Chapter 1

Real Numbers and Sequences

What is a real number? We start our investigation with the set of integers, denoted by $\mathbb{Z}$, which is the set of whole numbers and their "opposites"[1], $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$. From these we will form the rational numbers $\mathbb{Q}$, which fill in much of the gaps between the integers. In fact, as you will prove, between any two rational numbers there is another one, so the "rational number line" has no sizable gaps. However, the rational number line still has many "missing points" that are needed for the purposes of measuring lengths[2]—we will spend considerable time in §1.1.2 showing that indeed rational numbers do not suffice. This leads us to make a formal statement—the Axiom of Completeness—that the real number line, which can be used to measure any length, is not missing any points. That is, the real numbers $\mathbb{R}$ are *the numbers of measurement*.

1.1. What are the real numbers?

1.1.1. The rational numbers. The rational numbers are obtained from the integers by forming fractions and declaring which ones we will consider equivalent:

Definition 1.1. The set $\mathbb{Q}$ of rational numbers consists of all symbols of the form $\frac{m}{n}$ where:

(1) n and m are integers and $n \neq 0$, and

(2) we agree that $\frac{m}{n} = \frac{p}{q}$ if and only if $mq = np$.

[1] And what does this mean? The opposite of a whole number n is a number m such that $n + m = 0$. We will write the number m as $-n$. Note that we have just *hypothesized* the existence of (defined into existence?) such a number—by writing the symbol $-n$ and giving it a defining property.

[2] It is not an obvious fact that the rational numbers are not sufficient! Hippasus, a student of Pythagoras, is sometimes credited with first proving there are numbers that are not rational. According to lore, this was so shocking to the Pythagoreans that they drowned him. Just to be safe, you might want to watch out for any Pythagoreans after completing Problem 1.21.

In set notation, this says:

$$\mathbb{Q} = \left\{ \frac{m}{n} \;\middle|\; n \neq 0 \quad \text{and} \quad \frac{m}{n} = \frac{p}{q} \iff mq = np \right\}$$

***Problem 1.2.** Show, using the definition, that "fractions can be reduced"—that is,

$$\frac{m}{n} = \frac{\frac{m}{\gcd(m,n)}}{\frac{n}{\gcd(m,n)}},$$

where $\gcd(m,n)$ is the greatest common divisor of m and n.

In fact, among a set of equivalent fractions this reduced fraction is unique. Sometimes it is called the "lowest terms" fraction.

We next define how to add and how to multiply numbers in $\mathbb{Q}$. You probably know how to add and multiply fractions already, and you may have your own favorite way of doing these calculations. We concentrate here on what it means to write down a mathematical definition of these procedures.

Definition 1.3. We define addition and multiplication of rational numbers by the following formulas:

(1) $\frac{m}{n} + \frac{p}{q} = \frac{qm+np}{nq}$

(2) $\frac{m}{n} \cdot \frac{p}{q} = \frac{mp}{nq}$

***Exercise 1.4.** Note that, despite your many years of experience adding and multiplying fractions in your own way, we now have a written (i.e., common) *definition* of each of these operations. We should check these definitions to see if they do what we want them to. (If not, we should go back and revise the definition!)

(1) Add the fractions $\frac{1}{2}$ and $\frac{1}{4}$ according to the definition. Did you get the same answer you would have relying on your previous experience?

(2) Perform addition and multiplication of $\frac{2}{3}$ and $\frac{3}{5}$ according to the definition.

(3) Perform addition and multiplication of the fractions $\frac{3}{8}$ and $\frac{5}{12}$ according to the definition. Then put the resulting fractions in reduced form if they are not already.

(4) Do you want to revise the definition of addition and multiplication? If so, make sure to write a revised version down explicitly. What, if any, benefits are there to keeping the above printed definition?

Just like the integers, the rational numbers can be given an ordering $\leq$. This means that any two rational numbers can be compared using the symbol $\leq$ in a consistent way. Complete the following definition:

***Definition 1.5.** We will write $\frac{m}{n} \leq \frac{p}{q}$ if and only if...

You may not have been asked to write your own mathematical definition before, but formulating precise definitions is an important mathematical task. Besides being

precise (i.e., not depending on a reader's interpretation), two other key properties of a good definition are that it is concise and that it applies to examples the way you want it to. Mathematicians often revise their definitions as they gain more insight into how they want to use them.

***Exercise 1.6.** Does your definition correctly identify that $\frac{2}{5} \leq \frac{3}{7}$? What about $\frac{4}{3} \leq \frac{12}{9}$? Make sure that you are testing *your definition*—not finding some other process to get you through. Are there other types of examples you should try before deciding that your definition is a good one? Revise your definition as needed.

***Exercise 1.7.** Show that between any two rational numbers there is another rational number.

We will admit without proof or justification that the previous definitions (plus the usual properties of the integers) allow us to work with the rational numbers as we are used to. All of this mathematical structure can be encoded in the following statement: the rational numbers, with $+$, $\cdot$, and $\leq$ defined as above, are an *ordered field*. That is, for all rational numbers x, y, and z, the following are true:

Axioms of an ordered field

A1: $x + y = y + x$.

A2: $(x + y) + z = x + (y + z)$.

A3: $x + 0 = x$.

A4: There exists $(-x)$ such that $x + (-x) = 0$.

M1: $x \cdot y = y \cdot x$.

M2: $x \cdot (y \cdot z) = (x \cdot y) \cdot z$.

M3: $1x = x$.

M4: If $x \neq 0$ there exists x^{-1} such that $x \cdot x^{-1} = 1$.

D: $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$.

O1: If $x \leq y$ and $y \leq z$ then $x \leq z$.

O2: If $x \leq y$ and $y \leq x$ then $x = y$.

O3: If $x \leq y$ then $x + z \leq y + z$.

O4: If $x \leq y$ and $0 \leq z$ then $x \cdot z \leq y \cdot z$.

From these axioms one can deduce, for example, that $0 \cdot x = 0$ for all numbers x and that $(-1) \cdot (-1) = 1$.

1.1.2. A need for more numbers. The real numbers must contain the rational numbers, which are very useful for measuring. Are there any other real numbers? It is not very clear that there are. You may be tempted to bring up, say, $\sqrt{2}$, but who is to say there is not a very fine rational number—with billions of digits in the numerator and denominator—that is equivalent to $\sqrt{2}$? In any case, if there are such numbers that are real but not rational, let's call them "irrational".

To begin our search for irrational numbers (or—let's not be greedy—even just one irrational number), consider the geometric problem of measuring lengths—say, with a measuring stick.[3] One end of our measuring stick we will call 0, and its length we will call a unit. Once we decide on the unit, we can make and measure with fractions and multiples of it (i.e., this determines where the rational numbers are). No pressure, but choose your unit mark carefully! Let's start with trying to measure just one real number x. Easy: if you place the unit length at x, it is clearly measurable with rational multiples of the unit. Now suppose we also want to measure a second number, y. Unless y happens to be particularly "nice" (with respect to x), it is not very obvious whether or not we can measure y with fractions and multiples of our measuring stick of length x. Maybe if we go back and pick a different unit length, both x and y will be clearly measurable with fractions and multiples? This would need some kind of argument. In order to investigate, we start with some definitions.

Definition 1.8. (1) A *divisor* of a non-zero number x is a number c such that $x = nc$ for some integer n (that is, "a measuring stick of length c measures x in integers").

(2) Two non-zero numbers x and y are *commensurable* if and only if they have a common divisor.

This last part of the definition includes the phrase "if and only if", which indicates that the statement "x and y have a common divisor" is a complete characterization of the term commensurable. That is, if x and y have a common divisor then we will call them commensurable, and if we call x and y commensurable then they must have a common divisor. Since this is very common in definitions (indeed, the point of defining a term is to give a complete characterization of it), we will abbreviate this phrase by the term "iff" in definitions.

The phrase "if and only if" also occurs in theorems, exercises, problems, etc.—including the following problem. This indicates that there are really two claims to address. The first, "A if B", means that if you assume B, then A must be true (or, "if B then A"). The other, "A only if B", literally means that there are no instances of A being true without B also being true. A more straightforward way of saying this is that if you assume A, then B must be true (or, "if A then B"). We will continue to use the entire phrase "if and only if" in these contexts.[4] It is a good idea to write out these two statements when you see an "if and only if" statement and show each one separately. (See Appendix A for more discussion of such logic and proof-related advice.)

***Problem 1.9.** Check that two non-zero numbers x and y are commensurable if and only if there are a third number c and integers n and m such that

$$x = nc \quad \text{and} \quad y = mc.$$

Note that in this exercise, x, y, and c can be any kind of numbers but n and m are required to be integers. The idea is that positive numbers x and y are commensurable if and only if there is a measuring stick (of length c) that measures both x and y in integers.

It is good mathematical practice to think about a new definition using examples.

[3] We discuss this kind of problem in more detail in §6.1.

[4] That is, we will write "iff" if and only if "if and only if" is in a definition.

***Exercise 1.10.** What are the divisors of $\sqrt{2}$?

***Exercise 1.11.** Are $x = \frac{1}{2}$ and $y = \frac{5}{6}$ commensurable? If so, what is a common measure? What is their *greatest* common measure?

Next, try to tackle a more general question:

***Problem 1.12.** Are any two rational numbers commensurable?

What if you don't know whether your two numbers are rational—could they still be commensurable? The following theorem provides a complete characterization of when two numbers are commensurable. Note that it is an "if and only if" statement, so you have two separate "if…then…" statements to prove.

***Theorem 1.13.** Any two real numbers x and $y \neq 0$ are commensurable if and only if $\frac{x}{y}$ is a rational number.

While this answers the question "when are two real numbers commensurable?", it does not say anything about the existence of *non-commensurable* pairs of numbers. We'll have to spend more time looking for that. In the meantime, check that this theorem does imply the following (also an "if and only if" statement!):

***Corollary 1.14.** A real number x is irrational if and only if x and the number 1 are not commensurable.

Aha! This gives us a new way to think about our search for an irrational number. Having different ways to think about the same thing is a very useful tool to a mathematician, as we will see.

*
* *

We now introduce a way to visualize the notion of commensurable quantities.[5] The key geometric concept is the following:

Definition 1.15. A rectangle $ABCD$ will be called a *grid rectangle* iff its vertices A, B, C, and D are grid points of *some* square grid on the plane.

***Problem 1.16.** This problem relates grid rectangles and commensurable lengths:

(1) Show that if a rectangle is a grid rectangle, then its side lengths are commensurable.★

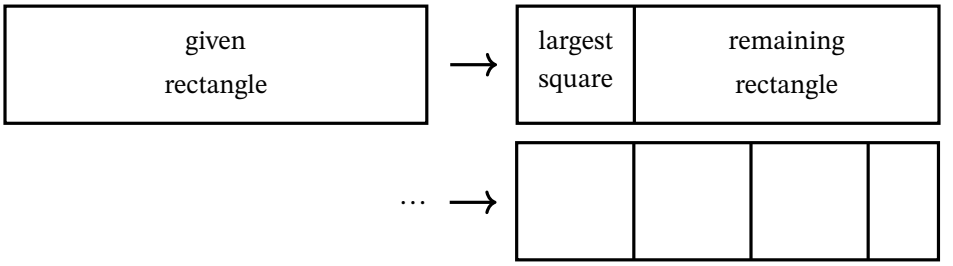
(2) Conversely, show that if $x > 0$ and $y > 0$ are commensurable, then there exists a grid rectangle with side lengths x and y.

A consequence of these results is that the existence of non-commensurable pairs of real numbers is equivalent to the existence of non-grid rectangles. So, our question is now: are there any non-grid rectangles?

To investigate this question we introduce the following *inscribed squares construction*: Given a rectangle with side lengths $x > y > 0$, inscribe a square of side length y

[5] Some of the topics introduced here will be explored further in §6.4.

sharing one of the sides with the rectangle. Then repeat with the remaining rectangle, always inscribing squares of side length y, until you run out of room, like this:

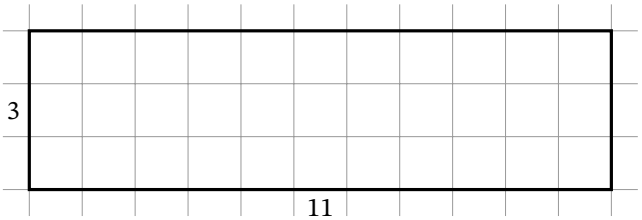


Stop after there is no room in the last remaining rectangle to inscribe a square of side length y.

Definition 1.17. The last remaining rectangle will be called the *remainder* of the original given rectangle.

***Exercise 1.18.** Now try it yourself on the 3-by-11 rectangle below.

Note: This rectangle is printed on a grid for convenience, but *the inscribed squares construction can be performed on any rectangle*, not just on grid rectangles.



Perhaps you noticed that there seems to be a close relationship between the inscribed squares construction and division with remainder. Indeed there is such a relationship! We will make it explicit in §6.4, but can you formulate it already?

We need to understand a bit more about the relationships between grids, given rectangles, and remainder rectangles.

***Problem 1.19.** Argue that if one starts with a rectangle on some grid, the inscribed squares and the remainder rectangle are on the same grid.

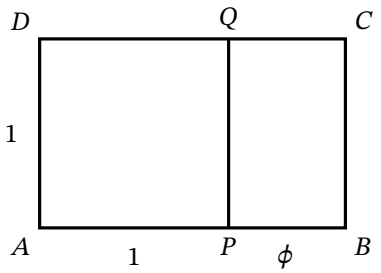
***Problem 1.20.** Denote by $\mathcal{R}_1$ and $\mathcal{R}_2$ two rectangles that are similar in the sense of geometry, and perform the inscribed squares construction on each of them.

(1) Argue that the number of inscribed squares in each is the same.

(2) Argue that the remainder rectangles are similar to each other.$\star$

With these results in hand, we revisit our quest for an irrational number.

***Problem 1.21.** The figure below represents a rectangle $ABCD$ to which the inscribed squares construction has been applied. Only one square, $APQD$, fits inside. The original rectangle $ABCD$ below is very special, because *the remainder rectangle $PBCQ$ is similar* (*in the geometric sense*) *to $ABCD$*.



(1) Find the numerical value of the length $\phi = |PB|$ if $|AD| = 1$.⋆

(2) Now imagine performing the inscribed squares construction on the remainder rectangle $\mathcal{R}_1 = PBCQ$. Call $\mathcal{R}_2$ the remainder rectangle of *that* construction. What can you say about $\mathcal{R}_2$ with respect to $\mathcal{R}_1$ and the original rectangle $ABCD$?⋆

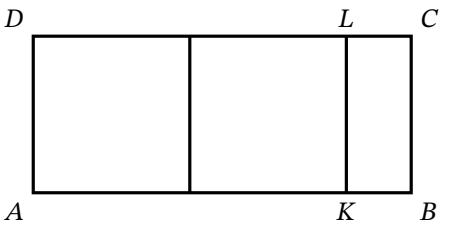
(3) Now imagine performing the inscribed squares construction on $\mathcal{R}_2$, then on the remainder rectangle of that, and so on. This process ends only if at some point the remainder rectangle is empty. Will this process ever end? Why?

(4) How does your answer to (3) imply that $ABCD$ cannot be a grid rectangle?

(5) How does your answer to (4) imply that ϕ is irrational?

In fact, we can use the setup of Problems 1.21 as a blueprint for finding more irrational numbers.

***Problem 1.22.** Consider the following rectangle, where now it is assumed that $BCLK$ is similar to $ABCD$.



(1) In the spirit of the previous exercise, if $|AD| = 1$ how much is $|KB|$? Prove that $|KB|$ is irrational.

(2) What would happen if instead you set $|AB| = 1$ and asked for the length $|AD|$? Can you explain how this compares to the result from part (1)?

(3) Construct other similar examples.

1.1.3. The Axiom of Completeness. We just showed that there are many numbers needed for purposes of measurement (and thus for algebra, geometry, calculus, etc.) that are not rational. In particular, there are lengths (i.e., points on the real line) that are not the ratio of two integers. How can we then describe the set of numbers that constitute the real line?

In fact, we are going to formulate an *axiom* to express the property that the real line does not have "holes". What does it mean to appeal to an axiom? Here is how Euclid described the role of axioms in his study of geometry:

> [Axioms] are certain general propositions, the truths of which are self-evident, and which are so fundamental, that they cannot be inferred from any propositions which are more elementary.

That is, we are claiming that it is "self-evident"—and unprovable from other more elementary notions—that the real number line has no holes in it, and we are going to study the implications of this statement. Our axiom is called the *Axiom of Completeness*. It is a fundamental tool in our analysis of the real numbers.

To motivate the version of the Axiom of Completeness that we will use (there are several equivalent possibilities), let's look at one way to approximate $\sqrt{2}$ by rational numbers. This method in some sense goes back to the Babylonians![6]

The general "Babylonian strategy" to find an approximation to a real number a is:

I. Find a transformation $F : \mathbb{R} \to \mathbb{R}$ such that $F(a) = a$ (a is called a *fixed point* of F).

II. Pick an "initial value" $a_0 \in \mathbb{Q}$, and let $a_k = F(a_{k-1})$ for $k = 1, 2, \ldots$.

III. Under some conditions on F and a_0, the number a_k will approximate a for large k.[7]

Note that we are using the notation a_k for the result of applying F k times to a_0 (e.g., $a_2 = F(F(a_0))$).

> The process of applying a transformation repeatedly is called *iterating* the transformation.

Iteration[8] is a central theme in this text.

***Investigation 1.23.** This problem is about approximating the number $a = \sqrt{2} - 1$ by rational numbers. (This of course will give us an approximation of $\sqrt{2}$.)

(1) Show that the number a satisfies the equation

$$x = \frac{1}{2 + x}$$

and argue that it is thus a *fixed point* of the transformation $F(x) = \frac{1}{2+x}$ (that is, $F(a) = a$).

(2) The table below contains decimal approximations to the iterates $a_k = F(a_{k-1})$ starting with $a_0 = 0$. The table also contains values of $w_k = (1 + a_k)^2$ for $k = 1, \ldots, 14$. What do you observe? More precisely, what can you conjecture (from the

[6] They iterated the transformation $F(x) = \frac{1}{2}\left(x + \frac{2}{x}\right)$. You can check that $x = F(x)$ if and only if $x^2 = 2$. It is unclear, however, whether the Babylonians knew that there are irrational numbers.

[7] We will not specify the necessary conditions on F and a_0, since they are beyond the current scope of our study. "Approximate" is a vague term that we will think more about soon.

[8] From the Latin *iteratio*, meaning *repetition*.

numerical evidence) about the numbers a_{2l} and a_{2l+1} for $l = 1, 2, \ldots$ (that is, a_k with k even and the following odd-indexed term) with respect to $\sqrt{2}-1$ and with respect to each other?

(3) Why must it be that $a_k \neq \sqrt{2}-1$ for all k?

(4) Show that if $a_k < \sqrt{2}-1$ then $a_{k+1} > \sqrt{2}-1$, and if $a_k > \sqrt{2}-1$ then $a_{k+1} < \sqrt{2}-1$.

k	a_k	$(1+a_k)^2$
0	0	1
1	0.5	2.25
2	0.4000000000	1.9600000000
3	0.4166666667	2.0069444444
4	0.4137931034	1.9988109394
5	0.4142857143	2.0002040816
6	0.4142011834	1.9999649872
7	0.4142156863	2.0000060073
8	0.4142131980	1.9999989693
9	0.4142136249	2.0000001768
10	0.4142135516	1.9999999697
11	0.4142135642	2.0000000052
12	0.4142135621	1.9999999991
13	0.4142135624	2.0000000002
14	0.4142135624	2.0000000000

It is very useful to be able to think about relationships such as these in multiple ways. For example, the relationships between numbers that you found above can also be described in statements about closed intervals.

