CHAPTER 2

Hausdorff metric and the space of fractals

The goal of this chapter and the next is to rigorously prove conditions under which the fractal for a given collage map actually exists. That is, we will show when a collage map $\mathcal{T}$ admits a set A for which $\mathcal{T}(A) = A$. As an added bonus, we will prove that the set A is the limit of $\mathcal{T}^n(S_0)$ for any viable initial set S_0. That insight gives rise to an efficient algorithm that may be implemented on a computer to generate images of fractals.

In order to do all of this in a mathematically sound fashion it is necessary to build a foundation, which is the purpose of this chapter. The foundation we require is an understanding of "the space of fractals" $\mathcal{H}(X)$ as a metric space under the "Hausdorff metric" d_H. There are a number of technical details we will need to address in order to make this precise.

Until further notice we will use X to mean some subset of $\mathbb{R}^n$ or $\mathbb{C}$ along with its usual metric, which we will denote by $d(x, y)$ and call the *standard Euclidean metric*. To be precise, if $x = (x_1, x_2, ..., x_n)$ and $y = (y_1, y_2, ..., y_n)$ are elements of $\mathbb{R}^n$, then we will use the following definition and notations interchangeably:

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

If $z = a + ib$ and $w = c + id$ are elements of $\mathbb{C}$ the metric is basically the same and can be written as

$$d(z, w) = \sqrt{a - c)^2 + (b - d)^2} = \sqrt{(z - w)\overline{(z - w)}} = |z - w|.$$

2.1. A tiny bit of point-set topology

We cannot avoid learning a little bit of terminology that is commonly found in the realm of "analysis" (the branch of mathematics that calculus lives in). The main thing we need is the idea of a *compact set*, which we will give a simplified definition of here, and which you will learn/have learned about in your Real Analysis course. Compact sets in $\mathbb{R}^n$ or $\mathbb{C}$ are sets that are closed and bounded, so we must define those ideas now.

DEFINITION 2.1. We say that a set $A \subset X$ is *closed* if it contains its limit points. That is, if a sequence $\{a_n\} = \{a_1, a_2, a_3, ...\}$ in A has the property that $\lim_{n\to\infty} a_n = a$, then $a \in A$ also.[a]

[a] Of course, if you have not had Real Analysis you may not know the precise definition of the limit of a sequence. For our purposes it will suffice for you to use your intuitive idea of a limit. However, you may be curious about the official definition, so here it is.

DEFINITION 2.2. Let $\{x_n\}$ be a sequence in X. We say $\lim_{n\to\infty} x_n = L$ if and only if for every $\epsilon > 0$ there is an $N \in \mathbb{N}$ such that for every $n \geq N$, $d(x_n, L) < \epsilon$.

EXERCISE 2.3. Throughought your mathematical career your instructors have been referring to intervals like $[3, 7]$ as closed intervals. It turns out they have always meant it in the precise sense of Definition 2.1.

(1) Convince yourself that $[3, 7]$ satisfies the definition of being closed.
(2) Make an example of an interval that is not closed. Show, using the definition, exactly why it fails to be closed.

EXERCISE 2.4. Can a finite set fail to be closed?

EXERCISE 2.5. Give several examples of sets that are or are not closed, in $\mathbb{R}$ and $\mathbb{C}$. Try to make examples that differ in interesting ways.

The intuitive idea behind a set being bounded is that no portion of it heads off to infinity. In $\mathbb{R}^2$ or $\mathbb{C}$ that is equivalent to saying you can draw a big circle around it. In $\mathbb{R}$ it means that you can put upper and lower bounds on its elements. In higher dimensions it means you can enclose the set in a sufficiently large "n-ball" (interval, circle, sphere,...). For concreteness in the definition we choose to say that the n-ball is centered at the origin.

DEFINITION 2.6. We say that a set $A \subset X$ is *bounded* if there is some number $r > 0$ such that $|a| \leq r$ for all $a \in A$. That is to say, there is an $r > 0$ for which $d(a, 0) \leq r$ for all $a \in A$.

EXERCISE 2.7. Can a finite set fail to be bounded?

