdots(S}_{k \text{ times}}(E))))$ is the kth iterate of S.

Proof. Let d_{H} be the Hausdorff metric (see Definition 2.2). Note that if $S_i(A) \subset N_\delta(S_i(B))$ for all i, then

$$\bigcup_{i=1}^{m} S_i(A) \subset \bigcup_{i=1}^{m} N_\delta(S_i(B)) = N_\delta\left(\bigcup_{i=1}^{m} S_i(B)\right).$$

(The same holds if you swap A with B.) Thus, if δ is such that $d_{\mathrm{H}}(S_i(A), S_i(B)) < \delta$ for all i, then $d_{\mathrm{H}}\left(\bigcup_{i=1}^m S_i(A), \bigcup_{i=1}^m S_i(B)\right) < \delta$. Therefore,

$$d_{\mathrm{H}}\left(\bigcup_{i=1}^m S_i(A), \bigcup_{i=1}^m S_i(B)\right) \le \max_{1\le i\le m} d_{\mathrm{H}}(S_i(A), S_i(B))$$

$$\le \left(\max_{1\le i\le m} c_i\right) d_{\mathrm{H}}(A, B), \tag{3.4}$$

since S_i is a contraction. Thus we have $d_{\mathrm{H}}(S(A), S(B)) \le (\max_i c_i)\, d_{\mathrm{H}}(A, B)$, and since $\max_i c_i < 1$, then S is also a contraction on Ω. From Theorem 2.3, we know that Ω is a complete metric space, and so by Theorem 2.6, S has a unique fixed point F in Ω.

It follows from the proof of Theorem 2.6 that $S^k(E)$ approaches F as k tends to infinity. Note that if $S(E) \subset E$, then $\left\{S^k(E)\right\}$ is a decreasing sequence of nonempty compact sets containing F. Therefore, their intersection $\bigcap_{i=1}^\infty S(E) = F$. □

This theorem holds for all contractions, not only contracting similarities. But if we restrict our focus to similarities, we can say a lot more about the attractors guaranteed by this theorem. In particular, we can easily calculate its Hausdorff dimension. We have seen an example of calculating the dimension of self-similar fractals at the end of Section 1, for the Cantor set. It was key to this calculation that the Cantor set is made up of two *disjoint* scaled copies of itself. However, we can slightly relax this condition and only require the following.

Definition 3.5. An IFS of similarity contractions $\{S_1, \ldots, S_m\}$ satisfies the *open set condition* if there exists a nonempty bounded open set V such that

$$V \supset \bigcup_{i=1}^m S_i(V),$$

with this union disjoint.

Note that the IFS that produces the Cantor set (see equations (3.2)) satisfies the open set condition, with $V = (0, 1)$.

Before we can prove a formula to calculate the dimension of self-similar sets, we will need the following lemma.

Lemma 3.6. *Let $\{V_i\}$ be a collection of disjoint open subsets of $\mathbb{R}^n$ such that each V_i contains a ball of radius $a_1 r$ and is contained in a ball of radius $a_2 r$. Then any ball B of radius r intersects at most $(1+2a_2)^n a_1^{-n}$ of the closures $\overline{V_i}$.*

Proof. If V_i intersects B, then $\overline{V_i}$ is certainly contained in a ball of radius

$$r + \operatorname{diam} V_i \le (1 + 2a_2)r$$

centered at the center of B. Suppose that q of the sets $\overline{V_i}$ intersect B. Then a single ball of radius $(1+2a_2)r$ contains all q sets $\overline{V_i}$ and thus contains q balls of radius $a_1 r$. Adding up the volumes of these q interior balls, we have

$$q(a_1 r)^n \le (1+2a_2)^n r^n$$
$$q \le (1+2a_2)^n a_1^{-n}.$$

□

We will also need the following lemma about mass distributions.

Definition 3.7. A measure μ on a set X is a *mass distribution* if $0 < \mu(X) < \infty$. For any subset A of X, we call $\mu(A)$ the *mass* of A.

Lemma 3.8 (Mass distribution principle)**.** *Let μ be a mass distribution on F such that for some s, there exists $c, \delta > 0$ such that*

$$\mu(U) \leq c(\operatorname{diam} U)^s \tag{3.9}$$

for all sets U with $\operatorname{diam} U \leq \delta$. Then $\mathcal{H}^s(F) \geq \mu(F)/c$ and $s \leq \dim_{\mathrm{H}} F$.

