CHAPTER 1

Introduction

In this course we will study the complicated geometric structures that arise from simple natural processes and are known as fractals. We will focus on fractals obtained in two distinct ways. One way, which is the subject of the first several sections of these notes, is when they arise as limit points of "iterated function systems". The other is when they arise as collections of points that exhibit certain behaviors under repeated iteration of functions. Examples of the former are shown next; the latter will be discussed in the Julia and Mandelbrot sections.

With mathematical precision we will discuss how fractals are constructed, and we will prove conditions that guarantee their existence. We will discuss geometric properties such as self-similarity and fractal dimensions. We will learn probabilistic algorithms that allow images to be generated efficiently on a computer, and we will spend time on the computer making our own pictures. To give a flavor of what sort of structures we will be discussing, we begin with three classic examples of fractals.

1.1. Classic examples

EXAMPLE 1.1. This example begins with the closed interval $C_0 = [0, 1]$ and proceeds in stages; the *middle-thirds Cantor set* $\mathcal{C}$ is the limit[a] of the process. In the first stage, the middle third is removed to obtain the set $C_1 = [0, 1/3] \cup [2/3, 1]$. In the second stage, the middle thirds of the remaining sets are removed to obtain the set $C_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1]$. To obtain the set C_3, remove the middle thirds of the remaining intervals in C_2. The process is illustrated geometrically in figure 1.

It is clear that not all of the points in $[0, 1]$ are removed, so $\mathcal{C}$ is not empty. It is also true, but not necessarily clear, that there are no intervals in $\mathcal{C}$, but yet that it is uncountable. Interestingly, no point $c \in \mathcal{C}$ is 'isolated' from the rest of it: in any interval around c there are other points from $\mathcal{C}$.

Although he did not discover the set, Georg Cantor introduced it to the mathematical public in a paper in 1883 as an example of a set that is totally disconnected but has no isolated points. This general definition is what is meant by topologists when they throw around the term "Cantor set".

[a] We will make the notion of convergence of sets more precise in Chapter 2.

EXAMPLE 1.2. The *Koch curve* lives in the unit square

$$S = [0,1] \times [0,1] = \{(x,y) \in \mathbb{R}^2 \text{ such that } 0 \leq x \leq 1, 0 \leq y \leq 1\}.$$

The usual way to construct the curve is to begin with a line segment, then replace it with four line segments that are each $1/3$ as long as the original and placed as in the figure marked K_1 below. To construct K_2, four copies of K_1 are made, scaled by $1/3$ and placed in the same configuration as the segments that made K_1. At the next stage, four copies of K_2 are made, scaled by $1/3$, and placed in the same configuration yet again. The first several steps appear in figure 2.

Helge von Koch introduced his curve in a paper in 1906 as an example of a curve that is everywhere continuous but nowhere differentiable. It can be parameterized in the form $x = f(t), y = g(t)$. We will see in Chapter 4 that it has infinite length yet zero volume, and a fractal dimension of $\ln 4/\ln 3$, which is strictly between 0 and 1, despite it being a parameterized curve.

EXAMPLE 1.3. The *Sierpinski triangle* lives in the set

$$\{(x,y) \in \mathbb{R}^2 \text{ such that } 0 \leq x \leq 1, 0 \leq y \leq 1, x + y \leq 1\}$$

and you can think about its construction in either the Cantor set or Koch curve ways. In the middle-thirds sense, one can see S_1 of figure 3 as obtained by removing the 'middle' triangle from S_0, to leave three triangles that are each half the size of the original. From S_1 we obtain S_2 by removing the middle triangles of the remaining triangles. We continue removing the middle triangles of the remainder ad infinitum to obtain the Sierpinski triangle $\mathcal{S}$.

Alternatively, we can think of constructing S_1 by making three copies of S_0, shrunk by a factor of two, and placed as shown. To make S_2, we shrink three copies of S_1 and place them as prescribed again. We continue to this process forever to obtain $\mathcal{S}$.

Waclaw Sierpinski introduced this triangle in 1915 using a construction as a curve that is distinct from the two methods discussed here. We will see that its fractal dimension is $\ln 3/\ln 2$, again between 1 and 2.

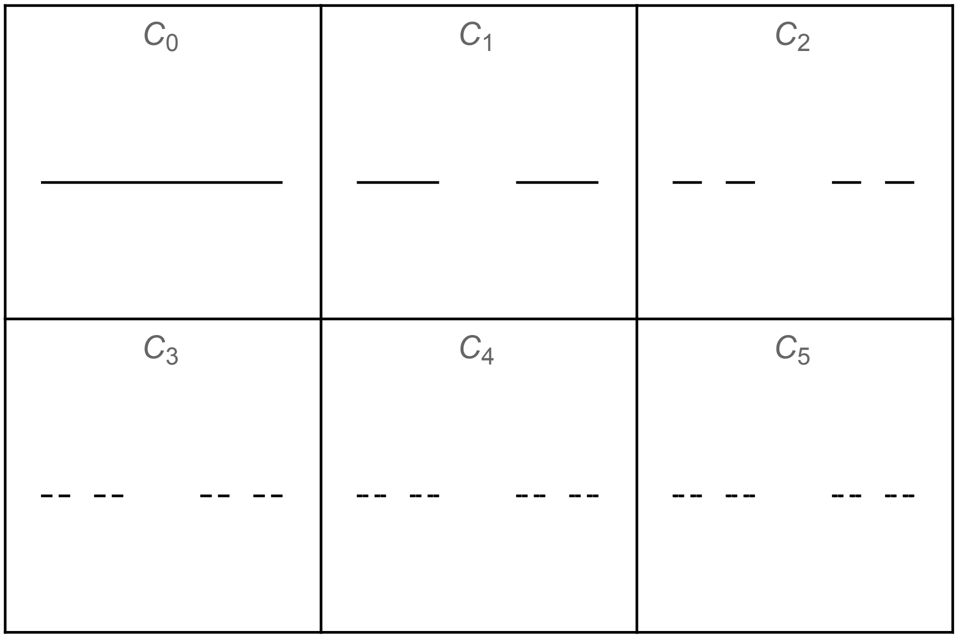


Figure 1. Constructing the middle-thirds Cantor set.