EXERCISE 2.8. Give several examples of sets that are or are not bounded, in $\mathbb{R}$ and $\mathbb{C}$. Try to make examples that differ in interesting ways.

What we really need to define the space of fractals is compactness. This is a profoundly useful property of sets that will be dealt with in detail in your real analysis course. What we are taking as a definition here is actually the celebrated Heine-Borel theorem, but that is something for you to tackle on another day.

DEFINITION 2.9. We say that a set $A \subset X$ is *compact* if it is closed and bounded.

There are a number of incredibly useful properties possessed by compact sets. One that is very important for fractal geometry is that it is possible to measure the distance between two compact sets unambiguously.

2.2. $\mathcal{H}(X)$, the space of fractals.

Let X be a compact subset of $\mathbb{R}^n$ (for some n) or $\mathbb{C}$, with the standard Euclidean metric d. The *space of fractals* is the set

$$\mathcal{H}(X) = \{A \subseteq X \text{ such that } A \text{ is compact}\}$$

An element of $\mathcal{H}(X)$ is therefore a compact subset of X. Take a moment to think about that carefully. Probably you will need several moments, because a "point" in the space of fractals $\mathcal{H}(X)$ is actually a "set" in the space X. Put another way, the space of fractals is the set of compact subsets of X.

EXERCISE 2.10. Let X be the unit interval $[0, 1]$. Give three specific examples of elements of $\mathcal{H}(X)$, complete with pictures. Try and make your examples as different from each other as possible.

EXERCISE 2.11. Give examples of subsets of $[0, 1]$ that are not elements of $\mathcal{H}(X)$. Draw pictures that explain why.

EXERCISE 2.12. Let X be the unit square in $\mathbb{R}^2$. Give three specific examples of elements of $\mathcal{H}(X)$, complete with pictures. Try and make your examples as different from each other as possible.

EXERCISE 2.13. Give examples of subsets of the unit square that are not elements of $\mathcal{H}(X)$. Draw pictures to help explain why.

Fractals are compact subsets of X, i.e. elements of $\mathcal{H}(X)$. They live in $\mathcal{H}(X)$ along with all the other compact subsets of X, but they satisfy special geometric properties that were discussed in the first chapter. We know that fractals can appear is as limit points of iterated function systems; next we need to develop the concept of metric spaces in order to make that idea precise.

2.3. Metric spaces

The only way to do geometry on a space is to first know how to measure the distance between points in that space. A property of an object in the space is considered "geometric" if it doesn't change when you move the object in a manner that preserves distances.

The domain of a metric on a space X is the set of all ordered pairs of elements of X. It takes the following notation and definition:

$$X \times X = \{(x, y) \text{ such that } x, y \in X\}.$$

DEFINITION 2.14. A metric on X is a function $d : X \times X \to \mathbb{R}$ satisfying the following conditions:

(1) $d(x, y) \geq 0$ for all $x, y \in X$,
(2) $d(x, y) = 0$ if and only if $x = y$,
(3) $d(x, y) = d(y, x)$ for all $x, y \in X$, and
(4) (triangle inequality) $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in X$.

Thus a metric is positive, symmetric, satisfies the triangle inequality, and distances between nonequal elements are never 0. An important detail to notice in this definition is the fact that the distance is not allowed to be ∞. That is because we have defined d with $d : X \times X \to \mathbb{R}$, and ∞ is not a real number.

One would hope that the standard Euclidean metric satisfies the official mathematical definition for a metric, and it does:

EXERCISE 2.15. Prove that when $X = \mathbb{R}$, the function $d(x, y) = |x - y|$ satisfies the definition of a metric. You may assume all standard properties of the absolute value function as you do your proof.

EXERCISE 2.16. For the standard Euclidean metrics in $\mathbb{R}$ and $\mathbb{R}^2$, determine conditions on x, y and z for which the triangle inequality is an equality.

EXERCISE 2.17. Consider the standard Euclidean metric in $\mathbb{R}^2$.

(1) Sketch a picture of $d(x, y)$ and explain it in terms of the Pythagorean Theorem.
(2) Sketch another picture in $\mathbb{R}^2$ that clearly shows why the triangle inequality is so named.