Proof. Let $\{U_i\}$ be a cover of F with $\operatorname{diam} U_i \leq \delta$. Then, by the properties of measures and equation (3.9),

$$\mu(F) \leq \mu\left(\bigcup_i U_i\right) \leq \sum_i \mu(U_i) \leq c \sum_i (\operatorname{diam} U_i)^s.$$

Taking the infimum and letting δ tend to 0, we have $\mathcal{H}^s(F) \geq \mu(F)/c$. Since $\mu(F) > 0$, we have $\dim_{\mathrm{H}} F \geq s$. □

We can now prove a formula to calculate the dimension of self-similar sets defined by IFSs. This will be similar to our calculation of the dimension of the Cantor set, in that we must determine lower and upper bounds for the Hausdorff measure of such a set.

Theorem 3.10. *Suppose an IFS of contracting similarities $\{S_1, \ldots, S_m\}$ satisfies the open set condition, and let c_i be the ratio of each similarity S_i. Let F be the attractor of this IFS. Then $\dim_{\mathrm{H}} F = s$, where*

$$\sum_{i=1}^{m} c_i^s = 1. \tag{3.11}$$

Also, we have $0 < \mathcal{H}^s(F) < \infty$.

Proof. Define s according to equation (3.11). Let $\mathcal{I}_k$ be the set of all sequences $(i_1, \ldots, i_k)$ where $i_j \in \{1, \ldots, m\}$. Also, for any set A and a given sequence $(i_1, \ldots, i_k)$ in $\mathcal{I}_k$, define $A_{i_1,\ldots,i_k} = S_{i_1} \circ \cdots \circ S_{i_k}(A)$. Then since F is a fixed point under the IFS, we have

$$F = \bigcup_{\mathcal{I}_k} F_{i_1,\ldots,i_k}.$$

Thus the sets $\{F_{i_1,\ldots,i_k}\}$ cover F. Note that the mapping $S_i \circ \cdots \circ S_k$ is a contracting similarity with ratio $c_{i_1} \cdots c_{i_k}$. Therefore,

$$\begin{aligned}\sum_{\mathcal{I}_k} (\operatorname{diam} F_{i_1,\ldots,i_k})^s &= \sum_{\mathcal{I}_k} (c_{i_1} \cdots c_{i_k})^s (\operatorname{diam} F)^s \\ &= \left(\sum c_{i_1}^s\right) \cdots \left(\sum c_{i_k}^s\right) (\operatorname{diam} F)^s \\ &= (\operatorname{diam} F)^s \qquad (3.12)\end{aligned}$$

by equation (3.11). For all $\delta > 0$, we can choose k large enough that $(\max_i c_i)^k \operatorname{diam} F \leq \delta$. Since $\operatorname{diam} F_{i_1,\ldots,i_k} \leq (\max_i c_i)^k \operatorname{diam} F$, the sets $\{F_{i_1,\ldots,i_k}\}$ cover F with $\operatorname{diam} F_{i_1,\ldots,i_k} \leq \delta$ for all sequences $(i_i, \ldots, i_k)$ in $\mathcal{I}_k$. Therefore, $\mathcal{H}^s(F) \leq \sum_{\mathcal{I}_k} (\operatorname{diam} F_{i_1,\ldots,i_k})^s = (\operatorname{diam} F)^s$ by equation (3.12).

We will now try to determine a lower bound for $\mathcal{H}^s(F)$. Let $\mathcal{I}$ be the set of all infinite sequences $(i_1, i_2 \ldots)$ with $i_j \in \{1, \ldots, m\}$, and let $I_{i_1,\ldots,i_k}$ be the subset of $\mathcal{I}$ consisting only of sequences starting with $i_1, \ldots, i_k$. Then define a mass distribution μ on $\mathcal{I}$ letting

$\mu(I_{i_i,\dots,i_k}) = (c_{i_1}\cdots c_{i_k})^s$. Proving that μ is indeed a measure follows immediately from $\sum_{i=1}^m \mu(I_{i_i,\dots,i_k,i}) = \mu(I_{i_i,\dots,i_k})$. It also follows from equation (3.11) that $\mu(\mathcal{I}) = 1$, and so μ is a mass distribution.

Define a point $x_{i_1,i_2,\dots}$ in F as $x_{i_1,i_2,\dots} = \bigcap_{k=1}^\infty F_{i_1,\dots,i_k}$. Then for a subset A in F, define a new mass distribution $\tilde{\mu}$ based on μ by letting $\tilde{\mu}(A) = \mu(\{I_{i_1,\dots,i_k} : x_{i_1,\dots,i_k,\dots} \in A\})$. Intuitively, the $\tilde{\mu}$-mass of a set is the μ-mass of the sequences that define the points in that set. Note that $\tilde{\mu}(F) = \mu(\mathcal{I}) = 1$ since F contains every point $x_{i_1,i_2,\dots}$ for each $(i_1, i_2, \dots)$ in $\mathcal{I}$.