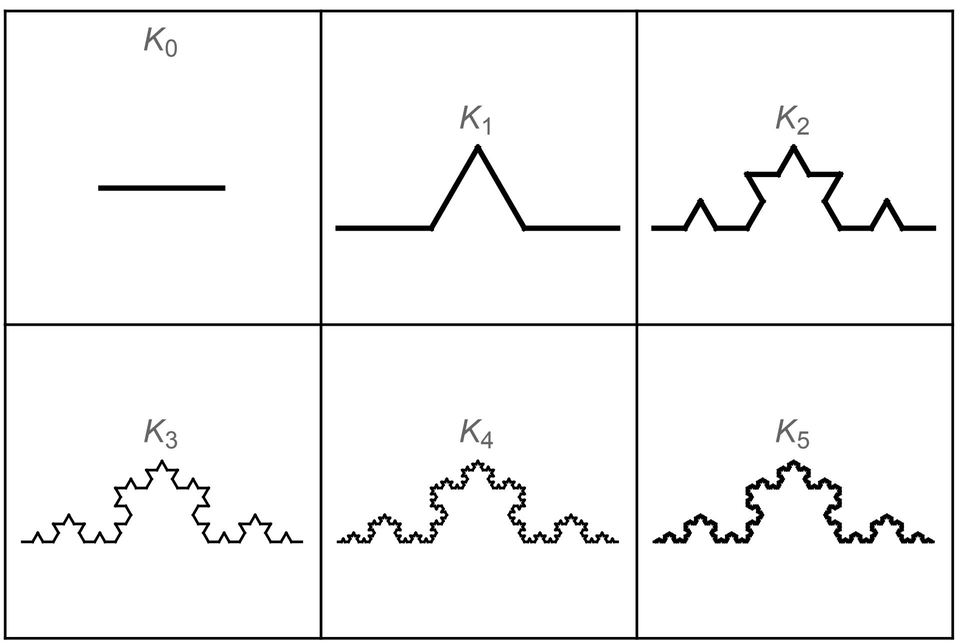


Figure 2. Constructing the Koch curve.

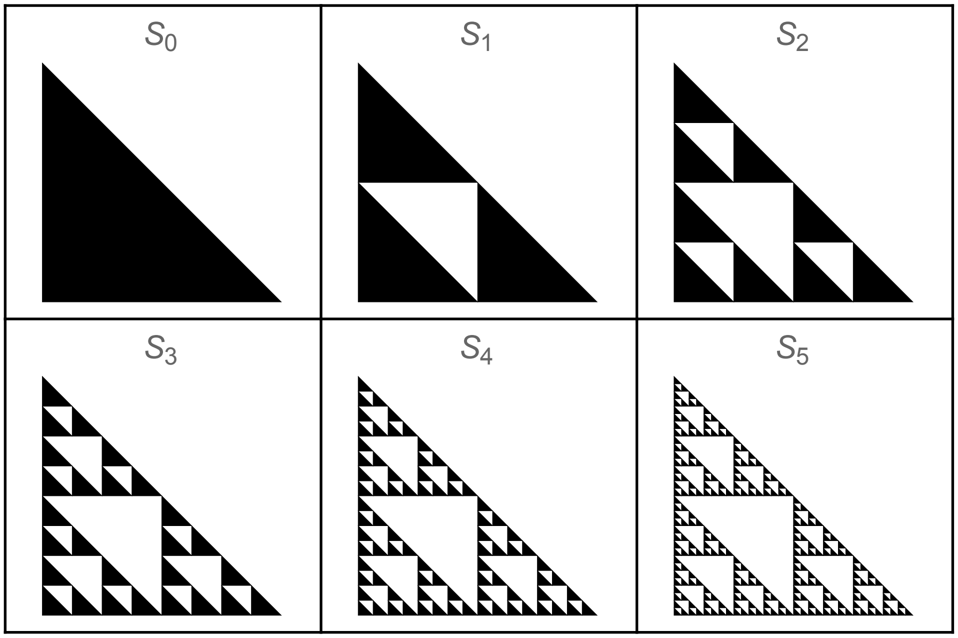


FIGURE 3. Constructing the Sierpinski triangle.

1.2. A geometric approach to transformations.

Transformations are essential in constructing fractals and are the fundamental building blocks of iterated function systems. In this section we give fundamental definitions to be used througout the course, and some one-dimensional examples are provided to help build your geometric intuition.

In general the symbol X will be used to denote a mathematical space where fractals live. Always X is a set; in the early going it will be always be $\mathbb{R}$ or $\mathbb{R}^2$ or $\mathbb{R}^n$ or $\mathbb{C}$ or a subset of those. Later on we will make X be a "compact metric space", but for now just imagine some subset of the real numbers. A transformation is a mapping defined from X to itself as follows.

DEFINITION 1.4. A *transformation* $T : X \to X$ is a function with domain X and codomain X. That is, T assigns to each element $x \in X$ an element $y = T(x) \in X$.

It is natural in fractal geometry to use certain notation and terminology from the world of dynamical systems, such as these next two definitions.

