Chapter 2
The Real Numbers

2.1 An Overview of the Real Numbers

Doing analysis in a rigorous way starts with understanding the properties of the real numbers. Readers will be familiar, in some sense, with the real numbers from studying calculus. A completely rigorous development of the real numbers requires checking many details. We attempt to justify one definition of the real numbers without carrying out the proofs.

Intuitively, we think of the real numbers as the points on a line stretching off to infinity in both directions. However, to make any sense of this, we must label all the points on this line and determine the relationship between them from different points of view. First, the real numbers form an algebraic object known as a field, meaning that one may add, subtract, and multiply real numbers and divide by nonzero real numbers. There is also an order on the real numbers compatible with these algebraic properties, and this leads to the notion of distance between two points.

All of these nice properties are shared by the set of rational numbers:

$$\mathbb{Q} = \left\{ \frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0 \right\}.$$

The ancient Greeks understood how to construct all fractions geometrically and knew that they satisfied all of the properties mentioned above. However, they were also aware that there were other points on the line that could be constructed but were not rational, such as $\sqrt{3}$. While the Greeks were focussed on those numbers that could be obtained by geometric construction, we have since found other reasonable numbers that do not fit this restrictive definition. The most familiar example is perhaps π, the area of a circle of radius one. Like the Greeks, we accept the fact that $\sqrt{3}$ and π are bona fide numbers that must be included on our real line.

We will define the real numbers to be objects with an infinite decimal expansion. A subtle point is that an infinite decimal expansion is used only as a name for a point and does mean the sum of an infinite series. It is crucial that we do not use limits to define the real numbers because we deduce properties of limits from the definition.

K.R. Davidson and A.P. Donsig, *Real Analysis and Applications: Theory in Practice*,
Undergraduate Texts in Mathematics, DOI 10.1007/978-0-387-98098-0_2,

This construction of the real numbers appears to be strongly dependent on the choice of 10 as the base. We are left with the nagging possibility that the number line we construct depends on the number of digits on our hands. For this reason, some purists prefer a base-independent method of defining the real numbers, albeit a more abstract one. (See Exercise 2.8.L.) Our construction does yield the same object, independent of choice of base; but the proof requires considerable work.

2.2 The Real Numbers and Their Arithmetic

We define a real number using an infinite decimal expansion such as

$$\begin{aligned}\frac{1}{3} &= 0.33333333333333333333333333333333333333333333333\dots\\ \sqrt{3} &= 1.73205080756887729352744634150587236694280525381038\dots\\ \pi &= 3.14159265358979323846264338327950288419716939937510\dots\end{aligned}$$

In general, an infinite decimal expansion has the form

$$x = a_0.a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}a_{11}a_{12}a_{10}a_{11}a_{12}a_{13}a_{14}a_{15}a_{16}a_{17}a_{18}\dots.$$

Formally, an **infinite decimal expansion** is a function $x(n) = a_n$ from $\{0\} \cup \mathbb{N}$ into $\mathbb{Z}$ such that for all $n \geq 1$, $a_n \in \{0, 1, \dots, 9\}$.

Be warned that, by this construction, the point usually thought of as $-5/4$ will be denoted by $\bar{2}.75$, for example, because we think of this as $-2 + .75$. The notation is simpler if we do this. After we have finished the construction, we will revert to the standard notation for negative decimals.

To relate infinite decimal expansions to our geometric idea of the real line, start with a line and mark two points on it; and call the left one 0 and the right one 1. Then we can construct points for every integer $\mathbb{Z}$, equally spaced along the line. Now divide each interval from an integer n to $n+1$ into 10 equal pieces, marking the cuts as $n.1$, $n.2$, ..., $n.9$. Proceed in this way, cutting each interval of length 10^{-k} into 10 equal intervals of length 10^{-k-1} and mark the endpoints by the corresponding number with $k+1$ decimals. In this way, all finite decimals are placed on the line.

To obtain a geometric version of the line, we postulate that for every infinite decimal $x = a_0.a_1a_2a_3\dots$, there will be a point (also called x) on this line with the property that for each positive integer k, x lies in the interval between the two rational numbers $y = a_0.a_1\dots a_k$ and $y + 10^{-k}$. For example,

$$3.141592653589 \leq \pi \leq 3.141592653590.$$

One difficulty with using infinite decimal expansions to define the real numbers is that some points have two names. For example consider the expansions $1.000000000\dots$ and $0.999999999\dots$. Call them 1 and z, respectively. Clearly these are *different* infinite decimal expansions. However, for each positive integer k,

$$1 - 10^{-k} = 0.\underbrace{99999999999999}_{k} \leq z \leq 1.$$

Thus the difference between z and 1 is arbitrarily small. It would create quite an unintuitive line if we decided to make z and 1 different real numbers. To fit in with our intuition, we must agree that $z = 1$. That means that some real numbers (precisely all those numbers with a finite decimal expansion) have two different expansions, one ending in an infinite string of zeros, and the other ending with an infinite string of nines. For example, $0.12500\ldots$ and $0.12499999\ldots$ are the same number.

Formally, this defines an equivalence relation on the set of infinite decimals by pairing off each decimal expansion ending in a string of zeros with the corresponding decimal expansion ending in a string of nines:

$$a_0.a_1a_2\ldots a_{k-1}a_k000\ldots = a_0.a_1a_2\ldots a_{k-1}(a_k - 1)999\ldots,$$

where $a_k \neq 0$. Each real number is an equivalence class of infinite decimal expansions given by this identification. The set of all real numbers is denoted by $\mathbb{R}$.

To recognize the rationals as a subset of the reals, we need a function F that sends a fraction a/b to an infinite decimal expansion. This is accomplished by long division, as you learned in grade school. For example, to compute $27/14$, divide 14 into 27 to obtain

$$F\left(\tfrac{27}{14}\right) = 1.928571428571428571428571428571 4\ldots.$$

Notice that this decimal expansion is *eventually periodic* because after the initial 1.9, the six-digit sequence 285714 is repeated ad infinitum. In the exercises, hints are provided to show that an infinite decimal represents a rational number if and only if it is eventually periodic.

We have a built-in order on the real line given by the placement of the points which extends the natural order on the finite decimals. When two infinite decimals $x = a_0.a_1a_2\ldots$ and $y = b_0.b_1b_2\ldots$ represent *distinct* real numbers, we say that $x < y$ if there is some integer $k \geq 0$ such that $a_i = b_i$ for $i < k$ and $a_k < b_k$. For example, if

$$x = 2.73421182845904523540000643383250288419716939 93\ldots,$$
$$y = 2.73421182845904523539999287471352249775724709 36\ldots,$$

then $y < x$ because

$$y < 2.734211828459045235399993 < 2.734211828459045235400000 < x.$$

For two real numbers x and y, either $x < y$, $x = y$, or $x > y$.

Next we extend the addition and multiplication operations on $\mathbb{Q}$ to all of $\mathbb{R}$. The basic idea is to extend addition and multiplication on finite decimals to $\mathbb{R}$ respecting the order properties. That is, if $w \leq x$ and $y \leq z$, then $w + y \leq x + z$, and if $x \geq 0$, then $xy \leq xz$. Some of the subtleties are explored in the exercises.

A basic fact about the order and these operations is known as the **Archimedean property** of $\mathbb{R}$: *for $x, y > 0$, there is always some $n \in \mathbb{N}$ with $nx > y$.* It is not hard to show this is equivalent to the following almost-obvious fact: *if $z > 0$, then there is some integer $k \geq 0$ so that $10^{-k} < z$.* To see this fact, observe that a decimal expansion of $z = z_0.z_1z_2\dots$ has a first nonzero digit, z_{k-1} and, since $z_{k-1} \geq 1$, we have $z \geq 10^{-(k-1)} > 10^{-k}$.

Finally, consider the distance between two points. The **absolute value function** is $|x| = \max\{x, -x\}$. Define the distance between x and y to be $|x-y|$. This is always nonnegative, and $|x-y| = 0$ only if $x - y = 0$, namely $x = y$.

Exercises for Section 2.2

A. Why, in defining the order on $\mathbb{R}$, did we insist that x and y be distinct real numbers?
HINT: consider a real number with two decimal expansions.

B. Prove that $|xy| = |x|\,|y|$ and $|x^{-1}| = |x|^{-1}$.

C. (a) Prove the **triangle inequality**: $|x+y| \leq |x| + |y|$.
HINT: Consider x and y of the same sign and different signs as separate cases.
(b) Prove by induction that $|x_1 + x_2 + \dots + x_n| \leq |x_1| + |x_2| + \dots + |x_n|$.
(c) Prove the **reverse triangle inequality**: $\big||x| - |y|\big| \leq |x-y|$.

D. (a) Prove that if $x < y$, then there is a rational number r with a finite decimal expansion and an integer k so that $x < r < r + 10^{-k} < y$.
(b) Prove that if $x < y$, then there is an irrational number z such that $x < z < y$.
HINT: Use (a) and add a small multiple of $\sqrt{2}$ to r.

E. (a) Explain how $x + y$ is worked out for

$$x = 2.1357\overbrace{999999\dots999999}^{10^7 \text{ nines}}\overbrace{0123456789\dots0123456789}^{10^{19} \text{ repetitions}}34524\dots,$$
$$y = 3.8642\underbrace{999999\dots999999}_{10^7 \text{ nines}}\underbrace{9876543210\dots9876543210}_{10^{19} \text{ repetitions}}39736\dots.$$

(b) How many digits of x and y must we know to determine the first 6 digits of $x+y$?
(c) How many digits of x and y must we know to determine the first 10^8 digits of $x+y$?

F. Describe an algorithm for adding two infinite decimals. You should work from 'left to right', determining the decimal expansion in order, as much as possible. When are you assured that you know the integer part of the sum? In what circumstance does it remain ambiguous?
HINT: Given infinite decimals a and b, define a carry function $\gamma : \{0\} \cup \mathbb{N} \to \{0,1\}$ and then define the decimal expansion of $a+b$ in terms of $a(n) + b(n) + \gamma(n)$.

G. Show that if x and y are known up to k decimal places, then the $x+y$ is known to within $2 \cdot 10^{-k}$, i.e., there is a finite decimal r with $r \leq x+y \leq r + 2 \cdot 10^{-k}$.