We now want to show that $\tilde{\mu}$ satisfies the conditions of Lemma 3.8, which will provide us with the desired lower bound. Let V be the open set guaranteed by the open set condition (see Definition 3.5), i.e., $V \supset \bigcup_{i=1}^m S_i(V)$ with this union disjoint. Letting $S(\overline{V}) = \bigcup_{i=1}^m S_i(\overline{V})$, it follows from Theorem 3.3 that $\left\{S^k(\overline{V})\right\}$ converges to F. Therefore, $\overline{V}_{i_1,\dots,i_k} \supset F_{i_1,\dots,i_k}$ for any sequence $(i_1, \dots, i_k)$.

Let B be a ball of radius $r < 1$ that intersects F. Also, truncate each sequence $(i_1, i_2, \dots)$ in $\mathcal{I}$ after the first i_k at which point

$$\left(\min_{1\le i\le m} c_i\right) r \le c_{i_1}c_{i_2}\cdots c_{i_k} \le r, \tag{3.13}$$

where $c_1, \dots, c_m$ are the ratios of the m contracting similarities. Let $\mathcal{Q}$ be the finite set of all such truncated sequences. Since $V_1, \dots, V_m$ are disjoint, $V_{i_1,\dots,i_k,1}, \dots, V_{i_1,\dots,i_k,m}$ are disjoint too, and thus all sets $V_{i_i,\dots,i_k}$ where $(i_1, \dots, i_k)$ is a sequence in $\mathcal{Q}$ are also disjoint. Furthermore,

$$F = \bigcup_{\mathcal{I}} F_{i_1,i_2,\dots} \subset \bigcup_{\mathcal{Q}} F_{i_1,\dots,i_k} \subset \bigcup_{\mathcal{Q}} \overline{V}_{i_1,\dots,i_k}. \tag{3.14}$$

Choose a_1 and a_2 such that V contains a ball of radius a_1 and is contained in a ball of radius a_2. Then for all $(i_1, \dots, i_k)$ in $\mathcal{Q}$, the set $V_{i_1,\dots,i_k}$ contains a ball of radius $c_i \cdots c_k a_1$ and is contained in a ball of radius $c_i \cdots c_k a_2$. By equation (3.13), every $V_{i_1,\dots,i_k}$ contains a ball of radius $(\min_i c_i r)a_1$ and is contained in a ball of radius ra_2. Let $\mathcal{Q}'$ be the subset of $\mathcal{Q}$ consisting only of sequences $(i_1, \dots, i_k)$ such that $\overline{V}_{i_1,\dots,i_k}$ intersects B. By Lemma 3.6, there are at most $q = (1 + 2a_2)^n(\min_i c_i a_1)^{-n}$ elements in $\mathcal{Q}'$.

Taking the $\tilde{\mu}$-mass of B,

$$\begin{aligned}\tilde{\mu}(B) &= \tilde{\mu}(F \cap B)\\ &= \mu(\{I_{i_1,\dots,i_k} : x_{i_1,\dots,i_k,\dots} \in F \cap B\})\\ &\le \mu\left(\left\{I_{i_1,\dots,i_k} : x_{i_1,\dots,i_k,\dots} \in \left(\bigcup_{\mathcal{Q}} \overline{V}_{i_1,\dots,i_k}\right) \cap B\right\}\right)\end{aligned}$$

since $F \subset \bigcup_{\mathcal{Q}} \overline{V}_{i_1,\dots,i_k}$ by equation (3.14). From the definition of $\mathcal{Q}'$, we have

$$\tilde{\mu}(B) \le \mu\left(\left\{I_{i_1,\dots,i_k} : x_{i_1,\dots,i_k,\dots} \in \bigcup_{\mathcal{Q}'} \overline{V}_{i_1,\dots,i_k}\right\}\right) = \mu\left(\bigcup_{\mathcal{Q}'} I_{i_1,\dots,i_k}\right).$$

Finally, it follows from the properties of the mass distribution μ and equation (3.13) that

$$\tilde{\mu}(B) \le \sum_{\mathcal{Q}'} \mu(I_{i_1,\dots,i_k}) = \sum_{\mathcal{Q}'} (c_{i_1}\cdots c_{i_k})^s \le \sum_{\mathcal{Q}'} r^s = qr^s.$$

It follows that, for any set U contained in a ball B_U of radius $\operatorname{diam} U$, we have $\tilde{\mu}(U) \leq \tilde{\mu}(B_U) \leq q(\operatorname{diam} U)^s$. Then by Lemma 3.8, $\mathcal{H}^s(F) \geq \mu(F)/q = 1/q > 0$. Thus $\mathcal{H}^s(F)$ is a positive real number and so $\dim_{\mathrm{H}} F = s$. □

This theorem provides quick ways to calculate dimension of self-similar fractals. Take the Cantor set for example. The IFS for which the Cantor set is an attractor is $S_1(x) = \frac{1}{3}x$ and $S_2(x) = 1 - \frac{1}{3}x$; as demonstrated above, this satisfies the open set condition with $V = (0, 1)$. Thus by Theorem 3.10, the Hausdorff dimension of the Cantor set is s such that $2\left(\frac{1}{3}\right)^s = 1$, or $s = \frac{\log 2}{\log 3}$. Note that this is similar to our previous calculation of the dimension of the Cantor set in Section 1.