There are numerous reasons why it is useful in real life to have different definitions of a metric on the same space. For example, imagine a ruler that measures distance in the English system (inches) on one side and in the metric system (centimeters) on the other. If I ask you the distance between two points, your answer is going to depend on which side of the ruler you are using, although it would be natural for you to provide the units with your answer.

EXERCISE 2.18. Consider $d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ given by the formula

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

Is d a metric on $\mathbb{R}^2$?

EXERCISE 2.19. For this exercise let $X = \mathbb{R}$.

(1) Make up an example of a function $d : X \times X \to \mathbb{R}$ such that condition (1) of a metric fails.
(2) Make up an example of a function $d : X \times X \to \mathbb{R}$ such that condition (1) of a metric holds but condition (2) fails.
(3) Make up an example of a function $d : X \times X \to \mathbb{R}$ such that conditions (1) and (2) of a metric hold but condition (3) fails.
(4) Make up an example of a function $d : X \times X \to \mathbb{R}$ such that conditions (1), (2), and (3) hold but (4) fails.

When you change the metric on a space you change its geometry in a fundamental way. Experiment with this idea using the metric(s) that you came up with here by considering *balls of radius ϵ centered at $x \in X$*, defined by and with the notation:

$$B(x, \epsilon) = \{y \in X \text{ such that } d(x, y) \leq \epsilon\}$$

The choice of metric determines the shape of the ball by determining which points fall into the ball and which do not. Put another way, the choice of the metric determines the geometry of the balls, and by extension the geometry of X. In the next few exercises you have the opportunity to see what happens when different metrics are applied.

EXERCISE 2.20. (1) For the standard Euclidean metric in $\mathbb{R}$, compute $B(3, 1/2)$ and $B(2, -1)$.
(2) Let $d_\alpha(x, y) = |\alpha x - \alpha y|$ for some $\alpha > 0$. Compute $B(3, 1/2)$ in this metric.

EXERCISE 2.21. Try to make up an example of a metric on $\mathbb{R}^2$ that isn't the standard Euclidean metric by altering the definition of the standard Euclidean metric just a little bit. If you are successful, try to compute $B(x, \epsilon)$ for some choice of x and ϵ and see what shape the ball is.

2.4. Hausdorff metric on $\mathcal{H}(X)$

The Hausdorff metric is a way to measure the distance $d_H(A, B)$ between two compact sets $A, B \in \mathcal{H}(X)$. We're going to look at the definition two different ways, both of which depend on the standard Euclidean metric and which are equivalent. The first way depends on the idea of *ϵ-thickening* of sets. The second way depends on maximizing the minimum distance between elements of A and B. In both cases care needs to be taken to ensure the metric is symmetric, i.e. that $d_H(A, B) = d_H(B, A)$, and in both cases this will be done by taking a maximum.

2.4.1. Hausdorff metric definition using ϵ-thickening. .

Let A be a compact subset of X and let $\epsilon \geq 0$. The ϵ-*thickening* of A is the set

$$A_\epsilon = \{x \in X \text{ such that } d(x, a) \leq \epsilon \text{ for some } a \in A\}.$$

That is, A_ϵ is the subset of X that contains all the points in X that are within some element of A. Alternatively, one could imagine taking the union of all of the ϵ-balls $B(a, \epsilon)$, using each element a of A as a center.

EXAMPLE 2.22. Suppose $A \subset \mathbb{R}$ is given by $A = \{1, 2\}$ and let $\epsilon = 1/3$. Then $A_\epsilon = [2/3, 4/3] \cup [5/3, 7/3]$. This is because every point in $[2/3, 4/3]$ is within ϵ of $1 \in A$ and every point in $[5/3, 7/3]$ is within ϵ of $2 \in A$.

EXERCISE 2.23. Let $A = [1, 2] \subset \mathbb{R}$ and let $\epsilon = 1/4$. Find A_ϵ.

EXERCISE 2.24. Let $\mathcal{C} \subset [0, 1]$ be the middle-thirds Cantor set and let $\epsilon = 1/9$. Find $\mathcal{C}_\epsilon$. Repeat with $\epsilon = 1/27$.