H. An infinite decimal $x = a_0.a_1a_2\dots$ is *eventually periodic* if there are positive integers n and k such that $a_{i+k} = a_i$ for all $i > n$. Show that any decimal expansion which is eventually periodic represents a rational number. HINT: Compute $10^{n+k}x - 10^n x$.

I. Prove that the decimal expansion of a rational number p/q is eventually periodic. We will use the Pigeonhole Principle, which states that if $n+1$ items are divided into n categories, then at least two of the items are in the same category.
(a) Assume $q > 0$. Let r_k be the remainder when 10^k is divided by q. Use the Pigeonhole Principle to find two different exponents $k < k+d$ with the same remainder.
(b) Express $p/q = 10^{-k}\big(a + b/(10^d - 1)\big)$ with $0 \leq b < 10^d - 1$.
(c) Write b as a d-digit number $b = b_1b_2\dots b_d$ even if it starts with some zeros. Show that the decimal expansion of p/q ends with the infinitely repeated string $b_1b_2\dots b_d$.

J. Explain how the associative property of addition for real numbers: $x+(y+z)=(x+y)+z$ follows from knowing it for for finite decimals.

K. Show that if r is rational and x is irrational, then $r+x$ and, if $r \neq 0$, rx are irrational.

L. Show that the two formulations of the Archimedean property of $\mathbb{R}$ are equivalent.

2.3 The Least Upper Bound Principle

After defining the least upper bound of a set of real numbers, we prove the Least Upper Bound Principle (2.3.3). This result depends crucially on our construction of the real numbers. It will be the basis for the deeper properties of the real line.

2.3.1. DEFINITION. A set $S \subset \mathbb{R}$ is **bounded above** if there is a real number M such that $s \leq M$ for all $s \in S$. We call M an **upper bound** for S. Similarly, S is **bounded below** if there is a real number m such that $s \geq m$ for all $s \in S$, and we call m a **lower bound** for S. A set that is bounded above and below is called **bounded**.

Suppose a nonempty subset S of $\mathbb{R}$ is bounded above. Then L is the **supremum** or **least upper bound** for S if L is an upper bound for S that is smaller than all other upper bounds, i.e., for all $s \in S$, $s \leq L$, *and* if M is another upper bound for S, then $L \leq M$. It is denoted by $\sup S$.

Similarly, if S is a nonempty subset of $\mathbb{R}$ which is bounded below, the **infimum** or **greatest lower bound**, denoted by $\inf S$, is the number L such that L is an lower bound *and* whenever M is another lower bound for S, then $L \geq M$.

The supremum of a set, if it exists, is unique. We have not defined suprema or infima for sets that are not bounded above or bounded below, respectively. For example, $\mathbb{R}$ itself has neither a supremum nor an infimum. For a nonempty set $S \subseteq \mathbb{R}$, sometimes we write $\sup S = +\infty$ if S is not bounded above and $\inf S = -\infty$ if S is not bounded below. Finally, by convention, $\sup \varnothing = -\infty$ and $\inf \varnothing = +\infty$.

Note that $\sup S = L \in \mathbb{R}$ if and only if L is a upper bound for S and for all $K < L$, there is $x \in S$ with $K < x < L$. There is an equivalent characterization for $\inf S$.

Recall that the **maximum** of a set $S \subset \mathbb{R}$, *if it exists*, is an element $m \in S$ such that $s \leq m$ for all $s \in S$. Thus, when the maximum of a set exists, it is the least upper bound. The situation for the **minimum** of a set and its infimum is the same. We use $\max S$ and $\min S$ to denote the maximum and minimum of S.

2.3.2. EXAMPLES.

(1) If $A = \{4, -2, 5, 7\}$, then any $L \leq -2$ is a lower bound for A and any $M \geq 7$ is an upper bound. So, $\inf A = \min A = -2$ and $\sup A = \max A = 7$.

(2) If $B = \{2, 4, 6, \dots\}$, then $\inf B = \min B = 2$ and $\sup B = +\infty$.

(3) If $C = \{\pi/n : n \in \mathbb{N}\}$, then $\sup C = \max C = \pi$. However, for any element of C, say π/n, we have a smaller element of C, such as $\pi/(2n)$. So C does not have a

minimum. Clearly, 0 is a lower bound and for all $x > 0$, there is some $\pi/n \in C$ with $\pi/n < x$, showing that 0 is the greatest lower bound.

(4) If $D = \{(-1)^n n/(n+1) : n \in \mathbb{N}\}$, then D has neither a maximum nor a minimum. However, D has upper and lower bounds, and $\inf D = -1$ and $\sup D = 1$. Neither 1 nor -1 belongs to D.

In proving the Least Upper Bound Principle, the definition of the real numbers as *all* infinite decimals is essential. The principle is not true for some subsets of the rational numbers. For example, $\{s \in \mathbb{Q} : s^2 < 2\}$ is bounded above but has no least upper bound in $\mathbb{Q}$.

2.3.3. Least Upper Bound Principle.

Every nonempty subset S of $\mathbb{R}$ that is bounded above has a supremum. Similarly, every nonempty subset S of $\mathbb{R}$ that is bounded below has an infimum.

Proof. We prove the second statement first, since it is more convenient. Let M be some lower bound for S with decimal expansion $M = m_0.m_1m_2\ldots$. Let s be some element of S with decimal expansion $s = s_0.s_1s_2\ldots$. Notice that since $m_0 \leq M$, we have that m_0 is a lower bound for S. On the other hand, $s < s_0 + 2$. So $s_0 + 2$ is not a lower bound. There are only finitely many integers between m_0 and $s_0 + 1$. Pick the largest of these that is still a lower bound for S, and call it a_0. Since $a_0 + 1$ is not a lower bound, we may also choose an element x_0 in S such that $x_0 < a_0 + 1$.

Next pick the greatest integer a_1 such that $y_1 = a_0 + 10^{-1}a_1$ is a lower bound for S. Since $a_1 = 0$ works and $a_1 = 10$ does not, a_1 belongs to $\{0, 1, \ldots, 9\}$. To verify our choice, pick an element x_1 in S such that $a_0.a_1 \leq x_1 < a_0.a_1 + 0.1$. Continue in this way recursively. Figure 2.1 shows how a_2 and x_2 would be chosen.

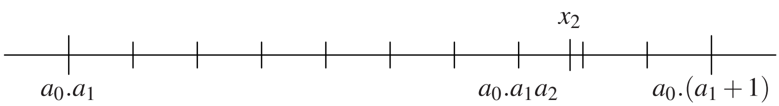

Fig. 2.1 The second stage ($k = 2$) in the proof.

At the kth stage, we have a lower bound $y_{k-1} = a_0.a_1\ldots a_{k-1}$ and an element $x_{k-1} \in S$ such that $y_{k-1} \leq x_{k-1} < y_{k-1} + 10^{1-k}$. Select the largest integer a_k in $\{0, 1, \ldots, 9\}$ such that $y_k = a_0.a_1a_2\ldots a_k$ is a lower bound for S. Since $y_k + 10^{-k}$ is not a lower bound, we also pick an element x_k in S such that $x_k < y_k + 10^{-k}$ to verify our choice.

We claim that $L = a_0.a_1a_2\ldots$ is $\inf S$. If $L = y_k$ for some k, then L is a lower bound for S. Otherwise, $L > y_k$ for all k and, in particular, for each k there is $l > k$ with $y_l > y_k$. If $s = s_0.s_1s_2\ldots$ is in S, then it follows that $s > y_k$ for each k. By the definition of the order, either $s_i = a_i$ for $1 \leq i \leq k$ or there is some j, $0 \leq j \leq k$, with $s_i = a_i$ for $1 \leq i < j$ and $s_j > a_j$. If the latter occurs for some k, then $s > L$; if the former occurs for every k, then $s = L$. Either way, L is a lower bound for S.

To see that L is the greatest lower bound, suppose $M = b_0.b_1b_2\ldots > L$. By the definition of the ordering, there is some first integer k such that $b_k > a_k$ and $b_i = a_i$ for all i with $0 \le i < k$. But then

$$M \ge a_0.a_1\ldots a_{k-1}b_k \ge y_k + 10^{-k} > x_k.$$

So M is not a lower bound for S. Hence L is the greatest lower bound.

A simple trick handles upper bounds. Notice that $S \subset \mathbb{R}$ is bounded above if and only if $-S = \{-s : s \in S\}$ is bounded below and that L is an upper bound for S precisely when $-L$ is a lower bound for $-S$. Further, $M < L$ if and only if $-M > -L$, so M is an upper bound of S less than L exactly when $-M$ is a lower bound of $-S$ greater than $-L$. Thus $\sup S = -\inf(-S)$, so $\sup S$ exists. ■

Exercises for Section 2.3

A. Suppose that $S \subset \mathbb{R}$ is bounded above. When does S have a maximum? Your answer should be expressed in terms of $\sup S$.

B. A more elegant way to develop the arithmetic properties of the real numbers is to prove the results of this section first and then define addition and multiplication using suprema. Let $\mathscr{D}$ denote the set of all finite decimals.
(a) Let $x, y \in \mathbb{R}$. Prove that $x + y = \sup\{a + b : a, b \in \mathscr{D},\ a \le x,\ b \le y\}$.
(b) Suppose that $x, y \in \mathbb{R}$ are positive. Show that $xy = \sup\{ab : a, b \in \mathscr{D},\ 0 \le a \le x,\ 0 \le b \le y\}$.
(c) How do we define multiplication in general?

C. With $\mathscr{D}$ as in the previous exercise, show that $\sup\{a \in \mathscr{D} : a^2 \le 3\} = \sqrt{3}$.

D. For the following sets, find the supremum and infimum. Which have a max or min?
(a) $A = \{a + a^{-1} : a \in \mathbb{Q},\ a > 0\}$.
(b) $B = \{a + (2a)^{-1} : a \in \mathbb{Q},\ 0.1 \le a \le 5\}$.
(c) $C = \{xe^{-x} : x \in \mathbb{R}\}$.

E. Show that the decimal expansion for the L in the proof of the Least Upper Bound Principle does not end in a tail of all 9's.