We can use this theorem to calculate the dimension of many other self-similar fractals. For example, to construct the von Koch curve, replace the middle third of a straight line segment with the other two sides of an equilateral triangle erected on this middle third (see Figure 2). For each iteration, repeat this procedure for every segment in the curve. (For the resulting fractal, see Figure 3.) Clearly, the contracting similarities are of ratio $\frac{1}{3}$. The open set condition holds with V as the interior of the tightest isosceles triangle that contains the von Koch curve, i.e., with the initial segment as the base and a height of $\frac{\sqrt{3}}{6}$ times the length of the base. Then the dimension s satisfies $4\left(\frac{1}{3}\right)^s = 1$, and so $s = \frac{\log 4}{\log 3}$.

FIGURE 2. The first iteration of the von Koch curve.

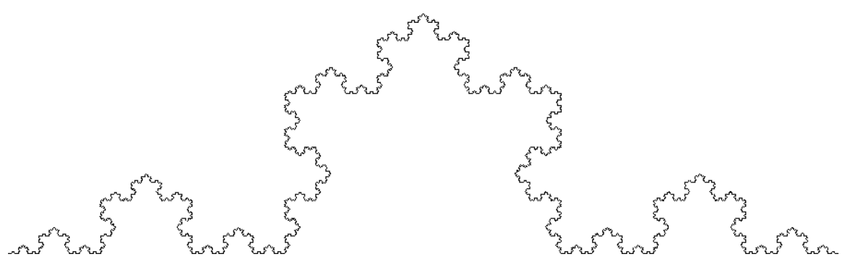

FIGURE 3. The von Koch curve.

We can also use this theorem in reverse. Given a number to be the dimension of a fractal, we can construct an IFS that has as its attractor a fractal with this dimension. We will choose $s = \frac{4}{3} = \frac{\log 16}{\log 8}$, or $16\left(\frac{1}{8}\right)^s = 1$. For aesthetics, we will choose the initial iteration displayed in Figure 4, which contains sixteen segments of length $\frac{1}{8}$ the total width, thus satisfying the requirements of our chosen dimension. As with the von Koch curve, replace each segment with a copy of the curve scaled by the chosen similarity ratio $\frac{1}{8}$, yielding the fractal in Figure 5.

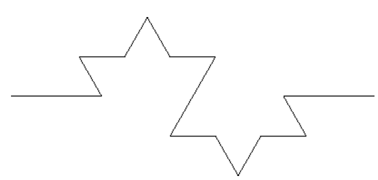

FIGURE 4. The first iteration of a fractal with dimension $\frac{4}{3}$.

(The images in Figures 2 through 5 were produced with the program included in the Appendix.)

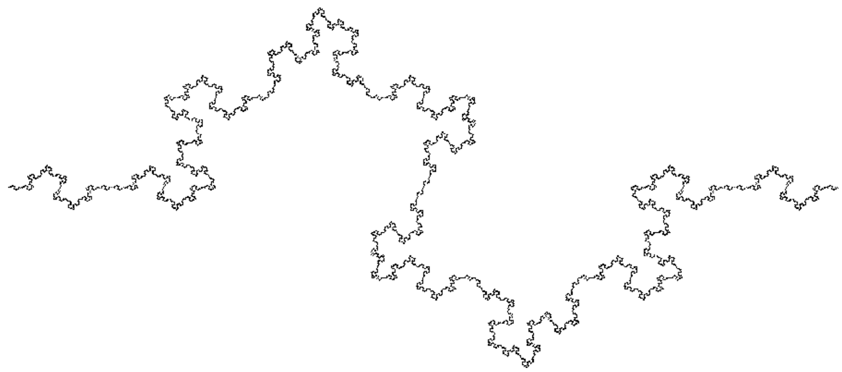

FIGURE 5. A fractal with dimension $\frac{4}{3}$.

4. FRACTALS IN DYNAMICAL SYSTEMS

Given the iterative nature of fractals, it is not surprising that they show up in dynamical systems, in which a function is iterated repeatedly on a set, with each iteration representing another step in time. To better explore the use of fractals in dynamical systems theory, we must set up the following definitions.