EXERCISE 2.25. Let $A \subset \mathbb{R}^2$ be the unit circle and let $\epsilon = 1/2$. Give a precise description of A_ϵ. Repeat for $\epsilon = 2$.

EXERCISE 2.26. Let A be the line segment in $\mathbb{R}^2$ connecting the origin to $(1, 1)$ and let $\epsilon = .1$. Make a fairly precise sketch of the set A_ϵ.

EXERCISE 2.27. Let $A \in \mathcal{H}(X)$ and let $\epsilon = 0$. What is A_ϵ?

In order to find the Hausdorff distance between two compact sets $A, B \subset X$ you will need to be able to find the smallest ϵ for which $B \subset A_\epsilon$. That is to say, you will need to be able to find the minimum amount of thickening A needs in order to cover all of B. Let's look at a few concrete examples first and then define what we mean by this precisely.

EXAMPLE 2.28. Let $A = \{1, 2\} \subset \mathbb{R}$ and let $B = [.75, 1.25]$. The smallest ϵ for which $B \subset A_\epsilon$ is .25. That's because $A_{.25} = [.75, 1.25] \cup [1.75, 2.25]$ and if ϵ is any smaller than .25, the interval from A_ϵ that intersects B is too small to contain it.

EXERCISE 2.29. For the A and B in the previous example, what is the smallest ϵ for which $A \subset B_\epsilon$?

Let's be really precise about our usage of the word "smallest" by defining what it means to be the minimum value in a set of real numbers.

DEFINITION 2.30. Let $A \subset \mathbb{R}$. The *minimum element of* A, if it exists, is denoted $\min A = \min\{x \text{ such that } x \in A\}$ and is defined to be the element $a \in A$ for which $a \leq x$ for all $x \in A$.

It is possible for the minimum element of a set not to exist, for instance the interval $(0, 1)$ has no minimum element.[1] In the space of fractals we do not have to worry about this problem because compact sets of real numbers always have a minimum element.

Exercises 2.28 and 2.29 show that one needs to consider two minimum epsilons: the one for which $B \subset A_\epsilon$ and the one for which $A \subset B_\epsilon$. If we do not consider both we run the risk of the metric we define using ϵ-thickenings to fail to be symmetric.

DEFINITION 2.31. Let $A, B \in \mathcal{H}(X)$. The *Hausdorff distance* between A and B is given by

$$d_H(A, B) = \min\{\epsilon \text{ such that } A \subset B_\epsilon \text{ and } B \subset A_\epsilon\} \tag{2.1}$$

It is possible to rewrite this definition as the maximum of two minimums:

$$d_H(A, B) = \max\left\{\min\{\epsilon \text{ such that } A \subset B_\epsilon\}, \min\{\epsilon \text{ such that } B \subset A_\epsilon\}\right\} \tag{2.2}$$

Although that may look more complicated, it may the more useful because you will calculate each ϵ separately and then just take the larger of the two. To see that it is equivalent, consider the ϵ defining the minimum in equation 2.1. That ϵ is greater than or equal to each of the minimums of equation 2.2 and thus is greater than or equal to their maximum. On the other hand, the maximum of the two epsilons from 2.2 is certain to be an epsilon for which both $A \subset B_\epsilon$ and $B \subset A_\epsilon$, so it is greater than or equal to that from 2.1. When two numbers are greater than or equal to each other they must be equal.

EXERCISE 2.32. Let $A = \{1, 2\} \subset \mathbb{R}$ and let $B = [.75, 1.25]$. Find $d_H(A, B)$.

EXERCISE 2.33. In this exercise we consider distances between compact subsets A and B of $\mathbb{R}^2$. Please make sketches to illustrate your answers.

(1) Let A be the unit square $[0, 1]\times[0, 1]$and let $B = \{(x, y) \text{ such that } x^2 + y^2 = 1\}$ (the unit circle). Find $d_H(A, B)$.
(2) Let A be the unit square and let B be the disk of radius $1/2$ centered at $(1/2, 1/2)$. Find $d_H(A, B)$.
(3) Let A be the unit square and let B be the line segment connecting $(0, 1)$ to $(1, 0)$. Find $d_H(A, B)$.
(4) Let A be the unit square and let B be the line segment connecting $(-1, 2)$ to $(0, 2)$. Find $d_H(A, B)$.
(5) Let A be the line segment from the origin to $(1, 0)$ and let B be the line segment from the origin to $(0, 1)$. Find $d_H(A, B)$.