2.4 Limits

The notion of a limit is *the* basic notion of analysis. Limits are the culmination of an infinite process. It is the concern with limits in particular that separates analysis from algebra. Intuitively, to say that a sequence a_n converges to a limit L means that eventually *all* the terms of the (tail of the) sequence approximate the limit value L to *any* desired accuracy. To make this precise, we introduce a subtle definition.

2.4.1. Definition of the Limit of a Sequence. A real number L is the **limit** of a sequence of real numbers $(a_n)_{n=1}^\infty$ if for *every* $\varepsilon > 0$, there is an integer $N = N(\varepsilon) > 0$ such that

$$|a_n - L| < \varepsilon \quad \text{for all} \quad n \ge N.$$

We say that the sequence $(a_n)_{n=1}^\infty$ **converges** to L, and we write $\lim\limits_{n\to\infty} a_n = L$.

The important issue in this definition is that for any desired accuracy, there is a point in the sequence such that *every* element after that point approximates the limit L to the desired accuracy. It suffices to consider only values for ε of the form $\frac{1}{2}10^{-k}$. The statement $|a_n - L| < \frac{1}{2}10^{-k}$ means that a_n and L agree to at least k decimal places. Thus a sequence converges to L precisely when for every k, no matter how large, eventually all the terms of the sequence agree with L to at least k decimals of accuracy.

2.4.2. EXAMPLE. Consider the sequence $(a_n) = (n/(n+1))_{n=1}^{\infty}$, which we claim converges to 1. Observe that $\left|\frac{n}{n+1} - 1\right| = \frac{1}{n+1}$. So if $\varepsilon = \frac{1}{2}10^{-k}$, we can choose $N = 2\cdot 10^k$. Then for all $n \geq N$,

$$\left|\frac{n}{n+1} - 1\right| = \frac{1}{n+1} \leq \frac{1}{2\cdot 10^k + 1} < \tfrac{1}{2}10^{-k} = \varepsilon.$$

We could also choose $N = 73\cdot 10^k$. It is not necessary to find the best choice for N. But in practice, better estimates can lead to better algorithms for computation.

2.4.3. EXAMPLE. Consider the sequence (a_n) with $a_n = (-1)^n$. Since this flips back and forth between two values that are always distance 2 apart, intuition says that it does not converge. To show this using our definition, we need to show that the definition of limit fails for *any* choice of L. However, for each choice of L, we need find *only one* value of ε that violates the definition. Observe that

$$|a_n - a_{n+1}| = |(-1)^n - (-1)^{n+1}| = 2$$

for all n, no matter how large. So let L be any real number. We notice that L cannot be close to both 1 and -1. To avoid cases, we use a trick. For any real number L,

$$|a_n - L| + |a_{n+1} - L| \geq |(a_n - L) - (a_{n+1} - L)| = |a_n - a_{n+1}| = 2.$$

Thus, for every $n \in \mathbb{N}$,

$$\max\{|a_n - L|, |a_{n+1} - L|\} \geq 1. \tag{2.4.4}$$

Now take $\varepsilon = 1$. If this sequence *did* converge, there would be an integer N such that $|a_n - L| < 1$ for all $n \geq N$. In particular, $|a_N - L|$ and $|a_{N+1} - L|$ are both less than 1, contradicting (2.4.4). Consequently, this sequence does not converge.

2.4.5. EXAMPLE. Consider the sequence $((\sin n)/n)_{n=1}^{\infty}$.The numerator oscillates, but it remains bounded between ± 1 while the denominator goes off to infinity. We obtain the estimates

$$-\frac{1}{n} \leq \frac{\sin n}{n} \leq \frac{1}{n}.$$

We know that $\lim\limits_{n\to\infty} 1/n = 0 = \lim\limits_{n\to\infty} -1/n$, since this is exactly like Example 2.4.2. Therefore, the limit can be computed using a familiar principle from calculus:

2.4.6. THE SQUEEZE THEOREM.
Suppose that three sequences (a_n), (b_n), and (c_n) satisfy

$$a_n \le b_n \le c_n \quad \text{for all} \quad n \ge 1 \qquad \text{and} \qquad \lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = L.$$

Then $\lim_{n\to\infty} b_n = L$.

PROOF. Let $\varepsilon > 0$. Since $\lim_{n\to\infty} a_n = L$, there is some N_1 such that

$$|a_n - L| < \varepsilon \quad \text{for all} \quad n \ge N_1,$$

or equivalently, $L-\varepsilon < a_n < L+\varepsilon$ for all $n \ge N_1$. There is also some N_2 such that

$$|c_n - L| < \varepsilon \quad \text{for all} \quad n \ge N_2$$

or $L-\varepsilon < c_n < L+\varepsilon$ for all $n \ge N_2$. Then, if $n \ge \max\{N_1, N_2\}$, we have

$$L-\varepsilon < a_n \le b_n \le c_n < L+\varepsilon.$$

Thus $|b_n - L| < \varepsilon$ for $n \ge \max\{N_1, N_2\}$, as required. ∎

Returning to our example $(\sin n/n)_{n=1}^{\infty}$, we have $\lim_{n\to\infty} \frac{1}{n} = \lim_{n\to\infty} \frac{-1}{n} = 0$. By the Squeeze Theorem,

$$\lim_{n\to\infty} \frac{\sin n}{n} = 0.$$

2.4.7. EXAMPLE. For a more sophisticated example, consider the sequence $\left(n\sin\left(\frac{1}{n}\right)\right)_{n=1}^{\infty}$. To apply the Squeeze Theorem, we need to obtain an estimate for $\sin\theta$ when the angle θ is small. Consider a sector of the circle of radius 1 with angle θ and the two triangles as shown in Figure 2.2.

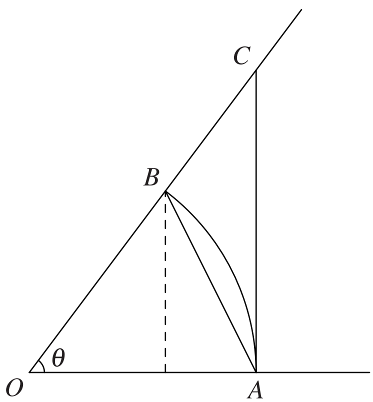

FIG. 2.2 Sector OAB between $\triangle OAB$ and $\triangle OAC$.

Since $\triangle OAB \subset$ sector $OAB \subset \triangle OAC$, we have the same relationship for their areas:

$$\frac{\sin\theta}{2} < \frac{\theta}{2} < \frac{\tan\theta}{2} = \frac{\sin\theta}{2\cos\theta}.$$

A manipulation of these inequalities yields

$$\cos\theta < \frac{\sin\theta}{\theta} < 1.$$

In particular, $\cos\frac{1}{n} < n\sin\frac{1}{n} < 1$. Moreover,

$$\cos\left(\tfrac{1}{n}\right) = \sqrt{1-\sin^2\left(\tfrac{1}{n}\right)} > \sqrt{1-\left(\tfrac{1}{n}\right)^2} > 1-\frac{1}{n^2}.$$

However,

$$\lim_{n\to\infty} 1-\frac{1}{n^2} = 1 = \lim_{n\to\infty} 1.$$

Therefore, by the Squeeze Theorem, $\lim\limits_{n\to\infty} n\sin\frac{1}{n} = 1$.

Exercises for Section 2.4

A. In each of the following, compute the limit. Then, using $\varepsilon = 10^{-6}$, find an integer N that satisfies the limit definition.

(a) $\lim\limits_{n\to\infty}\dfrac{\sin n^2}{\sqrt{n}}$ (b) $\lim\limits_{n\to\infty}\dfrac{1}{\log\log n}$ (c) $\lim\limits_{n\to\infty}\dfrac{3^n}{n!}$ (d) $\lim\limits_{n\to\infty}\dfrac{n^2+2n+1}{2n^2-n+2}$ (e) $\lim\limits_{n\to\infty}\sqrt{n^2+n}-n$

B. Show that $\lim\limits_{n\to\infty}\sin\frac{n\pi}{2}$ does not exist using the definition of limit.

C. Prove that if $a_n \le b_n$ for $n \ge 1$, $L = \lim\limits_{n\to\infty} a_n$, and $M = \lim\limits_{n\to\infty} b_n$, then $L \le M$.

D. Prove that if $L = \lim\limits_{n\to\infty} a_n$, then $L = \lim\limits_{n\to\infty} a_{2n}$ and $L = \lim\limits_{n\to\infty} a_{n^2}$.

E. Sometimes, a limit is defined informally as follows: "As n goes to infinity, a_n gets closer and closer to L." Find as many faults with this definition as you can.

(a) Can a sequence satisfy this definition and still fail to converge?
(b) Can a sequence converge yet fail to satisfy this definition?

F. Define a sequence $(a_n)_{n=1}^\infty$ such that $\lim\limits_{n\to\infty} a_{n^2}$ exists but $\lim\limits_{n\to\infty} a_n$ does not exist.

G. Suppose that $\lim\limits_{n\to\infty} a_n = L$ and $L \ne 0$. Prove there is some N such that $a_n \ne 0$ for all $n \ge N$.

H. Give a careful proof, using the definition of limit, that $\lim\limits_{n\to\infty} a_n = L$ and $\lim\limits_{n\to\infty} b_n = M$ imply that $\lim\limits_{n\to\infty} 2a_n + 3b_n = 2L+3M$.

I. For each $x \in \mathbb{R}$, determine whether $\left(\dfrac{1}{1+x^n}\right)_{n=1}^\infty$ has a limit, and compute it when it exists.

J. Let a_0 and a_1 be positive real numbers, and set $a_{n+2} = \sqrt{a_{n+1}} + \sqrt{a_n}$ for $n \ge 0$.

(a) Show that there is N such that for all $n \ge N$, $a_n \ge 1$.
(b) Let $\varepsilon_n = |a_n - 4|$. Show that $\varepsilon_{n+2} \le (\varepsilon_{n+1}+\varepsilon_n)/3$ for $n \ge N$.
(c) Prove that this sequence converges.