Definition 4.1. Given a continuous mapping $f : D \to D$ on a subset D of $\mathbb{R}^n$, let f^k denote the kth iterate of f. Then the system described by the iterations of f on D is a *discrete dynamical system*. For a given initial point x in D, the infinite sequence $\left\{f^k(x)\right\}$ is an *orbit*. If $f^k(x)$ settles into an orbit that is a finite set of points $\left\{w, f(w), \ldots, f^{p-1}(w)\right\}$, where p is the least positive number such that $f^p(w) = w$, then this set of points is an *orbit of period-p points*.

It may be the case that $f^k(x)$ settles into an orbit of period-1 points, i.e., settles upon a fixed point $f(w) = w$. The iterates $f^k(x)$ might also seem to jump about randomly, all the while approaching a certain set (indeed, often a fractal). To further discuss the sets to which iterates converge, we will need more terminology.

Definition 4.2. A subset F of D is an *attractor* for f if F is a closed set such that

- $f(F) = F$;
- for all x in an open set V containing F, the distance between $f^k(x)$ and F converges to zero as k tends to infinity; and
- for any other set G satisfying the above two conditions, $F \subset G$.

Likewise, F is a *repeller* if F satisfies the above conditions with the second condition modified so that the distance between $f^k(x)$ and F diverges as k tends to infinity.

Intuitively, an attractor is a sink. To use a power grid as an example, a house is an attractor, taking energy from a region of the grid surrounding it. This region may be small, depending on how close other houses are. Extending the analogy, the power station is a repeller.

A key concern of dynamical systems theory is the behavior of the function f, and in particular, whether it is chaotic. We will use the following definition.

Definition 4.3. Given a continuous mapping $f : D \to D$ on a subset D of $\mathbb{R}^n$, with attractor or repeller F, we say that f exhibits *chaotic behavior* on F if

- for some x in F, the orbit $\{f^k(x)\}$ is dense in F;
- the periodic points of f in F are dense in F; and
- f has a sensitive dependence on inital conditions, i.e., there exists $\delta > 0$ such that for all x in F, there exists y in F arbitrarily close to x such that $|f^k(x) - f^k(y)| \geq \delta$ for some k.

The choice of the word *chaotic* follows from the fact that f thoroughly "mixes up" points in F, and that a small change in the initial value of x can result in a large change in $f^k(x)$. The "butterfly effect"—that a butterfly flapping its wings in Brazil can produce a tornado in Texas—follows from this.

With our terminology now established, we will study the Cantor set as an example of fractals arising in dynamical systems. In this particular case, the Cantor set is both the *attractor* of an IFS and the *repeller* of a dynamical system. Furthermore, the function describing the dynamical system exhibits chaotic behavior in the Cantor set.

The function of interest is

$$f(x) = \frac{3}{2}(1 - |2x - 1|).$$

If we describe the IFS by the contractions $S_1, S_2 : [0,1] \to [0,1]$, given by $S_1(x) = \frac{1}{3}x$ and $S_2(x) = 1 - \frac{1}{3}x$, we find that f inverts both S_1 and S_2 on $[0,1]$, i.e.,

$$f(S_1(x)) = f(S_2(x)) = x.$$

Theorem 3.3 guarantees that there exists a unique attractor C for this IFS, and $C = \bigcap_{k=0}^{\infty} S^k([0,1])$, where $S = S_1 \cup S_2$. Indeed, C is the Cantor set.

We will first establish that C is a repeller for f. Since C is an attractor of the IFS, we have

$$\begin{aligned} C &= S(C) = S_1(C) \cup S_2(C) \\ f(C) &= f(S_1(C)) \cup f(S_2(C)) \\ &= C \cup C = C, \end{aligned}$$

and so C satisfies the first condition of being a repeller (see Definition 4.2). Furthermore, note that for $x < 0$, we have

$$f(x) = \frac{3}{2}(1 + (2x - 1)) = 3x,$$

and so $f^k(x) = 3^k x$. Therefore, $f^k(x) \to -\infty$ as $k \to \infty$. For $x > 1$, we have

$$f(x) = \frac{3}{2}(1 - (2x - 1)) = 3(1 - x).$$

Clearly, $f(x) < 0$ for all $x > 1$, and so just one iteration of f brings us to the previous case, in which $x < 0$. Thus, again, $f^k(x) \to -\infty$ as $k \to \infty$. For the final case, namely $x \in [0,1] \setminus C$, there exists some k such that $x \notin S^k([0,1])$, i.e.,

$$\begin{aligned} x &\notin \bigcup \{S_{i_1} \circ S_{i_2} \circ \cdots \circ S_{i_k}([0,1]) : i_j = 1,2\} \\ f(x) &\notin \bigcup \{S_{i_2} \circ S_{i_3} \circ \cdots \circ S_{i_k}([0,1]) : i_j = 1,2\} \\ &\vdots \\ f^k(x) &\notin [0,1]. \end{aligned}$$

We therefore return to one of the two previous cases, and so $f^k(x) \to -\infty$. In sum, f repels all $x \notin C$ to $-\infty$, and the second condition is satisfied.