[1] There is a related mathematical notion called the *infimum* of a set, which is the largest number that is not greater than any element of the set. The infimum of $(0, 1)$ is 0.

EXERCISE 2.34. Let $A = [0, 1/2]$ and consider the Cantor set collage map $\mathcal{T} = T_1 \cup T_2$ on $\mathcal{H}([0, 1])$ given by $T_1(x) = 1/3x$ and $T_2(x) = 1/3x+2/3$. Compute $d_H(A, \mathcal{T}(A))$ and $d_H(\mathcal{T}(A), \mathcal{T}^2(A))$.

EXERCISE 2.35. In exercise 1.26 you were assigned a set S_0 and asked to apply the collage map $\mathcal{T}$ to it two times. For that collage map and your particular S_0, compute $d_H(S_0, \mathcal{T}(S_0))$ and $d_H(\mathcal{T}(S_0), \mathcal{T}^2(S_0))$.

2.4.2. Hausdorff distance via the maximum of the minimum distances. We must build up to Hausdorff metric in stages in this method also. We begin by measuring the distance from a point $x \in X$ to a a set $A \in \mathcal{H}(X)$.

Consider the Euclidean distance $d(x, a)$ for each element of A; the minimum is defined to be *the distance from x to A* and we write

$$l(x, A) = \min\{d(x, a) \text{ such that } a \in A\}$$

As before, we don't have to worry about whether this minimum exists since A is a compact set.

EXERCISE 2.36. Let A be the unit circle in $\mathbb{R}^2$ and let $x = (1, 1)$. Find $l(x, A)$.

EXERCISE 2.37. Let $\mathcal{H}(X)$ be any space of fractals and let $A \in \mathcal{H}(X)$. If $x \in A$, what is $l(x, A)$?

A particularly nice consequence of the compactness of A is that not only is this minimum distance guaranteed to exist, it must be realized as the distance between x and at least one specific point in A.

LEMMA 2.38. *For any $A \in \mathcal{H}(X)$ and $x \in X$ there is an element $\hat{y} \in A$ for which $l(x, A) = d(x, \hat{y})$.*

For the proof of this lemma see [**Bar12**, p. 29].

EXERCISE 2.39. In each of the previous two exercises find $\hat{y}$.

Next we define the distance from one set $A \in \mathcal{H}(X)$ to another set $B \in \mathcal{H}(X)$ by considering all distances $l(a, B)$ over all $a \in A$. This maximum is, like the minimum, guaranteed to exist because of compactness. We define *the distance from A to B* as

$$l(A, B) = \max\{l(a, B) \text{ such that } a \in A\}$$

EXERCISE 2.40. Let $A = [0, 2]$ and let $B = [1, 1.5]$. Find $l(A, B)$.

There is a similar lemma saying that the maximum distance is attained by elements of A and B which we state here. The existence is a consequence of compactness.

LEMMA 2.41. *For any $A, B \in \mathcal{H}(X)$ there exists $\hat{x} \in A$ and $\hat{y} \in B$ such that $l(A, B) = d(\hat{x}, \hat{y})$.*

EXERCISE 2.42. Let $A = [0, 2]$ and let $B = [1, 1.5]$. Find $\hat{x}$ and $\hat{y}$ that satisfy the lemma. Are they unique?

So the function l seems like progress toward defining a metric on $\mathcal{H}(X)$. We immediately see that it is nonnegative. Consider the following exercise and then decide about conditions (2) and (3) of a metric.

EXERCISE 2.43. Let $A = [0, 2]$ and let $B = [1, 1.5]$. Find $l(B, A)$.