K. Show that the sequence $(\log n)_{n=1}^\infty$ does not converge.

2.5 Basic Properties of Limits

2.5.1. PROPOSITION. *If $(a_n)_{n=1}^\infty$ is a convergent sequence of real numbers, then the set $\{a_n : n \in \mathbb{N}\}$ is bounded.*

PROOF. Let $L = \lim_{n\to\infty} a_n$. If we set $\varepsilon = 1$, then by the definition of limit, there is some $N > 0$ such that $|a_n - L| < 1$ for all $n \geq N$. In other words,

$$L - 1 < a_n < L + 1 \quad \text{for all} \quad n \geq N.$$

Let $M = \max\{a_1, a_2, \ldots, a_{N-1}, L+1\}$ and $m = \min\{a_1, a_2, \ldots, a_{N-1}, L-1\}$. Clearly, for all n, we have $m \leq a_n \leq M$. ■

It is also crucial that limits respect the arithmetic operations. Proving this is straightforward. The details are left as exercises.

2.5.2. THEOREM. *If $\lim_{n\to\infty} a_n = L$, $\lim_{n\to\infty} b_n = M$, and $\alpha \in \mathbb{R}$, then*

(1) $\lim_{n\to\infty} a_n + b_n = L + M$,

(2) $\lim_{n\to\infty} \alpha a_n = \alpha L$,

(3) $\lim_{n\to\infty} a_n b_n = LM$, *and*

(4) $\displaystyle\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{L}{M}$ *if* $M \neq 0$.

In the sequence $(a_n/b_n)_{n=1}^\infty$, we ignore terms with $b_n = 0$. There is no problem doing this because $M \neq 0$ implies that $b_n \neq 0$ for all sufficiently large n (see Exercise 2.4.G). (We use "for all sufficiently large n" as shorthand for saying there is some N so that this holds for all $n \geq N$.)

Exercises for Section 2.5

A. Prove Theorem 2.5.2. HINT: For part (4), first bound the denominator away from 0.

B. Compute the following limits.

(a) $\displaystyle\lim_{n\to\infty} \frac{\tan\frac{\pi}{n}}{n\sin^2\frac{2}{n}}$ (b) $\displaystyle\lim_{n\to\infty} \frac{2^{100+5n}}{e^{4n-10}}$ (c) $\displaystyle\lim_{n\to\infty} \frac{\csc\frac{1}{n}}{n} + \frac{2\arctan n}{\log n}$

C. If $\lim_{n\to\infty} a_n = L > 0$, prove that $\lim_{n\to\infty} \sqrt{a_n} = \sqrt{L}$. Be sure to discuss the issue of when $\sqrt{a_n}$ makes sense. HINT: Express $|\sqrt{a_n} - \sqrt{L}|$ in terms of $|a_n - L|$.

D. Let $(a_n)_{n=1}^\infty$ and $(b_n)_{n=1}^\infty$ be two sequences of real numbers such that $|a_n - b_n| < \frac{1}{n}$. Suppose that $L = \lim_{n\to\infty} a_n$ exists. Show that $(b_n)_{n=1}^\infty$ converges to L also.

E. Find $\displaystyle\lim_{n\to\infty} \frac{\log(2+3^n)}{2n}$. HINT: $\log(2+3^n) = \log 3^n + \log\frac{2+3^n}{3^n}$

F. (a) Let $x_n = \sqrt[n]{n} - 1$. Use the fact that $(1+x_n)^n = n$ to show that $x_n^2 \leq 2/n$.
HINT: Use the Binomial Theorem and throw away most terms.

(b) Hence compute $\lim\limits_{n\to\infty} n^{1/n}$.

G. Show that the set of rational numbers is **dense** in $\mathbb{R}$, meaning that every real number is a limit of rational numbers.

H. (a) Show that $\frac{b-1}{b} \le \log b \le b-1$. HINT: Integrate $1/x$ from 1 to b.
(b) Apply this to $b = \sqrt[n]{a}$ to show that $\log a \le n(\sqrt[n]{a}-1) \le \sqrt[n]{a}\log a$.
(c) Hence evaluate $\lim\limits_{n\to\infty} n\big(\sqrt[n]{a}-1\big)$.

I. Suppose that $\lim\limits_{n\to\infty} a_n = L$. Show that $\lim\limits_{n\to\infty} \dfrac{a_1+a_2+\cdots+a_n}{n} = L$.

J. Show that the set $S = \{n+m\sqrt{2} : m,n \in \mathbb{Z}\}$ is dense in $\mathbb{R}$. HINT: Find infinitely many elements of S in $[0,1]$. Use the Pigeonhole Principle to find two that are close within 10^{-k}.

2.6 Monotone Sequences

We now consider some consequences of the Least Upper Bound Principle (2.3.3).

A sequence (a_n) is **(strictly) monotone increasing** if $a_n \le a_{n+1}$ ($a_n < a_{n+1}$) for all $n \ge 1$. Similarly, we define (strictly) monotone decreasing sequences.

2.6.1. Monotone Convergence Theorem.

A monotone increasing sequence that is bounded above converges.
A monotone decreasing sequence that is bounded below converges.

Proof. Suppose $(a_n)_{n=1}^\infty$ is an increasing sequence that is bounded above. Then by the Least Upper Bound Principle, there is a number $L = \sup\{a_n : n \in \mathbb{N}\}$. We will show that $\lim\limits_{n\to\infty} a_n = L$.

Let $\varepsilon > 0$ be given. Since $L-\varepsilon$ is not an upper bound for A, there is some integer N such that $a_N > L-\varepsilon$. Then because the sequence is monotone increasing,

$$L-\varepsilon < a_N \le a_n \le L \quad \text{for all} \quad n \ge N.$$

So $|a_n - L| < \varepsilon$ for all $n \ge N$ as required. Therefore, $\lim\limits_{n\to\infty} a_n = L$.

If (a_n) is decreasing and bounded below by B, then the sequence $(-a_n)$ is increasing and bounded above by $-B$. Thus the sequence $(-a_n)_{n=1}^\infty$ has a limit $L = \lim\limits_{n\to\infty} -a_n$. Therefore $-L = \lim\limits_{n\to\infty} a_n$ exists. ■

2.6.2. Example. Consider the sequence given recursively by

$$a_1 = 1 \quad \text{and} \quad a_{n+1} = \sqrt{2+\sqrt{a_n}} \quad \text{for all} \quad n \ge 1.$$

Evaluating $a_2, a_3, \dots, a_9$, we obtain 1.7320508076, 1.8210090645, 1.8301496356, 1.8310735189, 1.831166746, 1.8311761518, 1.8311771007, 1.8311771965. It appears that this sequence increases to some limit.

To prove this, first we show by induction that

$$1 \le a_n < a_{n+1} < 2 \quad \text{for all} \quad n \ge 1.$$

Since $1 = a_1 < \sqrt{3} = a_2 < 2$, this is valid for $n = 1$. Suppose that it holds for some n. Then

$$a_{n+2} = \sqrt{2 + \sqrt{a_{n+1}}} > \sqrt{2 + \sqrt{a_n}} = a_{n+1} \ge 1,$$

and

$$a_{n+2} = \sqrt{2 + \sqrt{a_{n+1}}} < \sqrt{2 + \sqrt{2}} < 2.$$

This verifies our claim for $n+1$. Hence by induction, it is valid for each $n \ge 1$.

Therefore, (a_n) is a monotone increasing sequence. So by the Monotone Convergence Theorem (2.6.1), it follows that there is a limit $L = \lim\limits_{n\to\infty} a_n$. It is not clear that there is a nice expression for L. However, once we know that the sequence converges, it is not hard to find a formula for L. Notice that

$$L = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} \sqrt{2 + \sqrt{a_n}} = \sqrt{2 + \sqrt{\lim_{n\to\infty} a_n}} = \sqrt{2 + \sqrt{L}}.$$

We used the fact that the limit of square roots is the square root of the limit (see Exercise 2.5.C). Squaring both sides gives $L^2 - 2 = \sqrt{L}$, and further squaring yields

$$0 = L^4 - 4L^2 - L + 4 = (L-1)(L^3 + L^2 - 3L - 4).$$

Since $L > 1$, it must be a root of the cubic $p(x) = x^3 + x^2 - 3x - 4$ in the interval $(1,2)$. There is only one such root. Indeed,

$$p'(x) = 3x^2 + 2x - 3 = 3(x^2 - 1) + 2x$$

is positive on $[1,2]$. So p is strictly increasing. Since $p(1) = -5$ and $p(2) = 2$, p has exactly one root in between. (See the Intermediate Value Theorem (5.6.1).)

For the amusement of the reader, we give an explicit algebraic formula:

$$L = \frac{1}{3}\left(\sqrt[3]{\tfrac{79+\sqrt{2241}}{2}} + \sqrt[3]{\tfrac{79-\sqrt{2241}}{2}} - 1\right).$$

Notice that we proved first that the sequence converged and then evaluated the limit afterward. This is important, for consider the sequence given by $a_1 = 2$ and $a_{n+1} = (a_n^2 + 1)/2$. This is a monotone increasing sequence. Suppose we let L denote the limit and compute

$$L = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} (a_n^2 + 1)/2 = (L^2 + 1)/2.$$

Thus $(L-1)^2 = 0$, which means that $L = 1$. This is an absurd conclusion because this sequence is monotone increasing and greater than 2. The fault lay in assuming that the limit L actually exists, because instead it diverges to $+\infty$ (see Exercise 2.6.A).

The following easy corollary of the Monotone Convergence Theorem is again a reflection of the completeness of the real numbers. This is just the tool needed to establish the key result of the next section, the Bolzano–Weierstrass Theorem (2.7.2).

Again, the corresponding result for intervals of rational numbers is false. See Example 2.7.6. The result would also be false if we changed closed intervals to open intervals. For example, $\bigcap_{n\geq 1}(0,\frac{1}{n}) = \varnothing$.

2.6.3. Nested Intervals Lemma.

Suppose that $I_n = [a_n, b_n] = \{x \in \mathbb{R} : a_n \leq x \leq b_n\}$ are nonempty closed intervals such that $I_{n+1} \subseteq I_n$ for each $n \geq 1$. Then the intersection $\bigcap_{n\geq 1} I_n$ is nonempty.