To prove the final condition, that C is minimal, suppose that $C = G_1 \cup G_2$ with this union disjoint and with G_1 a repeller under f. Note that for all i, we have $S_i(G_2) \cap G_1 = \varnothing$, since if $S_i(G_2)$ and G_1 did share some point x, then both $f(x) \in f(S_i(G_2)) = G_2$ and $f(x) \in f(G_1) = G_1$, which is impossible since G_1 and G_2 are disjoint. Since $S(G_2) \subset C$ but $S(G_2) \cap G_1 = \varnothing$, it must be that $S(G_2) \subset G_2$. But $G_2 \supset \bigcup_{k=1}^{\infty} S^k(\overline{G_2}) = C$ by Theorem 3.3. Contradiction.

We also wish to prove that f is chaotic in C. Let E in Ω be such that $S_i(E) \subset E$ for $i = 1, 2$. Recall that from Theorem 3.3,

$$C = \bigcap_{k=0}^{\infty} S^k(E)$$

for all such E. Thus, for any $x \in C$, we have $x \in S_{i_1} \circ \cdots \circ S_{i_k}(E)$ for some sequence $(i_1, \ldots, i_k)$ with $i_j = 1, 2$. Extending this sequence gives us a way to express x, namely,

$$x = \bigcap_{k=1}^{\infty} S_{i_1} \circ \cdots \circ S_{i_k}(E).$$

We will denote this by $x = x_{i_1,i_2,\ldots}$. Observe, first, that this notation is independent of E, and, second, that for any two points $x_{i_1,i_2,\ldots}$ and $x_{i'_1,i'_2,\ldots}$ in C such that $i_1 = i'_1, \ \ldots\ , i_k = i'_k$, we have

$$|x_{i_1,i_2,\ldots} - x_{i'_1,i'_2,\ldots}| \le 3^{-k}.$$

The latter follows from the fact that S_{i_1} is applied last, S_{i_2} is applied second-to-last, and so forth. Also note that $x_{i_1,i_2,\ldots} = S_1(x_{i_2,i_3,\ldots})$, and so $f(x_{i_1,i_2,\ldots}) = x_{i_2,i_3,\ldots}$,

For $\ell_j = 1, 2$, let $(\ell_1, \ell_2, \ldots)$ be a sequence containing every n-term sequence of 1s and 2s for all n, e.g., $(1, \quad 2, \quad 1,1, \quad 1,2, \quad 2,1, \quad 2,2, \quad 1,1,1, \quad 1,1,2, \quad \ldots)$. Then for any point $x_{i_1,i_2,\ldots}$ in C and for all positive integers q, there exists a k such that $i_1 = \ell_{k+1}, \ \ldots, \ i_q = \ell_{k+q}$. Therefore,

$$|x_{i_1,i_2,\ldots} - x_{\ell_{k+1},\ell_{k+2},\ldots}| \le 3^{-q}$$
$$|x_{i_1,i_2,\ldots} - f^k(x_{\ell_1,\ell_2,\ldots})| \le 3^{-q}.$$

That is to say, for any point $x_{i_1,i_2,\ldots}$ in C, the iterates $f^k(x_{\ell_1,\ell_2,\ldots})$ are arbitrarily close to $x_{i_1,i_2,\ldots}$ if q and k are sufficiently large. Therefore, $f^k(x_{\ell_1,\ell_2,\ldots})$ make up dense orbits in C, satisfying the first condition of chaotic behavior.

In addition, the periodic points $x_{\ell_1,\ldots,\ell_p,\ell_1,\ldots,\ell_p,\ell_1,\ldots}$ are dense in C, since for any $x_{i_1,i_2,\ldots}$ in C, one can construct a periodic point $x_{\ell_1,\ell_2,\ldots,\ell_k,\ldots,\ell_p,\ell_1,\ldots}$ such that $i_1 = \ell_1, \ \ldots, \ i_k = \ell_k$, and therefore

$$|x_{i_1,i_2,\ldots} - x_{\ell_1,\ell_2,\ldots}| \le 3^{-k}.$$

Periodic points can thus be made arbitrarily close to any other point in C, and thus the periodic points are, too, dense in C.