Definition 1.24. A *closed interval from a to b*, where a and b are real numbers such that $a \leq b$, is the set $[a, b]$ consisting of a, b, and all of the numbers between a and b—that is,

$$[a, b] \coloneqq \{\, x \in \mathbb{R} \mid a \leq x \leq b \,\}.$$

Using the numbers from Investigation 1.23, define

$$I_k = [a_{2k}\,,\, a_{2k+1}]$$

for each $k = 0, 1, 2, \ldots$. Then $I_{k+1} \subset I_k$[9] for all k (can you explain why this is true based on your observations in part (2) of Investigation 1.23?). By part (4), the number $\sqrt{2}-1$ is in *all* of the intervals I_k. Moreover, it appears that the lengths of the intervals I_k tend to zero, so that as k grows the intervals are shrinking in on $\sqrt{2}-1$.

Yet another way to think about this process is graphically.

***Investigation 1.25.** Here is a procedure for iterating a transformation graphically. The figure below shows this procedure for the transformation $F(x) = \frac{1}{2+x}$. Note that

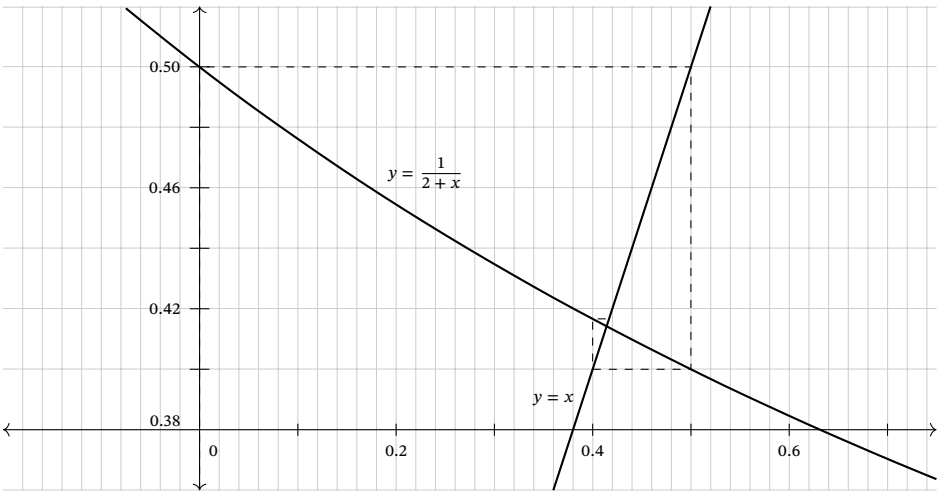
[9] This is set notation that means I_{k+1} is contained in I_k. For more on sets and set notation, see Appendix B.

the horizontal axis has been shifted up to $y = 0.38$ in order to zoom in on the interesting part of the picture.

I. Plot together the graph of the function F and the diagonal, $y = x$.

II. Pick the initial value (or "seed") a_0 on the x-axis and find the point (a_0, a_0) on the diagonal. (In the example below $a_0 = 0$ and so (a_0, a_0) is the origin, which is outside the plot.)

III. For $k = 1, 2, \ldots$ construct the point (a_k, a_k) as follows:
   i. Draw a vertical line segment starting at (a_{k-1}, a_{k-1}) and ending at the graph of F. Call this intersection point P_{k-1}.
   ii. Draw a horizontal line segment from P_{k-1} to the diagonal. This intersection point is (a_k, a_k).



(1) Locate the points P_0, P_1, P_2, and P_3 in the plot above.

(2) Explain why a_k, constructed graphically as above, is the kth iterate of F, starting from a_0. ⋆

(3) Use the graph to locate at least two of the intervals $I_k = [a_{2k}\,,\, a_{2k+1}]$ on the x-axis. What do you observe? What do you predict happens for larger k's?

We will use intervals as in the previous example to state the Axiom of Completeness. Before doing that, we need a few definitions.

Definition 1.26. A *sequence of closed intervals* is a rule that associates to each positive integer $k = 1, 2, \ldots$ a closed interval $I_k \subset \mathbb{R}$. We will denote such a sequence by (I_k). A sequence of closed intervals (I_k) is said to be *nested* iff for each positive integer k one has

$$I_{k+1} \subset I_k.$$

***Exercise 1.27.** Suppose that $(I_k = [x_k,\ y_k])$ is a sequence of closed intervals. Translate the condition that the sequence is *nested* into statements about the two infinite strings of numbers $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$.

Now we finally state

> **The Axiom of Completeness.** If (I_k) is any sequence of *nested closed intervals on the real line, then there is a(t least one) real number x such that $x \in I_k$ for all k.*

There is a lot to unpack in this axiom! First, notice the phrases "any", "there is at least one", "such that", and "for all".[10] Whenever you come across such phrases, it is important to closely observe the order of phrases in the statement.

***Exercise 1.28.** Compare and contrast the statements

(1) there is a real number x such that $x \in I_k$ for all k, and

(2) for each k there is a real number $x \in I_k$.

It is very important to understand the difference between these two!

The phrase "there is at least one real number x…" is often written simply as "there is a real number x…". In everyday language you might interpret a phrase such as "there is a bird in that tree" to mean there is one bird and not more, but in its mathematical usage this phrase says nothing about the possible existence of more. If there are ten birds in the tree, then the statement "there is a bird in that tree" is still true. In fact, in the Axiom of Completeness there may very well be more than one x with the stated property.

***Exercise 1.29.** Write down an example of a sequence of nested closed intervals such that there is *more than one* real number in I_k for all k.

In some circumstances, though, the Axiom of Completeness does pick out *exactly* one real number. Intuitively, if the lengths of the intervals I_k *shrink to zero* as $k \to \infty$, then there should be a *unique* point (or real number) that is in all I_k. This turns out to be the case in the previous investigation, and the one point that is in all the I_k is the real number equal to $\sqrt{2}-1$. In order to rigorously prove this, we need to know what "the lengths shrink to zero" means precisely (to be addressed in §1.2). However, intuition will serve well enough for the following problem.

Theorem 1.30. *If the lengths of the nested closed intervals goes to 0, then the real number in all of the intervals is unique.*

***Proof (to be completed).** We argue by contradiction: Suppose x and y are distinct real numbers that are both in I_k for all k. [*Show that then the lengths of I_k cannot go to zero. Since that contradicts the hypothesis of the theorem, there must not be two distinct numbers in all of the I_k, which means that the x guaranteed by the Axiom of Completeness is unique.*] □

Here is another problem for which the Axiom of Completeness is useful:

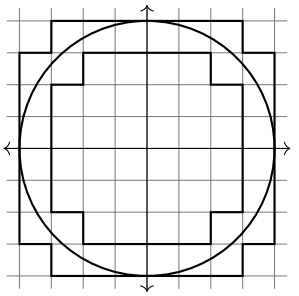
***Problem 1.31.** Let $D = \{(x, y) \mid x^2 + y^2 \leq 1\}$ be the disk centered at the origin with radius 1. For each integer $k \geq 0$, consider the square grid G_k built on the coordinate

[10] For more discussion on these phrases, see Appendix A.

system with side length $\frac{1}{2^k}$. For each k, define m_k and M_k as follows:

m_k = the total area of the grid squares in G_k completely contained in D, and

M_k = the total area of the grid squares in G_k that intersect D.



The disk D and boundaries of the regions whose areas give m_k and M_k for $k = 2$.

(1) To check your understanding of the definitions, compute m_2 and M_2 using the figure above.

(2) Argue that the closed intervals $I_k = [m_k, M_k]$ form a nested sequence of intervals whose lengths tend to zero.★

(3) What is the real number that is in *all* of the I_k? Explain.

1.1.4. The Archimedean Axiom. In order to proceed in our study of the real numbers, we need one more axiom that establishes a relationship between real numbers and integers.[11]

Archimedean Axiom. Given any real number a there exists an integer n such that $n > a$.

Despite its simplicity, this axiom has some very useful and clarifying implications. For example, one consequence of this axiom is that infinity is not a real number, since there would have to be an integer larger than it (contradicting the very notion of infinity).

Corollary 1.32. The symbol ∞ does not represent a real number.

While there is no infinitely large real number, there are *arbitrarily large* real numbers, in the sense of the Archimdedean Axiom. The symbol ∞ is still useful as notation for this idea—in fact, we will use it in important notation in §1.2—but it should not be given the status of a real number by, for instance, trying to carry out addition or multiplication with it.

[11] Another common version of the Axiom of Completeness is that every set bounded above has a least upper bound. This least upper bound statement is in fact stronger than our nested intervals axiom—it is equivalent to the nested intervals axiom *and* the Archimedean Axiom together. You will encounter the least upper bound statement in § 2.2.

Another consequence of the Archimedean Axiom is the Archimedean Property of the real numbers:

***Corollary 1.33** (Archimedean Property)**.** Given two positive real numbers, x and y, there exists a positive integer n such that $nx > y$.

The Archimedean Property can be leveraged to argue statements similar to the previous ones about infinity but for small numbers. For example, just as there is no infinitely large real number, so too there is no infinitesimally small (but still non-zero) real number.

***Exercise 1.34.** Use the statement of the Archimedean Property to argue that there is no non-zero "infinitesimal" real number.

While there is no infinitesimally small non-zero real number, there are *arbitrarily small* positive real numbers, as can be shown using the Archimedean Property. The phrase "goes to 0" in Theorem 1.30 can be understood in this sense. For example, consider bisecting the unit interval and retaining one half of it, and then iterating the process. This yields a sequence of nested closed intervals with lengths $\frac{1}{2^n}$. The following proposition shows that these lengths actually fo "go to 0".

***Proposition 1.35.** Given any number $x \geq 0$, if it has the property that for all $n \in \mathbb{N}$

$$x \leq \frac{1}{2^n},$$

then $x = 0$.

When you have completed a proof of this proposition, try to reframe it as a proof that "$\frac{1}{2^n}$ becomes arbitrarily small".

Interestingly, there are number systems that *do* have infinitesimal (and/or infinite) numbers—for example, the hyperreal numbers or the surreal numbers. For a mind-blowing experience, look these up!

*
* *

To conclude, the real numbers can be characterized as follows:

> The real numbers are an ordered field satisfying the Axiom of Completeness and the Archimedean Axiom.

1.1.5. Additional material. The following problems relate to the material of this section but are not central to developing the text narrative.

In Problem 1.22 you showed that $\sqrt{2}$ is irrational using a geometric argument about similarity of certain rectangles. Now let's try a different proof, based on the Fundamental Theorem of Arithmetic: any integer greater than 1 can be factored uniquely[12] as a product of prime numbers.

[12] Uniquely up to the order in which you write the factors.

***Problem 1.36.** Give a proof using the Fundamental Theorem of Arithmetic that $\sqrt{2}$ is not a rational number. That is, for all $x \in \mathbb{Q}$, $x^2 \neq 2$.

Proof (to be completed). Suppose, by way of contradiction, that there is an $x \in \mathbb{Q}$ such that $x^2 = 2 \ldots \star$ □

This argument is more powerful than just for solving this problem: it can be generalized. Generalizing an argument often involves taking a particular object with an obvious name (e.g., 2) and replacing it with a symbol that can represent a whole set of objects. This likely means you will need to introduce some notation in your argument. What do you think makes for good notation?

***Theorem 1.37.** If p is a prime number, then $\sqrt{p}$ is not a rational number.

In fact, even more is true!

***Theorem 1.38.** If n is a positive integer that is not a perfect square, then $\sqrt{n}$ is irrational.

We note that, taken in reverse order, these three theorems are a good example of finding an easier problem—also a very useful tool.

Taking roots of integers produces just a few of the infinitely many irrational numbers (for instance, π and e are two other well-known irrational numbers not produced in this way[13]). Despite the sense that rational numbers get as fine as one wishes and thus should fill out any measurement stick, there are good arguments that "most" numbers needed for measurement are *not* rational! An idea toward such an argument is described after Problem 1.40.

In Exercise 1.7 you showed that the rational numbers are "infinitely fine" (between any two rational numbers there is another rational number). Now try the following:

***Problem 1.39.** Argue that the rational and irrational numbers are "completely intertwined". That is,

(1) between any two distinct rational numbers there is an irrational number,$\star$ and

(2) between any two distinct irrational numbers there is a rational number.$\star$

***Problem 1.40.** The goal of this problem is to show that the Axiom of Completeness guarantees that any decimal expansion represents a real number.[14]

Recall place values for decimal numbers: given any infinite sequence of digits d_1, d_2, d_3, d_4, …, the corresponding decimal expansion is the "infinite sum"

$$\frac{d_1}{10} + \frac{d_2}{10^2} + \frac{d_3}{10^3} + \frac{d_4}{10^4} + \cdots = .d_1 d_2 d_3 \ldots \text{ (decimal)}.$$

But what does this mean, precisely? How do we know it is a real number? We are adding up infinitely many terms, and we saw in §1.1.4 that ∞ is not a real number. We need a clear analysis of this situation.

[13] But proving that they are irrational is much harder!

[14] The converse of this statement, that any real number has a decimal expansion, is addressed in §6.2.

Let $d_1, d_2, \ldots$ be *any* (infinite) sequence of digits—that is, for all n, d_n is an integer such that $0 \le d_n \le 9$. For each $k = 1, 2, \ldots$, define

$$s_k = \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_k}{10^k}$$

(a finite sum and so definitely a real number), and form the sequence of closed intervals

$$I_k = \left[s_k,\ s_k + \frac{1}{10^k} \right].$$

(1) Show that the sequence (I_k) is nested.

(2) Argue that there is *exactly one* number, w, such that $w \in I_k$ for *all* k.
Definition. The number represented by $.d_1d_2d_3\ldots$ (decimal) is that number w.

(3) Use this definition to show that

$$.\overline{9} = .99999\ldots = 1.$$

In §6.2, you will show that rational numbers are the ones whose decimal expansion eventually repeats (this includes "finite" decimal expansions, since those formally end with repeating 0's). With this in mind, suppose you roll a 10-sided dice infinitely many times, recording the results as $d_1d_2d_3\ldots$. What do you think the likelihood is of ending up with an infinitely repeating string? Those correspond to the rational numbers, which are, in these terms, exceedingly rare among the real numbers.

1.2. A first look at sequences

In the previous section, we generated numbers a_k that approximate $\sqrt{2} - 1$ and we found that as k gets larger, the approximation a_k gets better. In this section, we formally define such a list of numbers as a *sequence* and the notion of approximation as *convergence*. As we will see, the ideas of sequences and their convergence are very broadly applicable.

Definition 1.41. A sequence of real numbers (a_k) is a rule that associates to each positive integer k a real number a_k.

Another way to say this is that a sequence is a function $a : \mathbb{N} \to \mathbb{R}$ with domain the natural numbers $\mathbb{N}$. We will often call a sequence of real numbers simply "a sequence" (the "of real numbers" part being understood).

Example 1.42. Individual sequences can be defined in many ways. Here are a few:

(1) One can define a sequence by an explicit function of the index variable, such as $a_k = 1 + \frac{1}{k}$, or $b_k = \sin(1/k)$, etc.

(2) Or one can define a sequence recursively, for example

$$a_0 = 1, \quad a_{k+1} = \frac{1}{1 + a_k}.$$

(We have seen this kind of definition already.) In this case one does not have a "closed" or explicit formula for a_k just in terms of k, but one can compute a_k for any given k. For example:

$$a_0 = 1, \quad a_1 = \frac{1}{2}, \quad a_2 = \frac{1}{3/2} = \frac{2}{3}, \quad a_3 = \frac{1}{5/3} = \frac{3}{5}, \quad a_4 = \frac{1}{8/5} = \frac{5}{8}, \quad \ldots$$

and therefore the sequence is well-defined.

(3) Given a sequence of digits (d_n), $n = 1, 2, \ldots$ (that is, each d_n is a digit in $\{0, 1, \ldots, 9\}$, one can form the *sequence of decimal approximations*

$$a_k = \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_k}{10^k}.$$

This is an example of an important class of sequences called *series*. For example, if $d_n = 1$ for all n, then

$$a_k = \frac{1}{10} + \frac{1}{10^2} + \cdots + \frac{1}{10^k}.$$

(4) One can also define a sequence simply by listing numbers:

$$1,\ 2,\ 3,\ 5,\ 8,\ 13,\ \ldots$$

In this case, a_k is the kth number listed. While sequences are not required to follow recognizable patterns, we do need a_k to represent a well-defined number for each integer $k = 1, 2, \ldots$. Since one cannot actually write an infinite list, in examples like the one above we need to be able to recognize a pattern in order for the "…" to completely define a sequence.

Note that any mechanism that generates an infinite list of numbers provides a sequence, even if the sequence cannot be represented by some formula or does not have a recognizable pattern.

An important characteristic of a sequence is whether it converges or not. We have developed some intuition of convergence as "approximating better and better" some real number, in the same sense as the iterates $a_{k+1} = \frac{1}{2+a_k}$ starting at $a_0 = 0$ approximate $\sqrt{2} - 1$ in Investigation 1.23. If a sequence converges, we will call the number it approximates the "limit".

There are problems with relying on this (or any) intuition, though. How do you know that what you mean by "approximates" is the same thing as what your classmates mean? Could you agree on what the behavior of a given sequence is? Could you *prove* that a certain sequence converges or does not converge? If we want to build a mathematical theory on top of a definition, it had better be a solid one. The next few activities guide you through taking an intuitive but subtle idea and using it to create a *mathematical definition*.

The following investigation is very important: you will use it to begin the process of creating a good definition of convergence. One criterion for a good definition is that

it captures the notion you intend to define. A good way to test this is to apply it to, or check it against, a variety of examples.[15]

***Investigation 1.43.** The following are examples of sequences. In each case determine whether you believe the sequence converges, and if so what you expect the limit to be (expressed as a real number). You can use a calculator to experiment, but you should support your assertion with a sentence or two.

(1) $a_k = 1 + \dfrac{1}{k}$

(2) $b_k = (-1)^k + \dfrac{1}{k}$

(3) $c_k = \dfrac{1}{1} + \dfrac{1}{2} + \dfrac{1}{3} + \cdots + \dfrac{1}{k}$

(4) $s_k = \dfrac{1}{10} + \dfrac{1}{10^2} + \cdots + \dfrac{1}{10^k}$

(5) $t_{k+1} = \dfrac{1}{1 + t_k}$, $t_0 = 1$

(6) $z_k = \dfrac{2k + 1}{3k - 10}$

Note that the process of creating a definition involves revisiting these examples in detail. Your understanding of certain sequences may change during this process—this is not a deficiency! The process of revision applies not only to written words but also to the understanding of ideas. Our goal in this process is to converge, so to speak, on a well-written definition of a well-defined idea.

Toward that goal, here are three characteristics of a good definition.

(1) As we have already mentioned, a good definition should *function the way you want it to*. Check whether the proposed definition captures the notion you intend to define by applying it to a variety of examples.

(2) A good definition should be *precise*. Consider whether the proposed definition leaves no room for different interpretations—i.e., it is "impossible to misunderstand". If you give the definition to someone else, could they have any valid misinterpretations of what the definition is supposed to say?

(3) A good definition should be *concise*. Is there any way you can condense the proposed definition? Is every part necessary? Interpretive or descriptive language should be relegated to writing about the definition and not be included in the definition itself.[16]

Here are some first attempts at writing a definition of convergence.

***Investigation 1.44.** Consider the following attempted definitions of convergence. Critique each of them against the three characteristics of a good definition given above.