Proof. Notice that since I_{n+1} is contained in I_n, it follows that

$$a_n \leq a_{n+1} \leq b_{n+1} \leq b_n.$$

Thus (a_n) is a monotone increasing sequence bounded above by b_1; and likewise (b_n) is a monotone decreasing sequence bounded below by a_1. Hence by Theorem 2.6.1, $a = \lim\limits_{n\to\infty} a_n$ exists, as does $b = \lim\limits_{n\to\infty} b_n$. By Exercise 2.4.C, $a \leq b$. Thus

$$a_k \leq a \leq b \leq b_k.$$

Consequently, the point a belongs to I_k for each $k \geq 1$. ■

Exercises for Section 2.6

A. Say that $\lim\limits_{n\to\infty} a_n = +\infty$ if for every $R \in \mathbb{R}$, there is an integer N such that $a_n > R$ for all $n \geq N$. Show that a divergent monotone increasing sequence converges to $+\infty$ in this sense.

B. Let $a_1 = 0$ and $a_{n+1} = \sqrt{5 + 2a_n}$ for $n \geq 1$. Show that $\lim\limits_{n\to\infty} a_n$ exists and find the limit.

C. Is $S = \{x \in \mathbb{R} : 0 < \sin(\frac{1}{x}) < \frac{1}{2}\}$ bounded above (below)? If so, find $\sup S$ ($\inf S$).

D. Evaluate $\lim\limits_{n\to\infty} \sqrt[n]{3^n + 5^n}$.

E. Suppose (a_n) is a sequence of positive real numbers such that $a_{n+1} - 2a_n + a_{n-1} > 0$ for all $n \geq 1$. Prove that the sequence either converges or tends to $+\infty$.

F. Let a, b be positive real numbers. Set $x_0 = a$ and $x_{n+1} = (x_n^{-1} + b)^{-1}$ for $n \geq 0$.

(a) Prove that x_n is monotone decreasing.
(b) Prove that the limit exists and find it.

G. Let $a_n = (\sum_{k=1}^{n} 1/k) - \log n$ for $n \geq 1$. **Euler's constant** is defined as $\gamma = \lim\limits_{n\to\infty} a_n$. Show that $(a_n)_{n=1}^{\infty}$ is decreasing and bounded below by zero, and so this limit exists.
Hint: Prove that $1/(n+1) \leq \log(n+1) - \log n \leq 1/n$.

H. Let $x_n = \sqrt{1 + \sqrt{2 + \sqrt{3 + \cdots + \sqrt{n}}}}$.

(a) Show that $x_n < x_{n+1}$.
(b) Show that $x_{n+1}^2 \leq 1 + \sqrt{2}x_n$. Hint: Square x_{n+1} and factor a 2 out of the square root.
(c) Hence show that x_n is bounded above by 2. Deduce that $\lim\limits_{n\to\infty} x_n$ exists.

I. (a) Let $(a_n)_{n=1}^{\infty}$ be a bounded sequence and define a sequence $b_n = \sup\{a_k : k \geq n\}$ for $n \geq 1$. Prove that (b_n) converges. This is the **limit superior** of (a_n), denoted by $\limsup a_n$.
(b) Without redoing the proof, conclude that the **limit inferior** of a bounded sequence (a_n), defined as $\liminf a_n := \lim\limits_{n\to\infty} \bigl(\inf_{k\geq n} a_k\bigr)$, always exists.
(c) Extend the definitions of $\limsup a_n$ and $\liminf a_n$ to unbounded sequences. Provide an example with $\limsup a_n = +\infty$ and $\liminf a_n = -\infty$.

J. Show that $(a_n)_{n=1}^{\infty}$ converges to $L \in \mathbb{R}$ if and only if $\limsup a_n = \liminf a_n = L$.

K. If a sequence (a_n) is not bounded above, show that $\sup\{a_n : n \geq k\} = +\infty$ for all k. What should $\limsup a_n$ be? Formulate and prove a similar statement if (a_n) is not bounded below.

L. Suppose $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are sequences of nonnegative real numbers and $\lim\limits_{n\to\infty} a_n \in \mathbb{R}$ exists. Prove that $\limsup a_n b_n = \lim\limits_{n\to\infty} a_n \bigl(\limsup b_n\bigr)$.

M. Suppose that $(a_n)_{n=1}^{\infty}$ has $a_n > 0$ for all n. Show that $\limsup a_n^{-1} = \bigl(\liminf a_n\bigr)^{-1}$.

N. Suppose $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are sequences of positive real numbers and $\limsup a_n/b_n < \infty$. Prove that there is a constant M such that $a_n \leq M b_n$ for all $n \geq 1$.

2.7 Subsequences

Given one sequence, we can build a new sequence, called a subsequence of the original, by picking out some of the entries. Perhaps surprisingly, when the original sequence does not converge, it is often possible to find a subsequence that does.

2.7.1. Definition. A **subsequence** of a sequence $(a_n)_{n=1}^{\infty}$ is a sequence $\bigl(a_{n_k}\bigr)_{k=1}^{\infty} = (a_{n_1}, a_{n_2}, a_{n_3}, \dots)$, where $n_1 < n_2 < n_3 < \cdots$.

For example, $(a_{2k})_{k=1}^{\infty}$ and $(a_{k^3})_{k=1}^{\infty}$ are subsequences, where $n_k = 2k$ and $n_k = k^3$, respectively. Notice that if we pick $n_k = k$ for each k, then we get the original sequence; so $(a_n)_{n=1}^{\infty}$ is a subsequence of itself.

It is easy to verify that if $(a_n)_{n=1}^{\infty}$ converges to a limit L, then $\bigl(a_{n_k}\bigr)_{k=1}^{\infty}$ also converges to the same limit. On the other hand, the sequence $(1, 2, 3, \dots)$ does not have a limit, nor does any subsequence, because any subsequence must diverge to $+\infty$. However, we will show that as long as a sequence remains bounded, it has subsequences that converge.

2.7.2. Bolzano–Weierstrass Theorem.
Every bounded sequence of real numbers has a convergent subsequence.

Proof. Let (a_n) be a sequence bounded by B. Thus the interval $[-B, B]$ contains the whole (infinite) sequence. Now if I is an interval containing infinitely many points of the sequence (a_n), and $I = J_1 \cup J_2$ is the union of two smaller intervals, then at least one of them contains infinitely many points of the sequence, too.

So let $I_1 = [-B, B]$. Split it into two closed intervals of length B, namely $[-B, 0]$ and $[0, B]$. One of these halves contains infinitely many points of (a_n); call it I_2. Similarly, divide I_2 into two closed intervals of length $B/2$. Again pick one, called

I_3, that contains infinitely many points of our sequence. Recursively, we construct a decreasing sequence I_k of closed intervals of length $2^{2-k}B$ such that each contains infinitely many points of our sequence. Figure 2.3 shows the choice of I_3 and I_4, where the terms of the sequence are indicated by vertical lines.

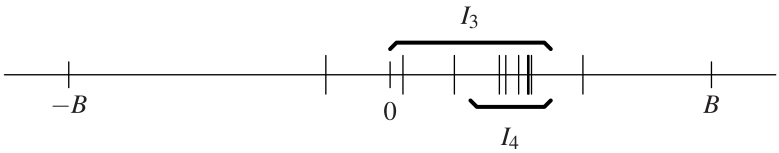

FIG. 2.3 Choice of intervals I_3 and I_4.

By the Nested Interval Lemma (2.6.3), we know that $\bigcap_{k\geq 1} I_k$ contains a number L. Choose an increasing sequence n_k such that a_{n_k} belongs to I_k. This is possible since each I_k contains infinitely many numbers in the sequence, and only finitely many have index less than n_{k-1}. We claim that $\lim\limits_{k\to\infty} a_{n_k} = L$. Indeed, both a_{n_k} and L belong to I_k, and hence

$$|a_{n_k} - L| \leq |I_k| = 2^{-k}(4B).$$

The right-hand side tends to 0, and thus $\lim\limits_{k\to\infty} a_{n_k} = L$. ■

2.7.3. EXAMPLE. Consider the sequence $(a_n) = (\operatorname{sign}(\sin n))_{n=1}^{\infty}$, where the sign function takes values ± 1 depending on the sign of x except for $\operatorname{sign} 0 = 0$. Without knowing anything about the properties of the sine function, we can observe that the sequence (a_n) takes at most three different values. At least one of these values is taken infinitely often. Thus it is possible to deduce the existence of a subsequence that is constant and therefore converges.

Using our knowledge of sine allows us to get somewhat more specific. Now $\sin x = 0$ exactly when x is an integer multiple of π. Since π is irrational, $k\pi$ is never an integer for $k > 0$. Therefore, a_n takes only the values ± 1. Note that $\sin x > 0$ if there is an integer k such that $2k\pi < x < (2k+1)\pi$; and $\sin x < 0$ if there is an integer k such that $(2k-1)\pi < x < 2k\pi$. Observe that n increases by steps of length 1, while the intervals on which $\sin x$ takes positive or negative values has length $\pi \approx 3.14$. Consequently, a_n takes the value $+1$ for three or four terms in a row, followed by three or four terms taking the value -1. Consequently, both 1 and -1 are limits of certain subsequences of (a_n).

2.7.4. EXAMPLE. Consider the sequence $(a_n) = (\sin n)_{n=1}^{\infty}$. As the angles n radians for $n \geq 1$ are marked on a circle, they appear gradually to fill in a dense subset. If this can be demonstrated, we should be able to show that $\sin\theta$ is a limit of a subsequence of our sequence for every θ in $[0, 2\pi]$.

The key is to approximate the angle 0 modulo 2π by integers. Let m be a positive integer and let $\varepsilon > 0$. Choose an integer N so large that $N\varepsilon > 2\pi$. Divide the circle into N arcs of length $2\pi/N$ radians each. Then consider the $N+1$ points $0, m, 2m, \dots, Nm$ modulo 2π on the circle. Since there are $N+1$ points distributed into only N arcs, the Pigeonhole Principle implies that at least one arc contains two points, say im and jm, where $i < j$. Then $n = jm - im$ represents an angle of absolute value at most $2\pi/N < \varepsilon$ radians up to a multiple of 2π. That is, $n = \psi + 2\pi s$ for some integer s and real number $|\psi| < \varepsilon$. In particular, $|\sin n| < \varepsilon$ and $n \geq m$. Moreover, since π is not rational, n is not an exact multiple of 2π.