The final condition, that f has a sensitive dependence on initial conditions, is easy to show. Consider two points $x_{i_1,i_2,\ldots}$ and $x_{i'_1,i'_2,\ldots}$ arbitrarily close to each other, as provided by $i_1 = i'_1, \ \ldots, \ i_k = i'_k$. Let $i_{k+1} = 1$ and $i'_{k+1} = 2$. Then

$$f^k(x_{i_1,i_2,\ldots}) = x_{1,i_{k+2},\ldots} \in \left[0, \tfrac{1}{3}\right] \quad \text{and} \quad f^k(x_{i'_1,i'_2,\ldots}) = x_{2,i'_{k+2},\ldots} \in \left[\tfrac{2}{3}, 1\right].$$

That is to say, there exists a $\delta = \frac{1}{3} > 0$ such that, given a point $x_{i_1,i_2,\ldots}$, there exists an arbitrarily close point $x_{i'_1,i'_2,\ldots}$ in C such that

$$|f^k(x_{i_1,i_2,\ldots}) - f^k(x_{i'_1,i'_2,\ldots})| = |x_{1,i_{k+2},\ldots} - x_{2,i'_{k+2},\ldots}| \geq \delta.$$

This satisfies the third condition of chaotic behavior.

Acknowledgments. It is a pleasure to thank my mentor, Max Engelstein, for his guidance and feedback during the research and writing that led to this paper. His enthusiasm for the material, challenging expectations, and constant patience and support were instrumental to writing this paper and making it an instructive and enjoyable experience. I would also like to thank Professor Peter May for his generous support in organizing the University of Chicago REU, which has made this all possible.

References

[1] Kenneth Falconer. *Fractal Geometry: Mathematical Foundations and Applications.* John Wiley & Sons, 2003.
[2] Frank Morgan. *Geometry Measure Theory.* Academic Press, 2009.
[3] Walter Rudin. *Principles of Mathematical Analysis.* McGraw-Hill, 1964.
[4] William P. Ziemer. *Modern Real Analysis.* http://www.scribd.com/doc/16597427/24/Hausdor%EF%AC%80-Measure
[5] Kenneth Falconer. *The Geometry of Fractal Sets.* Cambridge University Press, 1985.
[6] Shipan Lu. *The Hausdorff Dimension and Measure of Some Cantor Sets.* Real Analysis Exchange, Volume 25, Number 2 (1999).
[7] Jeff Henrikson. *Completeness and Total Boundedness of the Hausdorff Metric.* MIT Undergraduate Journal of Mathematics. www-math.mit.edu/phase2/UJM/vol1/HAUSF.PDF
[8] Michael Slone, Chi Woo. *Hausdorff metric inherits completeness.* PlanetMath.org. http://planetmath.org/HausdorffMetricInheritsCompleteness.html.
[9] Margherita Barile and Eric W. Weisstein. *Cantor Set.* MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/CantorSet.html