[15] When a researcher develops a survey designed to measure a particular thing they are interested in, they need to test it to make sure it measures the concept they designed it to measure. This is part of *validating* a survey. Testing your definition against examples is a similar form of *validation*.

[16] For example, compare and contrast the following two definitions of isosceles triangles: (a) An isosceles triangle is one having two congruent sides. (b) An isosceles triangle is one having two congruent sides and two congruent angles. Which one is better?

(1) A sequence converges if it gets closer to a number.
(2) A sequence (a_k) converges if $a_{k+1} - a_k$ goes to zero.
(3) A sequence (a_k) converges if it gets closer to a number and does not continue to have large jumps. The size of any jumps should be getting smaller.
(4) A sequence (a_k) converges to L if it gets infinitesimally close to L.
(5) A sequence (a_k) converges to L if $|a_k - L| = 0$ for large k.
(6) A sequence (a_k) converges to L if $|a_k - L| < \frac{1}{10}$ for all k.
(7) A sequence (a_k) converges to L if there is a very small number ϵ such that $|a_k - L| < \epsilon$ for a large k.

The following investigation is meant to be a bit messy: you should expect to need to go through a number of rounds of revision. Part of your process should involve getting a peer to interpret and use your definition.

***Investigation 1.45.** Try writing down a definition of what it means for a sequence to converge. You may start with a definition from the previous investigation and revise, or come up with your own version. Make sure to check the result against the three criteria above and revise as needed.

Definition of convergence. We say that a sequence (a_k) converges to a real number L iff...

We note that there are some notational conventions that are very widely followed in the mathematical community. For example, the ϵ used in Investigation 1.44 part (7) is a standard label for a "very small number". Your instructor can help you find appropriate notation for your definition.

***Problem 1.46.** Here goes a first exercise on the definition. Show that

$$\lim_{k\to\infty} \frac{1}{k} = 0.$$

Make sure you follow your definition!

Having now used your definition, does it work the way you want it to? If not, go back and revise.

The following problem asks you to now show that a sequence does *not* converge. Think of what you would need to show in order to have a counterexample to your definition of convergence.[17] Would it suffice to show that for some fixed notion of "close", there is a point of the sequence that is not "close" to the proposed limit L? Or do multiple points (or all points?) of the sequence have to fall outside of "close"? For

[17] A more thorough treatment of negation and quantifiers can be found in Appendix A.

one fixed L, or for any L? Can you work with just one fixed version of "close", or do you need to be able to argue this for any notion of "close"?

***Problem 1.47.** Consider the sequence $a_k = (-1)^k$.

(1) Show that (a_k) does *not* converge to 1. Similarly, show that it does not converge to (-1).

(2) Show that (a_k) does not converge to any real number (it "does not converge").

Here is another family of examples:

***Problem 1.48.** Pick a real number r, and define the sequence $a_k = r^k$ for $k = 1, 2, \ldots$. Does this sequence converge? Why or why not? Does the answer depend on the value of r? If the sequence does converge, can you say anything about its limit?

A sequence that does not converge is said to *diverge*. A sequence can diverge for two different reasons: it might wander around the real numbers and never settle on any number in particular (like the sequence in Problem 1.47), or it might wander off "to infinity" (like some of the sequences in Problem 1.48). A sequence that does the latter is said to "diverge to infinity", which can be defined in a similar spirit to the definition of a convergent sequence.

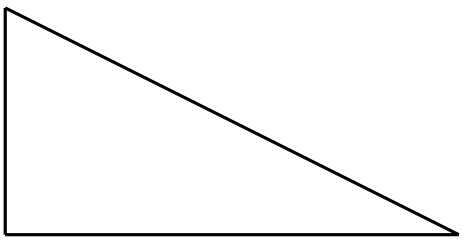
***Definition 1.49.** The sequence (a_k) diverges to infinity, written $\lim_{k\to\infty} a_k = \infty$, iff _____.

***Problem 1.50.** Show that $a_k = k + (-1)^k$ diverges to infinity.

You may wish to compare your definition of convergent sequence with the one given at the very end of Appendix D. In this text, we will proceed based on the definition given in the appendix.

1.3. A first look at series

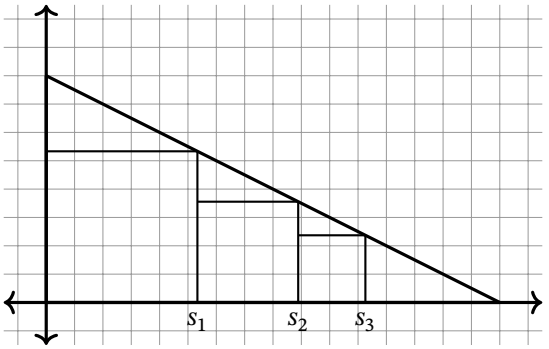
***Investigation 1.51.** We are going to analyze the following "doodle game".[18] Start with a right triangle like this one:



Then inscribe the largest *square* that fits inside the triangle, and then the largest square that fits in the triangle to the right of the first square, and so on, as in the following

[18] Take a look at Vi Hart's video `https://www.youtube.com/watch?v=DK5Z709J2eo` for doodling in action! It gets a bit out of hand, and leads to even more mathematics, but Vi returns to series at the end.

picture. The doodle game never ends; we can fit in infinitely many squares! We have put the picture on a coordinate grid because we are going to analyze it algebraically.



Suppose the hypotenuse is on the line $y = -mx + a$, with $m > 0$, and let $s_1, s_2, s_3, \ldots$ be the x coordinates of the lower right corners of the squares, as shown.

(1) Find a formula for s_1 in terms of m and a.

(2) Let $b_1, b_2, b_3, \ldots$ be the sequence of side lengths of the squares.
   (a) Find a formula for b_k.
   (b) Express s_n in terms of $b_1, \ldots, b_n$, and then in terms of m and a.
   (c) What is the ratio $\dfrac{b_{k+1}}{b_k}$ (in terms of m and a)?

(3) Does the sequence (s_n) converge? If so, what is $\lim\limits_{n\to\infty} s_n$?

In this investigation there are a lot of related but distinct objects to keep track of! Here are some definitions, to help organize them:

Definition 1.52. Let (b_k) be a sequence. The associated *infinite series* is the formal expression[19]

$$\sum_{k=1}^{\infty} b_k = b_1 + b_2 + b_3 + b_4 + \cdots$$

and the b_k are called the terms (or summands) of the series. Given (b_k), we define a related sequence (s_n), called the *sequence of partial sums*, by

$$s_n = b_1 + b_2 + \cdots + b_n = \sum_{k=1}^{n} b_k$$

for each $n = 1, 2, \ldots$. We say that the series $\sum\limits_{k=1}^{\infty} b_k$ *converges to a real number* S iff the sequence (s_n) converges to S, in which case we write $\sum\limits_{k=1}^{\infty} b_k = S$.

[19] A *formal expression* is one that we can write down, but it may not represent a real number.

There exist classes of series having their own specialized tools for determining if a series in the class converges. The following is an example is the class of geometric series.

Definition 1.53. Given real numbers c and r, the series

$$\sum_{k=1}^{\infty} c\, r^k$$

is called a *geometric series*. (Note that we are starting with the index $k = 1$, so the first term in this series is cr.)

***Problem 1.54.** Show that a series $\sum_{k=1}^{\infty} a_k$ is geometric if and only if the ratio of consecutive terms $\frac{a_{k+1}}{a_k}$ is independent of k. If that is the case, what are c and r?

The following theorem characterizes completely the convergence of geometric series. Prove it by using the setup of Investigation 1.51 and translating everything in the investigation into the geometric series notation of Definition 1.53 (i.e., using the variables r and c). (You may wish to first assume that $r > 0$, but you can then check that this assumption is not necessary.)

***Theorem 1.55.** A geometric series $\sum_{k=1}^{\infty} cr^k$ converges if $|r| < 1$, in which case

$$\sum_{k=1}^{\infty} c\, r^k = \frac{cr}{1-r}.$$

If $|r| \geq 1$, then the sequence (s_k) of partial sums diverges, and so the formal expression $\sum_{n=1}^{\infty} c\, r^n$ has no meaning as a real number.

***Problem 1.56.** Another common way to index a geometric series is starting with $k = 0$ instead of 1. Show that the formula you proved in the theorem above for $|r| < 1$ is equivalent to the statement

$$\sum_{k=0}^{\infty} c\, r^k = \frac{c}{1-r}.$$

It is useful to note that in each case, *the numerator on the right-hand side of the formula is the first term of the series.*

Another example of series with a particularly nice convergence result is the class of "telescoping" series.

***Problem 1.57.** Here's a clever idea that allowed Leibniz to find the limit of many series. His idea was that, in order to find the real number $\sum_{k=1}^{\infty} a_k$ (if it exists), one should look for a sequence (b_k) such that

$$a_k = b_k - b_{k+1} \qquad \text{for all } k.$$

(Unfortunately, it is often very hard or impossible to find such a sequence (b_k)—but when one can, it's a jewel!)

(1) As an example, use the fact that $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$ to evaluate $\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$.⋆

(2) Explain how, in general, knowledge of (b_k) with the property above helps one to compute the partial sums s_n and understand the behavior of the series $\sum_{k=1}^{\infty} a_k$.

(3) Use this method to find the value of $\sum_{k=5}^{\infty} \frac{1}{k^2 - 5k + 6}$.⋆

1.4. Properties of sequences and series

In this section we will investigate some other properties sequences can have and their relationship with convergence.

Definition 1.58. A sequence (a_k) is called

- *monotone increasing* if $a_k \leq a_{k+1}$ for all k;
- *monotone decreasing* if $a_k \geq a_{k+1}$ for all k;
- *monotone* if it is monotone increasing or monotone decreasing;
- *bounded* if there exist real numbers $m \leq M$ such that $m \leq a_k \leq M$ for all k; such an M (resp. m) is called an upper (resp. a lower) bound of the sequence.

Let's get a feel for these definitions by applying them to examples.

***Exercise 1.59.** For the following sequences, determine which are monotone (decreasing or increasing) and which are bounded.

(1) $a_n = 1 + \frac{1}{n}$

(2) $b_k = (-1)^k \frac{k}{k+1}$

(3) $c_k = \sin(k^2)$

(4) $d_n = \frac{n^2+1}{n}$

(5) $e_n = 10n - \frac{\sin^2(n)}{n}$

One relationship between convergence and the properties of Definition 1.58 is the following:

***Theorem 1.60.** If the sequence (a_k) converges, then it is bounded.⋆

The reverse implication is not true, however:

***Problem 1.61.** Write down the converse[20] of Theorem 1.60. Show that this statement is *false*.

[20] See Appendix A for a discussion of the converse of a statement, as well as other logically related statements.

While neither boundedness nor monotonicity on its own implies convergence, *together they do imply convergence*. This is a very important result, and we will often use it as a tool.

Monotone Convergence Theorem. If a sequence (a_k) is monotone and bounded, then it converges.

We present a proof of the Monotone Convergence Theorem (MCT, for short) in its entirety. Observe how bisection, iteration, and nested intervals work together in this proof; we will ask you to use these tools in similar situations later on.

Proof. Let's suppose that (a_k) is monotone increasing and bounded above by M (the proof is similar if (a_k) is monotone decreasing and bounded below by m).

> *The idea of our proof is to construct a sequence of nested intervals whose lengths tend to zero, and invoke the Axiom of Completeness to get a real number L that should be the limit. Then we will prove that indeed* $\lim_{k\to\infty} a_k = L$.

Let $I_1 = [a_1, M]$, and let $c_1 = \frac{a_1+M}{2}$ be the midpoint of I_1. In order to construct I_2, we consider two cases:

CASE 1: c_1 is also an upper bound for (a_k). Then let $I_2 = [a_1, c_1]$.
CASE 2: c_1 is not an upper bound for (a_k); then there must exist a natural number K such that $c_1 < a_K$. Let $I_2 = [a_K, M]$.

In either case,

(1) $I_2 \subseteq I_1$,

(2) the lengths of the intervals satisfy $|I_2| \leq \frac{1}{2}|I_1|$, and

(3) the left endpoint of I_2 is an element of the sequence (a_k) and the right endpoint of I_2 is an upper bound of (a_k). Let's write $I_2 = [a_{k_2}, M_2]$.

Now repeat: Let c_2 be the midpoint of I_2. There are two possibilities for c_2: it is either an upper bound for the sequence or it is not. Using exactly the same procedure as above, define a closed interval I_3. Continuing this procedure yields a sequence of closed intervals $(I_n)_{n=1}^{\infty}$ with the following properties:

(1) $I_{n+1} \subseteq I_n$;

(2) $|I_{n+1}| \leq \frac{1}{2}|I_n| \leq \frac{1}{2^n}|I_1|$; and

(3) $I_n = [a_{k_n}, M_n]$, where a_{k_n} is an element of the sequence (a_k) and M_n is an upper bound of (a_k).

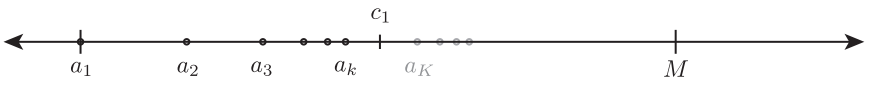


Figure 1.1. A bounded, monotone increasing sequence (a_k). In the proof above, $I_1 = [a_1, M]$ for some upper bound M. If the midpoint c_1 is also an upper bound, then $I_2 = [a_1, c_1]$. If c_1 is not an upper bound, then $I_2 = [a_K, M]$ for some $a_K > c_1$.

By properties (1) and (2), we have a nested sequence of closed intervals whose lengths go to zero. Therefore, by the Axiom of Completeness, there exists a unique real number L such that $L \in I_n$ for every n. We claim that L is the limit of the sequence!

Here is the proof that $\lim_{k\to\infty} a_k = L$. Let $\epsilon > 0$. Pick N sufficiently large so that the length of $I_N = [a_{k_N}, M_N]$ is less than ϵ. Then, for each $k \geq k_N$,

$$a_{k_N} \leq a_k \leq M_N$$

because (a_k) is monotone increasing and M_N is an upper bound for the sequence. Therefore $a_k \in I_N$. Additionally, $L \in I_N$ by the definition of L. Since both $a_k, L \in I_N$, it must be that

$$|a_k - L| \leq |I_N| < \epsilon.$$

□

In fact, this proof does a nice job of "finding" the limit of a sequence: once you supply an upper bound, the process of the proof "homes in" on the actual limit of the sequence. Compare this with the definition of convergence, where one has to know what the limit is beforehand.

The following problem asks you to run the machinery of the previous proof for a specific sequence.

***Problem 1.62.** Let

$$a_k = 1 - \frac{1}{k} = \frac{k-1}{k},$$

and let (I_n) be the nested sequence of closed intervals constructed by the bisection method as in the proof above.

(1) Explicitly construct the intervals I_n for $n \leq 5$ starting with the upper bound $M_1 = 3$.

(2) Repeat, starting with the upper bound $M_1 = 2$.

(3) Repeat, starting with the upper bound $M_1 = 1$.

In each case you should see how the I_n are "homing in" on the limit of the sequence, $L = 1$.

While following the proof of the MCT generates approximations of the limit of a monotone and bounded sequence, later on it will be important to have a different way to think about the limit:

***Problem 1.63.** Assume that (a_k) is bounded and monotone increasing, and denote its limit by L.

(1) Show that L is an upper bound of the sequence.⋆

(2) Any bounded sequence has *many* upper bounds. For instance, show that if M is an upper bound, then so is any number $R \geq M$.

(3) Show that L is less than or equal to any other upper bound of the sequence (a_k).⋆

This shows that L is the *least upper bound* of the sequence.[21]

*
* *

The remainder of this section is devoted to several "convergence tests" for series. These are theorems which can be used as tools to prove (or disprove) that a given series is convergent.

The first is a corollary of a result we have already seen (though it is up to you to figure out which one!).

***Corollary 1.64.** Let (a_k) be a sequence such that $a_k \geq 0$ for all k. If the sequence (s_n) of partial sums

$$s_n = a_1 + \cdots + a_n$$

is bounded, then the series $\sum_{k=1}^{\infty} a_k$ converges.

The next two theorems comprise a set of comparisons between two series.

***Theorem 1.65.** Consider two series $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} b_k$, where $0 \leq b_k \leq a_k$ for all k. Then if $\sum_{k=1}^{\infty} a_k$ is convergent, $\sum_{k=1}^{\infty} b_k$ is convergent as well and

$$\sum_{k=1}^{\infty} b_k \leq \sum_{k=1}^{\infty} a_k.$$

With the same setup, there is another comparison to make. (Recall your Definition 1.49 of a sequence diverging to infinity.)

***Theorem 1.66.** Under the same hypothesis as in Theorem 1.65, if $\sum_{k=1}^{\infty} b_k$ diverges to infinity then $\sum_{k=1}^{\infty} a_k$ diverges to infinity as well.

The next two problems address an important family of examples: the series $\sum_{k=1}^{\infty} \frac{1}{k^p}$ (for $p \in \mathbb{N}$), which diverges if $p = 1$ and converges if $p \geq 2$.

[21] We will revisit this concept in §2.2.1.

***Problem 1.67.** (1) In Problem 1.57, you showed that $\sum_{m=1}^{\infty} \frac{1}{m^2+m} = 1$. Use this result to show that

$$\sum_{k=2}^{\infty} \frac{1}{k^2-k} = 1.$$

(Note that the first sum starts at $m = 1$ while the second sum starts at $k = 2$.)

(2) Use the results from Theorem 1.65 and part (1) of this problem to show that $\sum_{k=1}^{\infty} \frac{1}{k^2}$ converges, and give an upper bound on its value.

(3) Show that $\sum_{k=1}^{\infty} \frac{1}{k^p}$ is convergent if $p \geq 2$.[22]

***Theorem 1.68.** $\sum_{k=1}^{\infty} \frac{1}{k} = \infty$ (i.e., the sequence of partial sums $s_n = \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n}$ diverges to infinity).★

The next theorem and problem address how convergence of a series is related to convergence of the sequence of summands.

***Theorem 1.69.** If $\sum_{k=1}^{\infty} a_k$ converges, then $\lim_{k\to\infty} a_k = 0$.★

***Problem 1.70.** State the converse of Theorem 1.69. Decide whether this statement is true or false, and give an argument for your claim.

We end by stating, without rigorous proof (but with some comments), two more theorems about convergence of series.

Theorem 1.71. *If $\sum_{k=1}^{\infty} |a_k|$ converges, then $\sum_{k=1}^{\infty} a_k$ converges.*

Informal argument: Why should this be true? Well, the fact that the sum of the absolute values of *all* the terms converges implies that the sum of only the positive a_k will converge (say, to $P \geq 0$), as will the sum of only the negative a_k (say, to $N \leq 0$). Since

$$\sum_{k=1}^{\infty} a_k = \sum(\text{positive } a_k) + \sum(\text{negative } a_k) = P + N,$$

this means that $\sum_{k=1}^{\infty} a_k$ converges. Warning: This is *not* a formal proof because we don't know that we can "split and rearrange an infinite sum". In fact, in general, one cannot!

Theorem 1.72. *If $a_k > 0$ and $b_k > 0$ for all k and*

$$\lim_{k\to\infty} \frac{a_k}{b_k} = c > 0,$$

[22] In fact it converges if and only if $p > 1$, but we are not equipped to deal with non-integer exponents yet.

then the series $\sum_{k=1}^{\infty} a_k$ *converges if and only if* $\sum_{k=1}^{\infty} b_k$ *converges* (*so the series are either both convergent or both divergent*).