So given $\theta \in [0, 2\pi]$, construct a subsequence as follows. Let $n_1 = 1$. Recursively we construct an increasing sequence n_k such that

$$|\sin n_k - \sin\theta| < \frac{1}{k}.$$

Once n_k is defined, take $\varepsilon = \frac{1}{k+1}$ and $m = n_k + 1$. As in the previous paragraph, there is an integer $n > n_k$ such that $n = \psi + 2\pi s$ and $|\psi| < \frac{1}{k+1}$. Thus there is a positive integer t such that $|\theta - t\psi| < \frac{1}{k+1}$. Therefore

$$|\sin(tn) - \sin(\theta)| = |\sin(t\psi) - \sin(\theta)| \leq |t\psi - \theta| < \frac{1}{k+1}. \tag{2.7.5}$$

Set $n_{k+1} = tn$. This completes the induction. The result is a subsequence such that

$$\lim_{k\to\infty} \sin(n_k) = \sin\theta.$$

To verify equation (2.7.5), recall the Mean Value Theorem (6.2.2). There is a point ξ between $t\psi$ and θ such that

$$\left|\frac{\sin(t\psi) - \sin(\theta)}{t\psi - \theta}\right| = |\cos\xi| \leq 1.$$

Rearranging yields $|\sin(t\psi) - \sin(\theta)| \leq |t\psi - \theta|$.

Therefore, we have shown that every value in the interval $[-1, 1]$ is the limit of some subsequence of the sequence $(\sin n)_{n=1}^{\infty}$.

2.7.6. Example. Consider the sequence $b_1 = 3$ and $b_{n+1} = (b_n + 8/b_n)/2$. Notice that

$$\begin{aligned} b_{n+1}^2 - 8 &= \frac{b_n^2 + 16 + (64/b_n^2) - 32}{4} = \frac{b_n^2 - 16 + (64/b_n^2)}{4} \\ &= \frac{(b_n - 8/b_n)^2}{4} = \frac{(b_n^2 - 8)^2}{4b_n^2}. \end{aligned}$$

It follows that $b_n^2 > 8$ for all $n \geq 2$, and $b_1^2 - 8 = 1 > 0$ also. Thus

$$0 < b_{n+1}^2 - 8 < \frac{(b_n^2 - 8)^2}{32}.$$

Iterating this, we obtain $b_2^2 - 8 < 32^{-1}$, $b_3^2 - 8 < 32^{-3}$, and $b_4^2 - 8 < 32^{-7}$. In general, we establish by induction that

$$0 < b_n^2 - 8 < 32^{1-2^{n-1}}.$$

Since b_n is positive and $b^2 - 8 = (b - \sqrt{8})(b + \sqrt{8})$, it follows that

$$0 < b_n - \sqrt{8} = \frac{b_n^2 - 8}{b_n + \sqrt{8}} < \frac{32^{1-2^{n-1}}}{2\sqrt{8}} < 6(32^{-2^{n-1}}).$$

Lastly, using the fact that $32^2 = 1024 > 10^3$, we obtain

$$0 < b_n - \sqrt{8} < 10 \cdot 10^{-3 \cdot 2^{n-2}}.$$

In particular, $\lim\limits_{n\to\infty} b_n = \sqrt{8}$. In fact, the convergence is so rapid that b_{10} approximates $\sqrt{8}$ to more than 750 digits of accuracy. See Example 11.2.2 for a more general analysis in terms of Newton's method.

Let $a_n = 8/b_n$. Then a_n is monotone increasing to $\sqrt{8}$. Both a_n and b_n are rational, but $\sqrt{8}$ is irrational. Thus the sets $J_n = \{x \in \mathbb{Q} : a_n \le x \le b_n\}$ form a decreasing sequence of nonempty intervals of *rational* numbers with empty intersection.

Exercises for Section 2.7

A. Show that $(a_n) = \left(\frac{n\cos^n(n)}{\sqrt{n^2+2n}}\right)_{n=1}^{\infty}$ has a convergent subsequence.

B. Does the sequence $(b_n) = \left(n + \cos(n\pi)\sqrt{n^2+1}\right)_{n=1}^{\infty}$ have a convergent subsequence?

C. Does the sequence $(a_n) = (\cos\log n)_{n=1}^{\infty}$ converge?

D. Show that every sequence has a monotone subsequence.

E. Use trig identities to show that $|\sin x - \sin y| \le |x - y|$.
HINT: Let $a = (x+y)/2$ and $b = (x-y)/2$. Use the addition formula for $\sin(a \pm b)$.

F. Define $x_1 = 2$ and $x_{n+1} = \frac{1}{2}(x_n + 5/x_n)$ for $n \ge 1$.

(a) Find a formula for $x_{n+1}^2 - 5$ in terms of $x_n^2 - 5$.
(b) Hence evaluate $\lim\limits_{n\to\infty} x_n$.
(c) Compute the first ten terms on a computer or a calculator.
(d) Show that the tenth term approximates the limit to over 600 decimal places.

G. Let $(x_n)_{n=1}^{\infty}$ be a sequence of real numbers. Suppose that there is a real number L such that $L = \lim\limits_{n\to\infty} x_{3n-1} = \lim\limits_{n\to\infty} x_{3n+1} = \lim\limits_{n\to\infty} x_{3n}$. Show that $\lim\limits_{n\to\infty} x_n$ exists and equals L.

H. Let $(x_n)_{n=1}^{\infty}$ be a sequence in $\mathbb{R}$. Suppose there is a number L such that every subsequence $\left(x_{n_k}\right)_{k=1}^{\infty}$ has a subsubsequence $\left(x_{n_{k(l)}}\right)_{l=1}^{\infty}$ with $\lim\limits_{l\to\infty} x_{n_{k(l)}} = L$. Show that the whole sequence converges to L. HINT: If not, you could find a subsequence bounded away from L.

I. Suppose $(x_n)_{n=1}^{\infty}$ is a sequence in $\mathbb{R}$, and that L_k are real numbers with $\lim\limits_{k\to\infty} L_k = L$. If for each $k \ge 1$, there is a subsequence of $(x_n)_{n=1}^{\infty}$ converging to L_k, show that some subsequence converges to L. HINT: Find an increasing sequence n_k such that $|x_{n_k} - L| < 1/k$.

J. (a) Suppose that $(x_n)_{n=1}^{\infty}$ is a sequence of real numbers. If $L = \liminf x_n$, show that there is a subsequence $\big(x_{n_k}\big)_{k=1}^{\infty}$ such that $\lim\limits_{k\to\infty} x_{n_k} = L$.

(b) Similarly, prove that there is a subsequence $\big(x_{n_l}\big)_{l=1}^{\infty}$ such that $\lim\limits_{l\to\infty} x_{n_l} = \limsup x_n$.

K. Let $(x_n)_{n=1}^{\infty}$ be an arbitrary sequence. Prove that there is a subsequence $\big(x_{n_k}\big)_{k=1}^{\infty}$ which converges or $\lim\limits_{k\to\infty} x_{n_k} = \infty$ or $\lim\limits_{k\to\infty} x_{n_k} = -\infty$.

L. Construct a sequence $(x_n)_{n=1}^{\infty}$ such that for every real number L, there is a subsequence $\big(x_{n_k}\big)_{k=1}^{\infty}$ with $\lim\limits_{k\to\infty} x_{n_k} = L$.

2.8 Cauchy Sequences

Can we decide whether a sequence converges *without* first finding the value of the limit? To do this, we need an intrinsic property of a sequence which is equivalent to convergence that does not make use of the value of the limit. This intrinsic property shows which sequences are ‘supposed’ to converge. This leads us to the notion of a subset of $\mathbb{R}$ being *complete* if all sequences in the subset that are ‘supposed’ to converge actually do. As we shall see, this completeness property has been built into the real numbers by our construction of infinite decimals.

To obtain an appropriate condition, notice that if a sequence (a_n) converges to L, then as the terms get close to the limit, they are getting close to each other.

2.8.1. Proposition. *Let $(a_n)_{n=1}^{\infty}$ be a sequence converging to L. For every $\varepsilon > 0$, there is an integer N such that*

$$|a_n - a_m| < \varepsilon \quad \textit{for all} \quad m,n \geq N.$$

Proof. Fix $\varepsilon > 0$ and use the value $\varepsilon/2$ in the definition of limit. Then there is an integer N such that $|a_n - L| < \varepsilon/2$ for all $n \geq N$. Thus if $m,n \geq N$, we obtain

$$|a_n - a_m| \leq |a_n - L| + |L - a_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

∎

In order for N to work in the conclusion, for every $m \geq N$, a_m must be within ε of a_N. It is not enough to just have a_N and a_{N+1} close (see Exercise 2.8.B).

We make the conclusion of this proposition into a definition. This definition retains the flavour of the definition of a limit, in that it has the same logical structure: *For all $\varepsilon > 0$, there is an integer N ....*

2.8.2. Definition. A sequence $(a_n)_{n=1}^{\infty}$ of real numbers is called a **Cauchy sequence** provided that for *every* $\varepsilon > 0$, there is an integer N such that

$$|a_m - a_n| < \varepsilon \quad \text{for all} \quad m,n \geq N.$$

2.8.3. PROPOSITION. *Every Cauchy sequence is bounded.*

PROOF. The proof is basically the same as Proposition 2.5.1. Let $(a_n)_{n=1}^\infty$ be a Cauchy sequence. Taking $\varepsilon = 1$, find N so large that

$$|a_n - a_N| < 1 \quad \text{for all} \quad n \geq N.$$

It follows that the sequence is bounded by $\max\{|a_1|, \dots, |a_{N-1}|, |a_N| + 1\}$. ■

Since the definition of a Cauchy sequence does not require the use of a potential limit L, it permits the following definition.

2.8.4. DEFINITION. A subset S of $\mathbb{R}$ is said to be **complete** if every Cauchy sequence (a_n) in S (that is, $a_n \in S$) converges to a point in S.

This brings us to an important conclusion about the real numbers themselves, another property that distinguishes the real numbers from the rational numbers.