DEFINITION 1.5. Let $T : X \to X$ be a transformation and let $x \in X$. We define $T^2(x) = T \circ T(x) = T(T(x))$ and, in general,

$$T^n(x) = T \circ T \circ \cdots \circ T(x) = T(T(\cdots T(x))\cdots) \text{ (n times)}.$$

We define the *orbit* of x to be the sequence

$$\mathcal{O}(x) = \{x, T(x), T^2(x), T^3(x), ...\}.$$

DEFINITION 1.6. Let $T : X \to X$ be a transformation. We say $x \in X$ is a *fixed point* of T if $T(x) = x$.

In the following exercise we are going to consider a transformation that has a geometric interpretation you've known for years: its graph is a line. However, we need a different geometric interpretation, which is to think of the transformation as taking points from the domain back into the domain and visualizing the orbits.

EXERCISE 1.7. Define the transformation $T : [0, 1] \to [0, 1]$ by $T(x) = \frac{-1}{2}x + \frac{3}{4}$.

(1) Compute the first four elements of the orbit of $x = 0$.
(2) Visualize the orbit as follows: draw yourself a fairly large copy of the unit interval $[0, 1]$ and place the points from the orbit on it. Connect each point to the one that follows it with an arrow in a follow-the-bouncing-ball sort of way. (Your arrows will live above or below your unit interval, depending on your artistic choices.)
(3) Repeat the previous parts using $x = 1$.
(4) Compute the fixed point(s) of T algebraically.[a]
(5) Reflect briefly on how the orbits seem to relate to the fixed point.

[a]The word "algebraically" tends to mean to do a computation or use mathematical symbols in some other way. Algebraic arguments tend to feel very precise and rigorous. I will also often ask you for "geometric" arguments; these can include graphs or sketches or they can be constructed from the language of geometric objects.

The transformation of exercise 1.7 is special in several ways, the most obvious being that its graph is a line. But it also has the property that it brings points closer together as it is applied.

DEFINITION 1.8. Let X be a subset of $\mathbb{R}, \mathbb{R}^n$, or $\mathbb{C}$.[a] The transformation $T : X \to X$ is called a *contraction* if there is some constant $c \in [0, 1)$ such that

$$|T(x) - T(y)| \leq c|x - y| \text{ for all } x, y \in X.$$

The number c, which is not unique, is called a *contractivity factor* for T.

[a]This definition applies to transformations on any metric space, so when we learn about those it will apply retroactively.

Exercise 1.9. Write as formal a proof as you can muster that the transformation from exercise 1.7 is a contraction. Note: in your "scratch work" to prepare for the proof, you should identify a good value for the contractivity factor c.

Another useful feature of the transformation we have been considering is that it can be broken down as the composition of two even simpler transformations: $T_1(x) = \frac{-1}{2}x$ and $T_2(x) = x + \frac{3}{4}$. Take a moment on scratch paper right now to verify that $T(x) = T_2 \circ T_1(x)$ algebraically. What this means is that you can see two separate geometric actions that move your point x from its location to the location specified by $T(x)$: first the size of x is halved and it is flipped across the origin, then it is moved over to the right by $\frac{3}{4}$. This insight might not seem like a big deal to you right now, but it is particularly useful when you are applying your transformation to entire subsets of X rather than just points.

Definition 1.10. Let $T : X \to X$ be a transformation and let $A \subseteq X$. We define the *image of A under T* to be

$$T(A) = \{T(a) \text{ such that } a \in A\}.$$

Similarly, for $n = 2, ...$ we define *image of A under T^n* to be

$$T^n(A) = \{T^n(a) \text{ such that } a \in A\}.$$

Put another way, $T(A)$ is the **set** you get by applying T to every element of A. It is useful to be able to think of a transformation applied to a set geometrically as well as algebraically, as in this next exercise.

Exercise 1.11. Let T be as in exercise 1.7 and let $A = [0, 1]$. Compute and sketch $T(A)$ in the following two different ways, the first more algebraic and the second more geometric.

(1) Plug the endpoints of A into T. If that is enough for you to know what $T(A)$ is, write the answer now. If not, choose some other points $a \in A$ and compute $T(a)$. Try to express $T(A)$ in a mathematically familiar way.
(2) Second, think of T as the composition $T_2 \circ T_1$ as discussed in the paragraph before definition 1.10. Construct a sketch for what T_1 does to A, then make a sketch what T_2 does to that.
(3) Repeat the process (both ways) to compute $T^2(A)$ and $T^3(A)$.

1.3. Collage maps: the building blocks of iterated function systems.

In Chapter 3 we will give the formal definition of an iterated function system, but we can start working with them right away by defining the collage map given by a finite set of transformations. We can use the transformations together to make a mapping that takes **subsets of X** to **subsets of X**:

DEFINITION 1.12. Let $T_i : X \to X$ be a transformation for all $i = 1, 2, ..., k$ and let $A \subseteq X$. The *collage map* defined by these transformations is given by

$$\mathcal{T}(A) = \bigcup_{i=1}^{k} T_i(A) = T_1(A) \cup T_2(A) \cup \cdots \cup T_k(A)$$

We may write $\mathcal{T} = T_1 \cup T_2 \cup \cdots \cup T_k$.

Notice that since $A \subseteq X$ and each T_i takes X to itself it must be true that $T_i(A) \subseteq X$ for each i. Since $\mathcal{T}(A)$ is a union of such subsets of X it must also be a subset of X. That means that it is perfectly appropriate to define $\mathcal{T}^2(A) = \mathcal{T} \circ \mathcal{T}(A)$, and $\mathcal{T}^n(A) = \mathcal{T} \circ \mathcal{T} \circ \cdots \circ \mathcal{T}(A)$ (n times) as before.

EXAMPLE 1.13. Let $X = [0, 10]$, $T_1(x) = \dfrac{x}{5}$, and $T_2(x) = \dfrac{x}{5} + 5$, and define the collage map $\mathcal{T} = T_1 \cup T_2$. Suppose that $A = [5, 10]$. Then $T_1(A) = [1, 2]$ and $T_2(A) = [6, 7]$, hence $\mathcal{T}(A) = [1, 2] \cup [6, 7]$. The set $\mathcal{T}(A)$ is pictured at the top of figure 4.