Informal argument: Note that the hypothesis $\lim_{k\to\infty} \frac{a_k}{b_k} = c > 0$ is a symmetric relationship between the two sequences (a_k) and (b_k): since $c \neq 0$ one also has that

$$\lim_{k\to\infty} \frac{b_k}{a_k} = \frac{1}{c} > 0.$$

This relationship roughly means that these two sequences "behave the same way at infinity"—that is, for large enough indices k, we have $a_k \approx c \cdot b_k$ for some real number c. Therefore, it should not be surprising that the behavior (convergence or divergence) of the series $\sum a_k$ and $\sum b_k$ is the same.

Now that we have a bunch of tools for showing convergence/divergence of series, it is time to put them to use.

***Problem 1.73.** Determine if each of the series below is convergent, divergent to infinity, or simply divergent. Compute, when you can, the limit of the convergent series (this is not always possible even if you know the series converges). Prove your answers by appealing to the theorems above.

(1) $\sum_{n=5}^{\infty} (0.25)^n$

(2) $\sum_{n=1}^{\infty} \frac{n}{n+1000000}$

(3) $\sum_{n=1}^{\infty} \frac{n+1}{n^3+1}$

(4) $\sum_{n=0}^{\infty} (-\frac{1}{3})^n \frac{n}{n^2+1}$

(5) $\sum_{n=1}^{\infty} \sqrt{\frac{1}{1+n}}$

1.5. Subsequences and the Bolzano–Weierstrass Theorem

In the previous section, we observed that while all convergent sequences are bounded, a bounded sequence may not be convergent. However, that does not mean a bounded sequence can have any kind of behavior—for example, such a sequence cannot diverge to infinity. Examples of bounded but divergent sequences include

$$a_k = (-1)^k,$$

$$b_k = (-1)^k + \frac{1}{k},$$

$$c_k : 1, 2, 3, 1, 2, 3, \ldots.$$

These sequences do not converge because they "oscillate" between different points (or intervals) on the number line. In particular, these sequences contain "subsequences" that *do converge*. Does this happen for every bounded sequence? It may be clear for the sequences above, but it is much less obvious for a sequence such as

$$d_k = \sin(k^2).$$

To study this question we first need to introduce the notion of a subsequence. This term may be intuitive enough, but as we have seen before, it is important to formulate a good, precise definition. As a motivating example, consider the sequences

$$a: \quad \frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \cdots$$

and

$$b: \quad \frac{1}{1}, \frac{1}{3}, \frac{1}{5}, \frac{1}{7}, \frac{1}{9}, \ldots.$$

Clearly the second one is "contained" in the first. More precisely, the terms of the second sequence are every other term of the first. Using formulas, these sequences are

$$a_k = \frac{1}{k} \quad \text{and} \quad b_n = \frac{1}{2n-1},$$

and the b sequence is a subsequence of a because, for all $n = 1, 2, \ldots$,

$$b_n = a_{2n-1}.$$

This motivates the following definition:

Definition 1.74. Let (a_k) be a sequence. A *subsequence* of (a_k) is any sequence (b_n) of the form

$$b_n = a_{k_n}$$

where (k_n) is a strictly increasing sequence of integers

$$k_1 < k_2 < \cdots.$$

Example 1.75. If $a_k = \frac{1}{k} + (-1)^k$ (one of our first examples), then the sequence $b_n = 1 + \frac{1}{2n}$ is a subsequence, since if we let $k_n = 2n$ then we get $b_n = a_{k_n}$. Notice that (b_n) is convergent while (a_k) is not!

***Problem 1.76.** Use the definition of a subsequence to justify your answers to the following:

(1) Explicitly write two subsequences (b_n) and (c_n) of the sequence $a_k = \frac{2^k}{2^k+1}$.

(2) For which values of p is $b_n = \frac{1}{n^p}$ a subsequence of $a_k = \frac{1}{k^3}$?

(3) If (b_m) is a subsequence of (a_k) and (c_n) is a subsequence of (b_m), how are (a_k) and (c_n) related?

What happens if we take a subsequence of a *convergent* sequence? In this case, *all* subsequences converge to the same limit:

***Theorem 1.77.** Let (b_n) be a subsequence of (a_k). If the sequence (a_k) converges to L, then (b_n) also converges to L.

***Problem 1.78.** State the converse of Theorem 1.77. Decide whether this statement is true or false and give an argument for your claim.

The following theorem answers our question from the introduction to this section: any bounded (not necessarily convergent) sequence is guaranteed to have a subsequence that converges. As was the case in the Axiom of Completeness, the indefinite article "a" here indicates *at least one*. That is, the bounded sequence may have multiple convergent subsequences, and these subsequences may even converge to different limits.

Bolzano–Weierstrass (B-W) Theorem. Every bounded sequence contains a convergent subsequence.

The Bolzano–Weierstrass Theorem, like the Monotone Convergence Theorem, can be proven with the tools of bisection, iteration, and nested intervals. The following provides some scaffolding to help you carry out the proof.

***Proof (to be completed).** Let (a_k) be a bounded sequence, so there exist $m, M \in \mathbb{R}$ such that $m \leq a_k \leq M$ for all k. Consider the closed interval $[m, M]$, and let $c = \frac{m+M}{2}$ be the midpoint of the interval. [*Explain why at least one of the closed intervals* $[m, c]$ *and* $[c, M]$ *contains infinitely many* a_k.[23]]

Pick one of the intervals that contains infinitely many a_k, and call it I_1. [*Now write down a procedure that defines a sequence of closed intervals, and argue that* (1) *they are nested and* (2) *the lengths of these intervals go to* 0.]

[*Finally, write down a subsequence of* (a_k) *that converges and prove that it converges.*]$\star$ □

What do you think happens in the proof above if (a_k) has multiple subsequences that converge to different limits? The following problem provides an example of a sequence that has a particularly rich set of subsequences.

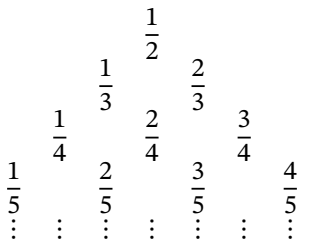
***Problem 1.79.** Consider the following infinite pyramid of rational numbers:

(1) Describe the pattern in the pyramid and show that every rational number in $(0, 1)$ appears in it at some point (in fact every rational number appears multiple times).

(2) Let $a = (a_k)$ denote the sequence obtained by "reading" the numbers in the table starting at the top. That is, the sequence begins with the string of numbers

$$\frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \ldots.$$

[23] More precisely, for at least one of these intervals, I, the set $\{k \in \mathbb{N} \mid a_k \in I\}$ is infinite



According to the Bolzano–Weierstrass Theorem this sequence has a convergent subsequence. Explicitly find three convergent subsequences, two of which are not constant, all three with different limits.

(3) Show that for *any* $L \in (0,1)$ there is a subsequence of (a_k) that converges to L.★

We end this section with a problem exploring other relationships between properties of sequences and subsequences.

***Problem 1.80.** Give an example of each of the following, or prove that no such example exists:

(1) An unbounded sequence with a convergent subsequence.

(2) An unbounded monotone sequence with a convergent subsequence.

(3) A sequence with a bounded subsequence but without a convergent subsequence.

1.6. Sequences and convergence in higher dimensions

A sequence of real numbers can be thought of geometrically as an ordered list of points on the line. But we can also think of points in spaces other than the line—for example, on the plane or in three-dimensional space. Can points in spaces other than the real line be used to make a sequence? In this section, we consider sequences in the plane ($\mathbb{R}^2$) and in three-dimensional space ($\mathbb{R}^3$). In fact, there is no need to stop our generalizations at three dimensions, and so we will consider sequences in $\mathbb{R}^d$ for any natural number $d \geq 1$.

This notation also gives us a way to be concise in our statements—$\mathbb{R}^d$ for $d \geq 1$ includes the cases of the real number line ($d = 1$), the plane ($d = 2$) and three-dimensional space ($d = 3$). We will state results as generally as possible (i.e., often in $\mathbb{R}^d$). Proofs and problems can also be completed in full generality, but if you wish for less abstraction, you can complete this section substituting $d = 2$ and considering only sequences on the plane.

Definition 1.81. $\mathbb{R}^d$ is the set of ordered d-tuples of real numbers,

$$\mathbb{R}^d = \{X = (x_1, x_2, \ldots, x_d) \mid x_i \in \mathbb{R}\},$$

with $X = Y$ if and only if $x_i = y_i$ for all $i = 1, \ldots, d$ (the order matters!). In $\mathbb{R}^2$ we often use x for the first real number and y for the second:

$$\mathbb{R}^2 = \{(x, y) \mid x \in \mathbb{R} \text{ and } y \in \mathbb{R}\},$$

where $(x, y) = (a, b)$ iff $x = a$ and $y = b$.

In $\mathbb{R}^2$, the formula for the distance from $A = (a_1, a_2)$ to $B = (b_1, b_2)$ is determined by Pythagoras' Theorem:

$$d(A, B) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2}.$$

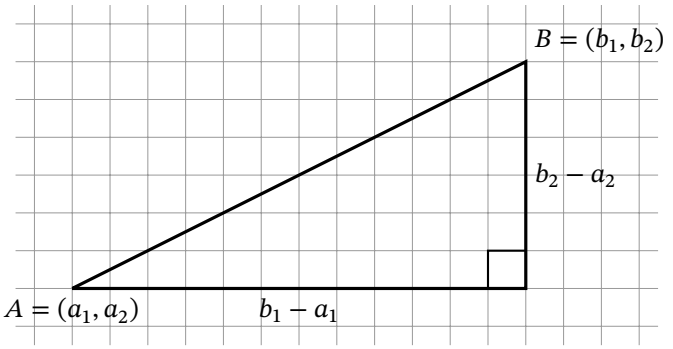


Figure 1.2. The distance formula.

This generalizes to higher dimensions $\mathbb{R}^d$, where the distance formula is

$$d(X, Y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_d - y_d)^2}.$$

We have already used the distance between two points x and y in $\mathbb{R}^1$, where

$$d(x, y) = \sqrt{(x - y)^2} = |x - y|.$$

***Exercise 1.82.** How would you compute the distance between two opposite corners of a cereal box (e.g., front-lower-left to back-upper-right)? Does it agree with the distance formula above?

The notion of distance allows us to define the *open ball of radius r and center O* as all of the points within distance r of the point O:

$$B_r(O) = \{ X \in \mathbb{R}^d \mid d(X, O) < r \}.$$

***Exercise 1.83.** In $\mathbb{R}^2$, the set $B_r(O)$ is often called an *open disk*. Show that if $O = (a, b)$, $B_r(O)$ is the same as the set $\{ (x, y) \mid (x - a)^2 + (y - b)^2 < r^2 \}$.

***Exercise 1.84.** What geometric object is $B_r(O)$ in dimension $d = 3$? Give a characterization as in Exercise 1.83 of the points (x, y, z) that are in $B_r(O)$.

***Problem 1.85.** The *triangle inequality*, valid in all dimensions, is the statement that for any three points A, B, $C \in \mathbb{R}^d$ one has

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

Argue geometrically why this must be true. Can you prove it algebraically?

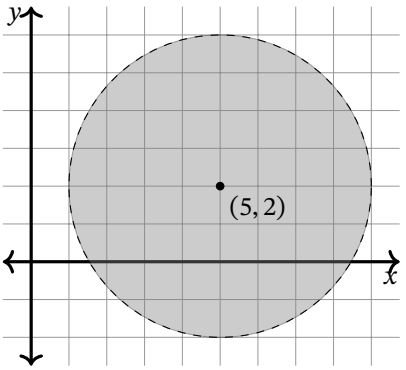


Figure 1.3. The shaded region (without the circumference) is the set of all (x, y) such that $(x-5)^2 + (y-2)^2 < 16$.

1.6.1. Sequences and convergence. Next we define sequences in any dimension and their convergence:

Definition 1.86. A sequence on $\mathbb{R}^d$ is a rule that assigns to each natural number k a point P_k in $\mathbb{R}^d$. If (P_k) is a sequence of points in $\mathbb{R}^d$ and $Q \in \mathbb{R}^d$, we say that

$$\lim_{k\to\infty} P_k = Q$$

iff the distance from P_k to Q goes to zero—i.e.,

$$\lim_{k\to\infty} d(P_k, Q) = 0.$$

This definition says that (P_k) converges to Q if a particular sequence of real numbers (the distances between P_k and Q) converges to 0. We can also write an equivalent definition referencing only objects in $\mathbb{R}^d$:

***Exercise 1.87.** Fill in the following blanks to complete an equivalent definition of convergence in $\mathbb{R}^d$. (There is more than one correct way to do it—you do not have to use both blanks, and the length of each blank does not indicate how many words you might use.)

If (P_k) is a sequence of points in $\mathbb{R}^d$ and Q is a point in $\mathbb{R}^d$, we say that

$$\lim_{k\to\infty} P_k = Q$$

iff _____ the points P_k lie in the ball of center Q and radius ϵ _____.

***Problem 1.88.** For each of the following sequences on the plane $\mathbb{R}^2$, sketch the first several points and determine (informally) whether the sequence converges:

(1) $P_k = \left(1 + \frac{1}{k}, \frac{1}{k^2}\right)$

(2) $Q_k = \left(\frac{1}{k}, k\right)$

(3) $R_k = \left(\frac{1}{k}, \frac{1}{k}\sin(k)\right)$

While working on the previous problem, you very well may have considered the coordinate sequences separately and used their limits (as sequences of real numbers) to put together the limit point of the sequence in $\mathbb{R}^2$. Note that this is not the definition of convergence in $\mathbb{R}^2$, which indicates that you need to use the $\mathbb{R}^2$ distance formula! But, in fact, this is a fine strategy, as the following theorem justifies.

***Theorem 1.89.** A sequence of points in $\mathbb{R}^d$ converges to Q if and only if the sequences of the coordinates of the points in the sequence converge to the corresponding coordinates of Q.

Explicitly, in two dimensions, let (P_k) be a sequence of points on the plane with $P_k = (a_k, b_k)$. Then

$$\lim_{k\to\infty} P_k = Q = (\ell, m)$$

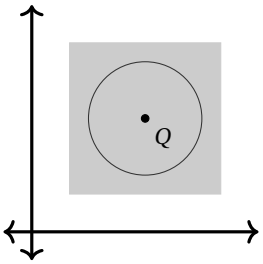
if and only if

$$\lim_{k\to\infty} a_k = \ell \qquad \text{and} \qquad \lim_{k\to\infty} b_k = m.$$

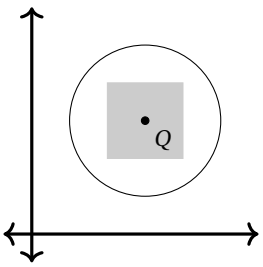
We will prove the theorem in two dimensions; in higher dimensions it follows similarly (but with more notation).

Proof (to be completed). This is an "if and only if" statement, so we break it into two parts.

(a) Assume that $\lim_{k\to\infty} P_k = Q$. [*Write out what you need to show, and consider how the figure below* (*along with the assumption*) *helps.*]



(b) Assume that $\lim_{k\to\infty} a_k = \ell$ and $\lim_{k\to\infty} b_k = m$, with $(\ell, m) = Q$. [*Write out what you need to show and consider how the figure below* (*along with the assumption*) *helps.*]



□

1.6.2. The Bolzano–Weierstrass Theorem in higher dimensions. How do subsequences work in higher dimensions? Do "bounded" sequences have convergent subsequences? First, we must define what "bounded" means.

***Definition 1.90.** A sequence (P_k) in $\mathbb{R}^d$ is *bounded* iff _____.

The following theorem is the B-W Theorem for higher dimensions. You might start with proving the theorem in the $d = 2$ case before trying to prove it in higher dimensions.

***Theorem 1.91** (Bolzano–Weierstrass Theorem in $\mathbb{R}^d$)**.** Every bounded sequence in $\mathbb{R}^d$ has a convergent subsequence.⋆

1.6.3. Additional material. We can generalize more of the ideas from previous sections of this chapter to higher dimensions:

***Definition 1.92.** A sequence (P_k) in $\mathbb{R}^d$ diverges to infinity, written $\lim_{k\to\infty} P_k = \infty$, iff _____.

***Problem 1.93.** Define the notion of "unbounded sequence" in $\mathbb{R}^d$, and compare it to the notion of "diverging to infinity".

We end this section with a new result. The result will not be used in subsequent chapters in this book, but it is a "big deal" result in analysis.

We begin with a definition:

Definition 1.94. A sequence (P_k) in $\mathbb{R}^d$ is a *Cauchy sequence* iff for all $\epsilon > 0$ one has

$$d(P_j, P_k) < \epsilon \qquad \text{for all sufficiently large } j \text{ and } k.$$

That is, a Cauchy sequence is one where the terms are getting arbitrarily close to each other.

***Theorem 1.95.** A sequence in $\mathbb{R}^d$ is convergent if and only if it is a Cauchy sequence.

This is a "big deal" theorem because the property of being Cauchy does not make reference to any potential limit. So, if one can show that a given sequence is Cauchy, one can conclude that it is convergent without having to guess what the limit is!

Proof (to be completed). (a) [*If a sequence converges, then its terms are approaching the limit, so naturally they are getting close to each other. This is the "easy" implication.*]

(b) [*For the converse (the substantive part!), first show that a Cauchy sequence is bounded (try taking $\epsilon = 1$). Then apply the B-W Theorem.*] □

Note: In more abstract settings (for more general spaces with a distance function but where the notion of "interval" makes no sense), the Axiom of Completeness is that every Cauchy sequence converges!

Chapter 5

Discrete Dynamical Systems

In this chapter we explore further topics of iteration. In Chapter 1, we generated some sequences of numbers by iterating a function—i.e., picking a number a_0 and setting $a_n = f^n(a_0)$ for a given function f (where $f^n = f \circ f \circ \cdots \circ f$ is the composition of f with itself n times). This chapter contains further analyses of this idea for various functions f. The behavior of all such sequences generated by a function f—i.e., for various initial a_0's—will collectively be called "the dynamics of f".

As an initial investigation of iteration and the various behaviors it can generate, consider the following activity for a group of n students:

***Investigation 5.1.** Get into groups of four (or three when necessary) in a way that the $\lceil n/4 \rceil$ groups form a circle. Introduce yourself! Make sure you know how to spell everyone's name in your group. Figure out who has the third name in alphabetical order (first name, then last if necessary). That person will move one group counterclockwise. Repeat this procedure. Repeat this procedure again. Repeat this procedure again...

What do you think happens in the long run? Which people never move groups? Does this depend on the value of n?

5.1. Iterating functions and types of orbits

As in the previous investigation, we will be interested in what long-term behaviors are generated in a dynamical system (e.g., what kinds of patterns—or lack of patterns—individual students fall into). We will also be interested in how these behaviors change as we change the parameters that define a system (e.g., how the set of possible student behaviors might change as the total number n of students changes).

We note that many of the following investigations, exercises, and problems are quite amenable to the use of computer code. If you have experience with a computer

programming language or mathematics software, or are interested to learn, we encourage you to explore these topics with an active command line. Many basic example codes (relevant to this material) are available with a quick search online.

***Investigation 5.2.** Iterating functions.

(1) Pick your favorite polynomial, call it f. Compute $f(1)$, and call this value x_1. Now, compute $f(x_1)$, and call this value x_2. Then, compute $x_3 = f(x_2)$, $x_4 = f(x_3)$, and $x_5 = f(x_4)$. What happens? Can you come up with a polynomial for which the orbit of $x = 1$ has a different long-term behavior than the one you just observed?