2.8.5. COMPLETENESS THEOREM.
Every Cauchy sequence of real numbers converges. So $\mathbb{R}$ is complete.

PROOF. Suppose that $(a_n)_{n=1}^\infty$ is a Cauchy sequence. By Proposition 2.8.3, $\{a_n : n \geq 1\}$ is bounded. By the Bolzano–Weierstrass Theorem (2.7.2), this sequence has a convergent subsequence, say

$$\lim_{k\to\infty} a_{n_k} = L.$$

Let $\varepsilon > 0$. From the definition of Cauchy sequence for $\varepsilon/2$, there is an integer N such that

$$|a_m - a_n| < \frac{\varepsilon}{2} \quad \text{for all} \quad m, n \geq N.$$

And from the definition of limit using $\varepsilon/2$, there is an integer K such that

$$|a_{n_k} - L| < \frac{\varepsilon}{2} \quad \text{for all} \quad k \geq K.$$

Pick any $k \geq K$ such that $n_k \geq N$. Then for every $n \geq N$,

$$|a_n - L| \leq |a_n - a_{n_k}| + |a_{n_k} - L| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

So $\lim\limits_{n\to\infty} a_n = L$. ■

2.8.6. REMARK. This theorem is not true for the rational numbers. Define the sequence $(a_n)_{n=1}^\infty$ by

$$a_1 = 1.4, \quad a_2 = 1.41, \quad a_3 = 1.414, \quad a_4 = 1.4142, \quad a_5 = 1.41421, \dots$$

and in general, a_n is the first $n+1$ digits in the decimal expansion of $\sqrt{2}$. If n and m are greater than N, then a_n and a_m agree for at least first $N+1$ digits. Thus

$$|a_n - a_m| < 10^{-N} \quad \text{for all} \quad m,n \geq N.$$

This shows that $(a_n)_{n=1}^{\infty}$ is a Cauchy sequence of rational numbers. (Why?)

However, this sequence has no limit *in the rationals*. In our terminology, $\mathbb{Q}$ is not complete. Of course, this sequence does converge to a real number, namely $\sqrt{2}$. This is one way to see the essential difference between $\mathbb{R}$ and $\mathbb{Q}$: the set of real numbers is complete and $\mathbb{Q}$ is not.

2.8.7. Example. Let α be an arbitrary real number. Define $a_n = [n\alpha]/n$, where $[x]$ is the nearest integer to x. Then $\big|[n\alpha] - n\alpha\big| \leq 1/2$. So

$$|a_n - \alpha| = \frac{\big|[n\alpha] - n\alpha\big|}{n} \leq \frac{1}{2n}.$$

We claim $\lim\limits_{n\to\infty} a_n = \alpha$. Indeed, given $\varepsilon > 0$, choose N so large that $\frac{1}{N} < \varepsilon$. Then for $n \geq N$, $|a_n - \alpha| < \varepsilon/2$. Moreover, if $m,n \geq N$,

$$|a_n - a_m| \leq |a_n - \alpha| + |\alpha - a_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus this sequence is Cauchy.

2.8.8. Example. Consider the infinite **continued fraction**

$$\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{2+\cdots}}}}$$

To make sense of this, it has to be interpreted as the limit of the finite fractions

$$a_1 = \frac{1}{2}, \qquad a_2 = \frac{1}{2+\frac{1}{2}} = \frac{2}{5}, \qquad a_3 = \frac{1}{2+\frac{1}{2+\frac{1}{2}}} = \frac{5}{12}, \qquad \cdots.$$

We need a better way of defining the general term. In this case, there is a recursive formula for obtaining one term from the preceding one:

$$a_1 = \frac{1}{2}, \qquad a_{n+1} = \frac{1}{2+a_n} \quad \text{for} \quad n \geq 1.$$

In order to establish convergence, we will show that (a_n) is Cauchy. Consider

$$a_{n+1} - a_{n+2} = \frac{1}{2+a_n} - \frac{1}{2+a_{n+1}} = \frac{a_{n+1} - a_n}{(2+a_n)(2+a_{n+1})}.$$

Now $a_1 > 0$, and it is readily follows that $a_n > 0$ for all $n \geq 2$ by induction. Hence the denominator $(2+a_n)(2+a_{n+1})$ is greater than 4. So we obtain

$$|a_{n+1} - a_{n+2}| < \frac{|a_n - a_{n+1}|}{4} \quad \text{for all} \quad n \geq 1.$$

Since $|a_1 - a_2| = 1/10$, we may iterate this inequality to estimate

$$|a_2 - a_3| < \frac{1}{10 \cdot 4}, \quad |a_3 - a_4| < \frac{1}{10 \cdot 4^2}, \quad |a_n - a_{n+1}| < \frac{1}{10 \cdot 4^{n-1}} = \tfrac{2}{5}(4^{-n}).$$

The general formula estimating the difference may be verified by induction.

Now it is straightforward to estimate the difference between arbitrary terms a_m and a_n for $m < n$:

$$\begin{aligned} |a_m - a_n| &= \big|(a_m - a_{m+1}) + (a_{m+1} - a_{m+2}) + \cdots + (a_{n-1} - a_n)\big| \\ &\leq |a_m - a_{m+1}| + |a_{m+1} - a_{m+2}| + \cdots + |a_{n-1} - a_n| \\ &< \tfrac{2}{5}(4^{-m} + 4^{-m-1} + \cdots + 4^{1-n}) < \frac{2 \cdot 4^{-m}}{5(1 - \frac{1}{4})} = \frac{8}{15} 4^{-m} < 4^{-m}. \end{aligned}$$

This tells us that our sequence is Cauchy. Indeed, if $\varepsilon > 0$, choose N such that $4^{-N} < \varepsilon$. Then

$$|a_m - a_n| < 4^{-m} \leq 4^{-N} < \varepsilon \quad \text{for all} \quad m, n \geq N.$$

Therefore by the Completeness Theorem 2.8.5, it follows that $(a_n)_{n=1}^{\infty}$ converges; say, $\lim\limits_{n \to \infty} a_n = L$. To calculate L, use the recurrence relation

$$L = \lim_{n \to \infty} a_n = \lim_{n \to \infty} a_{n+1} = \lim_{n \to \infty} \frac{1}{2+a_n} = \frac{1}{2+L}.$$

It follows that $L^2 + 2L - 1 = 0$. Solving yields $L = \pm\sqrt{2} - 1$. Since $L > 0$, we see that $L = \sqrt{2} - 1$.

We have accumulated five different results for $\mathbb{R}$ that distinguish it from $\mathbb{Q}$.

(1) the Least Upper Bound Principle (2.3.3),
(2) the Monotone Convergence Theorem (2.6.1),
(3) the Nested Intervals Lemma (2.6.3),
(4) the Bolzano–Weierstrass Theorem (2.7.2),
(5) the Completeness Theorem (2.8.5).

It turns out that they are all equivalent. Indeed, each of the proofs of items (2) to (5) relies only on the previous item in our list. To show how the Completeness Theorem implies the Least Upper Bound Principle, go through our proof to obtain an increasing sequence of lower bounds, y_k, and a decreasing sequence of elements $x_k \in S$ with $x_k < y_k + 10^{-k}$. Show that the sequence $x_1, y_1, x_2, y_2, \dots$ is Cauchy. The limit L will be the greatest lower bound. Fill in the details yourself (Exercise 2.8.G).

Exercises for Section 2.8

A. Let (x_n) be Cauchy with a subsequence (x_{n_k}) such that $\lim_{k\to\infty} x_{n_k} = a$. Show that $\lim_{n\to\infty} x_n = a$.

B. Give a sequence (a_n) such that $\lim_{n\to\infty} |a_n - a_{n+1}| = 0$, but the sequence does not converge.

C. Let (a_n) be a sequence such that $\lim_{N\to\infty} \sum_{n=1}^{N} |a_n - a_{n+1}| < \infty$. Show that (a_n) is Cauchy.

D. If $(x_n)_{n=1}^{\infty}$ is Cauchy, show that it has a subsequence (x_{n_k}) such that $\sum_{k=1}^{\infty} |x_{n_k} - x_{n_{k+1}}| < \infty$.

E. Suppose that (a_n) is a sequence such that $a_{2n} \le a_{2n+2} \le a_{2n+3} \le a_{2n+1}$ for all $n \ge 0$. Show that this sequence is Cauchy if and only if $\lim_{n\to\infty} |a_n - a_{n+1}| = 0$.

F. Give an example of a sequence (a_n) such that $a_{2n} \le a_{2n+2} \le a_{2n+3} \le a_{2n+1}$ for all $n \ge 0$ which does not converge.

G. Fill in the details of how the Completeness Theorem implies the Least Upper Bound Principle.

H. Let $a_0 = 0$ and set $a_{n+1} = \cos(a_n)$ for $n \ge 0$. Try this on your calculator (use radian mode!).

(a) Show that $a_{2n} \le a_{2n+2} \le a_{2n+3} \le a_{2n+1}$ for all $n \ge 0$.
(b) Use the Mean Value Theorem to find an explicit number $r < 1$ such that $|a_{n+2} - a_{n+1}| \le r|a_n - a_{n+1}|$ for all $n \ge 0$. Hence show that this sequence is Cauchy.
(c) Describe the limit geometrically as the intersection point of two curves.

I. Evaluate the continued fraction

$$1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cdots}}} \;.$$

J. Let $x_0 = 0$ and $x_{n+1} = \sqrt{5 - 2x_n}$ for $n \ge 0$. Show that this sequence converges and compute the limit. HINT: Show that the even terms increase and the odd terms decrease.

K. Consider an infinite binary expansion $(0.e_1e_2e_3\dots)_{\text{base }2}$, where each $e_i \in \{0,1\}$. Show that $a_n = \sum_{i=1}^{n} 2^{-i} e_i$ is Cauchy for every choice of zeros and ones.

L. One base-independent construction of the real numbers uses Cauchy sequences of rational numbers. This exercise asks for the definitions that go into such a proof.

(a) Find a way to decide when two Cauchy sequences should determine the same real number without using their limits. HINT: Combine the two sequences into one.
(b) Your definition in (a) should be an equivalence relation. Is it? (See Appendix 1.3.)
(c) How are addition and multiplication defined?
(d) How is the order defined?

2.9 Countable Sets

Cardinality measures the size of a set in the crudest of ways—by counting the numbers of elements. Obviously, the number of elements in a set could be 0, 1, 2, 3, 4, or some other finite number. Or a set can have infinitely many elements. Perhaps surprisingly, not all infinite sets have the same cardinality. We distinguish only between sets having the smallest infinite cardinality (countably infinite sets) and all larger cardinalities (uncountable sets). We use the term countable for sets that are either countably infinite or finite.