To compute $\mathcal{T}^2(A) = \mathcal{T}([1, 2] \cup [6, 7])$, we need to consider what T_1 and T_2 each do to the set $\mathcal{T}(A) = [1, 2] \cup [6, 7]$, then union the result. Now $T_1([1, 2] \cup [6, 7]) = [1/5, 2/5] \cup [6/5, 7/5]$, and $T_2([1, 2] \cup [6, 7]) = [5 + 1/5, 5 + 2/5] \cup [5 + 6/5, 5 + 7/5]$. Thus $\mathcal{T}^2(A)$ is the union of four intervals, shown in the middle line of figure 4. In the bottom of figure 4 we show $\mathcal{T}^3(A)$, which consists of eight intervals. At each stage, the leftmost interval's left endpoint is getting closer to 0. Can you determine where some other endpoints seem to be tending?

EXERCISE 1.14. In general, the collage map technically isn't a transformation from a space X to itself even though its component maps T_i all are. Why is that? (Hint: given the set $\{x\}$ containing a single point $x \in X$, what kind of object is $\mathcal{T}(\{x\})$?)

The collage map is actually a transformation on the "space of fractals" $\mathcal{H}(X)$, which we will define properly in chapter 2.

EXERCISE 1.15. Let $X = [0, 1]$ and define the transformations $T_1(x) = \dfrac{1}{3}x$ and $T_2(x) = \dfrac{1}{3}x + \dfrac{2}{3}$. Consider the collage map given by $\mathcal{T} = T_1 \cup T_2$.

(1) Let $A = X$ (i.e. the whole unit interval). Compute the sets $\mathcal{T}(A)$ and $\mathcal{T}^2(A)$ algebraically.
(2) Make sketches of $A, \mathcal{T}(A)$, and $\mathcal{T}^2(A)$.
(3) Repeat with $A = \left\{\dfrac{1}{2}\right\}$ (i.e. a set with one point in it).
(4) In each case, consider whether $\mathcal{T}^n(A)$ seems to tend to a specific set A_0 as n goes to infinity. If so, how do the limit sets compare?

EXERCISE 1.16. Let $X = [0, 6]$, $T_1(x) = \dfrac{x}{2}$, and $T_2(x) = \dfrac{2x}{3} + 2$, and define the collage map $\mathcal{T} = T_1 \cup T_2$.

(1) Let $A = X$. Compute the sets $\mathcal{T}(A)$ and $\mathcal{T}^2(A)$ algebraically.
(2) Make sketches of $A, \mathcal{T}(A)$, and $\mathcal{T}^2(A)$.
(3) Repeat with $A = \{6\}$.
(4) In each case, consider whether $\mathcal{T}^n(A)$ seems to tend to a specific set A_0 as n goes to infinity. If so, how do the limit sets compare?

EXERCISE 1.17. Let $X = [0, 1]$, $T_1(x) = \dfrac{x}{4}$, and $T_2(x) = \dfrac{x}{2} + \dfrac{1}{4}$, and define the collage map $\mathcal{T} = T_1 \cup T_2$.

(1) Let $A = X$. Compute the sets $\mathcal{T}(A)$ and $\mathcal{T}^2(A)$ algebraically.
(2) Make sketches of $A, \mathcal{T}(A)$, and $\mathcal{T}^2(A)$.
(3) Repeat with $A = \{0\}$. For this one you should try to go to $\mathcal{T}^4(A)$.
(4) In each case, consider whether $\mathcal{T}^n(A)$ seems to tend to a specific set A_0 as n goes to infinity. If so, how do the limit sets compare?

EXERCISE 1.18. Let $X = [0, 10]$, $T_1(x) = \dfrac{x}{10} + 1, T_2(x) = \dfrac{x}{10} + 4$, and $T_3(x) = \dfrac{x}{10} + 7$, and define the collage map $\mathcal{T} = T_1 \cup T_2 \cup T_3$.

(1) Let $A = X$. Compute the sets $\mathcal{T}(A)$ and $\mathcal{T}^2(A)$ algebraically.
(2) Make sketches of $A, \mathcal{T}(A)$, and $\mathcal{T}^2(A)$.
(3) Repeat with $A = \{7\}$.
(4) In each case, consider whether $\mathcal{T}^n(A)$ seems to tend to a specific set A_0 as n goes to infinity. If so, how do the limit sets compare?

EXERCISE 1.19. Let $X = [0, 1]$, $T_1(x) = \dfrac{x}{2}$, and $T_2(x) = \dfrac{x}{2} + \dfrac{1}{2}$, and define the collage map $\mathcal{T} = T_1 \cup T_2$.

(1) Let $A = X$. Compute the sets $\mathcal{T}(A)$ and $\mathcal{T}^2(A)$ algebraically.
(2) Make sketches of $A, \mathcal{T}(A)$, and $\mathcal{T}^2(A)$.
(3) Repeat with $A = \left\{\dfrac{1}{3}\right\}$. For this one you should try to go to $\mathcal{T}^4(A)$, at least with sketches.
(4) In each case, consider whether $\mathcal{T}^n(A)$ seems to tend to a specific set A_0 as n goes to infinity. If so, how do the limit sets compare?

1.4. Affine transformations in two dimensions: a geometric approach

Fractals are more fun in two dimensions and so we need to get serious about studying $\mathbb{R}^2$-transformations. Although we have many types to choose from, as in the one-dimensional case we choose to stick with simple "affine" transformations. In one dimension, an affine transformation is any transformation of the form $T(x) =$

$ax + b$, in other words a linear transformation followed by a translation. Every example from the previous sections was of this form, and the shrink-and-move intuition we developed there will be useful in two and higher dimensions also. To make a precise definition of an affine map in two dimensions we first need to set some notation.