(2) Let $g(x) = \lceil -\frac{1}{50}x^2 + 2x \rceil$ (where $\lceil x \rceil$ means the smallest integer $N \geq x$—i.e., "rounds up"), and pick a number x_0 between 1 and 99. Compute $x_n = g^n(x_0)$ up to $n = 10$. What happens?

The topic of discrete dynamical systems is the mathematical study of the effects of iteration. Above, you iterated functions $f : \mathbb{R} \to \mathbb{R}$ that were given in terms of a formula. Hopefully, you saw that depending on what function you pick (and also which initial value x_0 you pick), the effects of iteration can be very different.

Definition 5.3. A *discrete dynamical system* is a map $T : X \to X$ from a set X to itself. This allows us to *iterate* T. We will write $T^2(x) = T(T(x))$ and $T^n(x) = T(T(\dots T(x)\dots))$ (where T is applied n times). The set X is called the *phase space*. Given a point x_0 in X, the *orbit* of x_0 is the sequence

$$x_0,\ x_1 = T(x_0),\ x_2 = T(x_1), \dots$$

Note that the term *discrete* does not refer to the set X but rather to the fact that the iterations T^n are indexed by natural numbers $n = 1, 2, \dots$, which have the interpretation of (discrete) time points.

In this chapter, we will mostly deal with dynamical systems where $X \subseteq \mathbb{R}$. For now, you can consider $X = \mathbb{R}$.

In Chapter 1, we saw how to iterate a function graphically. As a reminder, here is the procedure.

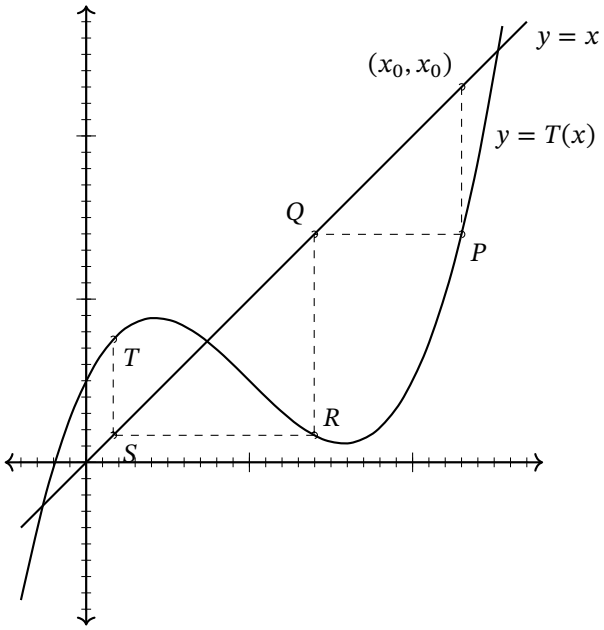
Given a function $f : \mathbb{R} \to \mathbb{R}$ and an initial value x_0, form the following "cobweb":

I. Draw the graph of f and the line given by $y = x$.

II. Draw a vertical line from the point (x_0, x_0) to the graph of f, and call this intersection P_1.

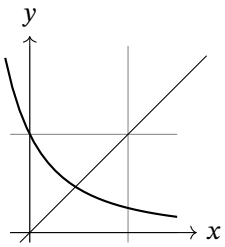
III. For $n \geq 1$, draw a horizontal line from the point P_n to the line $y = x$ and then a vertical line from there to the graph of f; label the point of intersection P_{n+1}. Repeat.

***Exercise 5.4.** In the graph below, label the coordinates of each point P, Q, R, S, and T using the numbers x_0, $f(x_0)$, $f^2(x_0)$, and $f^3(x_0)$. How does the geometric iteration of the cobweb algorithm model the algebraic iteration of the function f?

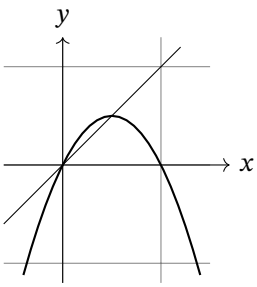


***Exercise 5.5.** Use the cobwebbing method to study various orbits of the following functions for various initial conditions. Can you classify which initial conditions have what kind of eventual behavior? Pay attention especially to things that you might call fixed points, periodic points, or divergent orbits. Larger versions of these graphs can be found in Appendix C.

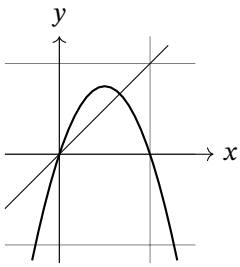
(1) $f(x) = \dfrac{1}{(1+x)^2}$



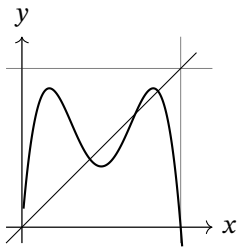
(2) [1] $f(x) = 2x(1-x)$



(3) $f(x) = 3x(1-x)$

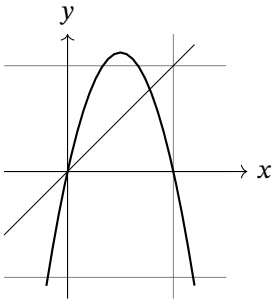


(4) $f^2(x) = f(f(x))$, where $f(x) = 3x(1-x)$

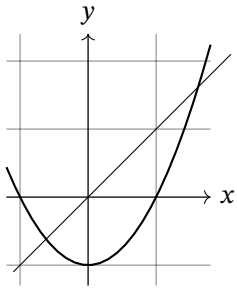


[1] Compare the behavior of this dynamical system with that of Investigation 5.2 part (2).

(5) $f(x) = 4.5x(1-x)$

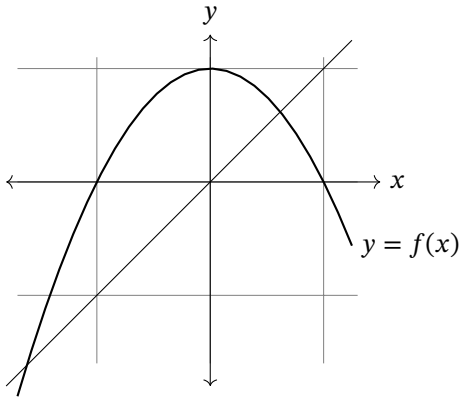


(6) $f(x) = x^2 - 1$



Note that a dynamical system consists not only of the map T but also of a set X that is both the domain and the codomain[2] for T. A given map can be part of multiple dynamical systems defined with different domains. In order to be able to iterate the map, however, the domain and codomain must be the same.

***Exercise 5.6.** Let $f(x) = 1 - x^2$.



(1) Consider the dynamical system $f : \{x_0\} \to \{x_0\}$, where $x_0 = \frac{\sqrt{5}+1}{2}$. What kinds of orbits are there?

(2) Consider the dynamical system $f : [0,1] \to [0,1]$. What kinds of orbits are there?

(3) Consider the dynamical system $f : \mathbb{R} \to \mathbb{R}$. What kinds of orbits are there?

(4) Does f with the domain $[-1,1]$ define a dynamical system? Why or why not?

(5) Does f with the domain $[0,2]$ define a dynamical system? Why or why not?

A dynamical system usually has infinitely many orbits, each of which consists of an infinite sequence (of not necessarily distinct points). How, then, can one get a global picture of the behavior of a dynamical system? One strategy is to first categorize possible behaviors of orbits and then look for these types of orbits in a given system. The rest of this section introduces special types of orbits that are often useful in this regard.

[2] See Appendix B for more on these terms.

Definition 5.7. Consider a dynamical system $T : X \to X$ and a point $x_0 \in X$. The point x_0 is called:

- *fixed* if $T(x_0) = x_0$;
- *eventually fixed* if $T^n(x_0) = T^{n-1}(x_0)$ for some $n > 1$;
- *periodic of (minimal) period N* if $T^N(x_0) = x_0$ and $T^i(x_0) \neq x_0$ for $0 < i < N$.

Recall that a function is invertible if and only if it is injective (one-to-one) and surjective (onto). In particular, if T is invertible, then there is a *function* T^{-1} with the property that $T^{-1}(T(x)) = x$ for all x in the domain of T and $T(T^{-1}(y)) = y$ for all y in the codomain. Many dynamical systems are *not* invertible (for example, the ones in §§5.4, 5.3, 6.2, and 6.3). In fact, invertibility can be an obstruction to some of the properties we are interested in:

***Theorem 5.8.** Let T be an invertible function. Then x_0 is an eventually fixed point of T if and only if x_0 is a fixed point of T.

Just as points in non-invertible dynamical systems can become *eventually* fixed, so too they may eventually enter a periodic orbit.

***Exercise 5.9.** Consider the term *eventually periodic point*. What do you think this should mean? Can you provide a definition in the style of Definition 5.7 for eventually periodic points?

***Exercise 5.10.** Find the fixed points of the function $f(x) = x^2 - 1$. Can you find the period 2 orbits of f?

***Exercise 5.11.** The points $\frac{3}{7}$ and $\frac{6}{7}$ form a period 2 orbit for the map $f(x) = 3.5x(1-x)$. Find an eventually period 2 point (that is not a period 2 point). How many eventually period 2 points can you find?

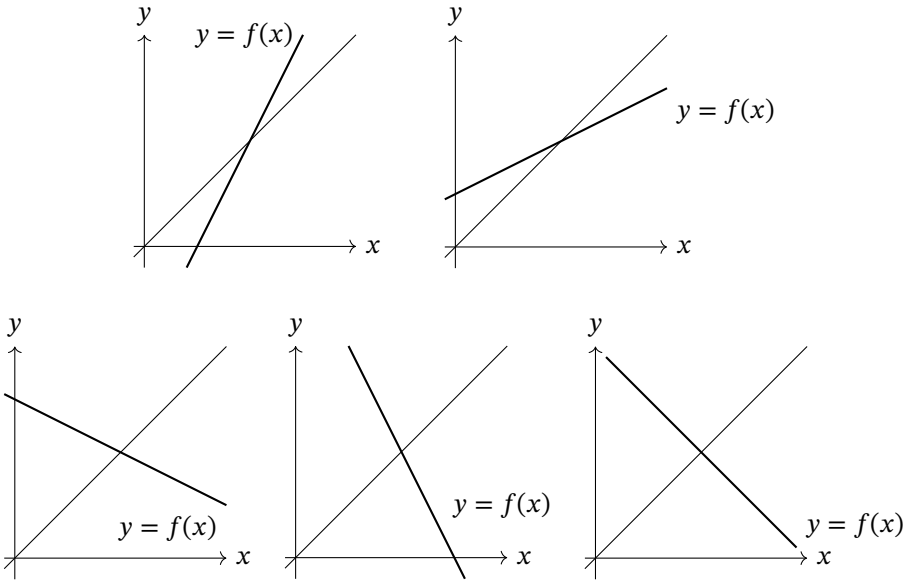
A map can have different kinds of local behavior around a fixed point.

Definition 5.12. Let x_0 be a fixed point of $T : X \to X$. Then x_0 is called:

- an *attracting* fixed point if the orbit of any sufficiently near point tends to x_0;
- a *repelling* fixed point if no matter how close one looks near x_0, the orbits of nearby points diverge from x_0.

Recall that if a function f is differentiable at a point x_0, then its graph looks linear when one zooms in far enough on the point $(x_0, f(x_0))$. Since whether a fixed point x_0 is attracting or repelling depends only on what happens near x_0, the following exercise indicates behavior that occurs in a much more general setting than it may seem at first glance.

***Exercise 5.13.** Use the cobweb method of iteration to determine whether the fixed point for each of the linear functions below is attracting or repelling (or neither!). Make a conjecture about what feature of the function determines the corresponding property of the fixed point.



The following theorem is very useful for determining whether a fixed point is attracting or repelling. It is somewhat involved to prove rigorously with all of the correct notation; you should give an informal argument about why it is true.[3]

***Theorem 5.14.** Let x_0 be a fixed point of a dynamical system $f : X \to X$ with $X \subset \mathbb{R}$ and f differentiable at x_0. Then:

(1) If $|f'(x_0)| < 1$, then x_0 is an attracting fixed point.

(2) If $|f'(x_0)| > 1$, then x_0 is a repelling fixed point.

***Exercise 5.15.** Find all fixed points of the map $f(x) = x^3 - 4x^2 + 4x$. Classify each fixed point as attracting or repelling.

***Problem 5.16.** Show that each of the following functions has a fixed point at $x = 1$. For each function, investigate the stability of the fixed point at $x = 1$.

(1) $f(x) = x^2 - x + 1$

(2) $g(x) = -x^3 + 3x^2 - 2x + 1$

(3) $h(x) = -x^3 + 3x^2 - 4x + 3$

Why is the case $|f'(x))| = 1$ not included in Theorem 5.14?

***Investigation 5.17.** What should it mean for a period N orbit to be attracting or repelling? Formulate a definition in the style of Definition 5.12 and a criterion for determining the stability in the style of Theorem 5.14. Do you think it is possible for some points in a periodic orbit to be attracting and other points in the same orbit repelling? Why or why not?

[3] See Definition 2.108 and Problem 2.110 for ideas on how to make a more formal argument.

5.2. The logistic map, modeling, and bifurcations

Iterating functions is often useful in modeling physical phenomena that evolve over time. We will discuss below models of a population of fish in a lake as a function of time. The models will be based on some assumptions that involve certain *parameters*: the amount of food available, how quickly fish reproduce, how potential predators interact with the population, etc. Mathematically, we will be led to studying the following family of dynamical systems involving a single parameter $r > 0$:

$$\boxed{f_r(x) = rx(1-x)}$$

This is called the *logistic family*.

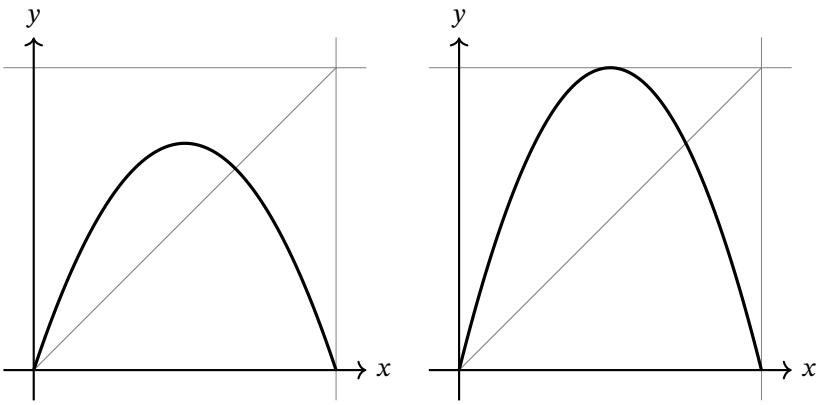


Figure 5.1. The f_3 and f_4 maps.

Let us begin with a very simple model. Suppose you have a small population of fish in a huge lake, with no predators and plenty of food (i.e., infinite resources). You observe the population of fish every unit of time (say, every week), starting at time 0. Let y_n be the number of fish at your nth observation.

If no fish die, then

$$y_{n+1} = y_n + ay_n,$$

where a is the (constant) reproduction rate (the number of offspring per fish per unit time). Taking into consideration that some fish will die, the model becomes

$$y_{n+1} = y_n + ay_n - by_n,$$

where b is the death rate (proportion of fish dying off per unit time). By changing our bookkeeping, we can wrap this up into an equation involving just one parameter:

$$y_{n+1} = ry_n, \qquad \text{where } r = 1 + a - b.$$

***Exercise 5.18.** (1) If you start with 50 fish, find a formula (depending on r) for how many fish there are after n weeks.

(2) Do the same but now starting with some number y_0 of fish. (Note that your answer now has *three* parameters: the model parameter r; the number of weeks you are interested in, n; and how many fish you start off with, y_0.)

(3) What happens in the long run (as n goes to infinity)? How does the answer depend on r? For what parameters, if any, does this model predict a sustainable population?

Alas, in the real world there is competition and resources are finite. The lake is only so big, and it will only support a maximum of N fish. To model this more realistic situation, we introduce a factor $(N - y)$ into the rate of growth:

$$y_{n+1} = y_n + ky_n(N - y_n).$$

***Exercise 5.19.** Check that the model above has the following properties, provided k has the appropriate sign:

(1) If $y_n = N$, then $y_{n+1} = y_n$ (i.e., no more growth is possible).

(2) If $y_n < N$ but y_n is close to N, then the growth $y_{n+1} - y_n$ is small.

(3) If $y_n > N$, then $y_{n+1} < y_n$ (i.e., fish will die off if there are too many).

Which sign should k have for the model to be realistic?

It turns out that we again can wrap up all these parameters into one, which will simplify our mathematical analysis. Note that

$$\begin{aligned} y_{n+1} &= y_n + ky_n(N - y_n) \\ &= (1 + kN)y_n - ky_n^2 \\ &= (1 + kN)y_n - (1 + kN)y_n\left(\frac{k}{1 + kN}y_n\right) \\ &= (1 + kN)y_n\left(1 - \frac{k}{1 + kN}y_n\right). \end{aligned}$$

Multiplying both sides of this identity by $\frac{k}{1+kN}$ yields

$$\frac{k}{1 + kN}y_{n+1} = (1 + kN)\frac{k}{1 + kN}y_n\left(1 - \frac{k}{1 + kN}y_n\right),$$

and by making the change of variable $x_n = \frac{k}{1+kN}y_n$ and letting $r = (1+kN)$ one obtains

$$x_{n+1} = rx_n(1 - x_n).$$

This is the logistic family mentioned at the beginning of the section! Thus, we are led to studying the dynamics of the function $f_r(x) = rx(1 - x)$, where r is a parameter incorporating the birth and death rates and the carrying capacity N of the lake. You should check that the meaningful range of x for this model is $x \in [0, \frac{kN}{1+kN}] \subset [0, 1]$.

Let's use some of the language and tools developed in the previous section to study this dynamical system.

***Proposition 5.20.** The logistic map $f_r(x)$ has two fixed points:

- a fixed point at $x = 0$, which is attracting if $0 < r < 1$ and repelling if $r > 1$; and
- a fixed point at $x = 1 - \frac{1}{r}$, which is attracting if $1 < r < 3$ and repelling if $0 < r < 1$ or $r > 3$.

According to this proposition, there are major changes in the dynamics of the family at the parameter values $r = 1$ and $r = 3$. We call such behavior, where a very small change to the parameter value creates a sudden qualitative change to the dynamical system, a *bifurcation*[4]. The logistic family has a fixed-point bifurcation at $r = 1$ and at $r = 3$.

Recall that for the population model, $x \in [0, 1]$. The maximum of f on $[0, 1]$ is attained at $x = 1/2$ (the midpoint between the two roots of f), and the value is $f(1/2) = r/4$. So for the population model one must have $0 < r \leq 4$. (We will see later what happens mathematically when $r > 4$.)

In order to analyze the family further we have to distinguish several cases:

- If $0 < r < 1$: We already know that $x = 0$ is an attracting fixed point, and the other fixed point is not in $[0, 1]$. It turns out that all orbits tend to zero—the population goes extinct no matter what the size of the initial population. Recalling that $r = 1 + kN$, this case corresponds to $k < 0$, so it is not a surprising result.
- If $1 < r < 3$: The origin is a repelling fixed point. The other fixed point $1 - 1/r$ is now in $[0, 1]$, and it is attracting. It turns out that the orbit of any non-zero population will tend to the fixed point, i.e., to a steady value.
- If $3 < r \leq 4$: Now both fixed points are repelling. Where will the orbits go?