So we don't quite have a metric yet. But it turns out that you can fix both of the issues this example presented in a very simple way by defining the *Hausdorff metric* to be

$$d_H(A, B) = \max\{l(A, B), l(B, A)\} \tag{2.3}$$

It is a technical exercise to prove that this version of the definition is equivalent to the one given in terms of ϵ-thickenings. A viable strategy for the proof is to consider ϵ to be the value given by the first definition and ϵ' the value given by the second. Then you would prove that $\epsilon \leq \epsilon'$ and $\epsilon' \leq \epsilon$. This shows they are equal.

EXERCISE 2.44. A corollary to Lemma 2.41 is that there is an $\hat{x} \in A$ and a $\hat{y} \in B$ such that $d_H(A, B) = d(\hat{x}, \hat{y})$. Prove this corollary.

EXERCISE 2.45. Let $A = [0, 2]$ and let $B = [1, 1.5]$. Find $d_H(A, B)$ and find the $\hat{x}$ and $\hat{y}$ that represent this distance.

EXERCISE 2.46. Let A be the disk of radius 2 centered at $(3, 0)$ and let B be the rectangle with corners at $(2, -2), (3, -2), (3, 4)$, and $(2, 4)$.

(1) Make a sketch that uses Lemma 2.41 to show $l(A, B)$. Be sure to label $\hat{x}$ and $\hat{y}$.
(2) Repeat the previous part for $l(B, A)$.
(3) Find $d_H(A, B)$.

EXERCISE 2.47. Let X be the unit square and let $\mathcal{T}$ be the collage map of the transformations $T_i, i = 1, .., 4$ defined as follows.

$$T_1(x) = x/2 \qquad T_2(x) = x/2 + (1/2, 0)$$
$$T_3(x) = x/2 + (0, 1/2) \qquad T_4(x) = x/2 + (1/2, 1/2)$$

Let $A \in \mathcal{H}(X)$ be given by $A = \{(0, 0)\}$.

(1) Find $d_H(A, X), d_H(\mathcal{T}(A), X)$, and $d_H(\mathcal{T}^2(A), X)$.
(2) Find a formula for $d_H(\mathcal{T}^n(A), X)$.
(3) We can consider the sequence $\{\mathcal{T}^n(A)\}_{n=0}^{\infty}$ of elements of $\mathcal{H}(X)$. Discuss the evidence for the existence of the limit $\lim_{n\to\infty} \mathcal{T}^n(A)$.

2.5. Exercises

EXERCISE 2.48. Consider the subset of $\mathbb{R}^2$ given by $A = \{(x, \sin(\pi/x))$ such that $x \in (0, 1)\}$.

(1) Make a fairly accurate sketch of this subset of $\mathbb{R}^2$.
(2) Is this set bounded?
(3) A is not a closed set. Find the limit points of A that are not in A.
(4) For two of the limit points of A that you found in the previous part, exhibit a sequence of elements of A that converge to it.

EXERCISE 2.49. The "taxicab metric" is a natural metric to use in $\mathbb{R}^2$ and is defined by:

$$d_t(x, y) = |x_1 - y_1| + |x_2 - y_2|$$

(1) Prove that d_t satisfies the conditions to be a metric.
(2) Explain in words, perhaps using a drawing to explain your thinking, why the word "taxicab" has been chosen to describe the metric.

EXERCISE 2.50. (1) For the metric $d_s(x, y) = |x^3 - y^3|$ on $\mathbb{R}$, compute $B(0, 1/8)$ and $B(3, 1/8)$. Compare and contrast to each other and to the balls you would get using the standard Euclidean metric.
(2) For the taxicab metric in $\mathbb{R}^2$, calculate $B((0, 0), 1)$. Compare and contrast to what you get for $B((0, 0), 1)$ using Euclidean metric.

EXERCISE 2.51. Prove that if $\epsilon \geq 0$, then $B \subseteq B_\epsilon$.

EXERCISE 2.52. Prove, using whichever definition of d_H you like, properties (1), (2), and (3) of a metric.

EXERCISE 2.53. Prove that if $B \subset A$, then $l(B, A) = 0$.

EXERCISE 2.54. Let $X = \mathbb{R}^n$ and consider A and B to be compact subsets of X. Prove the following two facts:

(1) $A \cup B$ is a compact subset of X.
(2) $A \cap B$ is a compact subset of X.