Let us denote elements of $\mathbb{R}^2$ in any of three ways, which we will use interchangeably:

$$\vec{x} = (x_1, x_2) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

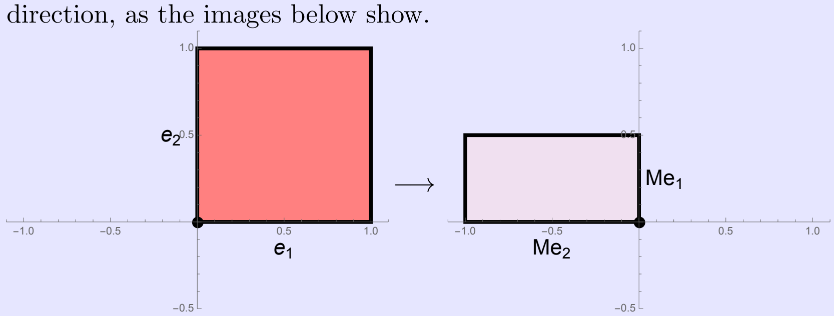
A *linear* transformation from $\mathbb{R}^2$ to itself is given by matrix multiplication and can be expressed in many ways, including these:

$$M\vec{x} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (ax_1 + bx_2, cx_1 + dx_2) \tag{1.1}$$ [1]

When thinking about linear transformations it is particularly useful to think about how they affect the standard basis vectors, which in this course we will denote by $\vec{e_1} = (1, 0)$ and $\vec{e_2} = (0, 1)$. Take a moment to use equation 1.1 to verify that $M\vec{e_1} = (a, c)$ and $M\vec{e_2} = (b, d)$. That is to say, the first column of M tells us what M does to the first standard basis vector, and the second column of M tells us what it does to the second.

A quick and dirty way to visualize matrix multiplication is to always do the following. Consider the unit square, which can be seen as having corners at the origin, $\vec{e_1}, \vec{e_1} + \vec{e_2}$, and $\vec{e_2}$. Then M sends the unit square to the parallelogram having corners at $M(0, 0), M\vec{e_1}, M(\vec{e_1} + \vec{e_2})$,and $M\vec{e_2}$. Take a moment to verify that if $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then these corners are $(0, 0), (a, c), (a + b, c + d)$, and (b, d), respectively.

Example 1.20. Suppose that $M = \begin{pmatrix} 0 & -1 \\ 1/2 & 0 \end{pmatrix}$. Then we can see that the unit square is rotated by $\pi/2$ counterclockwise and squished by half in one direction, as the images below show.



[1] If you have not seen matrix multiplication before you should take this as a definition.

Exercise 1.21. Consider the following matrices:

$$\begin{pmatrix} 0 & -1/2 \\ 1/2 & 0 \end{pmatrix} \quad \begin{pmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{pmatrix} \quad \begin{pmatrix} 0 & 1/3 \\ 1/3 & 0 \end{pmatrix} \quad \begin{pmatrix} 1/2 & 1/2 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1/2 & 1 \\ 0 & 1/2 \end{pmatrix}$$

(1) Sketch how each matrix acts on the unit square, like in example 1.20.
(2) Make a winking smiley face (or some other asymmetric image of your choice) inside your unit square and show how each matrix transforms it.
(3) For each matrix, determine whether the transformation given by $T(\vec{x}) = M\vec{x}$ is a contraction. If it is, give a contraction factor. If not, exhibit two vectors $\vec{x}$ and $\vec{y}$ that are not brought closer by the transformation.

We know by the remark immediately after definition 1.1 that the first column of a matrix is where it sends the first standard basis vector, and that the second column is where it sends the second. This means that if you know where a matrix sends the basis vectors then you know the matrix itself. That is, you can use geometric descriptions of linear transformations to make matrices for them. This next exercise uses words, but you can translate pictures into matrices using this principle also.

Exercise 1.22. In each of the following determine the matrix M that has the following effect on $\mathbb{R}^2$. **Find the columns by figuring out where $\vec{e_1}$ and $\vec{e_2}$ go.** You can verify your answer by multiplying your matrix by vectors of your choice and seeing if they go where they are supposed to.

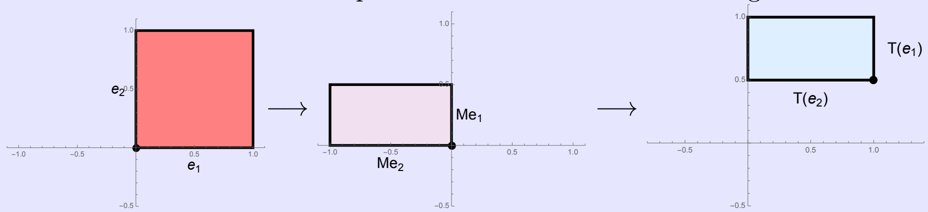
(1) M rotates vectors clockwise by $\pi/3$.
(2) M rotates vectors clockwise by $\pi/3$, then rescales the x direction by a factor of 2.
(3) M reflects across the line $y = -x$.
(4) M reflects across the line $y = -x$, then rescales the whole vector by a factor of 1/2.

An *affine* transformation from $\mathbb{R}^2$ to itself is given by matrix multiplication followed by translation. It can be expressed in many ways, including these:

$$T(\vec{x}) = M\vec{x} + (e, f) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} = (ax_1 + bx_2 + e, cx_1 + dx_2 + f) \tag{1.2}$$

Geometrically, affine transformations can be visualized by what they do to the unit square. This is best done in two stages: visualizing the matrix part of the transformation first, then translating.