As it turns out, the attracting fixed point that exists when r is a little less than 3 is replaced by an attracting *orbit* of period 2 when r is a little greater than 3. In order to see this, note that some arithmetic yields

$$\begin{aligned} f_r(x) &= rx(1-x), \text{ so} \\ f_r^2(x) &= r^2x(1-x)(1-rx+rx^2) \\ &= -r^3x^4 + 2r^3x^3 - r^3x^2 - r^2x^2 + r^2x. \end{aligned}$$

Period 2 points are those that are fixed by f_r^2 but not by f_r—that is, solutions to the equation $f_r^2(x) - x = 0$ that are not $x = 0$ or $x = 1 - 1/r$. Factoring the polynomial $f_r^2(x) - x$, discarding the factors corresponding to fixed points of f_r, and employing the quadratic formula, one finds that the potential points of period 2 are

$$x = \frac{1}{2}\left[(1 + r^{-1}) \pm \frac{\sqrt{(r-3)(r+1)}}{r}\right].$$

More precisely, whenever the square root on the right-hand side is real, this expression gives two period 2 points which make one period 2 orbit.

Note that for many of the following problems it is useful to make computations using a computer program.

[4]The dictionary definition of bifurcation is "the division of something into two branches or parts".

***Exercise 5.21.** Use the results of the calculations above to determine if there is a period 2 orbit for the following parameters and, if so, where the period 2 points are:

(1) $r = 2$
(2) $r = 3$
(3) $r = 3.2$
(4) $r = 3.5$

***Problem 5.22.** For each parameter value in the previous exercise that has a period 2 orbit, determine whether the period 2 orbit is attracting or repelling.

***Investigation 5.23.** Here is a method for using a spreadsheet program to find an attracting periodic orbit of f_r for different parameter values of r. In a spreadsheet program, enter the parameter value 3.3 (r) in cell A1 and your starting value 0.5 (x_0) in A2. In A3, type `= A$1 * A2 * (1 - A2)` (the dollar sign in this formula will keep the reference to your r value fixed). This should return the value 0.825. Now, fill this formula down for about 1000 cells. Looking at the end of this, you can see the points in the periodic orbit that the initial value 0.5 is attracted to (and, in particular, the period).

(1) Find the period of the attracting periodic orbit for parameter values $r = 3.3$, $r = 3.5$, $r = 3.55$, and $r = 3.565$.
(2) Try other parameter values between $r = 3.40$ and $r = 3.57$. Can you approximate the parameter value at which the attracting period 2 orbit is replaced by an attracting period 4 orbit? What about the parameter value at which the attracting period 4 orbit is replaced by an attracting period 8 orbit? Etc.

The logistic family has a series of "period doubling" bifurcations (as the parameter r changes), some of which you have just investigated. Figure 5.2 shows a diagram documenting these bifurcations (and others) between $r = 3.00$ and $r = 4.00$.[5] Can you find your data from Investigation 5.23 in this figure? What further questions does the figure inspire? A closer-up diagram for the parameter values $r = 3.5$ through $r = 4.0$ is displayed in Figure 5.3.

5.3. The doubling map and chaos

A baker kneads dough by stretching out the dough to double its original length, then cutting it at the middle, stacking the two halves on top of each other, and compressing them into one piece. Then, they repeat the process.

Let us consider this process just in one dimension, the direction of the stretching. This can be modeled mathematically by iterating the following function, which is called the *doubling map*[6]:

$$\boxed{T_2 : [0,1] \to [0,1], \quad T_2(x) = 2x \bmod 1}$$

[5] Figures 5.2 and 5.3 were created by Jacob Brown using Python.
[6] Or the "1-D baker's map". There is a similar $\mathbb{R}^2 \to \mathbb{R}^2$ map that is commonly called "the baker's map" or "the baker's transformation".

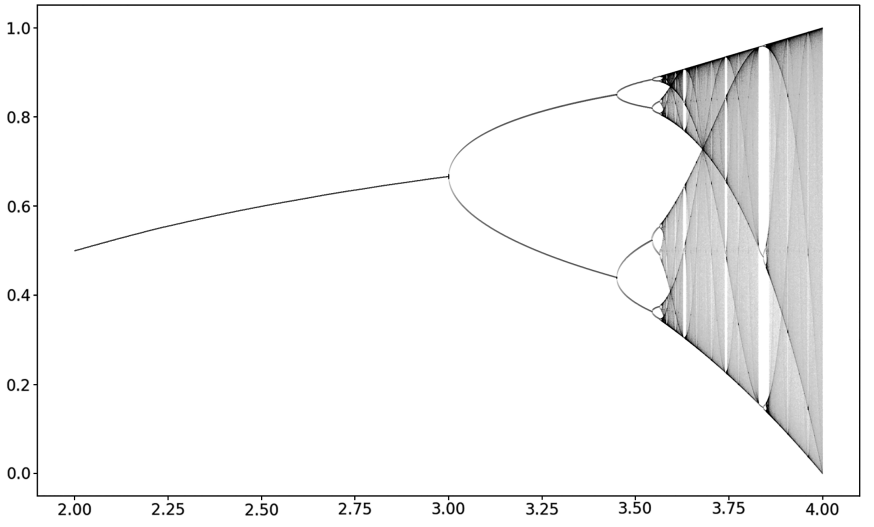


Figure 5.2

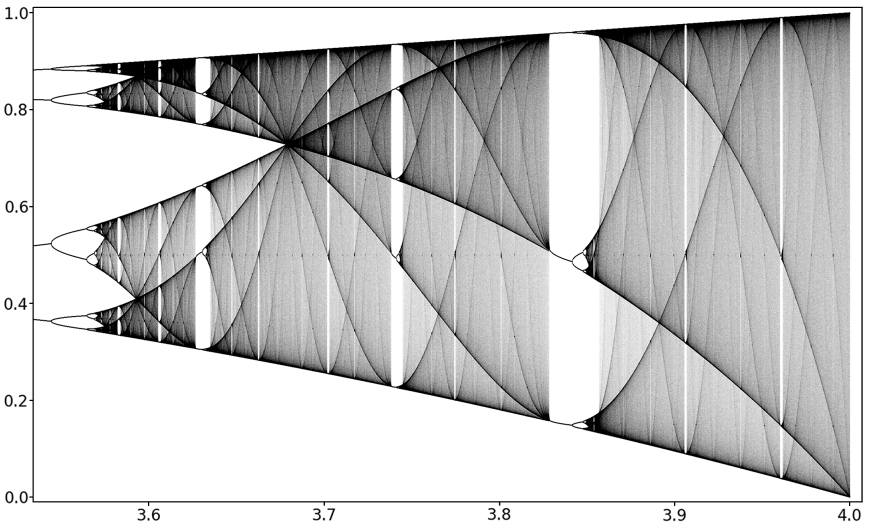


Figure 5.3

Specifically, the meaning of "mod 1" is the following:

$$\forall x \in [0,1] \quad T_2(x) = \begin{cases} 2x & \text{if } 2x < 1, \\ 2x - 1 & \text{if } 2x \geq 1. \end{cases}$$

The structure of the doubling map is most transparent if we represent real numbers in base 2. Let x_0 be a real number such that $0 \leq x_0 \leq 1$. The *binary expansion of* x_0 is

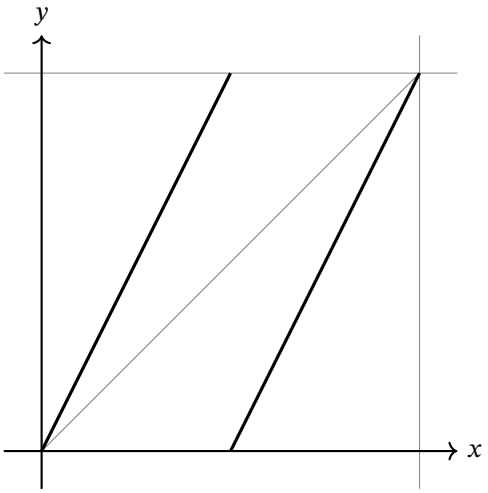


Figure 5.4. The doubling map.

a sequence of digits $d_1, d_2, \ldots \in \{0, 1\}$ such that

$$x_0 = \frac{d_1}{2} + \frac{d_2}{2^2} + \frac{d_3}{2^3} + \cdots.$$

The following two problems argue that any $x_0 \in [0, 1]$ has a binary expansion. The first uses the doubling map to produce a sequence of binary digits (d_i) from x_0, and the second shows that this sequence actually forms a binary expansion of x_0. We will further explore properties of base 2 expansions in §6.2.

***Problem 5.24.** Let $x_0, x_1, x_2, \ldots$ be the orbit of x_0 under the doubling map T_2. For each $n = 1, 2, \ldots$, let

$$d_n = \begin{cases} 0 & \text{if } 0 \le x_{n-1} < \frac{1}{2}, \\ 1 & \text{if } \frac{1}{2} \le x_{n-1} \le 1. \end{cases}$$

That is, we keep track of which half of the interval the iterates of x_0 land in. Let's check this in some examples.

(1) Compute d_n for $x_0 = \frac{2}{3}$, $n = 1, 2, 3, \ldots$.

(2) Show, using the geometric series formula (see Theorem 1.55 or Problem 1.56), that

$$\frac{2}{3} = \frac{d_1}{2} + \frac{d_2}{2^2} + \frac{d_3}{2^3} + \cdots.$$

(3) Repeat the last two questions with $x_0 = \frac{6}{7}$.

In order to show that the sequence $d_1, d_2, \ldots$ in fact is a base 2 representation of x_0, we need to show that the infinite sum $\frac{d_1}{2} + \frac{d_2}{2^2} + \frac{d_3}{2^3} + \cdots$ converges to x_0 in general. Recall your definition of convergence in §1.2 and Definition 1.52 for what it means for a series to converge.

***Problem 5.25.** Let the numbers d_n be defined as above. This problem shows the convergence $\sum_{n=1}^{\infty} \frac{d_n}{2^n} = x_0$ by way of induction.

(1) First, show that

$$0 \le x_0 - \frac{d_1}{2} \le \frac{1}{2}.$$

You will need to distinguish the two cases $0 \le x_0 < \frac{1}{2}$ and $\frac{1}{2} \le x_0 \le 1$.

(2) Then, complete the inductive step: for each $n = 1, 2, \ldots$,

$$0 \le x_0 - \left(\frac{d_1}{2} + \frac{d_2}{2^2} + \frac{d_3}{2^3} + \cdots + \frac{d_n}{2^n}\right) \le \frac{1}{2^n}.$$

In the inductive step you will need to distinguish the two cases $0 \le x_n < \frac{1}{2}$ and $\frac{1}{2} \le x_n \le 1$.

Let us use the following notation for binary expansions: we will write $[x]_2 = 0.d_1d_2d_3\ldots$, where $d_n \in \{0, 1\}$ for all n, to mean that $x = \frac{d_1}{2} + \frac{d_2}{2^2} + \frac{d_3}{2^3} + \cdots$. It turns out that when we write numbers in their binary representation, the map T_2 has a particularly simple form.

***Problem 5.26.** (1) Using the notation $[x]_2 = 0.d_1d_2d_3\ldots$, compute $[T_2(x)]_2$. Use this to explain why we can think of T_2 as a "shift map" when applied to numbers written in their base 2 form.

(2) Find a point whose orbit has period 1 (that is, a fixed point), one of (smallest) period 2, another of period 3, and so on. Conclude that the map T_2 has periodic points of all periods!

In addition to periodic points of any period, there are other points whose T_2 orbits have notable properties.

***Problem 5.27.** Let's construct a special number in base 2. The construction is a little complicated, so read carefully.

(1) For some integer ℓ, let's define "a block of length ℓ" to be any string of binary digits of length ℓ. For example, 0111 and 0101 are blocks of length 4. How many blocks of length ℓ are there?

(2) Here is how to construct the special number, which we will call s:

$$[s]_2 = .\{\text{put here, in any order, all blocks of length 1}\}$$

$$\{\text{next put here, in any order, all blocks of length 2}\}$$

$$\{\text{next put here, in any order, all blocks of length 3}\}\ldots$$

and so on, forever! Begin to write down s. Go as far as your page lets you.

(3) (…Seriously—write down a lot of digits.)

(4) There are actually infinitely many different numbers this process could produce. Why? This means that s is not well-defined by the process above, but any such s produced in this manner will work for our purposes, so we will not worry about the non-uniqueness.

***Problem 5.28.** In fact, the orbit of the number s under the doubling map T_2 is *dense* in $[0,1)$. This means that given any number y in the interval $[0,1)$, the orbit of s gets arbitrarily close to y.

(a) Write down $[\frac{6}{7}]_2$.

(b) Show that the orbit of s gets arbitrarily close to $\frac{6}{7}$.

(c) Argue that, in the previous question, $\frac{6}{7}$ is not in any way special. That is, for any other number $y \in [0,1)$, the orbit of s gets arbitrarily close to y.

***Problem 5.29.** Are there points $x \in [0,1)$ that neither are eventually periodic (this includes fixed, periodic, etc.) nor have a dense orbit under T_2? If so, can you write one down?

As we saw in Problem 5.26, the map T_2 has periodic orbits of any period. In fact, there is an even stronger result about "how common" periodic orbits are:

***Problem 5.30.** Show that the set of periodic points is dense in $[0,1)$. That is, given an arbitrary point $[x]_2 = 0.d_1d_2d_3\ldots$ and an arbitrarily small distance, say 2^{-N}, there is a point p such that $|x-p| < 2^{-N}$ and p is a periodic point for T_2.$\star$

***Problem 5.31.** In this problem, we investigate *how many periodic orbits of a given period* there are.

(1) Find a period 3 orbit of T_2. Can you find all period 3 orbits of T_2? How many are there?

(2) Find a period 4 orbit of T_2. Can you find all period 4 orbits of T_2? How many are there?

(3) How many period 5 orbits of T_2 are there?

(4) How many period n orbits of T_2 are there?

We are discovering that the dynamics of T_2 are pretty diverse: there are periodic orbits of any period, the set of periodic orbits is dense in [0,1], but there are also single orbits that are dense in [0,1]. Moreover, orbits that start off close together may end up very far apart, in the following sense:

Definition 5.32. A map $f : X \to X$ has *sensitive dependence on initial conditions* if there exists $c > 0$ such that given any $x \in X$ and any $\epsilon > 0$ there exists $y \in X$ and $N > 0$ such that

$$|x-y| < \epsilon \quad \text{and} \quad |f^N(x) - f^N(y)| > c.$$

That is, a map has sensitive dependence on initial conditions if there is a constant $c > 0$ such that for any point x there exists a point arbitrarily close to x whose orbit eventually separates from the orbit of x by a distance greater than c.

Theorem 5.33. *The doubling map T_2 has sensitive dependence on initial conditions.*

***Proof (to be completed).** We will prove that we can take $c = \frac{1}{2}$. Let $x \in [0,1)$ with $[x]_2 = .a_1a_2\ldots$ and let $\epsilon > 0$. Let $N \in \mathbb{N}$ be such that

$$\frac{1}{2^N} < \epsilon.$$

Now, define a number $y \in [0,1)$ with $[y]_2 = .b_1b_2\ldots$ such that $b_i = a_i$ for any $i \neq N$ and

$$b_N = \begin{cases} 0 & \text{if } a_N = 1, \\ 1 & \text{if } a_N = 0. \end{cases}$$

[*Show that* $|b-a| \leq \frac{1}{2^N} < \epsilon$ *and* $|f^N(b) - f^N(a)| > \frac{1}{2}$.] □

The doubling map T_2 is a deterministic system because given an exact initial condition, the map completely determines its future orbit. The theorems above, however, indicate that when an initial condition is known only approximately, the behavior of its orbit is highly unpredictable. This kind of highly unpredictable behavior in a deterministic system is known as *chaos*.

There is no universally used mathematical definition of a chaotic dynamical system, but one that is often cited is the following:

Definition 5.34. A dynamical system $T : X \to X$ is called *chaotic* if the following three conditions hold:

(1) There exists at least one dense orbit (this is also called "topological transitivity").

(2) The set of periodic points is dense in X.

(3) The system has sensitive dependence on initial conditions.

We have shown these properties for the doubling map: Problem 5.28 shows that condition (1) holds, Problem 5.30 shows that condition (2) holds, and Theorem 5.33 shows that condition (3) holds. Thus, we have

Theorem 5.35. *The doubling map T_2 is chaotic.*

5.4. The tent map and fractals

In §5.2, we stopped our analysis of the logistic family at the parameter value $r = 4$. For values of $r > 4$, the graph of the map f_r extends outside of the unit square. Therefore, f_r with the domain $[0,1]$ is not a dynamical system according to Definition 5.3. How can this be salvaged?

In this section, we study a simpler but similar map whose graph also extends beyond the unit square:

$$\boxed{T_3(x) = \begin{cases} 3x & \text{if } x < \frac{1}{2}, \\ 3 - 3x & \text{if } \frac{1}{2} \leq x. \end{cases}}$$

This function is known as the "tent" map, for reasons that are apparent when you look at its graph (compare it with Figure 5.1 in §5.2):

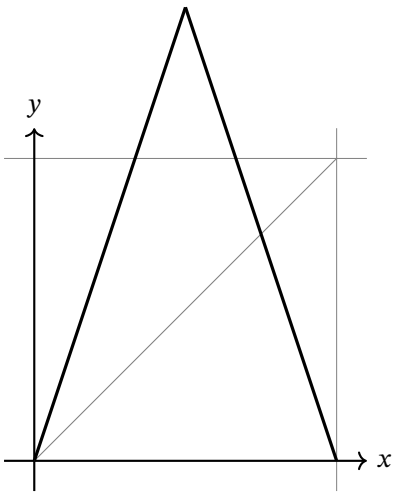


Figure 5.5. The tent map T_3.

As is clear from its graph, T_3 does not map $[0,1]$ into $[0,1]$, so it is not a dynamical system on the unit interval either. Hence, we begin our analysis by extending the domain and codomain to $\mathbb{R}$—that is, we consider $T_3 : \mathbb{R} \to \mathbb{R}$ so that we can iterate it. We will be particularly interested in points in $[0,1]$ whose orbits exit $[0,1]$. The following problem asks you to show that the orbits of such points tend to $-\infty$.

***Problem 5.36.** (1) Show that if $x_0 < 0$, then the orbit of x_0 under T_3 tends to $-\infty$.

(2) Show that if $x_0 > 1$, then the orbit of x_0 under T_3 also tends to $-\infty$.

What happens to points inside the unit interval? As alluded to before, there is some open interval—we'll call it A_1—in the middle of $[0,1]$ that is mapped by T_3 outside of $[0,1]$. And once these points are outside of $[0,1]$, Problem 5.36 shows that T_3 will send them off to $-\infty$. Not too complicated.

But are there points in $[0,1]$ that are mapped *into* A_1? An examination of the graph of T_3 shows that indeed there are such points; let's denote the set of these points by A_2. Points in A_2 will be mapped outside of $[0,1]$ by T_3^2 (but not by T_3), so the orbits of *those* points also tend to $-\infty$.

And so on! We recursively define

$$A_{n+1} = f^{-1}(A_n),$$

which are the points mapped outside $[0,1]$ by T_3^n but not by T_3^{n-1}.

***Problem 5.37.** (1) Determine the real numbers a_1 and b_1 that define the ends of the interval $A_1 = (a_1, b_1)$.

(2) Argue that A_2 consists of exactly two intervals. Determine these two intervals, as well as their lengths. (You may find the figure below useful.)

(3) How many intervals does A_3 consist of? Determine each of these intervals, as well as their lengths.