Appendix

The images in Figures 2 through 5 were produced with the following Python program.

```
# This program was designed and written by Christopher Natoli.
# It makes use of the Pyglet library to handle graphics and user interface.
# It can be run with Python 2.7.

import math
import pyglet
from pyglet.gl import *


################################################################################
# helper functions

def tuple_scale(tup, ratio):
    return tuple([elem*ratio for elem in tup])

def increment_pair(pair, val0, val1):
    return (pair[0]+val0, pair[1]+val1)

def increment_list(li, val0, val1):
```

```
    return [increment_pair(pair, val0, val1) for pair in li]


################################################################################
# geometric functions

def find_angle(a,b):
    if a[0]==b[0] and a[1]==b[1]: # same point
return -100
    return math.atan2(b[1]-a[1], b[0]-a[0])

def rotate_pt(pt, theta): # uses a rotation matrix
    return ( pt[0]*math.cos(theta)-pt[1]*math.sin(theta),
             pt[0]*math.sin(theta)+pt[1]*math.cos(theta) )

################################################################################
# fractal class ("Frack") and some examples

class Frack:
    def __init__(self, vertices, segments):
        self.vertices = vertices
        self.segments = segments

    def scale(self, ratio):
        return Frack([tuple_scale(point, ratio) for point in self.vertices],
                     self.segments)

    def rotate(self, theta):
        return Frack([rotate_pt(point, theta) for point in self.vertices],
                     self.segments)

quad_koch = Frack( [ (0.0, 0.25), (0.25, 0.25), (0.25, 0.0),
                     (0.5, 0.0), (0.5, 0.25), (0.5, 0.5),
                     (0.75, 0.5), (0.75, 0.25), (1.0, 0.25) ],
                   [ (0,1), (1,2), (2,3), (3,4), (4,5), (5,6), (6,7), (7,8) ]
                   )
treelike = Frack( [ (0.0, 0.333), (0.333, 0.333), (0.333, 0.666),
                    (0.666, 0.333), (0.666, 0.0), (1.0, 0.333) ],
                  [ (0,1), (1,2), (1,3), (3,4), (3,5) ]
                  )
koch = Frack( [ (0.0, 0.0),(0.333, 0.0),(0.5, 0.2887),(0.666, 0.0),(1.0,0.0) ],
              [ (0,1), (1,2), (2,3), (3,4) ]
              )
square_koch = Frack( [ (0.0, 0.0), (0.333, 0.0), (0.333, 0.25),
                     (0.666, 0.25), (0.666, 0.0), (1.0, 0.0) ],
                   [ (0,1), (1,2), (2,3), (3,4), (4,5) ]
                   )
```

```
seong = Frack( [ (0.0, 0.2165), (0.125, 0.2165), (0.25, 0.2165),
                 (0.1875, 0.32476), (0.3125, 0.32476), (0.375, 0.433),
                 (0.4375, 0.32476), (0.5625, 0.32476), (0.5, 0.2165),
                  (0.4375, 0.10825), (0.5625, 0.10825), (0.625, 0.0),
                 (0.6875, 0.10825), (0.8125, 0.10825), (0.75, 0.2165),
                 (0.875, 0.2165), (1.0, 0.2165)
                 ],
               [ (0,1), (1,2), (2,3), (3,4), (4,5), (5,6), (6,7), (7,8),
                 (8,9), (9,10), (10,11), (11,12), (12,13), (13,14), (14,15),
                 (15,16) ]
               )

################################################################################
# fractal functions

def center_generator(gen):
#lower all points in generator by y-value of first pt
    return Frack(increment_list(gen.vertices, 0, -gen.vertices[0][1]),
                 gen.segments)

def iterate(generator, iteration):
    new_vertices = []
    new_segments = []
    for seg in iteration.segments:
        a = iteration.vertices[seg[0]]
        b = iteration.vertices[seg[1]]
        theta = find_angle(a,b)
        if theta==-100:
            print("Error in finding angle.")
            return
        leng = math.sqrt((a[0]-b[0])**2 + (a[1]-b[1])**2)
        copy = center_generator(generator.scale(leng)).rotate(theta)
        new_segments.extend(increment_list(copy.segments,
                                           len(new_vertices),
                                           len(new_vertices)))
        new_vertices.extend(increment_list(copy.vertices, a[0], a[1]))
    return Frack(new_vertices, new_segments)

################################################################################
# drawing and graphics functions

width = 900
height = 700

window = pyglet.window.Window(width,height)
glClearColor(1,1,1,1) # white background
```

```
def draw_line(a,b):
    glColor4f(0,0,0,1.0)
    glBegin(GL_LINES)
    glVertex2f(a[0],a[1])
    glVertex2f(b[0],b[1])
    glEnd()

def max_height(frack):
    max_so_far = 0
    for point in frack.vertices:
        if point[1] > max_so_far:
            max_so_far = point[1]
    return max_so_far

def center_frack(frack):
    y_offset = (height/2) - (max_height(frack)/2)
    return Frack(increment_list(frack.vertices,0,y_offset),frack.segments)

def draw_frack(frack): # draws single iteration only
    for seg in frack.segments:
        draw_line(frack.vertices[seg[0]],frack.vertices[seg[1]])

def draw_fractal(gen):
    current_iter = center_frack(gen.scale(width))
    for i in range(4): # user can choose number of iterations here
        glClear(GL_COLOR_BUFFER_BIT)
        current_iter = iterate(gen, current_iter)
        draw_frack(current_iter)

@window.event
def on_draw():
    glClear(GL_COLOR_BUFFER_BIT)

    ############################
    # user can choose which fractal to draw by uncommenting only
    # the desired  fractal

    #draw_fractal(quad_koch)
    #draw_fractal(treelike)
    #draw_fractal(koch)
    #draw_fractal(square_koch)
    draw_fractal(seong)
    ############################

    pyglet.image.get_buffer_manager().get_color_buffer().save('screenshot.png')
    # the image will be saved as a file called 'screenshot.png'
```

```
pyglet.app.run()
```

This program produces fractals from user-specified generators (i.e., the first iterations of fractals). To produce the next iteration, it scales and then copies the generator onto each segment of the previous iteration. The contractions of the IFS are determined by the lengths of the segments of the generator.

This program is designed to be flexible so that users may create their own fractals. To do so, users must specify the details of the generator, namely, the locations of the vertices and which vertices are connected by segments. This is done by writing an instance of the `Frack` class. The variables `quad_koch`, `treelike`, `koch`, `square_koch`, and `seong` are all examples of generators (`koch` produced the fractal in Figure 3 and `seong` produced the fractal in Figure 5). The total width of the generator must be 1. There are no restrictions on the height, although it is expected that the fractal will be wider than it is high (this can be changed, however, by changing the dimensions of the image as defined by the `width` and `height` variables). The lowest point on the generator must have a y-coordinate of 0.0. The `segments` list in the `Frack` class determines which points of the `vertices` list are connected. Note that this program was written with fractals only produced from segments in mind. Also note that the program may take some time to draw the iterations. One can reduce this time by reducing the number of iterations drawn (it is currently set at 4).