2.9.1. DEFINITION. Two sets A and B have the same **cardinality** if there is a *bijection* f from A onto B. Write $|A| = |B|$ in this case. We say that the cardinality of A is at most that of B (write $|A| \leq |B|$) if there is an *injection* f from A into B.

The definition says simply that if all of the elements of A can be paired, one-to-one, with all of the elements of B, then A and B have the same size. If A fits inside B in a one-to-one manner, then A is smaller than or equal to B. It is natural to ask whether $|A| \leq |B|$ and $|B| \leq |A|$ imply $|A| = |B|$. The answer is yes, but this is not obvious for infinite sets. The Schroeder–Bernstein Theorem establishes this, but we do not include a proof.

2.9.2. EXAMPLES.

(1) The cardinality of any finite set is the number of elements, and this number belongs to $\{0,1,2,3,4,\dots\}$. This property is, essentially, the definition of finite set.

(2) Many sets encountered in analysis are infinite, meaning that they are not finite. The sets of natural numbers $\mathbb{N}$, integers $\mathbb{Z}$, rational numbers $\mathbb{Q}$, and real numbers $\mathbb{R}$ are all infinite. Moreover, we have the containments $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$. Therefore $|\mathbb{N}| \leq |\mathbb{Z}| \leq |\mathbb{Q}| \leq |\mathbb{R}|$. Notice that the integers can be written as a list $0,1,-1,2,-2,3,-3,\dots$. This amounts to defining a bijection $f:\mathbb{N}\to\mathbb{Z}$ by $f(2n-1) = 1-n$ and $f(2n) = n$ for $n \geq 1$. Therefore, $|\mathbb{N}| = |\mathbb{Z}|$.

2.9.3. DEFINITION. A set A is a **countable set** is it is finite or if $|A| = |\mathbb{N}|$. If $|A| = |\mathbb{N}|$, we say that A is **countably infinite**. The cardinal $|\mathbb{N}|$ is also denoted by $\aleph_0$, pronounced **aleph nought**. Aleph is the first letter of the Hebrew alphabet.

An infinite set that is not countable is called an **uncountable set**.

Equivalently, A is countable if the elements of A may be listed as $a_1, a_2, a_3, \dots$. Indeed, the list itself determines a bijection f from $\mathbb{N}$ to A by $f(k) = a_k$. It is a basic fact that countable sets are the smallest infinite sets.

Notice that two uncountable sets might have different cardinalities.

2.9.4. LEMMA. *Every infinite subset of* $\mathbb{N}$ *is countable. Moreover, if* A *is an infinite set such that* $|A| \leq |\mathbb{N}|$*, then* $|A| = |\mathbb{N}|$.

PROOF. Any nonempty subset X of $\mathbb{N}$ has a smallest element. This follows from induction: if X does not have a smallest element, then $1 \notin X$ and $1,\dots,n$ all not in X imply $n+1 \notin X$. By induction, X is empty, a contradiction.

Let B be an infinite subset of $\mathbb{N}$. List the elements of B in increasing order as $b_1 < b_2 < b_3 < \cdots$. This is done by choosing the smallest element b_1, then the smallest of the remaining set $B \setminus \{b_1\}$, then the smallest of $B \setminus \{b_1, b_2\}$, and so on. The result is an infinite list of elements of B in increasing order. It must include every element $b \in B$ because $\{n \in B : n \leq b\}$ is finite, containing say k elements. Then $b_k = b$. As noted before the proof, this implies that $|B| = |\mathbb{N}|$.

Now consider an infinite set A with $|A| \leq |\mathbb{N}|$. By definition, there is an injection f of A into $\mathbb{N}$. Let $B = f(A)$. Note that f is a bijection of A onto B. Thus B is an infinite subset of $\mathbb{N}$. So $|A| = |B| = |\mathbb{N}|$. ■

2.9.5. Proposition. *The set $\mathbb{N} \times \mathbb{N}$ is countable.*

Proof. Rather than starting with the formula of a bijection from $\mathbb{N}$ to $\mathbb{N} \times \mathbb{N}$, note that each 'diagonal set' $D_n = \{(i,j) \in \mathbb{N} \times \mathbb{N} : i + j = n+1\}$, $n \geq 1$, is finite. Thus, if we work through these sets in some methodical way, any pair (i,j) will be reached in finitely many steps. See Figure 2.4.

Noting that $|D_n| = n$ and $1 + 2 + \ldots + n = n(n+1)/2$, we define our bijection for $m \in \mathbb{N}$ by first picking n such that $n(n-1)/2 < m \leq n(n+1)/2$. Letting $k = m - n(n-1)/2$, we define $\varphi(m)$ to be $(k, n+1-k)$. It is routine, if tedious, to verify that φ is a bijection, i.e., one-to-one and onto. ■

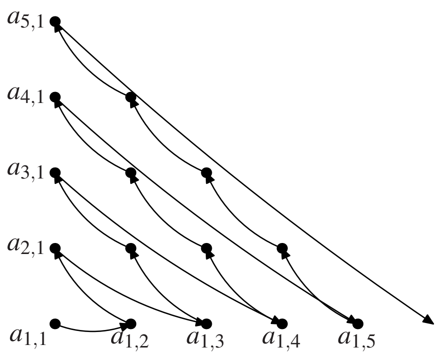

Fig. 2.4 The ordering on $\mathbb{N} \times \mathbb{N}$.

2.9.6. Corollary. *The countable union of countable sets is countable.*

Proof. Let $A_1, A_2, A_3, \ldots$ be countable sets. To avoid repetition, let $B_1 = A_1$ and $B_i = A_i \setminus \cup_{k=1}^{i-1} A_k$. Each B_i is countable, so list its elements as $b_{i,1}, b_{i,2}, b_{i,3}, \ldots$. Map $A = \cup_{i \geq 1} A_i = \cup_{i \geq 1} B_i$ into $\mathbb{N} \times \mathbb{N}$ by $f(b_{ij}) = (i,j)$. This is an injection; therefore $|A| \leq |\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$. Hence the union is countable. ■

2.9.7. Corollary. *The set $\mathbb{Q}$ of rational numbers is countable.*

PROOF. Observe that $\mathbb{Z}\times\mathbb{N}$ is countable, since we can take the bijection $f:\mathbb{N}\to\mathbb{Z}$ of Example 2.9.2 (2) and use it to define $g:\mathbb{N}\times\mathbb{N}\to\mathbb{Z}\times\mathbb{N}$ by $g(n,m)=(f(n),m)$, which you can check is a bijection.

Define a map from $\mathbb{Q}$ into $\mathbb{Z}\times\mathbb{N}$ by $h(r)=(a,b)$ if $r=a/b$, where a and b are integers with no common factor and $b>0$. These conditions uniquely determine the pair (a,b) for each rational r, and so h is a function. Clearly, h is injective since r is recovered from (a,b) by division. Therefore, h is an injection of $\mathbb{Q}$ into a countable set. Hence $\mathbb{Q}$ is an infinite set with $|\mathbb{Q}|\leq|\mathbb{N}|$. So $\mathbb{Q}$ is countable by Lemma 2.9.4. ■

There are infinite sets that are not countable. The proof uses a **diagonalization** argument due to Cantor.

2.9.8. THEOREM. *The set $\mathbb{R}$ of real numbers is uncountable.*

PROOF. Suppose to the contrary that $\mathbb{R}$ is countable. Then all real numbers may be written as a list $x_1,x_2,x_3,\ldots$. Express each x_i as an infinite decimal, which we write as $x_i=x_{i0}.x_{i1}x_{i2}x_{i3}\ldots$, where x_{i0} is an integer and x_{ik} is an integer from 0 to 9 for each $k\geq 1$. Our goal is to write down another real number that does not appear in this (supposedly exhaustive) list. Let $a_0=0$ and define $a_k=7$ if $x_{kk}\in\{0,1,2,3,4\}$ and $a_k=2$ if $x_{kk}\in\{5,6,7,8,9\}$. Define a real number $a=a_0.a_1a_2a_3\ldots$.

Since a is a real number, it must appear somewhere in this list, say $a=x_k$. However, the kth decimal place a_k of a and x_{kk} of x_k differ by between 3 and 7. This cannot be accounted for by the fact that certain real numbers have two decimal expansions, one ending in zeros and the other ending in nines because this changes any digit by either 1 or 9. So $a\neq x_k$, and hence a does not occur in this list. It follows that there is no such list, and thus $\mathbb{R}$ is uncountable. ■

Exercises for Section 2.9

A. Prove that the set $\mathbb{Z}^n$, consisting of all n-tuples $\mathbf{a}=(a_1,a_2,\ldots,a_n)$, where $a_i\in\mathbb{Z}$, is countable.

B. Show that $(0,1)$ and $[0,1]$ have the same cardinality as $\mathbb{R}$.

C. Show that if $|A|\leq|B|$ and $|B|\leq|C|$, then $|A|\leq|C|$.

D. Prove that the set of all infinite sequences of integers is uncountable.
HINT: Modify the diagonalization argument.

E. A real number α is called an **algebraic number** if there is a polynomial with integer coefficients with α as a root. Prove that the set of all algebraic numbers is countable.
HINT: First count the set of all polynomials with integer coefficients.

F. A real number that is not algebraic is called a **transcendental number**. Prove that the set of transcendental numbers has the same cardinality as $\mathbb{R}$.

G. Show that the set of all *finite* subsets of $\mathbb{N}$ is countable.

H. Prove **Cantor's Theorem**: that for any set X, the power set $P(X)$ of all subsets of X satisfies $|X|\neq|P(X)|$. HINT: If f is an injection from X into $P(X)$, consider $A=\{x\in X : x\notin f(x)\}$.

I. If A is an infinite set, show that A has a countable infinite subset.
HINT: Use recursion to choose a sequence a_n of distinct points in A.

J. Show that A is infinite if and only if there is a proper subset B of A such that $|B|=|A|$.
HINT: Use the previous exercise and let $B=A\setminus\{a_1\}$.

Springer

http://www.springer.com/978-0-387-98097-3

Real Analysis and Applications
Theory in Practice
Davidson, K.R.; Donsig, A.P.
2010, XII, 513 p., Hardcover
ISBN: 978-0-387-98097-3