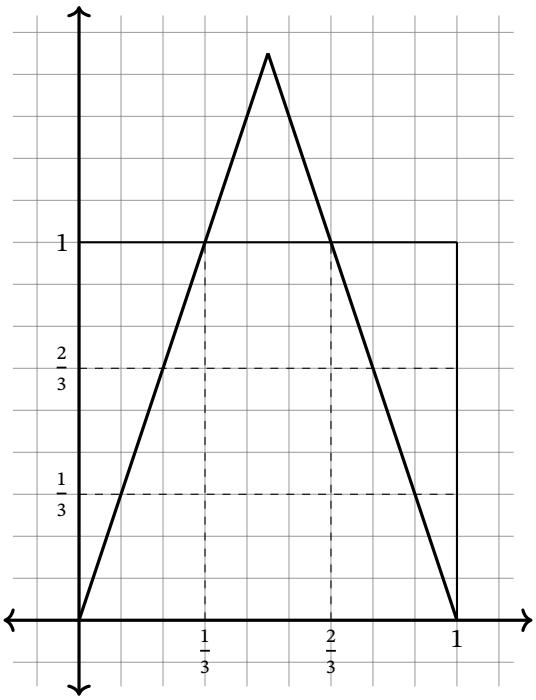
(4) How many intervals does A_n consist of? What are their lengths?

What can one say about the points whose orbits do *not* tend to $-\infty$? First let's check that such points exist:

***Exercise 5.38.** Show that the orbit of $\frac{1}{3}$ does not tend to $-\infty$. Find other points with the same property.

Definition 5.39. The *Cantor set* is the set

$$\mathcal{C} = [0,1] \setminus \bigcup_{n=1}^{\infty} A_n.$$



***Problem 5.40.** Argue that any point in $\mathbb{R}$ either tends to $-\infty$ or else stays in the set $[0,1]$ under iteration of T_3. Conclude that the Cantor set $\mathcal{C}$ can also be described as exactly the set of points in $[0,1]$ whose T_3 orbits are entirely contained in $[0,1]$.

The set $\mathcal{C}$ has a number of properties that present interesting conundrums. For example, how "big" is it? Should it have dimension 0, like a set of points, or dimension 1, like a line?

***Problem 5.41.** Let's compute the "length" of $\mathcal{C}$.

(1) What is the total length of the set A_n? (Note that it should depend on n.) We will denote this length by $\ell(A_n)$.

(2) Compute the value of the series

$$\sum_{n=1}^{\infty} \ell(A_n).$$

(3) Since this is the total length of the intervals taken out of $[0,1]$ to make $\mathcal{C}$, what is the "length" of $\mathcal{C}$?

***Problem 5.42.** The base 3 expansion of a number $x \in [0,1]$ is a sequence of ternary digits $d_i \in \{0,1,2\}$, written $[x]_3 = 0.d_1d_2d_3\ldots$, such that $x = \sum_{i=1}^{\infty} \frac{d_i}{3^i}$.

(1) How can you determine whether a number $x \in [0,1]$ is in A_1 by looking at its base 3 expansion?

(2) How can you tell if x is in A_2 by its base 3 expansion?

(3) How can you tell if x is in A_n by its base 3 expansion?

In fact, there is a little wiggle room in how you write down these rules, since some numbers in $[0,1]$ have two (equivalent) base 3 expansions. Similar to the case of repeating 9's for (base 10) decimals, any number in base 3 ending with the digit 1 is equivalent to the same number ending with the sequence $0\overline{2}$ (replacing the 1). In other words,

$$0.d_1d_2\ldots d_k\,1 = 0.d_1d_2\ldots d_k\,0\,\overline{2}.$$

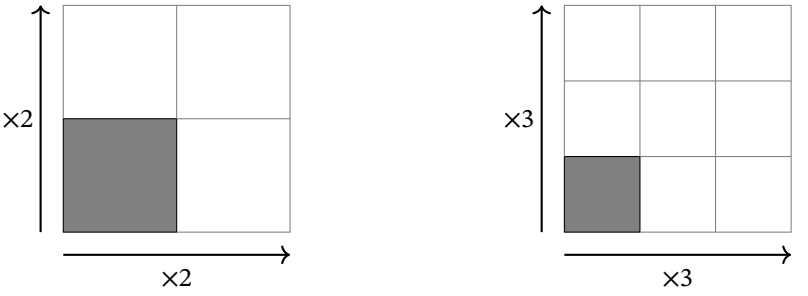
***Problem 5.43.** Use your characterizations from the previous problem and the note above about how some numbers have more than one base 3 expansions to argue that

$$\mathcal{C} = \{x \in [0,1] \mid [x]_3 \text{ can be written with only 0's and 2's}\}.$$

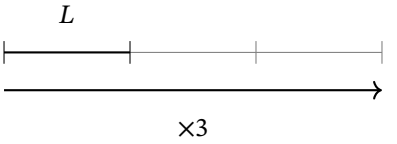
This last characterization of $\mathcal{C}$ has an interesting corollary. For any $x \in \mathcal{C}$, take its base 3 expansion (with only 0's and 2's) and replace every 2 by a 1. The image of this map is all of the sequences of 0's and 1's, a set that is equivalent to the unit interval when thought of in base 2 expansion. Thus, *the Cantor set has just as many points as the unit interval.*

Problem 5.41 seems to indicate that the Cantor set $\mathcal{C}$ should have dimension 0, while Problem 5.43 suggests that $\mathcal{C}$ should have dimension 1. Perhaps we had better think further about the notion of dimension!

Suppose we want to compute the dimension of some set $F \subset \mathbb{R}^n$. In order to do this, we scale F linearly in every direction by a factor of $c > 1$ so that we can cover the scaled image cF with a whole number of copies of the original object F. For example, if a square S is linearly scaled in every direction by $c = 2$, then 4 copies of the original version S make up the scaled square cS. Or, if we scale the original square by $c = 3$, then 9 copies of the original version S make up the scaled square cS.

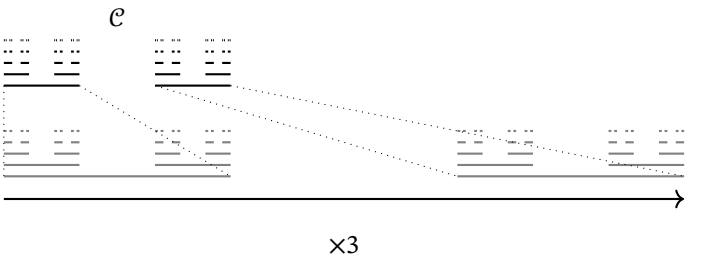


What happens if we apply this to a line segment L? If we scale L by $c = 3$, we find that cL consists of 3 copies of the original segment L.



***Exercise 5.44.** What quality of the scaled objects above indicates that the square is two-dimensional and the line segment is one-dimensional? Try your idea out on a cube: If you linearly scale the cube by a factor of, say, 3 in each direction, how many times will the original cube fit inside the scaled version? How does this indicate that the cube has dimension 3?

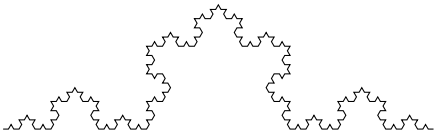
What happens when we linearly scale the Cantor set $\mathcal{C}$ by a factor of 3, as we did with the line above? In the image below, the Cantor set is represented by a stacked set of the first few A_n's—recall that the Cantor set is the limit of these sets.

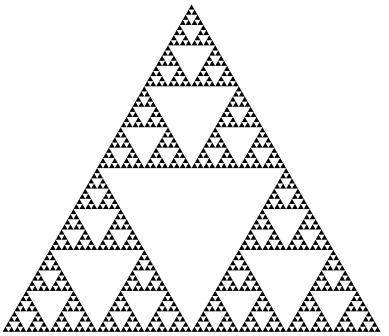


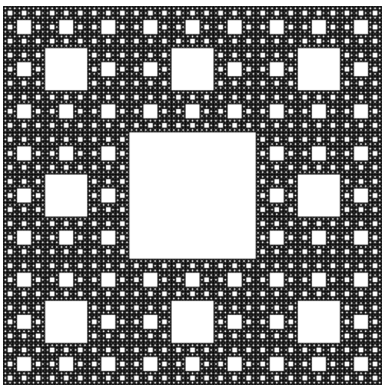
***Problem 5.45.** The Cantor set $\mathcal{C}$ is said to have "fractal dimension" $\dim(\mathcal{C}) = \frac{\log 2}{\log 3} \approx 0.6309$. Use the image above to explain why this is.★

So, it seems there is a compelling reason for why, when we forced it to take properties of whole-number-dimensional objects, the Cantor set varied between looking like a zero-dimensional object and a one-dimensional object.

***Problem 5.46.** Use the process of linear scaling to find the dimension of the following fractals:★

(1) 

(2) 

(3) 

***Investigation 5.47.** Can you create a fractal image that has dimension $\frac{\log 2}{\log 4}$? How about one that has dimension $\frac{\log 8}{\log 4}$? Or $\frac{\log n}{\log m}$ for other values of n and m?

We note that the process developed here for finding the dimension of a set only works if the set is self-similar under some scaling factor. There are other notions of dimension that accommodate more general sets.

5.5. The rotation map and Benford's Law

There is a famous "proof" whose punchline is $0 = 1$, an obvious impossibility that is supposed to spur the reader to revisit the provided logic in order to find the mistake. However, sometimes it is very useful to have $0 = 1$. Consider, for example, what happens if you take the unit interval $[0, 1]$ and identify the two ends, like a piece of string. You get a circle! This can be a useful shift of vantage point for thinking of the domain of a dynamical system: if $T : [0, 1] \to [0, 1]$ is a function with $T(0) = T(1)$, then we can identify 0 and 1 in the domain while retaining the property that T is a function, and thus consider T as a dynamical system on the circle $\mathbb{S}$. This is the case for the doubling map of §5.3. It is also the case for the family of maps $R_\alpha : [0, 1] \to [0, 1]$ that we study

in this section:

$$R_\alpha(x) = x + \alpha \mod 1$$

These maps are called rotation maps, for a reason that is clear when we identify 0 and 1 to get a circle.

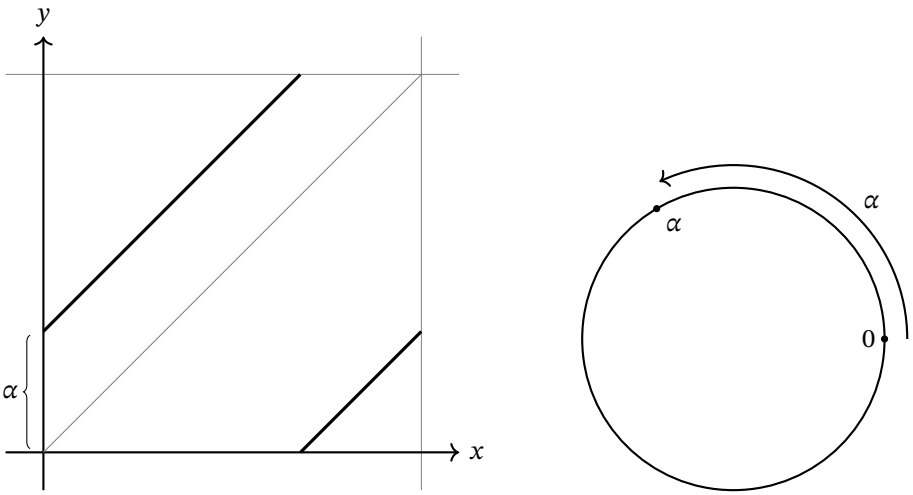


Figure 5.6. The rotation map R_α.

As was the case for the logistic family, the dynamics of R_α depend on the parameter α. But, in this case, the important distinction is not what interval the parameter is in but whether α is a rational or an irrational number.

***Problem 5.48.** Show that if $\alpha = \frac{p}{q} \in \mathbb{Q}$ in lowest terms, then every orbit is periodic of the same period, and determine the period.

In fact, there is a converse to this as well:

***Problem 5.49.** Show that if there is a periodic orbit of R_α, then α must be a rational number.

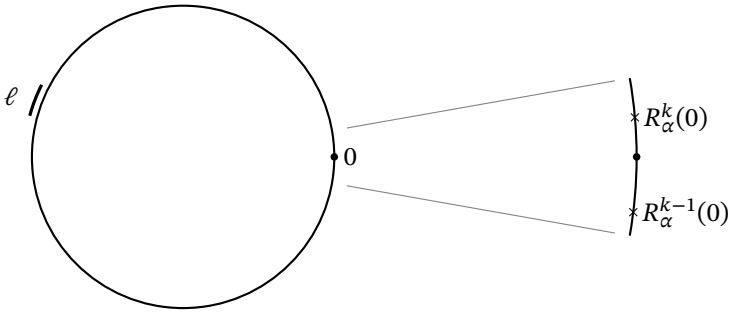
A consequence of Problem 5.49 is the contrapositive: if α is irrational, then R_α has no periodic orbits! There are still interesting things to say about how the orbits of these R_α's distribute, though.

In case you did not work on Theorem 2.21, we present here the definition of what it means for a set to be dense in another. This will give us another way to think about the density of an orbit.

Definition 5.50. Let $I \subset \mathbb{R}$ be an interval and $S \subset I$. We say that S is dense in I iff for every open interval $U \subset I$ there is a point $s \in S$ that is in U.

Recall that the orbit of a point x under a map $T : X \to X$ is the set $\mathcal{O}(x) = \{x, T(x), T^2(x), \ldots\}$. We say that the orbit of x is dense if the set $\mathcal{O}(x)$ is dense in X.

Theorem 5.51. *If α is irrational, then all orbits of R_α are dense in $\mathbb{S}$.*



***Proof (to be completed).** We will consider the orbit of the point 0. Let U be some open interval in S, and denote its length by ℓ. We want to show that $R_\alpha^n(0) \in U$ for some n.

[*First, argue that there is a k such that $R_\alpha^k(0)$ is between 0 and $R_\alpha(0)$.*]

We now iterate R_α^k. [*Adapt your previous argument to show that there is an iterate of $(R_\alpha^k)^{k_2}$ that is between 0 and $R_\alpha^k(0)$.*]

By continuing this process as needed, we can find an n such that $R_\alpha^n(0)$ is less than ℓ away from 0. [*Argue that the R_α^n orbit of 0 must have a point in U, and therefore the R_α orbit of 0 must have a point in U.*] □

In fact, not only does each orbit densely cover $\mathbb{S}$, it does so *evenly* as n increases. What does this mean? Suppose that over your life, you have seen everything in the town you live—that is, you have traveled your town "densely". Certainly you have not done so uniformly, though—you have spent much more of your time at home than, say, at the local movie theater. *Uniformly distributing* means that an orbit spends, in the long run, the same amount of time[7] in equal-sized regions. Or, the size of an interval (relative to the whole space) is equivalent to the amount of time (relative to the whole time) the orbit spends in it. Since the length of $\mathbb{S}$ is 1, the length of a sub-interval $J \subset \mathbb{S}$ is the proportion of $\mathbb{S}$ taken up by J.

Definition 5.52. The orbit of x under R_α distributes uniformly on $\mathbb{S}$ iff for any interval $J \subset \mathbb{S}$,

$$\lim_{N\to\infty} \frac{\#\{\, n \mid R_\alpha^n(x) \in J,\ 0 \le n \le N \,\}}{N} = \ell(J)$$

where $\ell(J)$ is the length of J.

This condition says that, as we keep track of longer and longer sets of iterates, the proportion of the orbit of x that is in the set J converges to the proportion of the size of J in the circle (i.e., the length of J, since we are setting the length of $\mathbb{S}$ to 1).

***Exercise 5.53.** What goes wrong if we do not include the limit in Definition 5.52?

Theorem 5.54. *If α is irrational, then all orbits of R_α distribute uniformly in $\mathbb{S}$.*

[7] It is often useful to think of the iteration index as time—i.e., every iterate takes one unit of time to complete.

We will not prove this theorem, but note that the ideas in the proof of Theorem 5.51 can be adapted with astute uses of limits to generate a proof. The result will be of use to us below.

5.5.1. Benford's Law. In 1938, the physicist Frank Benford published a paper titled "The Law of Anomalous Numbers" in the *Proceedings of the American Philosophical Society*. Across 20,229 observations of numbers coming from 20 different types of sources (including newspapers, molecular weights, U.S. populations, and American Baseball League data from 1936), he found that the leading digits of numbers followed a non-uniform distribution—that is, certain digits appeared more often than others as the leading digit of a number. This non-uniform distribution of leading digits in many sources of data is now known as Benford's Law[8]. In turn, it has been used as evidence of certain kinds of fraud, since people who make up numbers tend to use a more uniform distribution of digits.

Benford's distribution law occurs with beautiful precision in the world of theoretical mathematics. Suppose, for instance, you write out the sequence consisting of powers of 2:

$$2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, 4096, \ldots.$$

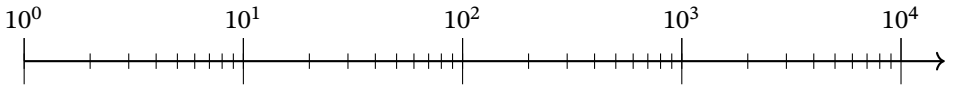
Suppose you keep track of only the first digits:

$$2, 4, 8, 1, 3, 6, 1, 2, 5, 1, 2, 4, \ldots.$$

Call this sequence S_n. Do all digits show up in S_n with equal probability, or does one occur more often than others? How often does a specific digit, say 3, occur? In order to answer these questions, we first need to develop more ways to think about how the powers of 2 are distributed on the real line.

Let's begin by considering the powers of 2 on a *logarithmic scale*[9]. You are used to plotting on number lines using a linear scale, on which the effect of adding a number k moves one down the number line the same length regardless of where one starts, while multiplying by k moves different numbers different lengths depending on where they start. On a log scale it is the opposite: multiplying by k produces the same length of movement regardless of where one is on the line, while adding k will move numbers different lengths depending on where they are.

***Exercise 5.55.** Each of the hash marks on the log scale number line below falls on an integer. Label some of them.



***Exercise 5.56.** Plot the powers of 2 on the log scale number line. What do you notice?

[8] As noted in the introduction to Chapter 4, mathematical results are often named after the second person to discover them: this distribution was described by Simon Newcomb in an 1881 paper "Note on the frequency of use of the different digits in natural numbers" in the *American Journal of Mathematics*.

[9] In base 10.

How can we distinguish numbers that, say, start with the digit 6? We want to be able to recognize 62 as well as 67,324. For example, x is a two-digit number that starts with a 6 if and only if

$$60 \leq x < 70.$$

A five-digit number y starts with a 6 if and only if

$$60000 \leq y < 70000.$$

On a linear scale, these two inequalities look very different, but on a log scale they are quite similar. Try highlighting on the log scale number line all of the numbers between 6 and 7, between 60 and 70, between 600 and 700, and between 6000 and 7000. What do you notice?

Let's formalize this with some arithmetic:

***Problem 5.57.** Take logarithms of the above inequality to show that, for a five-digit number x starting with the digit 6,

$$\log 6 + 4 \leq \log(x) < \log 7 + 4.$$

Generalize this to show that for an n-digit number x starting with the digit k,

$$\log k + (n-1) \leq \log(x) < \log(k+1) + (n-1).$$

Conclude that a natural number x starts with the digit k if and only if

$$\log k \leq \log(x) \bmod 1 < \log(k+1).$$

Hence, 2^n starts with the digit k if and only if $\log(2^n) \bmod 1 = n\log(2) \bmod 1$ is in the interval $[\log k, \log(k+1))$.