EXAMPLE 1.23. Make an affine transformation out of the matrix $M = \begin{pmatrix} 0 & -1 \\ 1/2 & 0 \end{pmatrix}$ from example 1.20 by defining $T(\vec{x}) = M\vec{x} + (1, .5)$. Then the transformation of the unit square can be seen in these two stages:



EXERCISE 1.24. For each matrix in exercise 1.21, sketch the image of the unit square under the transformation $T(\vec{x}) = M\vec{x} + (1, .5)$.

EXERCISE 1.25.

(1) For each matrix from example 1.22 sketch the affine transformation $T(\vec{x}) = M\vec{x} + (0, .5)$. Decorate your unit square with an asymmetric image so that you can see how it transforms in each case.
(2) Which transformations are contractions?
(3) Which can be considered transformations on the unit square (as opposed to all of $\mathbb{R}^2$)?

1.5. Collage maps in two dimensions

In two dimensions we see the power and beauty of iterated function systems even better than in one dimension. In this section we will experiment with various choices of affine transformations and see how these choices affect the underlying fractal. By now you are probably beginning to ask two very important questions: (1) What exactly do we mean when we say "the limit under the collage map"? and (2) What exactly is a fractal? The answer to (1) is the subject of Chapter 2. The answer to (2) has not yet been agreed upon by the mathematical community. In section 1.6 we will address the key points of agreement on that question.

EXERCISE 1.26. Let $X = [0,1] \times [0,1]$ be the unit square in $\mathbb{R}^2$ and let $C(x,y) = (.5x, .5y)$. Define the following three transformations from X to itself:

$$T_1(x,y) = C(x,y) + (.5, 0) \tag{1.3}$$

$$T_2(x,y) = C(x,y) + (0, .5) \tag{1.4}$$

$$T_3(x,y) = C(x,y) + (.5, .5) \tag{1.5}$$

Given an initial set S_0, do the following four exercises on graph paper.

(1) Compute and sketch $T_1(S_0), T_2(S_0)$, and $T_3(S_0)$ in separate unit squares. Label your scale and important points.
(2) Let $\mathcal{T}$ be the collage map given by $T_1 \cup T_2 \cup T_3$. Define the set $S_1 = \mathcal{T}(S_0)$ and sketch it in a fresh unit square.
(3) Sketch $S_2 = \mathcal{T}(S_1)$ in a separate unit square from that of S_1.
(4) On your graph paper, make a really big unit square and sketch a nice picture of $S_3 = \mathcal{T}(S_2) = \mathcal{T}^3(S_0)$.

The choices for S_0 are as follows:

$(a)X$, $(b)[0,.5]\times[.5,1]$, $(c)[0,1]\times[0,.5]$, or $(d)\{(x,y) \in X \text{ such that } y \geq x\}$

In the examples from the previous section and in this example you may be noticing a pattern: it does not seem to matter what the initial set is. That is not a fluke, but rather a common trait that collage maps constructed from contractions share. The fractal associated with the IFS from exercise 1.26 is shown in figure **??**.

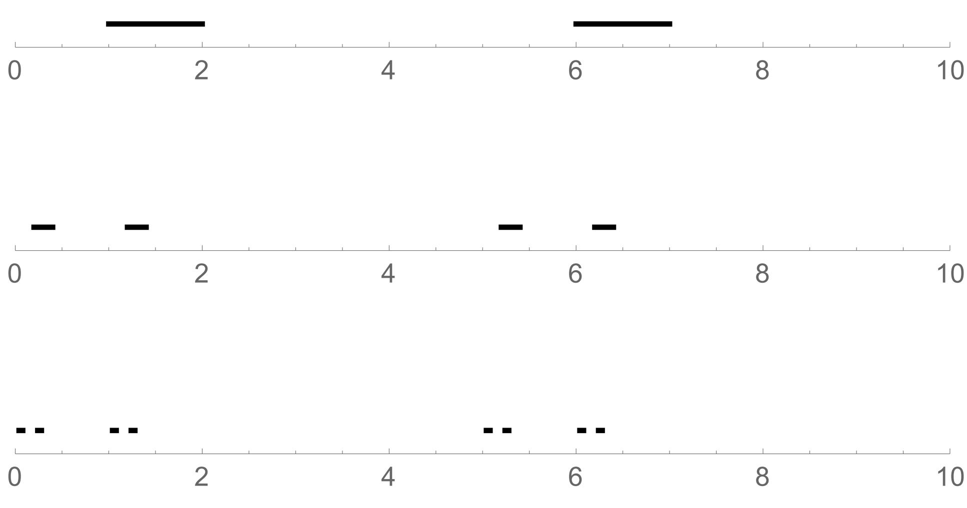


FIGURE 4. The collage map applied three times to A.

There are a few key things to notice about this fractal image, which we will call A. Each initial set S_0 was a subset of the unit square, as was $\mathcal{T}(S_0), \mathcal{T}^2(S_0)$, and all further images under the collage map. Naturally this makes A a subset of the unit square also. It is a very special subset in that it is a *fixed point of the collage* *map* in the sense that $\mathcal{T}(A) = A$. Let's analyze the statement that $\mathcal{T}(A) = A$ geometrically with a detailed look at figure **??**.