But $n\log(2) \bmod 1$ is just the orbit of the point $x = 0$ under the $R_{\log(2)}$ map. And because $\log(2)$ is irrational, Theorem 5.54 implies that the proportion of times the orbit spends in $J = [\log k, \log(k+1))$ is $\ell(J) = \log(k+1) - \log k$. That is, the proportion of time that 2^n starts with the digit k is the proportion of time the orbit $R^n_{\log 2}(0)$ spends in $J = [\log k, \log(k+1))$, which is $\ell(J) = \log(k+1) - \log k$. This proves Theorem 5.58.

Theorem 5.58. *In the sequence S_n, the digit k occurs with probability* $\log(k+1) - \log(k)$.

That is, the digit 1 occurs with probability $\log(2) - \log(1) \approx 0.301$, or approximately 30.1% of the time. Similarly, the other digits occur with the following probabilities:

Digit	Probability
1	30.1%
2	17.6%
3	12.5%
4	9.7%
5	7.9%
6	6.7%
7	5.8%
8	5.1%
9	4.6%

5.6. The billiard map and phase space

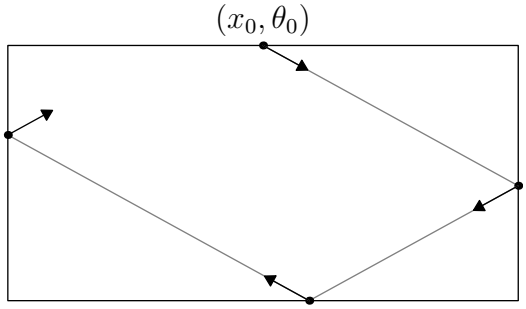
The previous sections of this chapter all dealt with one-dimensional dynamical systems. In this section, we briefly explore a two-dimensional dynamical system: the billiard map.

A mathematical billiard consists of a closed region (the "table") on the plane and the following set of rules that govern the motion of a single particle (a "ball"):

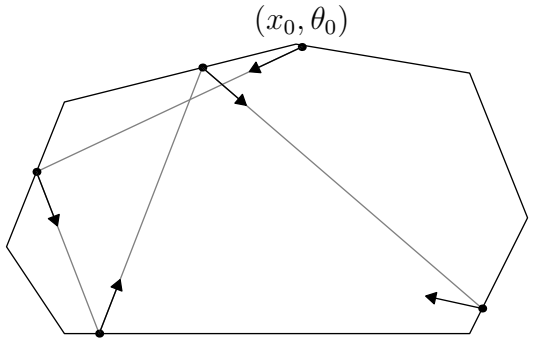
(1) in the interior of the table, a ball travels in a straight line at constant velocity; and

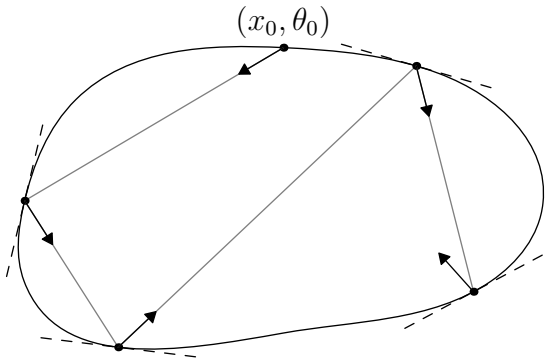
(2) when a ball hits the boundary of the table, its trajectory is reflected so that the incoming angle is equal to the outgoing angle.

That is, a mathematical billiard is a billiard table in which there is no friction or other type of force (except for the elastic collisions with the sides), and a ball is considered a single point. For example, here is a picture of the first few bounces of a billiard trajectory on a rectangular table:



Even though the word "billiard" may make you think of the rectangular table drawn above, a mathematical billiard table need not be rectangular—it may even have a curved boundary. In this case, the angles are calculated with respect to the tangent line to the table at the point where the billiard ball hits.



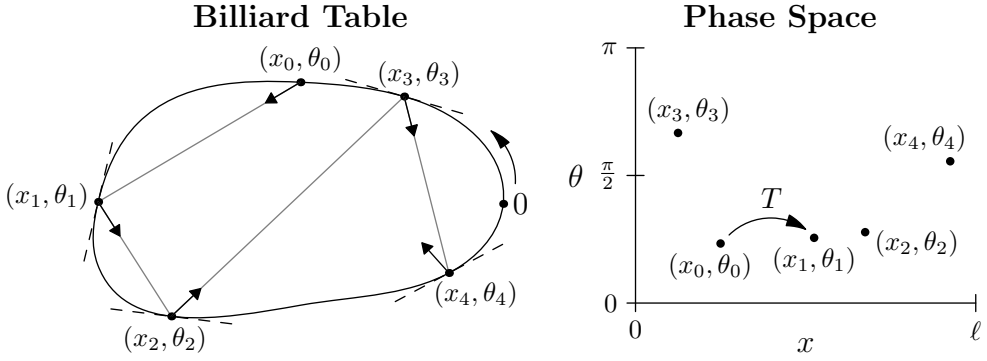


From a mathematical billiard one can extract a discrete dynamical system in the following way: We will ignore the path of the billiard ball as it traverses the table and only keep track of where and at what angle it hits (or, equivalently, bounces off) the boundary. This information determines the next point at which the trajectory hits the boundary, as well as the incoming (and hence outgoing) direction. Thus, the billiard rules above determine a billiard map $T : X \to X$, where X is the set of pairs consisting of boundary points and outgoing directions. X is called the *phase space*.

In order to analyze this map, we put coordinates on X as follows. Choose a starting point O on the boundary of the billiard table and label boundary points by their arc-length distance from O, measured counterclockwise. This gives us a coordinate $x \in [0, \ell)$ along the boundary, where ℓ is the length of the boundary. For a billiard ball trajectory leaving the boundary, let θ be the angle it makes with the tangent to the boundary oriented counterclockwise; this implies that $0 < \theta < \pi$. The pair (x, θ) are coordinates on X, and the billiard map is

$$\boxed{T : [0, \ell) \times (0, \pi) \to [0, \ell) \times (0, \pi),}$$

where T maps an (x, θ) to the next boundary point and (outgoing) angle.



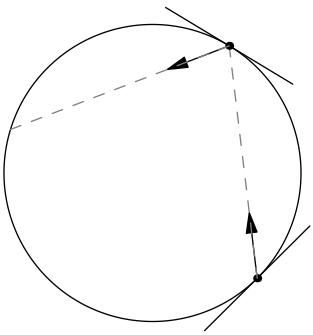
Note that the phase space is a two-dimensional region in which points move around by the rules of the billiard map—despite the two axes, this is *not* the graph of some function.

***Problem 5.59.** What would it mean for two points in the same orbit to be on the same vertical line in the phase space? Can you come up with a billiard table where this happens?

Let's investigate billiard maps on some basic shapes: the circle and the square.

***Problem 5.60.** Consider the billiard table whose boundary is the unit circle. Recall that the billiard map $T : [0, 2\pi) \times (0, \pi) \to [0, 2\pi) \times (0, \pi)$ is the map that takes a point and direction (x, θ) on the circle and sends it to the point and direction after the next hit along the billiard trajectory.

(1) Find a formula for $T(x, \theta)$ (the figure below might be useful).



(2) Plot some orbits for the circular billiard table in the phase space $X = [0, 2\pi) \times (0, \pi)$. What do you notice?

(3) If you restrict this map to a specific θ value, what do you get? In other words, what is $T(x, \theta_0)$ for a fixed θ_0? Can you recognize this as a map from one of the previous sections in this chapter?

The very regular behavior of these orbits is quite striking! We formalize it with the following definition.

Definition 5.61. A piece-wise continuous curve C in the phase space is called *invariant* if whenever x is on C, then $T(x)$ is also on C. A billiard map is *integrable* if the phase space is entirely made up of (or "foliated by") invariant curves.

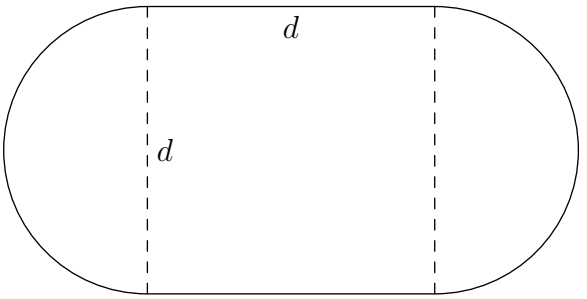
***Problem 5.62.** Let $\mathcal{B} = [0, 1] \times [0, 1]$ be the unit square billiard table. Let x be the distance measured counterclockwise around the square from the corner $(0, 0)$. The billiard map is a map $T : [0, 4) \times (0, \pi) \to [0, 4) \times (0, \pi)$. Find a formula for $T(\frac{1}{2}, \theta)$. $\star$

Recall from the definition that invariant curves in the phase space only need to be *piece-wise* continuous—i.e., continuous except at a finite number of points. This allows us to also say that the square has very regular billiard behavior:

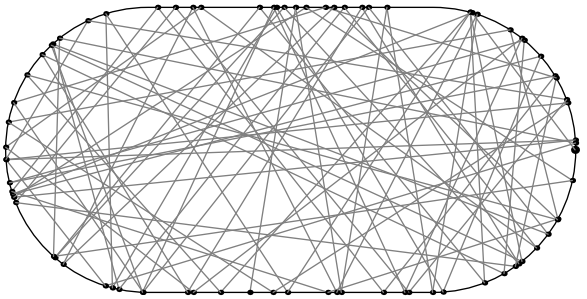
***Problem 5.63.** Plot some orbits of the billiard map on the square table. Argue that the square billiard is integrable. What do the invariant curves look like? Compare this with the circle billiard.

So both circle and square billiard tables have very regular behavior. What do you think happens if we combine these shapes into a single billiard table? Suppose we cut

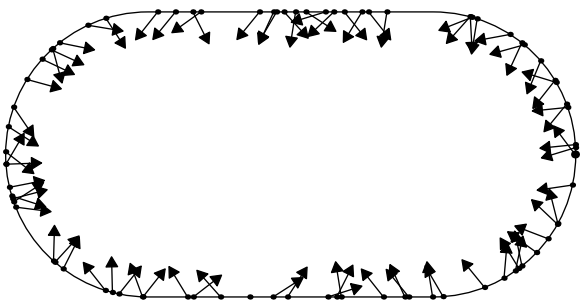
the circle of diameter d in half and attach the half-disks to opposite ends of a square of side length d, like this:

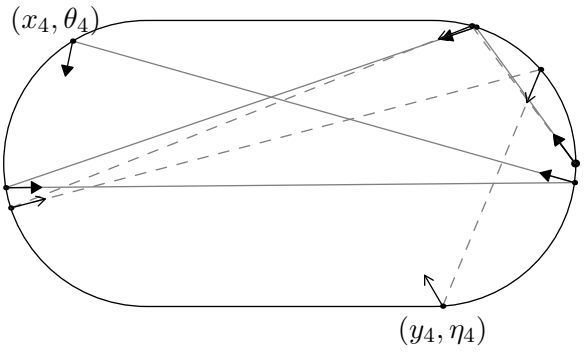


This billiard table is known as the "Bunimovich Stadium". Here is a picture of the first 80 segments of an orbit on the Bunimovich Stadium:



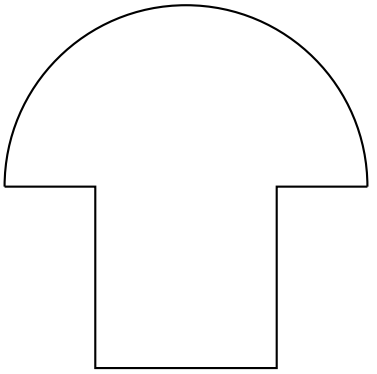
The orbit doesn't look very regular! In fact, it is not—the Bunimovich Stadium is a classic example of *chaotic* behavior. Recall from §5.3 our definition of chaos (Definition 5.34): there is a dense orbit, the set of periodic points is dense, and the system has sensitive dependence on initial conditions. These properties are all present in the Bunimovich Stadium, though we will not develop the tools to rigorously show this. Two of these properties are illustrated in the following figures, based on the same orbit as above. Can you describe which ones?



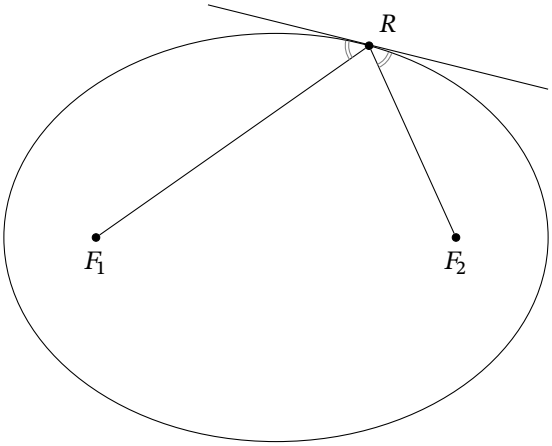


In fact, many billiard tables are neither completely integrable nor completely chaotic, but have a little bit of both behaviors. This is the case for a family of tables known as mushrooms.

***Problem 5.64.** Let $\mathcal{B}$ be the billiard table obtained from cutting the unit circle billiard table in half and attaching the unit square at the bottom in the middle. This billiard table has the property that in some areas of the phase space the dynamics are integrable, while in other areas of the phase space the dynamics are chaotic. Can you determine which areas are which? You might start by trying to draw some trajectories in $\mathcal{B}$.



Another class of billiard tables that can be seen as generalizations of the circle are elliptical billiards. We will see that the highly regular behavior of circular billiard tables is, to a great extent, shared by the more general elliptical billiard tables, although explicit formulas are hard to come by. Let us begin by reviewing a definition and some properties of ellipses.



An ellipse has two foci (plural of focus), the points F_1 and F_2 in the figure above, and has the property that for any point R on the ellipse the sum $\ell = d(F_1, R) + d(R, F_2)$ is a constant independent of R. When $F_1 = F_2$, the ellipse is a circle and the constant d is the diameter of the circle. Two ellipses having the same foci (and different values of ℓ) are said to be *confocal*.

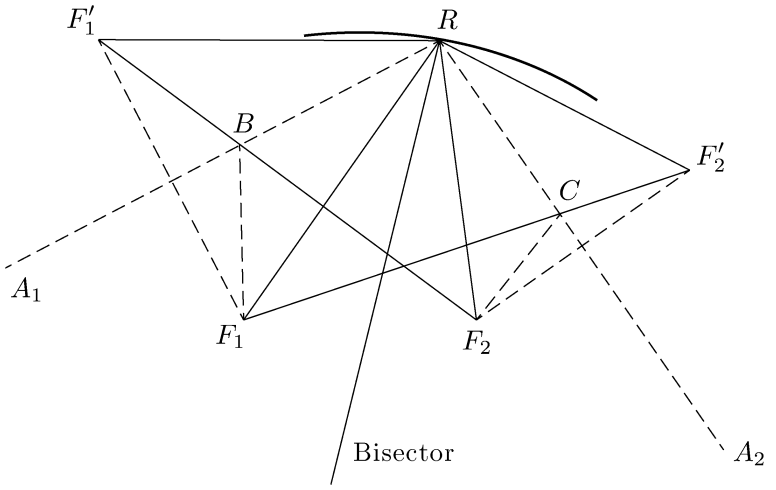
A key property of the ellipse is that every billiard trajectory that passes through one of the foci reflects at the boundary and then passes through the other focus; this is indicated in the figure above by the equality of the marked angles with the tangent. We will not prove this here, but there is a fun and pretty convincing intuitive mechanical argument: Think of a non-stretchable string of length ℓ, attached at its endpoints to the foci. With the tip of a sharp pencil, pull the string tight. Then the tip of the pencil will be a point on the ellipse. Since the pencil is not moving along the ellipse, the sum of the tension forces along F_1R and F_2R must be a vector normal (perpendicular) to the ellipse. Since the magnitudes of those forces are equal to each other, the direction of their vector sum is parallel to the bisector of $\angle F_1RF_2$. Voilà: the tangent at R is perpendicular to the bisector of $\angle F_1RF_2$.

Theorem 5.65. *If a billiard trajectory on an elliptical table does not cut the line segment between the two foci, then each straight-line segment of the trajectory is tangent to a fixed confocal ellipse.*

***Proof (to be completed).** Refer to the following figure. Let R be the reflection point of a trajectory A_1RA_2 on an ellipse with foci F_1 and F_2. We make the following construction:

(a) Reflect F_1 over A_1R to get F_1', and reflect F_2 over RA_2 to get F_2'.

(b) Take confocal ellipses (i.e., ellipses with foci F_1 and F_2) of growing size until one is tangent to A_1R. Label the point of tangency B and the ellipse $\mathcal{E}_B$. Do the same thing on the other side of the diagram to get an ellipse $\mathcal{E}_C$ with foci F_1 and F_2 that is tangent to RA_2 at the point C.



We want to show that $\mathcal{E}_B$ and $\mathcal{E}_C$ are the same ellipse. Note that these ellipses are confocal (by their definition), so it suffices to show that

$$d(F_1, B) + d(B, F_2) = d(F_1, C) + d(C, F_2).$$

Do this by carrying out the following arguments:

(1) F_1BF_2 is a billiard trajectory for $\mathcal{E}_B$ (why?). Use this to show that F_1', B, and F_2 are collinear. Note that the same argument works on the other side to show that F_1, C, and F_2' are collinear.

(2) Argue that $(d(F_1, B) = d(F_1', B))$ so that $(d(F_1, D) + d(B, F_2) = d(F_1', F_2)$.

(3) Find two congruent triangles, one with an edge $F_1'F_2$ and the other with an edge $F_2'F_1$, and prove they are congruent.

(4) Finish the proof!

□

A single segment (between bounces) of a billiard trajectory in an ellipse falls into exactly one of the following three categories:

(1) the segment passes through one of the foci (so that, according to the discussion preceding Theorem 5.65, the trajectory must pass through one of the foci between any two bounces); or

(2) the segment does not travel between the foci (so that, according to Theorem 5.65, there is a confocal ellipse such that between any two bounces the trajectory must be tangent to the confocal ellipse); or

(3) the segment intersects the line segment connecting the two foci.

Case (3) has a similar characterization to case (2): such a trajectory (passing between the foci) is tangent to a *confocal hyperbola* between any two bounces.

***Problem 5.66.** Argue that for any billiard trajectory that passes between the foci of an elliptical billiard table, there is a confocal hyperbola such that the trajectory is tangent to the confocal hyperbola between any two bounces.

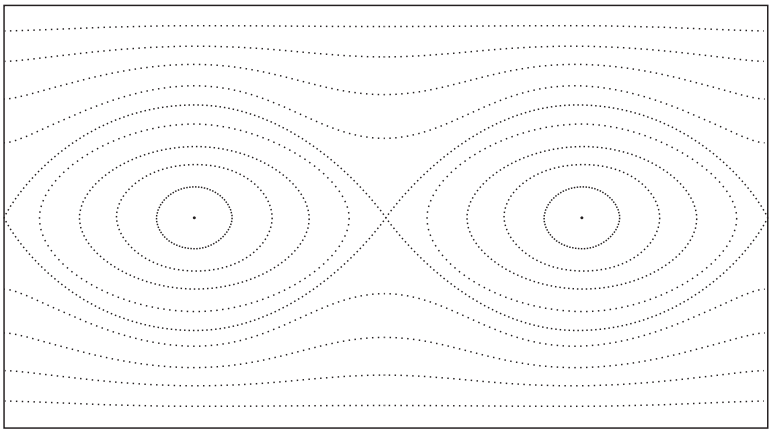


Figure 5.7. A sketch of some orbits in the phase space of the billiard map on an ellipse. Can you reconcile this figure with the orbits in the preceeding discussion? ⋆