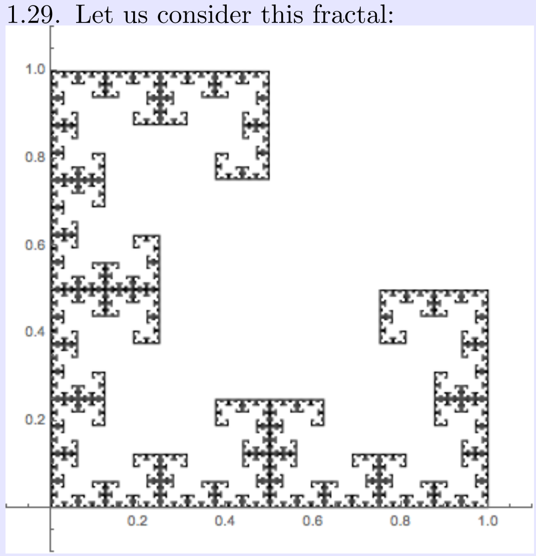
When T_1 is applied to A, it shrinks it by half and moves it to the right by $(.5, 0)$. You can see that A has a copy of itself in its lower-right corner, and that's the part of the collage map given by T_1. When T_2 is applied to A, it shrinks it by half and moves it up by $(0, .5)$. In that location you see that A has a copy of itself, which is the part of the collage map given by T_2. The map T_3 shrinks by half and moves diagonally by $(.5, .5)$; the third copy of A is there in the upper right. Thus when the collage map is applied to A, it makes three copies of itself whose union is A again. That is to say, $\mathcal{T}(A) = A$.

Exercise 1.27. Let X be the unit square in $\mathbb{R}^2$. Make up four affine transformations $T_i : X \to X, i = 1, ..., 4$ such that for their collage map $\mathcal{T}$ we have that $\mathcal{T}(X) = X$. Do you believe it is possible to do this in such a way that $T_i(X) \cap T_j(X) = \emptyset$ when $i \neq j$?

Exercise 1.28. Let $A = \{(1/2, 1/2)\}$ and consider the collage map you just made up for the previous question. Compute $\mathcal{T}^n(A)$ for some small values of n. Do you believe that it may be true (given proper definitions) that $\mathcal{T}^n(A) \to X$ as $n \to \infty$?

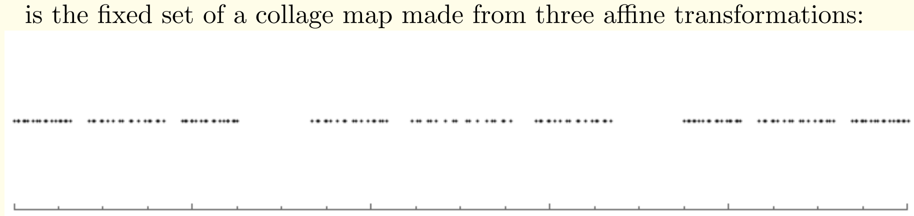
It is sometimes possible to look at a fractal such as the one in figure **??** and determine the collage it came from. To do this, one must parse the image into components that appear to be similar to the original. Then, for each piece of the image one must try to determine the affine transformation that takes the whole image into that piece.

EXAMPLE 1.29. Let us consider this fractal:



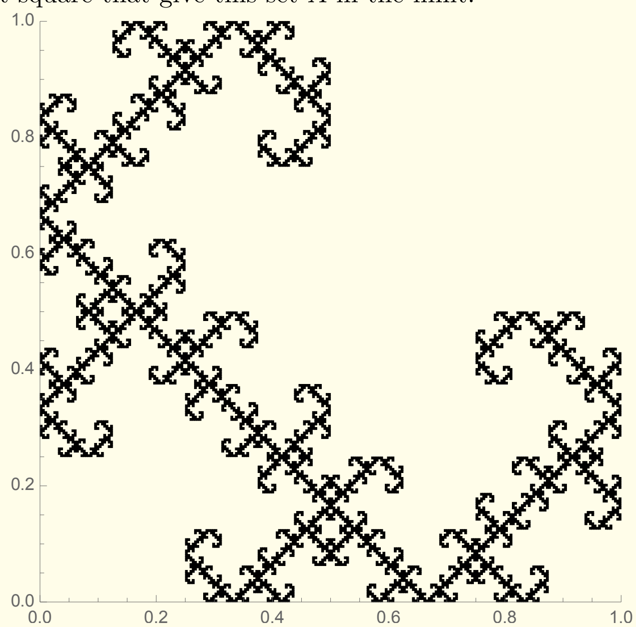
It is made up of three copies of itself, each half as big as the original. One copy contains the origin and is simply the image of the fractal under the map $T_1(\vec{x}) = \vec{x}/2$. The piece of the fractal between $x = .5$ and $x = 1$ is the entire image halved, rotated by $\pi/2$ counterclockwise, then moved to the right by $(1, 0)$. This corresponds to $T_2(\vec{x}) = \begin{pmatrix} 0 & -1/2 \\ 1/2 & 0 \end{pmatrix} \vec{x} + \vec{e_1}$. Finally the piece on the upper left, between $y = .5$ and $y = 1$, is the entire image halved, rotated by $\pi/2$ clockwise, and then moved up by $(0, 1)$. This corresponds to $T_3(\vec{x}) = \begin{pmatrix} 0 & 1/2 \\ -1/2 & 0 \end{pmatrix} \vec{x} + \vec{e_2}$. When we apply the collage map of these three transformations, we see that this fractal is fixed by it.

EXERCISE 1.30. Now it is your turn to try and decode the collage a fractal has come from, first in one dimension. Here we have a fractal A that is the fixed set of a collage map made from three affine transformations:



The fractal is contained in unit interval of the real line, but it is shown floating above the line so that you can see it. Find the three transformations $T_i : [0, 1] \to [0, 1]$ for which $\mathcal{T}(A) = A$. (You should be using the scale provided in the figure to help you determine precise contraction factors and translations).

EXERCISE 1.31. Now try to figure out the three affine transformations of the unit square that give this set A in the limit:



1.6. What is a fractal?

During the early study of fractals it became clear that a formal defintion was elusive, and indeed none of [**Bar08**, **Bar12**, **Fal06**, **PC09**] offer one. Instead, fractals are objects that are identified by two main properties: complicated geometric structure and self-similarity.

One way to describe a 'complicated geometric structure' is to describe what it is not: fractals are not simple objects like lines, circles, cones, or triangles. Rather, they are like clouds, trees, and coastlines: objects with obvious structure but no obvious way to describe or measure them. One way that this idea can be quantified is through dimension. In fact, originally Mandelbrot felt the definition of a fractal should depend on its fractal dimension.

The *box-counting dimension* Dim_B of a set is the foremost definition of a fractal dimension, and we will define it properly in chapter 4. For now, think of it as a number that captures how the mass of the set scales when it is expanded by a fixed amount. The box-counting dimension can be a number that is not an integer. By way of contrast, the *topological dimension* Dim_T of a set, which also has a technical definition we do not describe here, corresponds to our intuititve idea of dimension and can therefore only take integer values. In an early treatise on fractals [**Man77**], Mandelbrot brings us the following definition of what it means to be fractal:

> "The cases where $Dim_B = Dim_T$ include all of Euclid, and the cases where $Dim_B > Dim_T$ include every set I was ever tempted to call fractal... Hence, there is no harm in proposing the following definition:

> *A fractal will be defined as a set for which the Hausdorff-Besicovitch dimension*[2] *strictly exceeds the topological dimension.*"

In the fullness of time, it became clear that there were sets that failed this definition and yet still seemed to deserve to be called fractal [**Edg90**, p. 179]. However, sets with non-integer fractal dimension have the sort of complicated geometry that is often described as fractal.

The second main indicator of what it means for a set to be a fractal is *self-similarity.* The word 'similarity' can be construed in a number of ways, the strictest of which is the type taught in an elementary geometry course to describe things like similar triangles. In this definition two objects are similar if one is a rescaling of the other. A set is self-similar, then, if portions of the set are similar to the whole set. We see this version of self-similarity in all of the introductory examples of these notes.

However, it is useful to broaden the category of self-similarity to include sets for which portions are similar to some portion of the original, but perhaps not all of it. Moreover, we may wish to allow some flexibility in the word 'similar', perhaps allowing copies that are images under a contracting but nonlinear map. In chapters 5 and 6 we will learn about sets having this sort of self-similarity: Julia sets and the Mandelbrot set.

In [**Fal06**, p. xxv] an expansive list of properties is suggested, and we leave them as our final word on what it means to be fractal.

> "When we refer to a set F as a fractal, therefore, we will typically have the following in mind.
>
> (i) F has a fine structure, i.e. detail on arbitrarily small scales.
> (ii) F is too irregular to be described in traditional geometrical language, both locally and globally.
> (iii) Often F has some form of self-similarity, perhaps approximate or statistical.
> (iv) Usually, the 'fractal dimension' of F (defined in some way) is greater than its topological dimension.
> (v) In most cases of interest F is defined in a very simple way, perhaps recursively."

1.7. Exercises

Exercise 1.32. Let X be a subset of $\mathbb{R}, \mathbb{R}^n$, or $\mathbb{C}$, and let $T_1 : X \to X$ and $T_2 : X \to X$ be contraction mappings. Prove that $T_1 \circ T_2 : X \to X$ is also a contraction mapping.

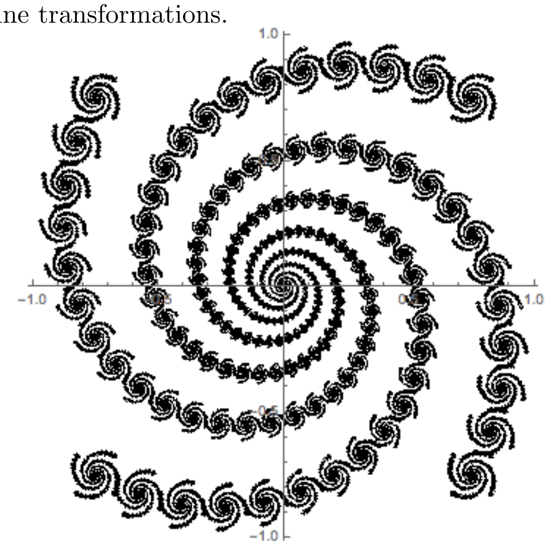
Exercise 1.33. Give an example of a contraction mapping $T : \mathbb{R} \to \mathbb{R}$ that has the property that $|T(x)| > |x|$ for some $x \in \mathbb{R}$. The ideal response will prove that T is a contraction, then exhibit a specific x that does not shrink under the transformation.

Exercise 1.34. Suppose that $T_1(\vec{x}) = (ax_1 + bx_2 + e, cx_1 + dx_2 + f)$ and $T_2 = (gx_1 + hx_2 + k, ix_1 + jx_2 + l)$ are affine transformations of $\mathbb{R}^2$. Prove that $T_1 \circ T_2$ is also an affine transformation of $\mathbb{R}^2$ and give its explicit formula.

[2] Pretend he said "box-counting dimension"

EXERCISE 1.35. (1) Find the formula of an affine transformation of $\mathbb{R}^2$ that takes the triangle with vertices at $(0,0),(1,0)$, and $(0,1)$ to the triangle with vertices at $(2,1),(3,3)$, and $(0,4)$.
(2) Find another affine transformation that accomplishes the same task.
(3) How many affine transformations are there in total that could do this?

EXERCISE 1.36. For this spiraling fractal do not attempt to find specific formulae for the affine transformations.



Instead, make a sketch of the domain $X = [-1,1] \times [-1,1]$ that shows your estimate of where each transformation takes X. (Hint: one of your transformations only contracts a little bit, and the rest contract a lot.)

EXERCISE 1.37. Find the affine transformations of the unit square whose collage map $\mathcal{T}$ gives this fractal: