DISCUSSION. *How would we do this from first principles? We need to show that for every $a \in \mathbb{R}$, for every $\varepsilon > 0$, we can always find a $\delta > 0$ such that for any $x \in \mathbb{R}$*

$$|x - a| < \delta \quad \Longrightarrow \quad |x^2 - a^2| < \varepsilon. \tag{5.19}$$

(Why don't we need to add the hypothesis $0 < |x - a|$?) The easiest way to do this is to write down a formula that, given a and ε, produces a δ satisfying (5.19).

As $x^2 - a^2 = (x - a)(x + a)$, if $|x - a|$ is less than some number δ (still unspecified), then $|x^2 - a^2|$ is less than $\delta|x + a|$. So we want

$$\delta|x + a| \leq \varepsilon. \tag{5.20}$$

We can't choose $\delta = \varepsilon/|x + a|$, because δ cannot depend on x. But if $|x - a| < \delta$, then

$$\begin{aligned} |x + a| &\leq |x| + |a| \\ &< |a| + \delta + |a| = 2|a| + \delta, \end{aligned}$$

so

$$|x^2 - a^2| < \delta(2|a| + \delta) \overset{?}{\leq} \varepsilon. \tag{5.21}$$

We must choose δ so that the last inequality holds. By the quadratic formula,

$$\delta(2|a| + \delta) \leq \varepsilon \quad \Longleftrightarrow \delta \leq \sqrt{|a|^2 + \varepsilon} - |a|.$$

So choose $\delta = \sqrt{|a|^2 + \varepsilon} - |a|$ and (5.19) holds. □

REMARK. A formally correct proof could have been reduced to:

PROOF. Let $a \in \mathbb{R}$ and $\varepsilon > 0$. Then letting $\delta = \sqrt{|a|^2 + \varepsilon} - |a|$ we have $|x - a| < \delta \implies |x^2 - a^2| < \varepsilon$. Q.E.D.

However, while a diligent reader could verify that this proof is correct, pulling δ out of a hat like this doesn't give the reader the insight that our much longer proof does. Remember, a proof has more than one function: not only must it convince the reader that the claimed result is true, but it should also help the reader understand *why* the result is true. A good proof should be describable in a few English

sentences, so that a knowledgeable listener can then go write down a more detailed proof fairly easily.

REMARK. One does not need to choose the largest value of δ so that the inequality $\overset{?}{\leq}$ in (5.21) holds — any positive δ that satisfies the inequality will work. This allows one to simplify the algebra. For example, let δ_1 be such that

$$|x - a| < \delta_1 \;\Rightarrow\; |x + a| < 2|a| + 1.$$

(Such a δ_1 exists from the continuity of the simpler function $x \mapsto x$). Then let

$$\delta \;=\; \min\left(\delta_1, \frac{\varepsilon}{2|a|+1}\right)$$

and (5.20) holds.

One could imagine repeating proofs like the above to show that x^3, x^4, and so on are continuous, but we want to take big steps. Can we show all polynomials are continuous?

First observe that because limits behave well with respect to algebraic operations (Theorem 5.9), and continuity is defined in terms of limits, then algebraic combinations of continuous functions are continuous.

PROPOSITION 5.22. *Suppose $f : X \to \mathbb{R}$ and $g : X \to \mathbb{R}$ are real functions that are continuous at $a \in X$. Let c and d be scalars*[2]*. Then $cf + dg$ and fg are both continuous at a, and so is f/g if $g(a) \neq 0$.*

PROOF. Exercise. □

Constant functions are continuous (Lemma 5.15), and the function $f(x) = x$ is continuous (Exercise 5.16). So one can prove by induction on the degree of polynomial, using Proposition 5.22, that that all polynomials are continuous (Exercise 5.27). Once you have proved that all polynomials are continuous, you may prove that rational functions are continuous wherever the denominator doesn't vanish.

This result is used so frequently that we will state it formally.

[2]A scalar is just a fancy word for a number.

PROPOSITION 5.23. *Every polynomial is continuous on* $\mathbb{R}$. *Every rational function is continuous wherever the denominator is non-zero.*

What about the exponential function

$$e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!} \ ?$$

Each partial sum is a polynomial, and hence continuous; so if we knew that the limit of a sequence of continuous functions were continuous, we would be done. This turns out, however, to be a subtle problem, which we address in the next Section.

5.3. Sequences of Functions

An infinite sequence of numbers $\langle a_n \rangle$ tends to a limit L if a_n approaches L as n tends to infinity. Try to write down a formal definition of this before reading further.

DISCUSSION. *Hint. We have already seen how to encode the statement "approaches L". The difficulty is to encode "as n tends to infinity". How might you do this?*

DEFINITION. $\lim_{n\to\infty} a_n$, converge, diverge The sequence $\langle a_n \rangle$ tends to the limit L as n tends to infinity, written

$$\lim_{n\to\infty} a_n \ = \ L,$$

if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$(\forall n \in \mathbb{N})\ n > N \quad \Longrightarrow \quad |a_n - L| \ < \ \varepsilon.$$

We say that the sequence $\langle a_n \rangle$ converges to L. If a sequence does not converge, we say it diverges.

EXAMPLE 5.24. Prove that the sequence

$$\langle \sin^2(n)/n \mid n \in \mathbb{N} \rangle$$

converges.

DISCUSSION. *It is generally easiest to prove that a sequence converges if we have an idea of its limit. To prove convergence of sequences without a candidate for the limit usually involves using the least upper bound property of* $\mathbb{R}$ *(which is covered in Chapter 8). It certainly seems that the terms in the sequence are getting closer to* 0*, so we try to show this rigorously.*

We observe that

$$(\forall n \in \mathbb{N}) \ \ | \sin^2(n) | \leq 1.$$

Hence

$$(\forall n \in \mathbb{N}) \ \ | \sin^2(n)/n | \leq | \ 1/n \ | \ .$$

Let $\varepsilon > 0$ and $N \in \mathbb{N}$ be such that $1/N \leq \varepsilon$. Then for any $n \geq N$,

$$| \sin^2(n)/n - 0 | \leq 1/n \leq \varepsilon.$$

Therefore

$$\lim_{n \to \infty} \frac{\sin^2(n)}{n} \ = \ 0.$$

EXAMPLE 5.25. For any $n \in \mathbb{N}$, let $a_n \ = \ (-1)^n$. Show that the sequence $\langle a_n \rangle$ diverges.

DISCUSSION. *Since the sequence alternates between* -1 *and* 1*, it is intuitively clear that the sequence does not tend to any particular number. We wish to show that a statement in the form*

$$(\exists L \in \mathbb{R})(\forall \varepsilon > 0)(\exists N \in \mathbb{N})(\forall n > N)(\ldots)$$

is false. So we must show that the negation of the statement is true. That is

$$(\forall L \in \mathbb{R})(\exists \varepsilon > 0)(\forall N \in \mathbb{N})(\exists n > N) \ \neg(\ldots).$$

For any $L \in \mathbb{R}$*, if we pick* $\varepsilon < 1$ *we will not be able to capture both* -1 *and* 1 *in the* ε*-neighborhoods of* L*. This will prove that the sequence diverges.*

Let $L \in \mathbb{R}$ and $\varepsilon < 1$. We show that for any $N \in \mathbb{N}$, there is $n > N$ such that

$$| \ (-1)^n - L \ | \geq \varepsilon.$$

Let $N \in \mathbb{N}^+$. We argue by cases.
Suppose $L < 0$. Then

$$| (-1)^{2N} - L | \geq 1 > \varepsilon.$$

Suppose $L \geq 0$. Then

$$| (-1)^{(2N+1)} - L | \geq 1 > \varepsilon.$$

Therefore the sequence $\langle a_n \rangle$ diverges.

EXAMPLE 5.26. For all $n \in \mathbb{N}$, let $a_n = \sum_{k=0}^{n} \left(\frac{k}{n}\right)^2 \frac{1}{n}$. What is the limit of the sequence $\langle a_n \rangle$?

DISCUSSION. *The terms of the sequence may be familiar to you as Riemann sums associated with the area under the parabola $f(x) = x^2$ between $x = 0$ and $x = 1$. We will use a combinatorial result we proved by induction in the last chapter.*

By Proposition 4.6,

$$\begin{aligned} \lim_{n\to\infty} \sum_{k=0}^{n} \left(\frac{k}{n}\right)^2 \frac{1}{n} &= \lim_{n\to\infty} \left(\frac{1}{n^3}\right) \sum_{k=0}^{n} k^2 \\ &= \lim_{n\to\infty} \left(\frac{1}{n^3}\right) \frac{(n)(n+1)(2n+1)}{6} \\ &= 1/3. \end{aligned}$$

Verification of the last equality is left to the reader as an exercise.

In the next section we are particularly interested in infinite sums.

DEFINITION. Infinite sum, partial sum, $\sum_{k=0}^{\infty} a_k$ Let $\langle a_k \mid k \in \mathbb{N} \rangle$ be a sequence of numbers. The n^{th} partial sum of the sequence is

$$s_n = \sum_{k=0}^{n} a_k.$$

The infinite sum of the sequence is

$$\sum_{k=0}^{\infty} a_k := \lim_{n\to\infty} s_n.$$

The infinite sum is the limit of the sequence of *partial* sums, $\langle s_n \rangle$.

EXAMPLE 5.27. Show that

$$\sum_{k=0}^{\infty} \frac{1}{2^k} \;=\; 2.$$

Let s_n be the n^{th} partial sum. We need to show that

$$\lim_{n\to\infty} s_n \;=\; 2.$$

Let $\varepsilon > 0$ and $N \in \mathbb{N}$ be such that $\frac{1}{2^N} < \varepsilon$. We show that if $n \geq N$, then

$$|\, s_n - 2 \,| < \varepsilon.$$

Since the series $\sum_{k=0}^{\infty} \frac{1}{2^k}$ is geometric, we know that

$$s_n \;=\; \frac{1 - 2^{-(n+1)}}{1 - 1/2} \;=\; 2 - 2^{-n}.$$

So if $n \geq N$, then

$$|\, s_n - 2 \,| \;=\; 2^{-n} \;<\; \varepsilon.$$

In analysis, one is often concerned with a sequence of functions f_n. For example, f_n might be the n^{th}-order Taylor polynomial of some function f, and one wants to know whether this sequence f_n converges to f; or the sequence f_n may represent functions whose graphs have a fixed boundary curve in $\mathbb{R}^3$ and have decreasing areas, and one wants to know if the sequence converges to the graph of a function with minimal area for that boundary. This sort of problem is so important that mathematicians study different ways in which a sequence of functions might converge. The most obvious way is pointwise:

DEFINITION. Pointwise convergence A sequence of functions f_n on a set X converges pointwise to the function f if, for all x in X, the sequence of numbers $\langle f_n(x) \rangle$ converges to $f(x)$.

In order for the definition to make sense, we require that

$$X \subseteq \bigcap_{n\in\mathbb{N}} \text{Dom}(f_n).$$

If $a \in X$, the pointwise convergence of a sequence of functions, $\langle f_n \rangle$, at the point a is dependent on the convergence of the sequence of numbers,

$\langle f_n(a) \rangle$. If you do not understand convergence of a sequence of numbers you cannot understand convergence of a sequence of functions.

EXAMPLE 5.28. Consider the functions $f_n(x) = x^n$. On the open interval $(-1, 1)$, these functions converge pointwise to 0. At the point 1, the functions converge to 1; at the point -1, the functions do not converge, because the values oscillate between $+1$ and -1. Outside of the set $(-1, 1]$ the sequence of functions diverges.

The preceding example illustrates the main problem with pointwise convergence: the sequence of continuous functions x^n on the set $[0, 1]$ converges, but the function to which it converges is not continuous. Even Cauchy made this mistake: he stated as a theorem in his 1821 book *Cours d'analyse* that if a sequence of continuous functions converges pointwise, then its limit is continuous[3]. To get around this problem, we introduce the notion of *uniform convergence*.

DEFINITION. Uniform convergence The sequence of real functions f_n defined on a set $X \subseteq \mathbb{R}$ is said to converge uniformly to the function f on X if, for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that, for every x in X, whenever $n > N$ then $|f_n(x) - f(x)| < \varepsilon$. In logical notation:

$$(\forall\, \varepsilon > 0)\ (\exists\, N \in \mathbb{N})\ (\forall\, x \in X)\ (\forall\, n > N) \quad |f_n(x) - f(x)| < \varepsilon.$$

Note the big difference between pointwise and uniform convergence: in pointwise convergence N can depend on x; in uniform convergence it cannot. The importance of uniform convergence stems from the following theorem.

THEOREM 5.29. *Let f_n be a sequence of continuous functions on X that converges uniformly to f on X. Then f is continuous on X.*

DISCUSSION. *We must show $|f(x) - f(a)|$ is small when x is close to a. We know that $|f_n(x) - f(x)|$ is small for* all *x; so we refine the*

[3]See the book [**4**] by Imre Lakatos for an interesting historical discussion of Cauchy's mistake and the discovery of uniform convergence, by Seidel and Stokes independently in 1847.

trick from p. 133, and add and subtract the same thing twice, writing

$$f(x) - f(a) \;=\; [f(x) - f_n(x)] \quad + \; [f_n(x) - f_n(a)] \quad + \; [f_n(a) - f(a)].$$

Then we try to make each of the three grouped pairs small, so their sum is small. This is sometimes called an $\varepsilon/3$ argument, because if we make each term smaller than $\varepsilon/3$, then their sum is smaller than ε.

PROOF. Fix some point $a \in X$, and let $\varepsilon > 0$. We must find $\delta > 0$ so that

$$|x - a| < \delta \quad \Longrightarrow \quad |f(x) - f(a)| < \varepsilon. \tag{5.30}$$

To do this, we split $f(x) - f(a)$ into three parts:

$$f(x) - f(a) \;=\; [f(x) - f_n(x)] \quad + \; [f_n(x) - f_n(a)] \quad + \; [f_n(a) - f(a)].$$

Choose N so that $n \geq N$ implies $|f_n(x) - f(x)| < \varepsilon/3$ for all x. Choose $\delta > 0$ so that $|f_N(x) - f_N(a)| < \varepsilon/3$ whenever $|x - a| < \delta$. Then by the triangle inequality, for $|x - a| < \delta$, we have

$$\begin{aligned} |f(x) - f(a)| &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(a)| + |f_N(a) - f(a)| \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \;=\; \varepsilon. \end{aligned}$$

□

QUESTION. Where did we use the hypothesis that the convergence was uniform?

We can use theorem 5.29, for example, to prove that the exponential function is continuous. We consider the exponential function as its Taylor series

$$e^x \;=\; \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Recall that the expression $\sum_{k=0}^{\infty} \frac{x^k}{k!}$ is a shorthand for

$$\lim_{n \to \infty} \sum_{k=0}^{n} \frac{x^k}{k!}.$$

For any real number a, $\sum_{k=0}^{\infty} \frac{a^k}{k!}$ is an infinite sum which converges if its corresponding sequence of partial sums converges. By the ratio test,

the exponential series converges for all real a. (For a formal proof of the ratio test, see Theorem 8.9).

PROPOSITION 5.31. *The exponential function is continuous on* $\mathbb{R}$.

PROOF. Let

$$p_n(x) \;:=\; \sum_{k=0}^{n} \frac{x^k}{k!}$$

be the n^{th}-order Taylor polynomial. We know each p_n is continuous, by Proposition 5.23. If we knew that $p_n(x)$ converged uniformly to e^x, we would be done by Theorem 5.29.

It is not true that p_n converges uniformly on $\mathbb{R}$ (why?). However, the sequence does converge uniformly on every interval $[-R, R]$, and this is good enough to conclude that e^x is continuous on $\mathbb{R}$ (why?).

To see this latter assertion, fix $R > 0$ and $\varepsilon > 0$. We must find N so that, for all $n > N$ and all $x \in [-R, R]$, we have $|e^x - p_n(x)| < \varepsilon$. Notice that

$$|\, e^x - p_n(x)| \;=\; \left| \frac{x^{n+1}}{(n+1)!} \;+\; \frac{x^{n+2}}{(n+2)!} \;+\; \dots \right|.$$

For each n, the right-hand side is maximized on $[-R, R]$ by its value at R (why?); and as n increases, this remainder decreases monotonically (because you lose more and more positive terms). As we know the exponential series for e^R converges, choose an N so that $e^R - p_N(R)$ is less than ε. Then for all x in $[-R, R]$ and all $n \geq N$, we have $|\, e^x - p_n(x)| < \varepsilon$, as desired. □

The sine and cosine functions can be defined in terms of their Taylor series too:

$$\begin{aligned} \sin(x) \quad &:= \quad \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \\ \cos(x) \quad &:= \quad \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}. \end{aligned}$$

They can be proved to be continuous by similar arguments.

REMARK. Notice that in our definitions of limits and continuity, we are using the absolute value just to measure distances. In other words, we are saying that f is continuous at a if, for all $\varepsilon > 0$, we can find $\delta > 0$, such that whenever the distance from x to a is less than δ, then the distance from $f(x)$ to $f(a)$ is less than ε. This definition makes perfectly good sense whenever one has a way of measuring distances on the domain and codomain. For example, if the function maps $\mathbb{R}^m$ to $\mathbb{R}^n$, one can measure distances in the usual Euclidean way. In even greater generality, mathematicians use something called *metrics* to measure distances, and once one has metrics, one can discuss the continuity of functions in a similar way to our discussion for real functions.

The mathematics of this chapter — a close look at the behavior of real functions — is called Analysis. This comprises one of the three major disciplines of pure mathematics; the other two are Geometry and Algebra. A good introduction to analysis is Walter Rudin's book [**?**].

5.4. Exercises

EXERCISE 5.1. Prove that the definitions of limit on pages 129 and 130 are the same.

EXERCISE 5.2. Prove Lemma 5.14, and the related assertion that $|c| - |d| \leq |c + d|$.

EXERCISE 5.3. For $n \in \mathbb{N}^+$, $a_i \in \mathbb{R}$ $(1 \leq i \leq n)$, prove that

$$\left| \sum_{i=1}^{n} a_i \right| \leq \sum_{i=1}^{n} | \, a_i \, | \, .$$

EXERCISE 5.4. Prove Lemma 5.15.

EXERCISE 5.5. Prove part *(v)* of Theorem 5.9.

EXERCISE 5.6. Give an example of two functions f and g that don't have limits at a point a but such that $f + g$ does. For the same pair of functions, can $f - g$ also have a limit at a?

EXERCISE 5.7. Assume that f is a real function and $\lim_{x\to a} f(x) = L$. Prove that if $X \subseteq \mathrm{Dom}(f)$, then

$$\lim_{X\ni x\to a} f(x) = L.$$

EXERCISE 5.8. Use Archimedes's method (the method of Riemann sums) to prove that

$$\int_0^1 x^2 dx = \frac{1}{3}.$$

(You will need to know a formula for $\sum_{k=0}^{n} k^2$ - see Proposition 4.6).

EXERCISE 5.9. Use Archimedes's method to prove that

$$\int_0^1 x^3 dx = \frac{1}{4}.$$

(See Exercise 4.16).

EXERCISE 5.10. Prove that the Heaviside function has both left and right-hand limits at 0.

EXERCISE 5.11. Prove that a function has a limit at a point if and only if it has both left and right limits at that point and their values coincide.

EXERCISE 5.12. Prove that Theorem 5.9 applies to restricted limits.

EXERCISE 5.13. The point a is a *limit point* of the set X if, for every $\delta > 0$, there exists a point x in $X \setminus \{a\}$ with $|x - a| < \delta$. Let f be a real-valued function on $X \subseteq \mathbb{R}$. Prove that if a is a limit point of X, then if f has a restricted limit at a it is unique. Prove that if a is not a limit point of X, then every real number is a restricted limit of f at a.

EXERCISE 5.14. Prove that $\lim_{x\to 0} \sin(x)/x = 1$.

EXERCISE 5.15. Prove Proposition 5.22.

EXERCISE 5.16. Prove that the function $f(x) = x$ is continuous everywhere on $\mathbb{R}$.

EXERCISE 5.17. A formula for the Fibonacci numbers is given in Exercise 4.12. Evaluate $\lim_{n\to\infty} F_{n+1}/F_n$.

EXERCISE 5.18. How large must n be to ensure that F_{n+1}/F_n is within 10^{-1} of the limit in Exercise 5.17? Within 10^{-2}? Within 10^{-k}?

EXERCISE 5.19. Define the function $\psi : \mathbb{R} \to \mathbb{R}$ by

$$\psi(x) \ := \ \begin{cases} 0 & x \notin \mathbb{Q} \\ 1 & x \in \mathbb{Q}. \end{cases}$$

Prove that ψ is discontinuous everywhere.

EXERCISE 5.20. Define the function $\phi : \mathbb{R} \to \mathbb{R}$ by

$$\phi(x) \ := \ \begin{cases} 0 & x \notin \mathbb{Q} \\ \frac{1}{n} & x \in \mathbb{Q} \setminus \{0\},\ x = \frac{m}{n},\ \gcd(m,n) = 1,\ n > 0 \\ 1 & x = 0. \end{cases}$$

Prove that ϕ is continuous at every irrational number and discontinuous at every rational number.

EXERCISE 5.21. Prove that a real-valued function f on an open interval I is continuous at any point where its derivative exists, *i.e.* where

$$\lim_{x\to a} \frac{f(x) - f(a)}{x - a}$$

exists. What is the converse of this assertion? Prove that the converse is not true.

EXERCISE 5.22. Prove that if the function f has the limit L from the right at a, then the sequence $f(a + \frac{1}{n})$ has limit L as $n \to \infty$. Show that the converse is false in general.

EXERCISE 5.23. Let f and g be real functions. Let $a \in \mathbb{R}$ and suppose that

$$\lim_{x\to a} g(x) \ = \ L_1$$

and

$$\lim_{x\to L_1} f(x) \ = \ L_2.$$

Prove that

$$\lim_{x\to a} f \circ g \ = \ L_2.$$

If g is continuous at a and f is continuous at $g(a)$, is $f \circ g$ continuous at a?

EXERCISE 5.24. Let f be a real function, $a \in \mathbb{R}$ and $\lim_{x\to a} f(x) = L$. If $\langle a_n \rangle$ converges to a, prove that $\langle f(a_n) \rangle$ converges to L.

EXERCISE 5.25. Complete Example 5.26. That is, prove that

$$\lim_{n\to\infty} \left(\frac{1}{n^3}\right) \frac{(n)(n+1)(2n+1)}{6} \;=\; 1/3.$$

EXERCISE 5.26. Evaluate

$$\lim_{n\to\infty} \sum_{k=0}^{n} \left(\frac{k}{n}\right) \frac{1}{n}.$$

Can you give a geometrical interpretation of this limit?

EXERCISE 5.27. Use induction to prove that every polynomial is continuous at every real number.

EXERCISE 5.28. Let $-1 < x < 1$. Prove that the geometric series with ratio x, $\sum_{k=0}^{\infty} x^k$, converges to $\frac{1}{1-x}$.

EXERCISE 5.29. Let the Fibonacci numbers F_n be defined as in Exercise 4.12. Consider the power series $F(x) = \sum_{n=1}^{\infty} F_n x^n$. Prove that the power series satisfies

$$F(x) \;=\; x^2F(x) + xF(x) + x. \tag{5.32}$$

Solve (5.32) for $F(x)$, decompose it by partial fractions, and use Exercise 5.28 to derive Formula 4.20. This technique to find a formula for F_n by studying the function F is often fruitful. The function F is called the generating function for the sequence.

EXERCISE 5.30. Suppose one defines a sequence with the same recurrence relation as the Fibonacci numbers, $F_{n+2} = F_{n+1} + F_n$, but with different starting values for F_1 and F_2. Find the generating function for the new sequence, and hence calculate a formula for the general term. Is $\lim_{n\to\infty} F_{n+1}/F_n$ always the same?

EXERCISE 5.31. Prove that sine and cosine are continuous functions on all of $\mathbb{R}$.

CHAPTER 6

Cardinality

In this chapter we use functions to explore the idea of the size of a set. The results we derive are deep and very interesting, especially when we consider the simplicity of the tools we are using. Of course, we shall have to use these tools somewhat cleverly.

Set theory comes in different flavors. The most difficult is axiomatic set theory. Many interesting results have been derived in formal axiomatic set theory, but the topic is advanced and not suitable for an introduction to higher mathematics. Instead, we shall study what is called naive set theory. The use of the word "naive" is not pejorative, but is meant to differentiate this approach from axiomatic set theory. Most mathematicians have studied naive set theory, but relatively few have worked extensively with set axioms.

6.1. Cardinality

We wish to compare the size of sets. The fundamental tool for our investigation is the bijection. In the case of finite sets, which can be exhaustively listed, this is easy. Given any two finite sets, X and Y, we could list the elements and count them. Provided that our lists have no redundancies, the larger set is the one with the higher count. The act of listing the elements in a set, where this is possible, is also defining a bijection from a natural number (interpreted as a set) to the set being counted. The idea of using functions to compare the size of sets can be generalized to arbitrary sets.

When it comes to comparing the size of infinite sets there are competing intuitions. On the one hand we have an intuition that if one set is a proper subset of another set, it should be smaller. On the other

hand if two sets are infinite, how can one be larger than the other? Using bijections, injections and surjections to define the relative size of sets allows us to see our way through this paradox.

DEFINITION. Equinumerous, cardinality Let X and Y be sets. We say that X and Y have the same cardinality if there is a bijection $f : X \rightarrowtail Y$. We can express that two sets have the same cardinality by

$$| X | = | Y | .$$

If $| X | = | Y |$, then we say that X and Y are equinumerous.

CLAIM. Equinumerosity is an equivalence relation.

(Prove this: Exercise 6.2).

Although we used the ideas of finite and infinite before now, we shall define the ideas in terms of bijections.

DEFINITION. Finite, infinite Let X be a set. X is finite if there exists some $n \in \mathbb{N}$ and a bijection $f : \ulcorner n \urcorner \rightarrowtail X$. In the case that $X = \emptyset$, we say that X is bijective with $\ulcorner 0 \urcorner$ via the empty function. If X is not finite, we say that X is infinite.

So a set is finite if it is bijective with a set $\ulcorner n \urcorner$ for some $n \in \mathbb{N}$. It is probably no surprise that a set cannot be bijective with different natural numbers.

PROPOSITION 6.1. *Let $m, n \in \mathbb{N}$. Then*

$$(\ | \ulcorner m \urcorner | = | \ulcorner n \urcorner | \ ) \iff (\ m = n \ ).$$

DISCUSSION. *We prove the non-trivial direction of this biconditional by induction on one of the integers in the statement.*

PROOF. $\Leftarrow$
Let $m = n$. Then it is obvious that

$$| \ulcorner m \urcorner | = | \ulcorner n \urcorner | .$$

$\Rightarrow$

We argue by induction on m.

Base case:

If $m = 0$ and $|\ulcorner n \urcorner| = |\ulcorner m \urcorner|$ then clearly $n = 0$.

Induction step:

Let $m \in \mathbb{N}$ and assume that

$$(\forall n \in \mathbb{N})\ [\ |\ulcorner m \urcorner| = |\ulcorner n \urcorner|\ ] \Rightarrow [m = n].$$

We show that

$$(\forall n \in \mathbb{N})\ [\ |\ulcorner m+1 \urcorner| = |\ulcorner n \urcorner|\ ] \Rightarrow [m+1 = n].$$

Assume that

$$|\ulcorner m+1 \urcorner| = |\ulcorner n \urcorner|\ .$$

Let

$$f : \ulcorner m+1 \urcorner \rightarrowtail \ulcorner n \urcorner.$$

Discussion. *A natural way to proceed with this argument is to restrict the domain of f to $\ulcorner m \urcorner$ and use the induction hypothesis. Unfortunately if $f(m) \neq n-1$ then $f|_{\ulcorner m \urcorner}$ is not a bijection from $\ulcorner m \urcorner$ to $\ulcorner n-1 \urcorner$, and the induction hypothesis will not directly apply. To address this issue, we shall define a permutation $g : \ulcorner m+1 \urcorner \to \ulcorner m+1 \urcorner$ that rearranges the elements of $\ulcorner m+1 \urcorner$ so that $f \circ g$ will be a bijection satisfying*

$$(f \circ g)(m) = n-1.$$

We define $g : \ulcorner m+1 \urcorner \to \ulcorner m+1 \urcorner$ as follows:

$$g(x) = \begin{cases} f^{-1}(n-1) & \text{if} \quad x = m \\ m & \text{if} \quad x = f^{-1}(n-1) \\ x & \text{otherwise.} \end{cases}$$

Let $h = f \circ g$. Then h is a bijection and

$$h(m) = (f \circ g)(m) = n-1.$$

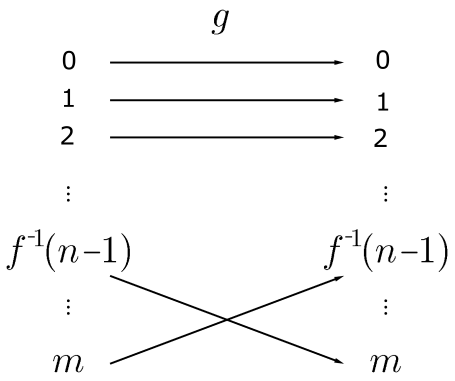

FIGURE 6.2. Picture of the permutation g

Therefore

$$h|_{\ulcorner m \urcorner} : \ulcorner m \urcorner \rightarrowtail \ulcorner n - 1 \urcorner.$$

By the induction hypothesis

$$m \ = \ n - 1.$$

Therefore

$$m + 1 \ = \ n.$$

By the induction principle,

$$(\forall m \in \mathbb{N})(\forall n \in \mathbb{N}) \ (| \ \ulcorner m \urcorner \ | = | \ \ulcorner n \urcorner \ | \ ) \Rightarrow (m \ = \ n).$$

□

COROLLARY 6.3. *If X is a finite set, there is exactly one $n \in \mathbb{N}$ such that $\ulcorner n \urcorner$ is bijective with X.*

DISCUSSION. *This is a standard uniqueness argument. We assume that a set is bijective with natural numbers $\ulcorner n \urcorner$ and $\ulcorner m \urcorner$, and we use that the composition of bijections is a bijection to show that $m = n$. This is not a proof by contradiction. Rather we are proving that any two names for natural numbers that are bijective with X must name the same natural number.*

PROOF. X is finite, so there is $n \in \mathbb{N}$ such that

$$| X | = | \ulcorner n \urcorner | .$$

Let $m \in \mathbb{N}$ and

$$| X | = | \ulcorner m \urcorner | .$$

Let $f : X \rightarrowtail \ulcorner n \urcorner$ and $g : X \rightarrowtail \ulcorner m \urcorner$. Then $g^{-1} : \ulcorner m \urcorner \rightarrowtail X$. The composition of bijections is a bijection, so

$$f \circ g^{-1} : \ulcorner m \urcorner \rightarrowtail \ulcorner n \urcorner.$$

By Proposition 6.1,

$$m \ = \ n.$$

□

DEFINITION. Finite cardinality If X is a finite set, we say that it has finite cardinality. Let $n \in \mathbb{N}$ be the unique natural number such that $\ulcorner n \urcorner$ is bijective with X. Then we say that X has cardinality n, or

$$| X | = \ n.$$

Corollary 6.3 guarantees that the cardinality of a finite set is well-defined.

6.2. Infinite Sets

Infinite sets are mysterious. Many classical paradoxes address historical confusions about the idea of infinity. At the same time, mathematicians from the ancient Greeks on have found it impossible to develop mathematical thinking without the use of infinity. Why is this so? From a metaphysical point of view, the idea of infinity is probably not necessary. From a physical point of view, there is no evidence for infinity. That is, the universe, as we understand it, is finite. Even from a theological point of view, infinity is to some extent the complement of the finite — and correspondingly gives rise to its own paradoxes.

Infinity has troubled some mathematicians and philosophers, and a few have tried to dispense with it. There aren't many adherents to this school. The idea of infinity is so useful that the mathematics student

will have to develop some comfort with the idea — and its logical consequences. At any rate, infinity clearly exists in the mathematical universe, whether or not it exists in the natural world, and using infinity has been crucial to developing a mathematical understanding of the natural world. In this section we begin an investigation of infinite sets.

We shall use injections and surjections to build some analytical machinery for comparing sets.

NOTATION. $\preceq$ *Let X and Y be sets. We write $X \preceq Y$ if there is an injection*

$$f : X \to Y.$$

This notation suggests that, under the conditions of the definition, we think of X as being "no bigger than" than Y. This makes sense, since we are able to associate to any element of X a distinct element of Y. If f in the definition is a surjection, then f is also a bijection and $| X | = | Y |$. Otherwise, f is a function that associates with each element of the range of f (which is a proper subset of Y) a unique element of X, and Y still has some elements unaccounted for by f. So Y might be "bigger" than X, but it certainly won't be "smaller". You might wish to consider this definition in the special case of finite sets X and Y. You will observe that

$$X \preceq Y \iff | X | \leq | Y | .$$

In Exercise 6.3 you are asked to prove that $\preceq$ is transitive and reflexive.

REMARK. Are any two sets comparable with respect to $\preceq$? Rather surprisingly, it requires a more advanced assumption, called the Axiom of Choice (see Appendix B), in order to guarantee the comparability of all pairs of sets. Virtually all mathematicians accept the Axiom of Choice. We shall assume the Axiom of Choice in this text.

If $X \preceq Y$ and $Y \preceq X$, we would hope that X and Y are the same size. This is indeed true, though the proof is a little tricky. The result is very useful, because it is often much easier to write down two injections than one bijection.

THEOREM 6.4. *Schröder-Bernstein Theorem Let X and Y be sets. If $X \preceq Y$ and $Y \preceq X$, then $| X | = | Y |$.*

DISCUSSION. *The idea behind this proof is as follows. We show that $| X | = | Y |$ by constructing a bijection $F : X \rightarrowtail Y$. We are given injections $f : X \to Y$ and $g : Y \to X$. We build F using the injections f and g as guides. That is, we wish to define F so that for each $x \in X$, either $F(x) = f(x)$ or $F(x) = g^{-1}(x)$. It is obvious that this cannot be accomplished blindly. For instance, if $x \in X \setminus g[Y]$, our hand is forced, and $F(x) = f(x)$. Similarly, if $y \in Y \setminus f[X]$, our only chance of achieving our objective is for $F(g(y)) = y$. If we make the wrong choice for $F(x)$, we shall lose the use of f and g as guides. We might consider F undecided about x since f and g^{-1} do not agree. The solution is to use f and g to move back and forth between X and Y until we find that our hand is forced.*

PROOF. Let

$$f : X \to Y$$

and

$$g : Y \to X$$

be injections. We may assume that X and Y are disjoint.

DISCUSSION. *If X and Y are not disjoint, we can replace X with $X \times \{0\}$ and Y with $Y \times \{1\}$. The existence of a bijection*

$$g : X \times \{0\} \rightarrowtail Y \times \{1\}$$

clearly implies the existence of a bijection from X to Y.

If $x \in X$ we say $y \in Y$ is the predecessor of x if $g(y) = x$. Analogously, if $y \in Y$ we say that $x \in X$ is the predecessor of y if $f(x) = y$. It is possible for an element not to have a predecessor. For example, if $x \in X \setminus g[Y]$, then x has no predecessor. However, if an element does have a predecessor, that predecessor is unique (since f and g are both injections).

Given an element w, let $m(w)$ be 0 if w does not have a predecessor. Otherwise, let $m(w)$ be the maximum number $N \geq 1$ such that there is a finite sequence $\langle z_n \mid 0 \leq n \leq N \rangle$ for some $N \geq 1$ satisfying

(1) $w = z_N$
(2) For $n < N$, z_n is the predecessor of z_{n+1},

if the maximum exists. If the maximum doesn't exist (*i.e.* if one can make arbitrarily long chains of predecessors), let $m(w) = \infty$.

Now define

$$\begin{aligned} X_e &= \{x \in X \mid m(x) \text{ is even}\} \\ X_o &= \{x \in X \mid m(x) \text{ is odd}\} \\ X_i &= \{x \in X \mid m(x) = \infty\} \\ Y_e &= \{y \in Y \mid m(y) \text{ is even}\} \\ Y_o &= \{y \in Y \mid m(y) \text{ is odd}\} \\ Y_i &= \{y \in Y \mid m(y) = \infty\} \end{aligned}$$

The collection

$$\{X_e, X_o, X_i\}$$

is obviously a partition of X. Similarly,

$$\{Y_o, Y_e, Y_i\}$$

is a partition of Y.

We are now in a position to define a bijection between X and Y. Let

$$F(x) = \begin{cases} f(x) & \text{if} \quad x \in X_i \\ f(x) & \text{if} \quad x \in X_e \\ g^{-1}(x) & \text{if} \quad x \in X_o. \end{cases}$$

Discussion. *We have some work left in this proof. We need to verify that F is a bijection from X to Y. The idea behind the definition of F may not be obvious, so let's investigate the motivation for the definition. Suppose that f and g fail to be surjections (if either of the functions is a surjection there would be nothing to prove, since it would also be a bijection). Let $x \in X \setminus g[Y]$ and $y \in Y \setminus f[X]$. Since $x \notin g[Y]$, the only possible choice for $F(x)$ is $f(x)$. Similarly, $y \notin f[X]$, and the*

only possible value of $F^{-1}(y)$ *is* $g(y)$. *But this does not solve all of our problems. The set* $X \setminus g[Y]$ *is made up of those members of* X *that have no predecessors, and* $Y \setminus f[X]$ *is composed of the members of* Y *with no predecessors. If we are to define* F *by piecing together* f *and* g, *we found that our hands were forced with these sets. Now suppose that* $x \in X$ *has exactly one antecedent. Then* $g^{-1}(x)$ *has no predecessor. As we observed earlier, we need to satisfy*

$$F^{-1}(g^{-1}(x)) \;=\; g(g^{-1}(x)) \;=\; x$$

and therefore we must satisfy

$$F(x) \;=\; g^{-1}(x).$$

Similarly, if $y \in Y$ *has exactly one antecedent, we must satisfy*

$$F^{-1}(y) \;=\; f^{-1}(y).$$

If an element w *of* $X \cup Y$ *has finitely many antecedents,* $F|_{A(w)}$ *will be determined by the constraint imposed by the antecedent with no predecessor.*

We claim that

$$F : X \rightarrowtail Y.$$

It is easily seen that F is well-defined. Since $X_o \subseteq g[Y]$ and g is an injection, $F|_{X_o} \;=\; g^{-1}|_{X_o}$ is well defined. That F is well defined on X_e and X_i is obvious. Furthermore

$$F[X_e] \;=\; f[X_e] \;=\; Y_o$$

$$F[X_o] \;=\; g^{-1}[X_o] \;=\; Y_e$$

and

$$F[X_i] \;=\; f[X_i] \;=\; Y_i.$$

Discussion. *Although we had no choice in the definition of* F *on* X_e *and* X_o, *we could have defined* F *so that* $F|_{X_i} \;=\; g^{-1}|_{X_i}$.

Therefore,

$$F[X] \;=\; F[X_e \cup X_o \cup X_i] \;=\; f[X_e] \cup g^{-1}[X_o] \cup f[X_i] \;=\; Y_o \cup Y_e \cup Y_i \;=\; Y.$$

So F is a surjection. We show that F is an injection. Let $x, z \in X$, and suppose $F(x) = F(z)$. If $x \in X_e$, then $F(x) \in Y_o$ and $z \in X_e$. Hence

$$F(x) = f(x) = f(z) = F(z).$$

Since f is an injection, so is $f|_{X_e}$. Therefore $x = z$.

If $x \in X_o$, then $F(x) \in Y_e$ and $z \in X_o$. So

$$F(x) = g^{-1}(x) = g^{-1}(z) = F(z).$$

The function g is an injection, therefore $g^{-1}|_{X_o}$ is an injection and so $x = z$.

Finally, if $x \in X_i$, then $F(x) \in X_i$ and $z \in X_i$. So

$$F(x) = f(x) = f(z) = F(z).$$

Since f is an injection, $x = z$.

Therefore F is an injection. Hence,

$$F : X \rightarrowtail Y$$

and

$$| X | = | Y | .$$

□

THEOREM 6.5. $\mathbb{N}$ *is an infinite set.*

DISCUSSION. *We show that any function with domain* $\ulcorner n \urcorner$*, for* $n \in \mathbb{N}$*, fails to be a surjection. Therefore* $\mathbb{N}$ *is not finite.*

PROOF. Assume $n \in \mathbb{N}$, and

$$f : \ulcorner n \urcorner \longrightarrow \mathbb{N}.$$

Let

$$a = 1 + \sum_{i=0}^{n-1} f(i) \in \mathbb{N}.$$

Clearly $a \notin f[\ulcorner n \urcorner]$, so f is not a surjection. Consequently, there is no $n \in \mathbb{N}$ which can be mapped surjectively onto $\mathbb{N}$. Therefore $\mathbb{N}$ is not finite. □

Not only is $\mathbb{N}$ an infinite set, it is in some sense the "smallest" infinite set.

THEOREM 6.6. *If X is infinite, then $\mathbb{N} \preceq X$.*

DISCUSSION. *We shall define an injection $f : \mathbb{N} \to X$ inductively, building it up one step at a time.*

PROOF. As X is infinite, it is non-empty, so must contain some element x_0. Define $f(0) = x_0$.

Now, suppose that $x_0 = f(0), x_1 = f(1), \ldots, x_n = f(n)$ have all been chosen, so that

$$f|_{\{0,1,\ldots,n\}} \;=\; f|_{\ulcorner n+1 \urcorner} \;:\; k \mapsto x_k$$

is injective. As X is infinite, the function $f|_{\ulcorner n+1 \urcorner}$ that we have defined cannot be surjective. So there exists some x_{n+1} in $X \setminus \{x_0, \ldots, x_n\}$. Define $f(n+1) = x_{n+1}$. Continuing in this way, we attain an injection f defined on all of $\mathbb{N}$. □

REMARK. The astute reader may have noticed that in the previous proof, we end up making an infinite number of choices of elements of X.

DEFINITION. Cardinality, $\aleph_0$ We use the expression $\aleph_0$ (read "aleph nought"[1]) for the size of $\mathbb{N}$. That is

$$\aleph_0 := |\, \mathbb{N} \,| \,.$$

The size of a set is called the cardinality of the set. Any set which is bijective with $\mathbb{N}$ has cardinality $\aleph_0$. A finite set has cardinality equal to the unique natural number with which it is bijective.

DEFINITION. Countable A set that is finite or has cardinality $\aleph_0$ is called a countable set.

We are not formally developing the idea of cardinality. This would require working with ordinals, which would distract us from more immediate mathematical interests. However we shall use the language

[1] $\aleph$ is the first letter of the Hebrew alphabet.

and conventions of cardinals where it is intuitive and does not interfere with our program.

6.3. Uncountable Sets

In this section we prove one of the most remarkable results of modern mathematics. There are sets which are not countable. When this result was first communicated in 1878 by Georg Cantor, it astonished the mathematical world. It follows from this result that, in a most meaningful sense, there are different sizes of infinity. Suppose X is not a countable set. By Theorem 6.6, $\mathbb{N} \preceq X$. By the Schröder-Bernstein Theorem, if $X \preceq \mathbb{N}$, then $|\,\mathbb{N}\,| = |\,X\,|$, and X would be countable. So if X is not countable, $X \not\preceq \mathbb{N}$, and $\mathbb{N} \prec X$. That is,

$$\aleph_0 < |\,X\,|\,.$$

DEFINITION. Uncountable A set that is not countable is called uncountable.

Of course, we have yet to show that there are any uncountable sets.

NOTATION. $\prec$ *Let X and Y be sets. Then $X \prec Y$ provided that*

$$X \preceq Y$$

and

$$|\,X\,| \neq |\,Y\,|\,.$$

We write $|\,X\,| \leq |\,Y\,|$ if $X \preceq Y$, and $|\,X\,| < |\,Y\,|$ if $X \prec Y$.

DEFINITION. Power set, $P(X)$ Let X be a set. Then

$$P(X) \;=\; \{Y \mid Y \subseteq X\}.$$

$P(X)$ is called the power set of X. It is the set of all subsets of X.

The next theorem, due to G. Cantor, is one of the most remarkable results of mathematics. It not only proves the existence of an uncountable set, it implies that the power set necessarily generates sets of larger cardinality and thereby provides a means of constructing infinitely many different infinite cardinalities.

THEOREM 6.7. *Let X be a set. Then*

$$| X | < | P(X) | .$$

DISCUSSION. *To prove this result we need to show that a bijection between a set and its power set is impossible. How does one show the impossibility of such a function? We can assume that such a bijection exists and derive a contradiction. Alternatively, we can show that any function from a set to its power set necessarily fails to be a surjection — which is nearly the same thing, and more elegant. We need to show that any function from a set to its power set "misses" some elements of the power set. We shall use a technique known as a diagonal argument to construct an element of the power set that is not in the range of the function. The domain, X, acts like an index to keep track of the elements of the range of the function (this is another use for functions). We construct an element $Y \in P(X)$ not in the range of the function by adding $x \in X$ to Y iff x is not in the element of $P(X)$ indexed by x (that is, x is not in the image of x under the function). It is easy to show that this subset Y of X is not in the range of the function, and the function therefore fails to be a surjection onto $P(X)$.*

PROOF. We observe that the function $g : X \to P(X)$ defined by

$$g(x) \ = \ \{x\}$$

is an injection. In the case the $X \ = \ \emptyset$, g is the empty function — that is, the function whose graph is the $\emptyset$. (You should check that the empty function is an injection.) Therefore

$$| X | \leq | P(X) | .$$

Let

$$f : X \to P(X).$$

We define

$$Y : = \ \{x \in X \mid x \notin f(x)\}.$$

DISCUSSION. *Recall that for every $x \in X$, $f(x)$ is a subset of X. Thus it makes sense to consider whether x is an element of $f(x)$. The*

self-referential flavor of this argument makes it challenging on the first reading!

Clearly,

$$Y \subseteq X.$$

Is $Y \in f[X]$? Suppose it were, so $Y = f(x_0)$ for some x_0 in X. But then, x_0 would be in Y iff x_0 were not in $f(x_0) = Y$. This is impossible, contradicting the assumption that Y is in $f[X]$. □

You might try to repair f by modifying it to include in its range the diagonal set we constructed. Applying the diagonal argument again will identify a new element missing from the range of the modified function. In fact most elements of the codomain are missing from the range of the function, although this is not immediately obvious from the proof.

You still might be confused by why this is called a diagonal argument. This will be obvious when we apply the technique to infinite binary sequences.

If X is finite the theorem is obvious. Indeed, if there is $n \in \mathbb{N}$ such that $\mid X \mid = n$, then $\mid P(X) \mid = 2^n$. Theorem 6.7 implies that any set, including an infinite set, is strictly smaller than its power set. In fact, by iterating the applications of the power set function to $\mathbb{N}$, it is easily seen that there are infinitely many infinite sets of distinct cardinality in the sequence of sets

$$\langle \mathbb{N}, P(\mathbb{N}), P(P(\mathbb{N})), \ldots \rangle.$$

(What is the cardinality of the union over this sequence? What is the cardinality of the power set of that union?)

We shall prove that two more sets are uncountable. Both of these are sets of mathematical interest. We shall show first that infinite binary sequences are uncountable. Infinite binary sequences are functions from $\mathbb{N}$ to $\ulcorner 2 \urcorner$. As we shall see, there is a very close relationship

between infinite binary sequences and the power set of $\mathbb{N}$. More generally, the collection of all functions from one set to another can be of mathematical interest. We introduce a notation for such collections.

NOTATION. Y^X *Let X and Y be sets. The set of all functions with domain X and codomain Y is written Y^X.*

Do not confuse this with exponentiation. However if X and Y are finite,

$$| Y^X | = | Y |^{|X|} .$$

The set of all functions from some set X into $\ulcorner 2 \urcorner$ is in bijective correspondence with $P(X)$:

PROPOSITION 6.8. *Let X be a set, and define $F : \ulcorner 2 \urcorner^X \to P(X)$ by: for $\chi \in \ulcorner 2 \urcorner^X$,*

$$F(\chi) \;=\; \chi^{-1}(1).$$

That is $F(\chi) = \{x \in X \mid \chi(x) = 1\}$. Then $F : \ulcorner 2 \urcorner^X \to P(X)$ is a bijection.

PROOF. The proof is left as an exercise. □

The existence of this bijection allows us to easily prove the following theorem.

THEOREM 6.9. *The set of infinite binary sequences is bijective with $P(\mathbb{N})$ and is therefore uncountable.*

PROOF. By Proposition 6.8

$$| \ulcorner 2 \urcorner^{\mathbb{N}} | = | P(\mathbb{N}) | .$$

By Theorem 6.7,

$$| \mathbb{N} | \prec | P(\mathbb{N}) | .$$

Therefore $\ulcorner 2 \urcorner^{\mathbb{N}}$ is uncountable. □

NOTATION. $2^{\aleph_0}$ *We use $2^{\aleph_0}$ for the cardinality of $\ulcorner 2 \urcorner^{\mathbb{N}}$.*

It is worth illustrating by an application to infinite binary sequences why the technique used to prove Theorem 6.7 is called a diagonal argument (sometimes called the second diagonal argument, to distinguish from the "first diagonal argument" in Section 6.4). Let $f : \mathbb{N} \to \ulcorner 2^{\urcorner \mathbb{N}}$. We prove that f is not a surjection by direct application of the diagonal argument in the proof of Theorem 6.7. We enumerate all the elements in the range of f; each one is a sequence of 0's and 1's.

$$\begin{array}{lcccccccc} f(0) & = & a_{00} & a_{01} & a_{02} & a_{03} & \ldots & a_{0j} & \ldots \\ f(1) & = & a_{10} & a_{11} & a_{12} & a_{13} & \ldots & a_{1j} & \ldots \\ f(2) & = & a_{20} & a_{21} & a_{22} & a_{23} & \ldots & a_{2j} & \ldots \\ f(3) & = & a_{30} & a_{31} & a_{32} & a_{33} & \ldots & a_{3j} & \ldots \end{array}$$

We now create a sequence by altering the diagonal elements of this infinite array. Let s be the sequence of diagonal elements

$$\langle 1 - a_{00}, 1 - a_{11}, \ldots, 1 - a_{ii}, \ldots \rangle.$$

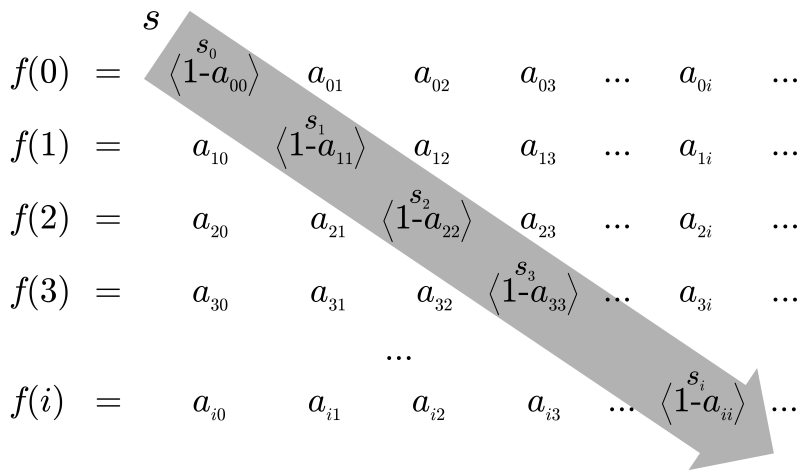

FIGURE 6.10. The second diagonal argument

The sequence s is an element of $\ulcorner 2^{\urcorner \mathbb{N}}$, and differs from every element in the range of f: indeed, s differs from $f(i)$ in at least the i^{th} slot. Hence,

$$s \notin f[\mathbb{N}],$$

and so f is not a surjection. We leave it as an exercise to show that s is the diagonal set Y constructed in the proof of Theorem 6.7, where $X = \mathbb{N}$. More precisely, s is the image of Y under the natural bijection from $P(\mathbb{N})$ to $\ulcorner 2^{\urcorner\mathbb{N}}$ of Proposition 6.8. □

We consider another set of mathematical interest, the set of all infinite decimal sequences. This set has a close relationship with the closed interval $[0, 1]$. Understanding this relationship requires a deeper, more formal understanding of the real numbers than most students have been exposed to in calculus, and we postpone the detailed discussion of this relationship until Section 8.9. With some modifications, the following theorem will prove that $[0, 1]$ is uncountable, and therefore $\mathbb{R}$ is uncountable (see Section 8.9).

THEOREM 6.11. *The set of infinite decimal expansions is uncountable. In fact,*

$$| \ulcorner 10^{\urcorner\mathbb{N}} | = 2^{\aleph_0}.$$

DISCUSSION. *The identity function on the infinite binary sequences into the infinite decimal sequences is clearly an injection. We shall construct an injection from the infinite decimal sequences to infinite binary sequences. The theorem will follow from the Schröder-Bernstein Theorem.*

PROOF. It is obvious that

$$| \ulcorner 2^{\urcorner\mathbb{N}} | \leq | \ulcorner 10^{\urcorner\mathbb{N}} | . \tag{6.12}$$

(Why?) We shall define an injection

$$f : \ulcorner 10^{\urcorner\mathbb{N}} \to \ulcorner 2^{\urcorner\mathbb{N}}.$$

Let $x \in \ulcorner 10^{\urcorner\mathbb{N}}$. So

$$x = \langle x_j \mid j \in \mathbb{N} \rangle$$

where x_j is the j^{th} member of the sequence x and

$$0 \leq x_j \leq 9.$$

We want to define a binary sequence $s(x)$ that "encodes" x. There are many ways to do it. One is to look at blocks of 10 bits (short for "binary digits"), and, in the j^{th} such block, have nine of the bits 0, and put a 1 in the x_j^{th} slot. Formally, given an infinite decimal sequence x, we define a binary sequence

$$f(x) \;=\; \langle y_i \mid i \in \mathbb{N} \rangle$$

so that $y_i \;=\; 1$ if there is $j \in \mathbb{N}$ such that

$$i \;=\; 10j + x_j.$$

Otherwise $y_i \;=\; 0$. We thereby define a function

$$f : \ulcorner 10 \urcorner \to \ulcorner 2 \urcorner.$$

We show that f is an injection. Let x and y be distinct elements of $\ulcorner 10 \urcorner^{\mathbb{N}}$. Then there is some $j \in \mathbb{N}$ such that

$$x_j \neq y_j.$$

Then

$$10j + x_j \neq 10j + y_j.$$

Let $i \;=\; 10j + x_j$. Then $f(x)$ and $f(y)$ differ in the i^{th} component. That is,

$$(f(x))_i \;=\; 1 \neq 0 \;=\; (f(y))_i.$$

Therefore f is an injection and

$$| \ulcorner 10 \urcorner^{\mathbb{N}} | \preceq | \ulcorner 2 \urcorner^{\mathbb{N}} | \,. \tag{6.13}$$

By the Schröder-Bernstein Theorem, (6.12) and (6.13) yield

$$| \ulcorner 10 \urcorner^{\mathbb{N}} | \;=\; 2^{\aleph_0}.$$

□

We prove in Section 8.9 that

$$| [0,1] | = | \ulcorner 10 \urcorner^{\mathbb{N}} |,$$

essentially by identifying a real number with its decimal expansion. If we assume this result, we can easily prove that the real numbers are uncountable.

COROLLARY 6.14. $\mathbb{R}$ *is uncountable.*

6.4. Countable Sets

The uncountable sets we have identified so far have a certain structural characteristic in common. We have shown that the set of all functions from a fixed infinite domain to a fixed codomain of at least two elements is uncountable. Cantor's theorem that the power set of an infinite countable set is uncountable can be interpreted this way as well. If X is a set, then $P(X)$ can be understood as $\ulcorner 2 \urcorner^X$, the set of all functions from X to $\ulcorner 2 \urcorner$. In the case of finite sets, X and Y, the set of all functions from X to Y, Y^X, has cardinality $\mid Y \mid^{\mid X \mid}$. That is, the cardinality of

$$\{f \subseteq X \times Y \mid f \text{ is a function}\}$$

is an exponential function of $\mid X \mid$. Of course, exponential functions grow relatively fast. For finite sets, the cardinality of the union of disjoint sets is the sum of the cardinalities of the sets. The cardinality of the direct product of two finite sets is the product of the cardinalities. What happens to the union or the direct product of countable infinite sets? Can the set operations of union and direct product generate uncountable sets from countable sets? We answer the questions for unions (addition) first.

The following proposition will simplify some of the technical details in the arguments which follow.

PROPOSITION 6.15. *Let X and Y be sets. Then there is a surjection $f : X \to Y$ iff $\mid Y \mid \leq \mid X \mid$.*

DISCUSSION. *We shall use the level sets of the surjection f to define the injection from Y to X. This uses the machinery of equivalence relations developed in Chapter 2 with the Axiom of Choice.*

PROOF. ($\Rightarrow$)
Let X, Y and f be as in the statement of the proposition. Let

$$\widehat{f} : X/f \to Y$$

be the canonical bijection associated with f that was defined in Section 2.3. We ask whether there is an injection $g : X/f \to X$ where $g([x]) \in [x]$. Recall that X/f is the collection of level subsets of X, with respect to f, and is a partition of X. Why not simply choose an element from each equivalence class and define g to be the function from X/f to X defined by these choices?

DISCUSSION. *The Axiom of Choice is the assertion that such "choice" functions exist.*

The function g is clearly an injection, so

$$g \circ \widehat{f}^{-1} : Y \to X$$

is an injection. Therefore if there is a surjection $f : X \to Y$, then $| Y | \leq | X |$.
($\Leftarrow$)
The proof of this implication is left as an exercise. □

THEOREM 6.16. *Cantor's Theorem Let $\{X_n \mid n \in \mathbb{N}\}$ be a family of sets such that X_n is countable for all $n \in \mathbb{N}$, and $X = \bigcup_{n\in\mathbb{N}} X_n$. Then*

$$| X | \leq \aleph_0.$$

DISCUSSION. *This Theorem, also due to G. Cantor, is the key result for proving that sets are countable. It is proved by a technique also called a diagonal argument (sometimes called the first diagonal argument). We use the index set $\mathbb{N}$ to construct an infinite array, and use that array to illustrate an enumeration of the union. This enumeration is a surjection from $\mathbb{N}$ to X.*

PROOF. For $n \in \mathbb{N}$, X_n is countable and by Proposition 6.15 there is a surjection

$$f_n : \mathbb{N} \to X_n.$$

Use the functions f_n to construct an infinite array. The 0^{th} column will contain all the elements of X_0, in the order $f_0(0), f_0(1), f_0(2), \ldots$. (It does not matter if the same element is listed multiple times). The next

column has the elements of X_1 in the order $f_1(0), f_1(1), f_1(2)$, *etc.* We define a function $g : \mathbb{N} \to X$ by traversing this array along the northeast to southwest diagonals, *viz.* $g(0) = f_0(0), g(1) = f_1(0), g(2) = f_0(1), g(3) = f_2(0), g(4) = f_1(1), g(5) = f_0(2), g(6) = f_3(0)$, and so on.

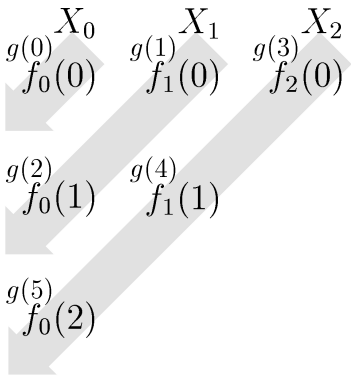

FIGURE 6.17. The first diagonal argument

Then g is a surjection, because every element of $\bigcup X_n$ occurs in the array, and is therefore in the range of g. By Proposition 6.15,

$$|X| \leq \aleph_0.$$

□

COROLLARY 6.18. *Let A be a countable set and $\{X_\alpha \mid \alpha \in A\}$ be a family of countable sets indexed by A. Then*

$$\left| \bigcup_{\alpha \in A} X_\alpha \right| \leq \aleph_0.$$

PROOF. Since A is countable, there is a surjection

$$f : \mathbb{N} \to A.$$

Therefore

$$\bigcup_{\alpha \in A} X_\alpha = \bigcup_{n \in \mathbb{N}} X_{f(n)}.$$

By Cantor's Theorem 6.16,

$$\left| \bigcup_{\alpha \in A} X_\alpha \right| \leq \aleph_0.$$

□

COROLLARY 6.19. *$\mathbb{Z}$ is countable.*

DISCUSSION. *Without too much effort, we could define a bijection from $\mathbb{N}$ to $\mathbb{Z}$. Instead we shall prove the existence of the bijection without explicitly defining a bijection.*

PROOF. Let $f : \mathbb{N} \to \mathbb{Z}$ be such that

$$f(n) \ = \ -n.$$

Then $f[\mathbb{N}]$ is countable. By Cantor's Theorem

$$\mathbb{Z} \ = \ \mathbb{N} \cup f[\mathbb{N}]$$

is countable. □

We turn our attention to direct products.

THEOREM 6.20. *If $n \in \mathbb{N}$, and $X_1, X_2, \ldots, X_n$ are countable sets, then*

$$X_1 \times X_2 \times \cdots X_n$$

is countable.

PROOF. We assume that all of the factors, $X_1, \ldots, X_n$ are non-empty. We argue by induction on the number of factors.
Base case: $n \ = \ 2$.

$$X_1 \times X_2 \ = \ \bigcup_{x \in X_2} X_1 \times \{x\}.$$

For each $x \in X_2$,

$$| \, X_1 \, | = | \, X_1 \times \{x\} \, | \, .$$

By Corollary 6.18, $X_1 \times X_2$ is countable.
Induction step:

Assume that for any collection of n countable sets $X_1, \dots X_n$, the product $X_1 \times \cdots \times X_n$ is countable. Let $X_1, \dots, X_{n+1}$ be countable nonempty sets. Then

$$X_1 \times \cdots \times X_{n+1} \;=\; (X_1 \times \cdots \times X_n) \times X_{n+1}.$$

By the induction hypothesis, $X_1 \times \cdots \times X_n$ is countable, and by the base case the direct product of two countable sets is countable. Therefore, $X_1 \times \cdots \times X_{n+1}$ is countable. □

COROLLARY 6.21. *$\mathbb{Q}$ is countable.*

PROOF. Let $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Q}$ be defined by

$$f(a,b) = \begin{cases} a/b & \text{if} \quad b \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Then f is a surjection, and by Proposition 6.15, $\mathbb{Q}$ is countable. □

We have evaluated the nested sequence of sets,

$$\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}.$$

These are important mathematical sets and, with the exception of $\mathbb{R}$, they are countable. We investigate the cardinality of one more set between $\mathbb{Q}$ and $\mathbb{R}$.

DEFINITION. Algebraic real number, $\mathbb{K}$ A real number α is algebraic if there is a polynomial p (not identically 0) with integer coefficients such that $p(\alpha) \;=\; 0$. We shall denote the set of all algebraic numbers by $\mathbb{K}$.

Any rational number $a/b \in \mathbb{Q}$ is algebraic, since a/b is a root of the polynomial

$$p(x) \;=\; bx - a.$$

Moreover, in Example 3.23, we showed that $\sqrt{2}$ is irrational, and it is clearly algebraic, since it is a root of $x^2 - 2$. Therefore we have

$$\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{K} \subseteq \mathbb{R}.$$

Finally we prove that $\mathbb{K} \neq \mathbb{R}$ by showing that $\mathbb{K}$ is countable.

THEOREM 6.22. $\mathbb{K}$ *is countable.*

DISCUSSION. *This result is proved by showing that the algebraic real numbers can be constructed by a countable procedure. That is,* $\mathbb{K}$ *may be built by adding to* $\mathbb{Q}$ *countably many elements at a time countably many times. Cantor's Theorem implies that any set so constructed will be countable.*

PROOF. Let $n \in \mathbb{N}$ and define $f : \prod_{i\,=\,0}^{n} \mathbb{Z} \to \mathbb{Z}[x]$ by

$$f(a_0, \ldots, a_n) \;=\; \sum_{i\,=\,0}^{n} a_i x^i.$$

By Corollary 6.19, $\mathbb{Z}$ is countable. By Theorem 6.20, $\prod_{i\,=\,0}^{n} \mathbb{Z}$ is countable. The range of f is the set of polynomials with integer coefficients with degree $\leq n$ (or the polynomial identically equal to 0). By Proposition 6.15, the range of a function with a countable domain is countable as well. Therefore the set of polynomials of degree $\leq n$ is countable.

Let P_n be the set of polynomials with integer coefficients of degree $\leq n$. Then

$$\mathbb{Z}[x] \;=\; \bigcup_{i=0}^{\infty} P_n.$$

By Theorem 6.16, $\mathbb{Z}[x]$ is countable. By Theorem 4.10, if $p(x)$ is a polynomial with real coefficients of degree n, it has at most n real roots. Let

$$Z_p \;=\; \{\alpha \mid p(\alpha) \;=\; 0\}.$$

So Z_p is finite for every polynomial p. Applying Cantor's Theorem (Theorem 6.16) again,

$$\mathbb{K} \;=\; \bigcup_{p \in \mathbb{Z}[x]} Z_p$$

is countable. □

COROLLARY 6.23. $\mathbb{K} \neq \mathbb{R}$

Since $\mathbb{K}$ is countable and $\mathbb{R}$ is uncountable, $\mathbb{K}$ is a proper subset of $\mathbb{R}$.

DEFINITION. Transcendental number A real number that is not algebraic is called a transcendental number.

Corollary 6.23 states that there are transcendental numbers. This is an existence claim in which no witness to the claim is produced. Rather it is an example of a counting argument (on infinite sets). There are too many real numbers for them all to be algebraic. By the end of the nineteenth century it was proved that π and e are transcendental, but these proofs are much more complicated than Cantor's existence proof above, which is, in essence, a very clever application of the pigeon-hole principle.

COROLLARY 6.24. *There are uncountably many transcendental numbers.*

PROOF. Let T be the set of transcendental numbers. As

$$| \mathbb{R} | = | T \cup \mathbb{K} | > \aleph_0,$$

and $\mathbb{K}$ is countable, T must be uncountable. □

So we have shown that

$$\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{K} \subsetneq \mathbb{R}.$$

However,

$$| \mathbb{N} | = | \mathbb{Z} | = | \mathbb{Q} | = | \mathbb{K} | < | \mathbb{R} | .$$

6.5. Functions and Computability

In Section 1.3 we made the off-hand comment that most functions are not defined by rules (by which we meant instructions for computing the function). We consider a rule to be an instruction (in some language) of finite length. Functions that are unambiguously defined by a rule of finite length are called computable, or recursive functions. Naturally there is a complicated mathematical definition of recursive functions, but we shall dispense with the formalities and say that a function is recursive, or computable, if there is an instruction (of finite

length) for finding the image of any element in the domain. How many computable functions are there?

We shall restrict our investigation to functions from $\mathbb{N}$ to $\mathbb{N}$. We consider functions as graphs of functions. That is, every subset of $P(\mathbb{N} \times \mathbb{N})$ that satisfies the definition of a function is a function in $\mathbb{N}^{\mathbb{N}}$. Are all such functions computable? It is obvious that

$$2^{\aleph_0} \leq | \mathbb{N}^{\mathbb{N}} | .$$

(Why?) In fact you can show that the sets are bijective. So there are uncountably many functions in $\mathbb{N}^{\mathbb{N}}$. How many instructions for computing functions are there? An instruction is a finite string, or sequence, of symbols. For instance, an instruction for the function that squares natural numbers is

$$f(x) \ = \ x^2.$$

This is a finite sequence of seven symbols. The instruction gives enough information to compute the image of any natural number. There are many other rules for computing this function. For instance the rule

$$f(x) \ = \ x \cdot x$$

obviously defines the same function, but the instruction is different — it contains one more symbol. Consider the set of all possible instructions for computing functions of natural numbers. How are the instructions formulated? One produces a finite sequence of symbols that forms an explicit guide for computing the image of any natural number.

Let X be the set of all symbols appearing in instructions for computing functions of natural numbers. The set X will include letters, digits, symbols for operations, symbols for relations and potentially any other symbol that you might see in a book on mathematics. How large is X? If you require that every symbol appear in some actual dictionary, it would clearly be finite. You will probably wish to allow any natural number to appear in the instruction. However, although there are infinitely many natural numbers, we need only ten symbols to name them all. It seems that we can reasonably require that X is

finite, but as it turns out, we can allow for X to be countably infinite without changing our conclusion.

If there is *any* language with countably many symbols in which the set of all instructions for computing functions could be written, then we may assume that X is countable. If F is an instruction or rule (and hence a finite sequence of symbols from X), then there is $N \in \mathbb{N}$ such that

$$F \in X^N.$$

So it is easily seen that the set of all possible instructions for elements of $\mathbb{N}^{\mathbb{N}}$, I, satisfies

$$I \preceq \bigcup_{N \in \mathbb{N}} X^N.$$

For $N \in \mathbb{N}$, X^N is the direct product of N factors of X, and by Theorem 6.20,

$$| X^N | \leq \aleph_0.$$

The set $\bigcup_{N \in \mathbb{N}} X^N$ is the countable union of countable sets, and by Theorem 6.16 is countable. Therefore there are uncountably many functions of natural numbers that are not defined by rules.

For a more thorough treatment of set theory, see the book [**5**] by Yiannis Moschovakis.

6.6. Exercises

EXERCISE 6.1. Let $f : X \rightarrowtail Y$ and $g : Y \rightarrowtail Z$. Prove that

$$g \circ f : X \to Z$$

is a bijection.

EXERCISE 6.2. Prove that equinumerosity is an equivalence relation.

EXERCISE 6.3. Prove that the relation on sets $\preceq$ is reflexive and transitive.

EXERCISE 6.4. In the proof of the Schröder-Bernstein Theorem, define a function

$$G(x) \;=\; \begin{cases} g^{-1}(x) & \text{if} \quad x \in X_i \\ f(x) & \text{if} \quad x \in X_e \\ g^{-1}(x) & \text{if} \quad X_o. \end{cases}$$

Prove that $G : X \rightarrowtail Y$.

EXERCISE 6.5. Let $n \in \mathbb{N}$. Prove that

$$|P(\ulcorner n \urcorner)| \;=\; 2^n.$$

EXERCISE 6.6. Let $X = \{0, 1, 2\}$. Write down some function $f : X \to P(X)$. For this particular f, what is the set Y of Theorem 6.7?

EXERCISE 6.7. Let X be a set and define a sequence of sets $\langle X_n \mid n \in \mathbb{N} \rangle$ by

$$X_0 \;=\; X$$

and

$$X_{n+1} \;=\; P(X_n).$$

Let $Y \;=\; \bigcup_{n\,=\,0}^{\infty} X_n$. Prove

$$(\forall n \in \mathbb{N}) \;\; | \, X_n \, | < | \, Y \, | \, .$$

EXERCISE 6.8. Let X and Y be finite sets. Prove that

$$| \, X^Y \, | = | \, X \, |^{|Y|} \, .$$

EXERCISE 6.9. Prove Proposition 6.8.

EXERCISE 6.10. Let $f : \mathbb{N} \to \ulcorner 2^{\urcorner \mathbb{N}}$ and for $i, j \in N^+$

$$a_{ij} \;=\; (f(i))_j.$$

(That is, a_{ij} is the j^{th} term of the i^{th} sequence.) Let s be the "diagonal" sequence

$$s \;=\; \langle 1 - a_{nn} \mid n \in N^+ \rangle.$$

We know that $s \notin f[\mathbb{N}]$. If $F : \ulcorner 2^{\urcorner \mathbb{N}} \rightarrowtail P(\mathbb{N})$ is the bijection in Proposition 6.8, then $F \circ f : \mathbb{N} \to P(\mathbb{N})$. Prove that that $F(s)$ is the "diagonal" set of Theorem 6.7 (where $X \;=\; \mathbb{N}$, and $F \circ f$ is the enumeration of subsets of $\mathbb{N}$), and hence that $F(s) \notin (F \circ f)[\mathbb{N}]$.

EXERCISE 6.11. Prove that if $X \subseteq Y$ and X is uncountable, then Y is uncountable.

EXERCISE 6.12. Let X be an uncountable set, Y be a countable set and $f : X \to Y$. Prove that some element of Y has an uncountable pre-image.

EXERCISE 6.13. Complete the proof of Proposition 6.15.

EXERCISE 6.14. Define an explicit bijection from $\mathbb{N}$ to $\mathbb{Z}$.

EXERCISE 6.15. Prove that $| \mathbb{K} \setminus \mathbb{Q} | = \aleph_0$.

EXERCISE 6.16. Prove that

$$e = \sum_{n=0}^{\infty} \frac{1}{n!}$$

is irrational. (Hint: Argue by contradiction. Assume $e = \frac{p}{q}$ and multiply both sides by $q!$. Rearrange the equation to get an integer equal to an infinite sum of rational numbers that converges to a number in the open interval $(0, 1)$.)

Remark: This was also Exercise 3.32. Is it easier now?

EXERCISE 6.17. Suppose that $a, b, c, d \in \mathbb{R}$, $a < b$ and $c < d$. Prove
a) The open interval (a, b) is bijective with the open interval (c, d).
b) The closed interval $[a, b]$ is bijective with the closed interval $[c, d]$.
c) The open interval $(0, 1)$ is bijective with the closed interval $[0, 1]$.
d) The open interval (a, b) is bijective with the closed interval $[c, d]$.
e) $| [0, 1] | = | \mathbb{R} |$.

EXERCISE 6.18. Construct explicit bijections for each of the pairs of sets in Exercise 6.17.

EXERCISE 6.19. Let $f(x)$ be a non-zero polynomial with integer coefficients, and suppose $\alpha \in \mathbb{R}$ is transcendental. Prove that $f(\alpha)$ is transcendental.

EXERCISE 6.20. Let $F : \mathbb{K} \to \mathbb{R}$ be defined by: If $x \in \mathbb{K}$, $F(x)$ is the lowest degree of a polynomial with integer coefficients for which x is a root. Is F well-defined?

EXERCISE 6.21. Let $a \in \mathbb{R}$ be a root of a polynomial with rational coefficients. Prove that a is a root of a polynomial with integer coefficients, and is therefore an algebraic number.

EXERCISE 6.22. For each of the following sets, state and prove whether it is bijective with $\mathbb{N}$, $P(\mathbb{N})$ or is larger than $P(\mathbb{N})$ (with respect to the relation $\prec$):
a) The set of finite subsets of $\mathbb{N}$
b) The set of all permutations of finite sets of natural numbers
c) The set of finite sequences of natural numbers
d) The set of finite sequences of integers
e) The set of finite sequences of algebraic numbers
f) The set of finite sequences of real numbers
g) The set of infinite sequences of natural numbers
h) The set of infinite sequences of real numbers
i) Countable subsets of $\mathbb{R}$.
h) $\mathbb{N}^{\mathbb{R}}$
k) $\mathbb{R}^{\mathbb{R}}$.
You may use the fact that $| \mathbb{R} | = 2^{\aleph_0}$.

EXERCISE 6.23. Prove that $| \mathbb{R}^{\mathbb{R}} | \geq | P(\mathbb{R}) |$.

CHAPTER 7

Divisibility

In this chapter we investigate divisibility. It may seem peculiar that we would investigate a topic that you have studied since elementary school, but don't be fooled by the apparent simplicity of the subject. The study of divisibility of integers is part of number theory. Geometry and number theory are the oldest areas of mathematical study, and they are still active fields of mathematical research.

7.1. Fundamental Theorem of Arithmetic

DEFINITION. Divides, factor Let $a, b \in \mathbb{Z}$. We say that a divides b, or a is a factor of b, if

$$(\exists\, c \in \mathbb{Z})\ a \cdot c \ = \ b.$$

We write this as $a \mid b$. If a does not divide b we write $a \nmid b$.

Divisibility is the central idea of number theory. It is precisely because one integer need not be a factor of another integer, or a pair of integers may fail to have non-trivial common factors, that divisibility provides insight into the structure of integers. Put another way, consider the definition of divisibility applied to rational numbers — you will find that it does not provide any insight at all since a nonzero rational number is a factor of any other rational number. Furthermore, many of the properties of integers with regard to divisibility generalize to other interesting mathematical structures. You will see an example of this in Section 7.5.

DEFINITION. Prime number Let $p \in \mathbb{N}$. We say that p is a prime number if $p > 1$ and the only positive factors of p are p and 1.

DEFINITION. Relatively prime Let $a, b \in \mathbb{Z}$. We say that a and b are relatively prime if they have no common factor greater than 1.

DEFINITION. Integer combination Let $a, b, c \in \mathbb{Z}$. Then c is an integer combination of a and b if

$$(\exists\, m, n \in \mathbb{Z})\ c \ = \ ma + nb.$$

PROPOSITION 7.1. *Let $a, b \in \mathbb{Z}$. If a and b are relatively prime, then $a - b$ and b are relatively prime.*

DISCUSSION. *We shall prove the contrapositive by showing that any common factor of $a - b$ and b is also a factor of a.*

PROOF. Let $c > 1$ be a common factor of b and $a - b$. So

$$(\exists\, m \in \mathbb{Z})\ b \ = \ cm$$

and

$$(\exists\, n \in \mathbb{Z})\ a - b \ = \ cn.$$

Then

$$c(m + n) \ = \ a$$

and so $c \mid a$. Therefore if a and b are relatively prime, then $a - b$ and b are relatively prime. □

PROPOSITION 7.2. *Let a and b be integers. If a and b are relatively prime, then*

$$(\exists\, m, n \in \mathbb{Z})\ ma + nb \ = \ 1.$$

DISCUSSION. *We shall argue for the case in which a and b are natural numbers. Given the proposition for all pairs of relatively prime natural numbers, we may easily extend it to arbitrary pairs of relatively prime integers by changing the sign of m or n in the integer combination. This assumption allows us to argue by induction on the sum of the integers. The base case for this argument by induction will be $a + b \ = \ 3$. If $a \ = \ 0 \ = \ b$, then a and b are not relatively prime. If $a + b \ = \ 1$, then a and b are relatively prime and the choice of m*

and n is obvious. If $a = b = 1$ then a and b are relatively prime and again the choice of m and n is obvious.

PROOF. We may assume that $a > b > 0$. We argue by induction on $a + b$.

Base case: $a + b = 3$.

Then $a = 2$ and $b = 1$. So

$$a - b = 1.$$

Induction step:

Assume that the result holds for all pairs of relatively prime natural numbers with sum less than $a + b$.

By Proposition 7.1, b and $a - b$ are relatively prime. By the induction hypothesis, there are $i, j \in \mathbb{Z}$ such that

$$i(a - b) + jb = 1.$$

DISCUSSION. *If $a - b = b$, we are not in the case where we have two distinct positive numbers. How do we handle this possibility?*

Let $m = i$ and $n = j - i$. Then

$$ma + nb = 1.$$

By the induction principle the result holds for all relatively prime pairs of natural numbers. □

DEFINITION. Greatest common divisor, $\gcd(a, b)$ Let $a, b \in \mathbb{Z}$. The greatest common divisor of a and b, written $\gcd(a, b)$, is the largest integer which divides both a and b.

So a and b are relatively prime iff $\gcd(a, b) = 1$.

PROPOSITION 7.3. *Let $a, b, c \in \mathbb{Z}$, and assume that $\gcd(a, b) = 1$. If $a \mid cb$, then $a \mid c$.*

PROOF. By Proposition 7.2 there are $m, n \in \mathbb{Z}$ such that

$$ma + nb = 1.$$

Therefore

$$cma + cnb \;=\; c.$$

Clearly $a \mid cnb$ (since $a \mid cb$) and $a \mid cma$. So

$$a \mid (cma + cnb),$$

and therefore $a \mid c$. □

PROPOSITION 7.4. *Let $a, b, c \in \mathbb{Z}$. If $\gcd(a, b) \;=\; 1$, $a \mid c$ and $b \mid c$, then*

$$ab \mid c.$$

PROOF. Let $m, n \in \mathbb{Z}$ be such that $am \;=\; c$ and $bn \;=\; c$. Then

$$a \mid bn.$$

By Proposition 7.3, $a \mid n$. Hence there is $k \in \mathbb{Z}$ such that

$$ak \;=\; n.$$

Therefore

$$akb \;=\; c$$

and

$$ab \mid c.$$

□

LEMMA 7.5. *Assume*

(1) $p \in \mathbb{N}$ *is prime*
(2) $N \geq 1$, *and* $a_1, \ldots, a_N \in \mathbb{Z}$
(3) $p \mid (\prod_{n \;=\; 1}^{N} a_n)$.

Then there is some $n \leq N$ such that $p \mid a_n$.

PROOF. Let p be a prime number. We argue by induction on N.
Base case: $N \;=\; 1$
The base case is obvious.
Induction step:
Let $N > 1$ and assume that the result holds for all products of fewer

than N factors.
Let

$$a = \prod_{n=1}^{N-1} a_n$$

and suppose that

$$p \mid (\prod_{n=1}^{N} a_n).$$

Then

$$p \mid a \cdot a_N.$$

If $p \mid a$, then by the induction hypothesis,

$$(\exists\, n < N)\ p \mid a_n.$$

Assume that p is not a factor of a; since p is prime, $\gcd(p, a) = 1$. By Proposition 7.3, $p \mid a_N$. □

THEOREM 7.6. *Fundamental Theorem of Arithmetic Let N be a natural number greater than 1. Then N may be uniquely expressed as the product of prime numbers (up to the order of the factors).*

DISCUSSION. *We permit a "product" with only one factor. So any prime number is its own unique prime factoring.*

PROOF. We argue by induction on the natural numbers greater than 1.
Base case: $(N = 2)$
By the discussion preceding the proof, 2 is its own prime factoring.
Induction step:
Assume that the result holds for all natural numbers greater than 1 and less than N. If N is prime, the result follows. If N is not prime, then there are $a, b \in \mathbb{N}$, $a < N$ and $b < N$, such that

$$a \cdot b = N.$$

By the induction hypothesis, a and b have unique prime factorings. The product of the factorings will be a prime factoring of N. Is the

factoring unique up to order? Suppose that

$$N \;=\; \prod_{i\,=\,1}^{m} p_i \;=\; \prod_{j\,=\,1}^{n} q_j$$

where p_i is prime for $1 \leq i \leq m$, and q_j is prime for $1 \leq j \leq n$. Then

$$p_1 \mid \prod_{j\,=\,1}^{n} q_j.$$

By Lemma 7.5,

$$(\exists\, j \leq n)\; p_1 \mid q_j.$$

We may reorder the factors $q_1, \ldots, q_n$ so that $p_1 \mid q_1$. Both p_1 and q_1 are prime, so

$$p_1 \;=\; q_1.$$

Therefore

$$\prod_{i\,=\,2}^{m} p_i \;=\; \prod_{j\,=\,2}^{n} q_j < N.$$

By the induction hypothesis, $p_2, \ldots, p_m$ is a unique prime factoring of $\prod_{i\,=\,2}^{m} p_i$, so $m \;=\; n$ and $q_2, \ldots, q_n$ is a reordering of $p_2, \ldots, p_m$. Therefore $q_1 \cdots q_n$ is a reordering of $p_1 \cdots p_m$ and the prime factoring of N is unique. □

Remark. Why is the number 1 not defined to be a prime? After all, it has no factors other than itself or 1! The reason is because it is very useful to have uniqueness in the Fundamental Theorem of Arithmetic. If 1 were considered prime, it could be included arbitrarily often in the factoring of N.

7.2. The Division Algorithm

The Division Algorithm, Theorem 7.11, is the result that guarantees that long division of natural numbers will terminate in a unique quotient and remainder with the remainder strictly smaller than the divisor. Long division is difficult and tedious for young students. Typically it is the most challenging computation that elementary school

students are expected to master. You may have revisited the algorithm again when you learned to divide polynomials. Here the Division Algorithm says that the quotient and remainder are unique and the remainder is either identically 0 or has degree strictly smaller than the divisor. We frequently compare the arithmetic of integers and the arithmetic of polynomials, and it is the Division Algorithm that makes this comparison useful.

Let's extend the link between integer combinations and greatest common divisors. According to Lemma 7.2, a pair of integers are relatively prime if there is an integer combination of the pair which equals 1. This result generalizes to greatest common divisors other than 1.

THEOREM 7.7. *Let $a, b \in \mathbb{Z}$. The set of integer combinations of a and b equals the set of integer multiples of* $\gcd(a, b)$.

PROOF. Let $c = \gcd(a, b)$ and

$$M = \{kc \mid k \in \mathbb{Z}\}.$$

Since c is a divisor of a and b, there are $i, j \in \mathbb{Z}$ such that

$$a = ic$$

and

$$b = jc.$$

Let

$$I = \{ma + nb \mid m, n \in \mathbb{Z}\}.$$

We show first that $I \subseteq M$.
If $m, n \in \mathbb{Z}$, then

$$ma + nb = mic + njc = (mi + nj)c.$$

Hence every integer combination of a and b is a multiple of c and

$$I \subseteq M.$$

Now we show that $M \subseteq I$. Let $kc \in M$ and

$$r = \gcd(i, j).$$

Then there are $m, n \in \mathbb{Z}$ such that

$$rmc \;=\; ic \;=\; a \tag{7.14}$$

and

$$rnc \;=\; jc \;=\; b. \tag{7.15}$$

So $rc \mid a$ and $rc \mid b$. Hence

$$\gcd(a, b) \geq rc \geq c.$$

However $\gcd(a, b) \;=\; c$, and thus $r \;=\; 1$. Therefore i and j are relatively prime.

By Proposition 7.2, there is an integer combination of i and j that equals 1. Let $u, v \in \mathbb{Z}$ be such that

$$ui + vj \;=\; 1.$$

Then

$$c(ui + vj) \;=\; c$$

and

$$kc \;=\; kc(ui + vj) \;=\; k(ua + vb)$$

by equations 7.14 and 7.15. Hence

$$kc \in I,$$

and as k was arbitrary,

$$M \subseteq I.$$

□

COROLLARY 7.8. *Let $a, b \in \mathbb{Z}$. Then $\gcd(a, b)$ is the smallest positive integer combination of a and b.*

Theorem 7.7 tells us that the integer combinations of a and b are precisely the integer multiples of $\gcd(a, b)$ (which happens to be the smallest positive integer combination of a and b). We think of $\gcd(a, b)$ as "generating" through multiplication the set of integer combinations of a and b.

PROPOSITION 7.9. *Let $a, b, k \in \mathbb{Z}$. Then*

$$\gcd(a, b) \;=\; \gcd(a - kb, b).$$

PROOF. If $c \in \mathbb{Z}$, $c \mid a$ and $c \mid b$, then $c \mid a - kb$. Therefore

$$\gcd(a, b) \leq \gcd(a - kb, b). \tag{7.10}$$

Likewise, if $c \mid a - kb$ and $c \mid b$, then $c \mid a$, so we get the reverse inequality of (7.10), so the two sides are equal. □

THEOREM 7.11. *Division Algorithm* *Let $a, b \in \mathbb{Z}$, $b \neq 0$. Then there are unique $q, r \in \mathbb{Z}$ such that*

$$a \;=\; qb + r$$

where $0 \leq r <\mid b \mid$.

DISCUSSION. *In the Division Algorithm a is called the dividend, b the divisor, q the quotient, and r the remainder.*

PROOF. Let $a, b \in \mathbb{Z}$ and $b \neq 0$. Define $I \subseteq \mathbb{N}$ by

$$I \;=\; \{a - kb \mid k \in \mathbb{Z}\} \cap \mathbb{N}.$$

I has a smallest element, $a - qb$, for some $q \in \mathbb{Z}$.
Claim: $0 \leq a - qb < \mid b \mid$.
Proof of Claim. We argue by cases.
Case 1: $b > 0$
If $a - qb \geq b$ then

$$a - (q + 1)b \geq 0.$$

Hence

$$a - (q + 1)b \in I.$$

However $a - qb$ is minimal in I, so this is impossible. Therefore

$$a - qb <\mid b \mid .$$

Case 2: $b < 0$
If $a - qb \geq\mid b \mid$, then

$$a - qb > a - (q - 1)b \geq 0.$$

As in the first case

$$a - (q-1)b \in I.$$

This is impossible since by assumption $a - qb$ is minimal in I. Therefore

$$a - qb <\mid b \mid .$$

$\lhd$

Thus if

$$r := a - qb,$$

we have $a = qb + r$ and $0 \leq r <\mid b \mid$. It remains to show that the quotient and remainder are unique. Suppose

$$a = mb + r = nb + s$$

where $0 \leq r, s < |b|$. If $r = s$ then $mb = nb$ and $m = n$. So we assume that $r \neq s$. Without loss of generality we assume that $r < s$. Then,

$$0 \leq s - r = (m-n)b <\mid b \mid .$$

So $m - n = 0$ and $r = s$, a contradiction. $\square$

Of course, q and r could be found by long division — that is, one may subtract multiples of b until the remainder is less than $\mid b \mid$.

7.3. Euclidean Algorithm

How do we find $\gcd(a, b)$, for $a, b \in \mathbb{N}$? One might invoke the Fundamental Theorem of Arithmetic and compare the prime decompositions of a and b. Suppose

$$a = \prod_{n=1}^{N} p_n^{r_n}$$

and

$$b = \prod_{n=1}^{N} p_n^{s_n}$$

where $r_n, s_n \in \mathbb{N}$ for $1 \leq n \leq N$. If $t_n = \min(r_n, s_n)$ for $1 \leq n \leq N$, then

$$\gcd(a, b) = \prod_{n=1}^{N} p_n^{t_n}.$$

However finding the prime decomposition of an integer can be quite difficult. We shall define an operation on pairs of integers that after a reasonable number of applications will yield the greatest common divisor of the integers.

If $a, b \in \mathbb{N}$, $a > b > 0$, define $E : \mathbb{N}^2 \to \mathbb{N}^2$ by

$$E(a, b) \;=\; (b, r)$$

where r is the unique remainder (when dividing a by b) whose existence was proved in the Division Algorithm. That is, if

$$a \;=\; qb + r$$

with $0 \leq r < b$, then define

$$E(a, b) :=\; (b, r).$$

If $b \;=\; 0$, then let

$$E(a, 0) \;=\; (a, 0).$$

Let $(a, b) \in \mathbb{N}^2$, $a > b > 0$. We define a sequence of elements in $\mathbb{N}^2$, $\langle E_i(a, b) \mid i \in \mathbb{N} \rangle$, by recursion:

$$E_0(a, b) \;=\; (a, b)$$

and if $n > 0$

$$E_n(a, b) \;=\; E(E_{n-1}(a, b)).$$

So long as $E_n(a, b)$ has non-zero components, the sequence of second components is strictly decreasing, so it is clear that the sequence must eventually become fixed on an ordered pair (see Exercise 4.11). By the Division Algorithm, this will occur when the second component equals 0. Let k be the smallest integer such that

$$E_k(a, b) \;=\; E_{k+1}(a, b).$$

Then we say that $\langle E_n(a, b) \mid n \in \mathbb{N} \rangle$ stabilizes at step k. For $n \geq k$,

$$E_n(a, b) \;=\; E_{n+1}(a, b) \;=\; E_k(a, b).$$

If $\langle E_n(a, b) \rangle$ stabilizes at step k, it is obvious that $k \leq b$. Typically, the sequence stabilizes much faster than this.

THEOREM 7.12. *Let $a, b \in \mathbb{N}$, $a > b > 0$. The non-zero component on which the sequence*

$$\langle E_n(a,b) | n \in \mathbb{N} \rangle$$

stabilizes is $\gcd(a,b)$.

PROOF. Let a be fixed, we argue by induction on the smaller of the integers, b.

Base case: $b = 1$

Then for any $a > 1$,

$$E(a,1) = (1,0)$$

and the sequence $\langle E_n(a,1) \rangle$ stabilizes at step 1 with non-zero component 1.

Induction step:

Let $b > 1$. Assume the result holds for all $c < b$ — that is, for any $(a,c) \in \mathbb{R}^2$, where $c < b < a$, the non-zero component of the ordered pair at which the sequence $\langle E_n(a,c) \rangle$ stabilizes is $\gcd(a,c)$. We show that the non-zero component of the ordered pair at which the sequence $\langle E_n(a,b) \rangle$ stabilizes is $\gcd(a,b)$. If $a > b > 0$ then

$$E(a,b) = (b, a - qb)$$

where $0 \leq a - qb < b$. By the induction hypothesis, the non-zero component of the ordered pair at which the sequence $\langle E_n(b, a-qb) \mid n \in \mathbb{N} \rangle$ stabilizes is $\gcd(b, a-qb)$. By Proposition 7.9

$$\gcd(a,b) = \gcd(b, a-qb).$$

So the non-zero component of the ordered pair at which the sequence

$$\langle E_n(a,b) \mid n \in \mathbb{N} \rangle$$

stabilizes is $\gcd(a,b)$. By the induction principle, the result holds for all ordered pairs $(a,b) \in \mathbb{N}^2$ where $a > b > 0$. □

An algorithm is a set of executable computational instructions. The Euclidean algorithm is the following set of instructions:

Given a pair of natural numbers, $a > b > 0$, compute the sequence

$\langle E_n(a,b) \mid n \in \mathbb{N} \rangle$ until the sequence stabilizes. The non-zero component of the ordered pair on which the sequence stabilizes is $\gcd(a,b)$.

EXAMPLE 7.13. Let $a = 29712375$ and $b = 119119$. Find the $\gcd(a,b)$. We use the Euclidean Algorithm. So

$$E_0(a,b) = (a,b)$$

$$E_1(a,b) = E(a,b) = (b, 51744)$$

$$E_2(a,b) = E(b,51744) = (51744, 4851)$$

$$E_3(a,b) = E(51744,4851) = (4851, 1078)$$

$$E_4(a,b) = E(4851,1078) = (1078, 539)$$

$$E_5(a,b) = E(1078,539) = (539, 0).$$

Therefore $\gcd(a,b) = 539$. If you employ the Fundamental Theorem of Arithmetic, with some work you can determine that

$$29,712,375 = (3^2)(5^3)(7^4)(11)$$

and

$$119,119 = (7^2)(11)(13)(17).$$

So $\gcd(a,b) = (7^2)(11) = 539$.

7.4. Fermat's Little Theorem

NOTATION. $\mathbb{Z}_n^*$ *Let* $n \in \mathbb{N}$, $n \geq 2$. *Then*

$$\mathbb{Z}_n^* = \mathbb{Z}_n \setminus \{[0]\}.$$

LEMMA 7.14. *Let* $a, n \in \mathbb{Z}$, $n \geq 2$, *be such that* $\gcd(a,n) = 1$. *Define* $\phi_a : \mathbb{Z}_n^* \to \mathbb{Z}_n^*$ *by*

$$\phi_a([b]) = [ab].$$

Then ϕ_a *is a permutation of* $\mathbb{Z}_n^*$.

PROOF. We show that $[a], [2a], \ldots, [(n-1)a]$ are distinct elements of $\mathbb{Z}_n^*$. Let $0 < i \leq j < n$ and suppose that $ia \equiv ja \mod n$. Then

$$n \mid ja - ia$$

and

$$n \mid (j-i)a.$$

We assume that $\gcd(n, a) = 1$, so by Proposition 7.3, $n|(j-i)$. However $0 \leq j - i < n$, so $j - i = 0$ and $i = j$. Hence, if $0 < i < j < n$,

$$[ia] \neq [ja].$$

It follows that ϕ_a is an injection from $\mathbb{Z}_n^*$ to $\mathbb{Z}_n^*$. Any injection from a finite set to itself is a surjection, so ϕ_a is a permutation of $\mathbb{Z}_n^*$. □

DEFINITION. Order, $o_p(a)$ Let p be a prime number and $a \in \mathbb{Z}$ not a multiple of p. The order of a in $\mathbb{Z}_p$ is the least $k \in \mathbb{N}^+$ such that $a^k \equiv 1 \mod p$. We write the order of a in $\mathbb{Z}_p$ as $o_p(a)$.

If a is a multiple of p, then the order of a in $\mathbb{Z}_p$ is undefined, since $a \equiv 0 \mod p$, and for all $k \in \mathbb{N}^+$,

$$a^k \equiv 0 \mod p.$$

The following proposition shows in particular that if a is not a multiple of p, then the order is well-defined (*i.e.* that there is some k with $a^k \equiv 1 \mod p$).

PROPOSITION 7.15. *Let $a \in \mathbb{Z}$, and p be a prime number such that $p \nmid a$. Then $o_p(a) < p$.*

PROOF. Let p be a prime number and $a \in \mathbb{Z}$ be such that a is not a multiple of p. By Lemma 7.5, as $p \nmid a$, then $p \nmid a^n$, and therefore $[a^n] \in \mathbb{Z}_p^*$ for any $n \in \mathbb{N}$. Since $|\mathbb{Z}_p^*| = p - 1$, the finite sequence

$$\langle [a^n] \mid 1 \leq n \leq p \rangle$$

must have a repetition. Let $1 \leq n < k \leq p$ be such that

$$a^n \equiv a^k \mod p.$$

Then

$$p \mid a^k - a^n.$$

Hence

$$p \mid a^n(a^{k-n} - 1).$$

However $p \nmid a^n$ and thus by Proposition 7.3,

$$p \mid a^{k-n} - 1.$$

Thus

$$a^{k-n} \equiv 1 \mod p.$$

Therefore

$$o_p(a) \leq k - n < p.$$

□

PROPOSITION 7.16. *Let $a \in \mathbb{Z}$ and p be a prime number such that a is not a multiple of p. Then the remainder classes $[1], [a], [a^2], \ldots, [a^{o_p(a)-1}]$ in $\mathbb{Z}_p$ are distinct.*

PROOF. Exercise. □

NOTATION. $S_a(n)$ *Fix a prime p for the remainder of this section. Let a be an integer such that $p \nmid a$. Then for any positive natural number n, we let $S_a(n)$ denote the set of equivalence classes $\{[n \cdot a^k] \mid k \in \mathbb{N}\}$ in $\mathbb{Z}_p$. (Although $S_a(n)$ depends on the choice of p, we suppress this in the notation and assume that p is understood).*

LEMMA 7.17. *Let $a \in \mathbb{Z}$ be such that $p \nmid a$. If $n \in \mathbb{N}^+$ is not a multiple of p, then*

$$\mid S_a(n) \mid = \; o_p(a).$$

PROOF. By Proposition 7.15, $o_p(a) < p$. Let $k = o_p(a)$. By Proposition 7.16 the remainder classes $[1], [a], [a^2], \ldots, [a^{k-1}]$ are distinct. Let ϕ_n be defined as in Lemma 7.14. Then ϕ_n is a permutation of $\mathbb{Z}_p^*$. Therefore the remainder classes $[n], [na^2], \ldots, [na^{k-1}]$ are distinct. But

$$na^k \equiv n \mod p,$$

so

$$S_a(n) \;=\; \{[n], [na^2], \ldots, [na^{k-1}]\}.$$

(Why?) Therefore

$$|\, S_a(n) \,| \;=\; o_p(a).$$

□

LEMMA 7.18. *Let $a \in \mathbb{Z}$ be such that $p \nmid a$. Then for any $m, n \in \mathbb{N}^+$ which are not multiples of p, the sets $S_a(m)$ and $S_a(n)$ are either equal or disjoint.*

PROOF. Suppose $S_a(m) \cap S_a(n) \neq \emptyset$. Let $m, n \in \mathbb{N}$, $\gcd(m, p) \;=\; 1$, $\gcd(n, p) \;=\; 1$ and

$$[ma^i] \in S_a(n).$$

Then there is $j \in \mathbb{N}$ such that

$$[ma^i] \;=\; [na^j].$$

We may assume that $i < j$, since there are infinitely many $j \in \mathbb{N}^+$ that satisfy the equation. Then

$$[m] \;=\; [na^{j-i}].$$

So

$$[m] \in S_a(n).$$

Therefore if $S_a(m)$ and $S_a(n)$ are not disjoint, we have

$$S_a(m) \subseteq S_a(n).$$

By symmetry, we also have

$$S_a(n) \subseteq S_a(m),$$

and so either

$$S_a(m) \;=\; S_a(n)$$

or

$$S_a(m) \cap S_a(n) \;=\; \emptyset.$$

□

THEOREM 7.19. *Fermat's Little Theorem If $a \in \mathbb{Z}$ and p is a prime number such that $p \nmid a$, then*

$$a^{p-1} \equiv 1 \mod p.$$

PROOF. Let $k = o_p(a)$. We show that $k \mid (p-1)$. Let $n \in \mathbb{N}$, where n is not a multiple of p. By Lemma 7.17

$$\mid S_a(n) \mid = k.$$

By Lemma 7.18, the sets

$$\{S_a(n) \mid n \in \mathbb{N}^+, p \nmid n\}$$

partition $\mathbb{Z}_p^*$ into sets of cardinality k. Therefore k divides $\mid Z_p^* \mid$. Since $\mid Z_p^* \mid = p-1$, we have

$$k \mid (p-1).$$

It follows that there is $j \in \mathbb{N}$ such that

$$a^{p-1} \equiv (a^k)^j \equiv 1^j \equiv 1 \mod p.$$

□

COROLLARY 7.20. *If $a \in \mathbb{Z}$ and p is a prime number such that $p \nmid a$, then*

$$a^p \equiv a \mod p.$$

Fermat's Little Theorem is an important result in the theoretical study of prime numbers, and determining primality. How might the theorem be used? Consider the problem of deciding whether a particular natural number n is prime. In order to determine whether n is prime, you may invoke the Fundamental Theorem of Arithmetic, and begin checking all the prime numbers up to $\sqrt{n}$ to determine whether any are non-trivial factors of n. We needn't check primes greater than $\sqrt{n}$ since the existence of such a factor entails the existence of a factor less then $\sqrt{n}$, and by the Fundamental Theorem of Arithmetic, a prime factor less than $\sqrt{n}$. This may require checking many candidates — in addition to requiring that you *know* all of the prime numbers smaller than $\sqrt{n}$, or are willing to check factors that are not prime. For large n

this is a formidable challenge. Alternatively, you can seek $a \in \mathbb{Z}$ such that $[a^n] \neq [a]$ in $\mathbb{Z}_n$ in order to determine that n is not prime.

For instance, is 12,871 prime? We assume that you have access to a computer (doing these computations by hand can be tedious). One approach is to check for factors among the prime numbers less than $\sqrt{12,871}$, that is the thirty prime numbers less than 114. Alternatively, for $a \in \mathbb{Z}$, we can check whether

$$a^{12,871} \equiv a \mod 12,871.$$

If the answer is no, then 12,871 is not prime. We shall try $a = 2$:

$$2^{12,871} \equiv 5732 \mod 12,871.$$

Therefore 12,871 is not prime. If you were to check primes sequentially, you would have to check 18 primes before finding that 61 is the smallest prime that divides 12,871.

If $a^{12,871} \equiv a \mod 12,871$ for a given choice of a, then we can draw no conclusion. In fact there are non-prime numbers, n, such that for any choice of a,

$$a^n \equiv a \mod n.$$

Numbers that satisfy the conclusion of Theorem 7.19, but are not prime are called Carmichael numbers. So Fermat's Little Theorem can be used to show that a number is not prime, but not to prove that a number is prime.

7.5. Divisibility and Polynomials

We apply some of the ideas on divisibility introduced in earlier sections of this chapter to polynomials with real coefficients, $\mathbb{R}[x]$. This requires us to treat polynomials algebraically. We begin by formally defining operations on $\mathbb{R}[x]$. Let $f, g \in \mathbb{R}[x]$,

$$f(x) = \sum_{n=0}^{N} a_n x^n$$

and

$$g(x) = \sum_{m=0}^{M} b_m x^m.$$

So f is a polynomial of degree at most N and g is a polynomial of degree at most M. In order to simplify our expressions, we subscribe to the convention that for the polynomials f and g, $a_n = 0$ for all $n > N$, and $b_m = 0$ for all $m > M$. That is, we may consider a polynomial as a power series in which all but finitely many of the coefficients equal 0.

REMARK. If a polynomial is identically equal to a non-zero constant, we say that the polynomial has degree zero. If the polynomial is identically zero, we do not define its degree. This is a notational convenience: a polynomial of degree 0 is a non-zero constant.

We define addition and multiplication in $\mathbb{R}[x]$ by

$$f(x) + g(x) := \sum_{i=0}^{\max(M,N)} (a_i + b_i)x^i$$

and

$$f(x) \cdot g(x) := \sum_{i=0}^{M+N} \left(\sum_{j=0}^{i} a_j \cdot b_{i-j} \right) x^i.$$

You should confirm that $0 \in \mathbb{R}[x]$ is the additive identity in $\mathbb{R}[x]$, and $1 \in \mathbb{R}[x]$ is the multiplicative identity in $\mathbb{R}[x]$. You should also verify that addition and multiplication in $\mathbb{R}[x]$ are

(1) associative
(2) commutative
(3) distributive (i.e. multiplication distributes over addition).

We shall prove that a version of the Division Algorithm holds for polynomials. Indeed, it is the reason that long division of polynomials is essentially similar to division of integers.

THEOREM 7.21. *Division Algorithm If $f, g \in \mathbb{R}[x]$, and $g \neq 0$, then there are unique polynomials q and r such that*

$$f = q \cdot g + r$$

and either $r = 0$ or the degree of r is less than the degree of g.

DISCUSSION. *We argue first for the existence of a quotient and remainder satisfying the statement of the theorem. We let g be an arbitrary real polynomial and argue by induction on the degree of f — for this particular divisor g. The induction principle will yield the result for the divisor g and any dividend. Since g is an arbitrary real polynomial, the existence of a quotient and remainder is guaranteed for any divisor and dividend. Uniqueness is proved directly.*

PROOF. Let $g \in \mathbb{R}[x]$. If g is a constant, then $q(x) = (1/g(x))(f(x))$ and $r = 0$ satisfy the statement of the theorem. Furthermore, any remainder must be the zero polynomial, since it is impossible to have degree smaller than the degree of g. Hence, $q(x) = (1/g(x))(f(x))$ is the unique quotient which satisfies the Division Algorithm.

Let g be a polynomial of degree greater than 0. We prove the result for all possible f (for this particular g) by induction on the degree of f. Let M be the degree of g and N be the degree of f.
Base case: $N < M$
Then $q = 0$ and $r = f$ satisfy the conclusion of the theorem.
Induction step: Let $N \geq M$ and assume that the result holds for all polynomials of degree less than N. We show that it holds for $f \in \mathbb{R}[x]$ of degree N. We assume that

$$f(x) = \sum_{n=0}^{N} a_n x^n$$

where $a_n \in \mathbb{R}$ (for $0 \leq n \leq N$) and $a_N \neq 0$. Let

$$g(x) = \sum_{m=0}^{M} b_m x^m$$

where $b_m \in \mathbb{R}$ (for $0 \leq m \leq M$) and $b_M \neq 0$. Let

$$h(x) = \left(\frac{a_N}{b_M}\right) x^{(N-M)}.$$

Then the degree of $f - h \cdot g$ is less than N or $f - h \cdot g$ is identically 0. So there is $s \in \mathbb{R}[x]$ such that

$$f = h \cdot g + s$$

where $s = 0$ or the degree of s is less than N. If $s = 0$, then let $q = h$ and $r = 0$.

Otherwise, by the induction hypothesis, there is some polynomial $\bar{q}$ such that

$$s = \bar{q} \cdot g + r$$

where $r = 0$ or the degree of r is less than M. Thus

$$f = hg + s = hg + \bar{q}g + r = (h + \bar{q})g + r.$$

If we let $q = h + \bar{q}$ then

$$f = qg + r.$$

So, by the principle of induction, for any $f \in \mathbb{R}[x]$, there are q and r such that

$$f = q \cdot g + r.$$

Since g was an arbitrary polynomial of degree greater than 0, the result holds for all f and g.

We prove that q and r are unique. Let

$$f = qg + r = \bar{q}g + \bar{r}$$

where the remainders, r and $\bar{r}$, have degree less than the degree of g or are the 0 polynomial. Then

$$qg + r - (\bar{q}g + \bar{r}) =$$

$$(q - \bar{q})g + (r - \bar{r}) = 0.$$

Let $Q = q - \bar{q}$ and $R = r - \bar{r}$. Assume that $Q \neq 0$. Then the degree of $Q \cdot g$ is no less than the degree of g. However the remainders r and $\bar{r}$ have degree less than the degree of g, or are the 0 polynomial. Thus the degree of R is strictly less than the degree of g, or $R = 0$. The sum of two polynomials of different degree cannot be identically 0. Hence it is impossible that $Q \neq 0$. If $Q = 0$ then $R = 0$. Therefore

$$q = \bar{q}$$

and

$$r = \bar{r}$$

and the quotient and remainder are unique. □

COROLLARY 7.22. *If $f \in \mathbb{R}[x]$ and $x_0 \in \mathbb{R}$, then there is $q \in \mathbb{R}[x]$ such that*

$$f(x) \;=\; (x - x_0) \cdot q(x) + f(x_0).$$

PROOF. Apply the Division Algorithm with $g(x) = x - x_0$. Then the remainder r is of degree 0, or identically zero, so is constant, and evaluating

$$f(x) \;=\; (x - x_0)q(x) + r(x)$$

at $x = x_0$ gives $r(x) = f(x_0)$. Therefore

$$f(x) \;=\; (x - x_0)q(x) + f(x_0).$$

□

We use these results to prove an algebraic property of polynomials.

DEFINITION. Ideal If $I \subseteq \mathbb{R}[x]$ and $I \neq \emptyset$, then we call I an ideal of $\mathbb{R}[x]$ provided the following conditions are satisfied:

(1) If $f, g \in I$ then $f + g \in I$.
(2) If $f \in I$ and $g \in \mathbb{R}[x]$ then $f \cdot g \in I$.

An ideal of $\mathbb{R}[x]$ is a set that is closed under addition of elements in the ideal, and multiplication by all elements of $\mathbb{R}[x]$, whether or not they are in the ideal. If you look closely at the definition of integer combination (Section 7.1), you will observe that the set of integer combinations of a pair of integers is closed under addition of elements in the set and multiplication by arbitrary integers. Of course this analogy between the integers and the polynomials is not accidental. If you generalize the idea of an integer combination to polynomials, you would say that the polynomial combinations of a pair of polynomials is an ideal of $\mathbb{R}[x]$. For the integers we were able to prove that the set of integer combinations of a pair of integers is precisely the integer multiples of the greatest common divisor of the integers. Can we prove an analogous result for polynomials?

DEFINITION. Principal ideal An ideal I in $\mathbb{R}[x]$ is principal if there is $f \in \mathbb{R}[x]$ such that

$$I = \{f \cdot g \mid g \in \mathbb{R}[x]\}.$$

In the definition of principal ideal, f is called a generator of I. Theorem 7.7 can be restated to say that the set of integer combinations of a pair of integers is the principal ideal (in $\mathbb{Z}$) generated by the greatest common divisor of the pair.

THEOREM 7.23. *Every ideal of* $\mathbb{R}[x]$ *is principal.*

PROOF. Let I be an ideal of $\mathbb{R}[x]$. Let f be a polynomial of lowest degree in I. We prove that f generates I. Let $h \in I$. It is sufficient to show that h is a multiple of f. By Theorem 7.21, there are $q, r \in \mathbb{R}[x]$, $r = 0$ or the degree of r less than the degree of f, such that

$$h = qf + r.$$

Since I is an ideal and $f \in I$,

$$qf \in I$$

and

$$h - qf = r \in I.$$

By assumption f is of minimal degree in I, so $r = 0$. Therefore

$$h = qf$$

and f generates I. □

This program seems to be moving us towards a result for polynomials that is analogous to the Fundamental Theorem of Arithmetic. A polynomial is irreducible if it cannot be written as the product of polynomials of lower degree. We shall prove in Theorem 9.48 that every polynomial in $\mathbb{R}[x]$ factors uniquely into the product of irreducible polynomials (up to the order of factors and multiplication by constants), and moreover that all irreducible polynomials are of degree at most 2.

Studying algebraic properties of polynomials is the most important motivating principle in Algebra. Good texts on Algebra include John Fraleigh's [**2**] and Israel Herstein's [**3**].

7.6. Exercises

EXERCISE 7.1. Let $n \in \mathbb{N}$. Prove that if n is not prime then n has a prime factor $p \leq \sqrt{n}$.

EXERCISE 7.2. Are $15,462,227$ and $15,462,229$ relatively prime?

EXERCISE 7.3. If $n \in \mathbb{N}$, under what conditions are n and $n+2$ relatively prime?

EXERCISE 7.4. Prove that every natural number may be written as the product of a power of 2 and an odd number.

EXERCISE 7.5. Find $\gcd(8243235, 453169)$.

EXERCISE 7.6. Find $\gcd(15570555, 10872579)$.

EXERCISE 7.7. Let a and b be integers and $m = \gcd(a,b)$. Prove that $\frac{a}{m}$ and $\frac{b}{m}$ are relatively prime integers.

EXERCISE 7.8. Let a and b be positive integers with prime decomposition given by

$$a = \prod_{n=1}^{N} p_n^{r_n}$$

and

$$b = \prod_{n=1}^{N} p_n^{s_n}$$

where $p_n, r_n, s_n \in \mathbb{N}$ and p_n is prime for $1 \leq n \leq N$. Prove that if $t_n = \min(r_n, s_n)$ for $1 \leq n \leq N$, then

$$\gcd(a,b) = \prod_{n=1}^{N} p_n^{t_n}.$$

EXERCISE 7.9. In the statement of Lemma 7.14, suppose that $\gcd(a,n) \neq 1$. Prove that the function ϕ_a is not a permutation of $\mathbb{Z}_n^*$.

EXERCISE 7.10. Prove Proposition 7.16.

EXERCISE 7.11. Is 4757 prime?

EXERCISE 7.12. Define an ideal of $\mathbb{Z}$ in the natural way: A set $I \subseteq \mathbb{Z}$ is an ideal of $\mathbb{Z}$ if for any $m, n \in I$ and $c \in \mathbb{Z}$,
1) $m + n \in I$
and
2) $mc \in I$.
If $a, b \in \mathbb{Z}$, prove that the set of integer combinations of a and b are an ideal of $\mathbb{Z}$.

EXERCISE 7.13. Prove that every ideal of $\mathbb{Z}$ is principal. (Hint: find the generator of the ideal.)

EXERCISE 7.14. Let p be prime and $\mathbb{Z}_p[x]$ be the set of polynomials with coefficients in $\mathbb{Z}_p$. What can you say about the roots of the polynomial $x^{p-1} - [1]$ in $\mathbb{Z}_p$? (We say that $[a] \in \mathbb{Z}_p$ is a root of a polynomial $f \in \mathbb{Z}_p[x]$ if $f([a]) = [0]$.)

EXERCISE 7.15. Prove that 0 is the additive identity in $\mathbb{R}[x]$ and 1 is the multiplicative identity in $\mathbb{R}[x]$. Use the formal definitions of addition and multiplication in $\mathbb{R}[x]$.

EXERCISE 7.16. Prove that the degree of the product of polynomials is equal to the sum of the degrees of the polynomials. Use the formal definition of multiplication in $\mathbb{R}[x]$.

EXERCISE 7.17. Let $p \in \mathbb{R}[x]$. Prove that p has an additive inverse in $\mathbb{R}[x]$. Prove that p has a multiplicative inverse iff p has degree 0. Use the formal definitions of addition and multiplication in $\mathbb{R}[x]$.

EXERCISE 7.18. Prove that addition and multiplication in $\mathbb{R}[x]$ are associative and commutative, and that the distributive property holds. Use the formal definitions of addition and multiplication in $\mathbb{R}[x]$.

EXERCISE 7.19. For $0 \leq n \leq N$, let $a_n \in \mathbb{R}$. If $f = \sum_{n=0}^{N} a_n x^n$ and $g(x) = x - 1$, find the unique quotient and remainder where f is the dividend and g is the divisor.

EXERCISE 7.20. Let $f, g, q \in \mathbb{R}[x]$, $g \neq 0$. Suppose that f is the dividend, g the divisor and q the quotient. Prove that the sum of the degree of g and the degree of q equals the degree of f.

EXERCISE 7.21. Is there a version of the Euclidean Algorithm for $\mathbb{R}[x]$?

CHAPTER 8

The Real Numbers

What are the real numbers and why don't the rational numbers suffice for our mathematical needs? Ultimately the real numbers must satisfy certain axiomatic properties which we find desirable for interpreting the natural world while satisfying the mathematician's desire for a formal foundation for mathematical reasoning.

Of course the real numbers must contain the rational numbers. We also require that the real numbers satisfy rather obvious algebraic properties which hold for the rational numbers, such as commutativity of addition or the distributive property. These axioms allow us to use algebra to solve problems. Additionally we must satisfy geometric properties like the triangle inequality which permit the interpretation of positive real numbers as distances. We need our number system to contain numbers that arise from the algebraic and geometrical interpretation of numbers. Unfortunately the rational numbers do not suffice for this limited objective. For instance, $\sqrt{2}$, which you know by Example 3.23 to be irrational, arises geometrically as the length of the diagonal of the unit square, and as the solution to the algebraic equation $x^2 = 2$.

The development of the limit gave rise to new questions about the real numbers. In particular, when are we assured that a sequence of numbers is convergent in our number system? The proof of convergence claims often use another property of the real numbers, the least upper bound property. Many of the powerful conclusions of calculus are consequences of this property. Loosely speaking, the least upper bound property implies that the real number line doesn't have any "holes". Put another way, if all the elements of one non-empty set of

real numbers are less than all elements of another non-empty set of real numbers, then there is a real number greater than or equal to all the elements of the first set, and less than or equal to all the elements of the second set. This property is called order-completeness, and is formally defined in Section 8.10. Order-completeness, and its desirable consequences, do not hold for the rational numbers.

How do we prove the existence of a set with order and operations that satisfies all these needs simultaneously? One cannot simply assume that such a structure exists. It is possible that the properties specified are logically inconsistent. We might attempt to construct the set. What are the rules for the construction of a mathematical object? This question prompted mathematicians of the late nineteenth and early twentieth century to develop the rules for such a construction — the axioms of set theory.

For this reason we build the real numbers with a set-theoretic construction. That is, we shall construct the natural numbers, integers, rational numbers and irrational numbers in turn, using basic sets, functions and relations. In so doing we shall construct a set with order and operations that contains the rational numbers (or a structure that behaves precisely like we expect the rational numbers to behave), satisfies the algebraic and order axioms, has the properties we need for calculus and is constructed with the tools that you developed in Chapters 1 and 2.

8.1. The Natural Numbers

When we introduced the natural numbers in Chapter 1 we were explicit that we were not defining the set. Instead we proceeded under the assumption that you are familiar with the natural numbers by virtue of your previous mathematical experience. Now we define the natural numbers in the universe of sets, constructing them out of the empty set.

DEFINITION. Successor function Let Y be a set. The successor function, S, is defined by

$$S(Y) := Y \cup \{Y\}.$$

DEFINITION. Inductive set Let S be the successor function and X be any collection of sets satisfying the following conditions:

(1) $\emptyset \in X$
(2) $[Y \in X] \Rightarrow [S(Y) \in X]$.

Then X is called an inductive set.

DEFINITION. Natural numbers Let X be any inductive set. The set of natural numbers is the intersection of all subsets of X that are inductive sets.

Are the natural numbers well-defined? That is, does the definition depend on the choice of the set X? If $\mathcal{F}$ is a family of sets, all of which are inductive, it is easily proved that the intersection over $\mathcal{F}$ is also inductive. If we are given sets X and Y that are inductive, will the sets give rise to the same set of "natural numbers"? Again it is easily seen that the answer is yes since $X \cap Y$ is a subset of both X and Y, and is inductive. The "natural numbers" defined in terms of X and Y will be the "natural numbers" defined in terms of $X \cap Y$ — they constitute the "smallest" inductive set. In order to define the natural numbers in the universe of sets, it must be granted that there exists an inductive set. It is an axiom of set theory that there is such a set, called the axiom of infinity (see Appendix B for a discussion of the axioms of set theory).

What does this set have to do with the natural numbers as we understand and use them in mathematics? Consider the function, i, defined by

$$i(0) \;=\; \emptyset$$

and

$$i(n+1) \;=\; i(n) \cup \{i(n)\}.$$

So

$$\begin{aligned} i(0) &= \emptyset \\ i(1) &= \{\emptyset\} \\ i(2) &= \{\,\emptyset, \{\emptyset\}\,\} \\ i(3) &= \{\,\emptyset, \{\emptyset\}, \{\,\emptyset, \{\emptyset\}\,\}\,\}. \end{aligned}$$

Then i gives a bijection between the natural numbers, as we understand them intuitively, and the minimal inductive set which we defined above.

Let us define $\ulcorner n \urcorner$ formally as the set one obtains by applying the successor function S to the empty set n times. So

$$0 = \emptyset$$

and for $n > 0$ the set

$$\ulcorner n \urcorner = \{\,\emptyset, \{\emptyset\}, \dots\}$$

has exactly n elements, and we shall identify it with the set

$$\{0, 1, \dots, n-1\}$$

that we earlier chose as the canonical set with n elements.

The set

$$\mathbf{N} := \{\ulcorner n \urcorner \mid n \in \mathbb{N}\} \tag{8.1}$$

is inductive, and therefore contains the natural numbers. Finally the reader may confirm that $\mathbf{N}$ has no proper subset that is inductive.

To summarize the construction so far, the axiom of infinity guarantees that there is a set that is inductive. Pick such a set, X. The intersection of all subsets of X that are inductive is $\mathbf{N}$, which we can identify with the natural numbers (conceived intuitively) by the bijection i. In order to continue the construction, we consider $\mathbb{N}$ and $\mathbf{N}$ to be the same set. We need $\mathbb{N}$ to have the operations $+$ and $\cdot$ as well as the relation $\leq$.

We shall define addition in $\mathbb{N}$ with basic set operations and cardinality. If $m, n \in \mathbb{N}$, then we define addition by

$$m + n :=| \left(\ulcorner m \urcorner \times \{\ulcorner 0 \urcorner\}\right) \cup \left(\ulcorner n \urcorner \times \{\ulcorner 1 \urcorner\}\right) | \, .$$

Recall that the cardinality of a finite set is the unique natural number that is bijective with the set — hence the complicated expression on the right hand side of the definition is a natural number. It is easy to confirm that addition defined in this manner agrees with the usual operation in $\mathbb{N}$. Why would we bother to define an operation you have understood for many years? We have defined addition of natural numbers as a *set operation*.

Multiplication is somewhat easier to define. If $m, n \in \mathbb{N}$, then

$$m \cdot n \ := \ | \ulcorner m \urcorner \times \ulcorner n \urcorner | \, .$$

(Of course, by $\ulcorner m \urcorner \times \ulcorner n \urcorner$ we mean the Cartesian product of the sets $\ulcorner m \urcorner$ and $\ulcorner n \urcorner$.) Finally if $m, n \in \mathbb{N}$

$$[m \leq n] \iff [\ulcorner m \urcorner \subseteq \ulcorner n \urcorner].$$

You should confirm that the operations and the relation agree with the usual $+$, $\cdot$ and $\leq$ on the natural numbers.

Having completed this construction it is reasonable to ask whether $\mathbb{N}$ is truly the set of natural numbers. It is certainly justifiable for you to conclude that no clarity about the number 2 is provided by identifying it with the set $\{\, \emptyset, \{\emptyset\} \,\}$. What we gain is a reduction of numbers to sets that will carry us through the construction of all real numbers, including numbers that are not easy to construct.

8.2. The Integers

We construct the integers out of the natural numbers. The algebraic purpose of the integers is to include additive inverses for natural numbers. Of course this naturally gives rise to the operation of subtraction.

Let $Z \ = \ \mathbb{N} \times \mathbb{N}$. Define an equivalence relation, $\sim$ on Z by

$$\langle m_1, n_1 \rangle \sim \langle m_2, n_2 \rangle \quad \iff \quad m_1 + n_2 \ = \ m_2 + n_1.$$

Then the integers are

$$\mathbf{Z} := Z/\sim .$$

We think of the ordered pair $\langle m, n\rangle \in \mathbf{Z}$ as being a representative of the integer $m - n$. We say that an integer is positive if $m > n$ and negative if $m < n$. It should be clear that the set of non-negative integers (that is $\mathbb{N}$) is

$$\{[\langle m, n\rangle] \mid m \geq n\} \;=\; \{[\langle m, 0\rangle] \mid m \in \mathbb{N}\}.$$

Let $\mathbb{Z}$ be the (intuitive) integers and let $i : \mathbf{Z} \to \mathbb{Z}$ be defined by

$$i([\langle m, n\rangle]) \;=\; m - n.$$

Then i is a bijection. As we did with the natural numbers, we shall construct operations and order on $\mathbf{Z}$ that agree with the usual operations and an order on $\mathbb{Z}$. Of course, we could use i and the usual definitions in $\mathbb{Z}$ to define operations and relations on $\mathbf{Z}$, but that would miss the spirit of the construction, and would neglect the desire for set-theoretic definitions. Analogous to the construction of the previous section, we define $\mathbb{Z}$ as $\mathbf{Z}$. Let $x_1, x_2 \in \mathbb{Z}$ where $x_1 \;=\; [\langle m_1, n_1\rangle]$ and $x_2 \;=\; [\langle m_2, n_2\rangle]$. Addition is defined by

$$x_1 + x_2 \;=\; [\langle m_1 + m_2, n_1 + n_2\rangle].$$

The additive inverse of $[\langle m, n\rangle]$ is $[\langle n, m\rangle]$ (i.e. the sum of these integers is $[\langle 0, 0\rangle]$ — the additive identity in $\mathbb{Z}$).

Multiplication is defined by

$$x_1 \cdot x_2 \;=\; [\langle m_1 \cdot m_2 + n_1 \cdot n_2,\; n_1 \cdot m_2 + m_1 \cdot n_2\rangle].$$

The linear ordering on $\mathbb{Z}$ is defined by

$$x_1 \leq x_2 \quad\Longleftrightarrow\quad m_1 + n_2 \leq n_1 + m_2.$$

Addition and multiplication have been defined for the natural numbers, and the operations and linear ordering on $\mathbb{Z}$ are defined with respect to operations and the linear ordering that were previously defined for $\mathbb{N}$. Note that all our definitions were given in terms of representatives of equivalence classes. To show that $+$, $\cdot$ and $\leq$ are well-defined,

we must show that the definitions are independent of the choice of representative — see Exercise 8.6.

8.3. The Rational Numbers

Rational numbers are ratios of integers, or nearly so. Of course, different numerators and denominators can give rise to the same rational number — indeed a good deal of elementary school arithmetic is devoted to determining when two distinct expressions for rational numbers are equal. We built the integers from the natural numbers with equivalence classes of "differences" of natural numbers. We construct the rational numbers from the integers analogously, with equivalence classes of "quotients" of integers. Algebraically this gives rise to division.

Let $Q = \mathbb{Z} \times \mathbb{N}^+$. We define an equivalence relation $\sim$ on Q. If $\langle a, b\rangle, \langle c, d\rangle \in Q$, then

$$\langle a, b\rangle \sim \langle c, d\rangle \quad \Longleftrightarrow \quad a \cdot d = b \cdot c.$$

We define the rational numbers, $\mathbf{Q}$, as the equivalence classes of Q with respect to the equivalence relation $\sim$. That is,

$$\mathbf{Q} := Q/\sim .$$

We associate the equivalence classes of $\mathbf{Q}$ with the intuitive rational numbers via the bijection $i : \mathbb{Q} \to \mathbf{Q}$ defined by

$$i\left(\frac{p}{q}\right) = [\langle p, q\rangle]$$

for $\langle p, q\rangle \in Q$.

We define the operations and linear ordering on $\mathbb{Q}$ in terms of the operations and linear ordering in $\mathbb{Z}$. Define addition by

$$[\langle a, b\rangle] + [\langle c, d\rangle] := [\langle ad + bc, bd\rangle]$$

and multiplication by

$$[\langle a, b\rangle] \cdot [\langle c, d\rangle] := [\langle a \cdot c, b \cdot d\rangle].$$

We define the linear ordering on $\mathbb{Q}$ by

$$[\langle a, b\rangle] \leq [\langle c, d\rangle] \quad \text{iff} \quad a \cdot d \leq b \cdot c.$$

Through the construction of the rational numbers, we have used set operations to build mathematical structures with which you are already familiar. Consequently you are able to check that the construction behaves as you expect. For instance, one can easily prove that the operations we have constructed agree with the usual operations of addition and multiplication on the rational numbers. Similarly one can easily check that the relation we have constructed on $\mathbb{Q}$ agrees with the usual linear ordering of the rational numbers. Constructing the real numbers is more complicated.

8.4. The Real Numbers

We complete our construction of the real numbers (we have the irrational numbers remaining) with the objective of proving the order-completeness of the real numbers, and deriving some important consequences of completeness. Many of the most powerful and interesting results of calculus depend on this property of the real numbers. If you have been asked to accept some of these theorems on faith, now it is time to reward your trust.

There are a couple of different ways to construct the real numbers from the rational numbers. One approach is to define real numbers as convergent sequences of rational numbers. The other common approach is to characterize real numbers as subsets of rational numbers that satisfy certain conditions.

DEFINITION. Dedekind cut A Dedekind cut L is a nonempty proper subset of $\mathbb{Q}$ that has no maximal element and satisfies

$$(\forall\, a, b \in \mathbb{Q})\ [a \in L\ \wedge\ b < a]\ \Rightarrow\ [b \in L].$$

Let L be a Dedekind cut. Then there is some rational number $a \in L$, and therefore all rational numbers less then a are in L. Let $R = \mathbb{Q} \setminus L$. Since $L \neq \mathbb{Q}$, there is $c \in R$ and every rational number

greater than c is in R. It is clear that $\{L, R\}$ is a partition of $\mathbb{Q}$ and that every element of L is less than every element of R. So Dedekind cuts "split" the rational numbers. We shall associate each Dedekind cut with a real number located at the split on the real number line.

REMARK. To help our mental picture of what is going on, we think of L as all rational numbers to the left of some fixed real number α, *i.e.* as $(-\infty, \alpha) \cap \mathbb{Q}$, and R as the rational numbers to the right, $[\alpha, \infty) \cap \mathbb{Q}$. Of course we don't yet know what exactly we mean by "the real number α", but this is the idea to keep in mind. Note that R will have a least element iff α is rational.

To understand how Dedekind cuts relate to numbers we construct an injection from the rational numbers to the Dedekind cuts. Let $\mathcal{D}$ be the set of Dedekind cuts. We define an injection $i : \mathbb{Q} \to \mathcal{D}$ by

$$i(a) \;=\; \{b \in \mathbb{Q} \mid b < a\}.$$

The function i is a well-defined injection that informs us of how $\mathbb{Q}$ fits into $\mathcal{D}$.

We shall define order and operations on $\mathcal{D}$ so that they agree with the usual linear ordering and operations on $\mathbb{Q}$ that are inherited in $i[\mathbb{Q}]$. That is, we shall define the linear order, addition and multiplication on $\mathcal{D}$ so that for $a, b \in \mathbb{Q}$,

$$[a \leq b] \iff [i(a) \leq i(b)] \tag{1}$$
$$i(a + b) \;=\; i(a) + i(b) \tag{2}$$
$$i(a \cdot b) \;=\; i(a) \cdot i(b) \tag{3}.$$

If we can do this, we can think of $\mathcal{D}$ as an extension of $\mathbb{Q}$. How do we do it?

For $L, K \in \mathcal{D}$, we define the relation $\leq$ in $\mathcal{D}$ by

$$[L \leq K] \iff [L \subseteq K].$$

You should confirm that $\leq$ is a linear ordering of $\mathcal{D}$ and that the relation $\leq$ on $i[\mathbb{Q}]$ satisfies (1). If $L \in \mathcal{D}$ and $L < i(0)$ we say that L is negative. If $L > i(0)$, we say that L is positive.

With a similar objective in mind we define addition and multiplication on $\mathcal{D}$. That is, we want the operations to satisfy certain properties of addition and multiplication and we want the operations defined on $i[\mathbb{Q}]$ to agree with the operations on $\mathbb{Q}$.

If $L, K \in \mathcal{D}$, then

$$L + K \ := \ \{a + b \mid a \in L \text{ and } b \in K\}.$$

Verify that $L + K$ is a Dedekind cut, and that (2) holds.

Multiplication takes a bit more effort to define. (Why can't we let $L \cdot K = \{ab \mid a \in L,\ b \in K\}$?) If L or K is $i(0)$, then

$$L \cdot K \ := \ i(0).$$

If $L, K \in \mathcal{D}$ are both positive, then

$$L \cdot K \ = \ \{a \cdot b \mid a \in L,\, b \in K,\, a > 0 \text{ and } b > 0\} \cup \{c \in \mathbb{Q} \mid c \leq 0\}.$$

Verify that $L \cdot K$ is a Dedekind cut, and that (3) holds for $a, b > 0$.

How do we define multiplication by "negative" Dedekind cuts? Let's start with defining multiplication by -1. Let $L \in \mathcal{D}$ and $R = \mathbb{Q} \setminus L$. We define $-L$ by

$$-L : = \ \{c \in \mathbb{Q} \mid (\exists\, r \in R) \ -c > r\}.$$

Now we can define multiplication on arbitrary elements of $\mathcal{D}$ to satisfy the properties we desire. If $L, K \in \mathcal{D}$ and both are negative, then

$$L \cdot K : = \ (-L \cdot -K).$$

If exactly one of L and K is negative, then

$$L \cdot K : = \ -(-L \cdot K).$$

Definition. Real numbers, $\mathbb{R}$ The real numbers are the Dedekind cuts, with addition, multiplication and $\leq$ defined as above. We denote the real numbers by $\mathbb{R}$ when we do not need to think of them explicitly as Dedekind cuts.

We have defined the real numbers as sets of rational numbers. Since the rational numbers were defined using basic ideas about sets, functions and relations, so are the real numbers. The properties of the real

numbers that we discussed at the beginning of this section are satisfied by the Dedekind cuts. For every rational number a, we identify a with the Dedekind cut $i(a)$.

THEOREM 8.2. *The real numbers as defined above satisfy:*
(i) Addition and multiplication are both commutative and associative.
(ii) $(\forall\, L \in \mathcal{D})\ L + 0 = L,\ L \cdot 1 = L.$
(iii) $(\forall\, L \in \mathcal{D})\ L + (-L) = 0.$
(iv) $(\forall\, L \in \mathcal{D} \setminus \{0\})(\exists\, K \in \mathcal{D})\ L \cdot K = 1.$
(v) $(\forall\, L, K, J \in \mathcal{D})\ \ L \cdot (K + J) = L \cdot K + L \cdot J.$

PROOF. Exercise. □

8.5. The Least Upper Bound Property

DEFINITION. Upper bound Let $X \subset \mathcal{D}$. We say that X is bounded above if there is $M \in \mathcal{D}$ such that

$$(\forall x \in X)\ x \leq M.$$

In this event we say that M is an upper bound for X.

DEFINITION. Least upper bound Let $X \subset \mathcal{D}$ be bounded above. Suppose M is an upper bound for X such that for any upper bound N for X, $M \leq N$. Then the number M is called the least upper bound for X.

Lower bound and greatest lower bound are defined analogously.

THEOREM 8.3. *Least Upper Bound Property If X is a non-empty subset of $\mathcal{D}$ and is bounded above, then X has a least upper bound. If it is bounded below, then it has a greatest lower bound.*

PROOF. Let $X \subset \mathcal{D}$ be bounded above. Let

$$M = \bigcup_{L \in X} L \subseteq \mathbb{Q}.$$

The set M is bounded above (why?), and hence $M \neq \mathbb{Q}$. Any element of M is an element of some $L \in X$, and consequently cannot be a

maximal element of L. Therefore M has no largest element. If $a \in M$, $c \in \mathbb{Q}$ and $c < a$ then $c \in M$. Therefore M is a Dedekind cut. For any $L \in X$, $L \subseteq M$ and hence

$$L \leq M.$$

That is, M is an upper bound for X.

Let $K < M$. Then there is $a \in M \setminus K$. So a is in some L_0 in X. Therefore L_0 is not contained in K and K is not an upper bound for X. It follows that M is the least upper bound for X.

We leave the argument for the existence of a greatest lower bound to the reader. □

The least upper bound property is the essential property of real numbers that permits the main theorems of calculus. It is the reason we use this large set, rather than, say, the algebraic numbers. It uniquely characterizes the real numbers as an extension of the rational numbers — see Theorem 8.23 for a precise statement.

Now that we have proved this key property, we shall use $\mathbb{R}$ to denote the set of real numbers, identifying a real number α with the Dedekind cut $(-\infty, \alpha) \cap \mathbb{Q}$. We shall no longer need to concern ourselves with Dedekind cuts *per se*.

8.6. Real Sequences

Recall that a sequence is a function with domain $\mathbb{N}$ (or $\mathbb{N}^+$). A real sequence is a real-valued sequence (that is, the range of the sequence is a subset of the real numbers).

DEFINITION. Subsequence Let $\langle a_n \mid n \in \mathbb{N} \rangle$ be a sequence and $f \in \mathbb{N}^{\mathbb{N}}$ be a strictly increasing sequence of natural numbers. Then

$$\langle a_{f(n)} \mid n \in \mathbb{N} \rangle$$

is a subsequence of $\langle a_n \mid n \in \mathbb{N} \rangle$.

EXAMPLE 8.4. Let s be the sequence

$$\langle 2n \mid n \in \mathbb{N} \rangle \;=\; \langle 0, 2, 4, 6, 8, \ldots \rangle.$$

Then the sequence t given by

$$\langle 6n \mid n \in \mathbb{N} \rangle \;=\; \langle 0, 6, 12, 18, \ldots \rangle$$

is a subsequence of s. In this example, $f(n) = 3n$ is the function that demonstrates that t is a subsequence of s. Another subsequence of s is the sequence

$$\langle 2^{5n+3} \mid n \in \mathbb{N} \rangle.$$

Recall that a sequence $\langle a_n \rangle$ is called *non-decreasing* if $a_{n+1} \geq a_n$ for all n. It is called non-increasing if the inequality is reversed. Everything that is true for a non-decreasing sequence is true, with inequalities reversed, for non-increasing sequences (why?), so rather than state everything twice, we can use the word *monotonic* to mean a sequence that is either non-increasing (everywhere) or non-decreasing.

LEMMA 8.5. *Every non-decreasing real sequence $\langle a_n \mid n \in \mathbb{N} \rangle$ that is bounded above converges to its least upper bound. Every non-increasing real sequence that is bounded below converges to its greatest lower bound.*

PROOF. We shall only prove the first assertion. Let M be the least upper bound of $\langle a_n \rangle$. Let $\varepsilon > 0$. Since M is the least upper bound, there is $N \in \mathbb{N}$ such that,

$$0 < M - a_N < \varepsilon.$$

Since the sequence is non-decreasing,

$$(\forall n \geq N)\; 0 < M - a_n < \varepsilon.$$

Therefore M is the limit of the sequence, as desired. □

THEOREM 8.6. *Bolzano-Weierstrass Theorem* *Let $[b, c]$ be a closed bounded interval of real numbers and $s = \langle a_n \mid n \in \mathbb{N} \rangle$ be a sequence of real numbers such that*

$$(\forall n \in \mathbb{N})\; a_n \in [b, c].$$

Then $\langle a_n \mid n \in \mathbb{N} \rangle$ has a convergent subsequence with limit in $[b, c]$.

DISCUSSION. We consider a nested sequence of intervals, all of which contain infinitely many elements of the range of the sequence s, with the radius of the intervals approaching 0. We construct a subsequence of s by sequentially selecting elements in the intersection of the range of s and the successive intervals. We then show that the subsequence we construct is convergent.

PROOF. We prove the theorem for the closed unit interval $[0, 1]$. It is straightforward to generalize this argument to arbitrary closed bounded intervals.

If the range of the sequence is a finite set, then at least one element of the range, a_n, must have an infinite pre-image. The pre-image of a_n gives a subsequence that converges to a_n. Therefore we assume that the range of the sequence is infinite. Let S be the range of the sequence $\langle a_n \rangle$.

We define a nested sequence of closed intervals, $I_n = \langle [b_n, c_n] \mid n \in \mathbb{N} \rangle$ satisfying

(1) $I_0 = [0, 1]$
(2) For all $n \in \mathbb{N}$, $I_{n+1} \subset I_n$
(3) $c_n - b_n = \frac{1}{2^n}$
(4) For all $n \in \mathbb{N}$, $I_n \cap S$ is infinite.

Let $I_0 = [0, 1]$. Assume that we have I_n satisfying the conditions above. At least one of the intervals $[b_n, b_n + \frac{1}{2^{n+1}}]$ and $[b_n + \frac{1}{2^{n+1}}, c_n]$ must contain infinitely many elements of S. Let $I_{n+1} = [b_n, b_n + \frac{1}{2^{n+1}}]$ if the intersection of this set with S is infinite; otherwise let $I_{n+1} = [b_n + \frac{1}{2^{n+1}}, c_n]$. Then I_{n+1} satisfies the conditions above.

The sequence of left end-points of the intervals I_n, $\langle b_n \mid n \in \mathbb{N} \rangle$ is non-decreasing. The sequence of right endpoints of the intervals I_n, $\langle c_n \mid n \in \mathbb{N} \rangle$ is non-increasing. Furthermore, for any $m, n \in \mathbb{N}$,

$$b_m < c_n.$$

The set $\{b_n \mid n \in \mathbb{N}\}$ is bounded above, so by the Least Upper Bound Property the set has a least upper bound, β. Similarly the set $\{c_n \mid$

$n \in \mathbb{N}\}$ has a greatest lower bound γ. By Lemma 8.5

$$\begin{aligned} \lim_{n\to\infty} b_n &= \beta \\ \lim_{n\to\infty} c_n &= \gamma. \end{aligned}$$

By the triangle inequality, for any $n \in \mathbb{N}$,

$$| \beta - \gamma | \leq | \beta - b_n | + | b_n - c_n | + | c_n - \gamma | .$$

All three terms on the right hand side of the inequality tend to 0 as n approaches infinity, so for any $\varepsilon > 0$,

$$| \beta - \gamma | < \varepsilon.$$

Hence $\beta = \gamma$.

We now want to define a subsequence that converges to β, by choosing a point in each interval I_n in turn. Formally we do this by defining $f \in \mathbb{N}^{\mathbb{N}}$ recursively by

$$f(0) = 0$$

and $f(n+1)$ is the least $k \in \mathbb{N}$ such that

$$[k > f(n)] \wedge [a_k \in I_{n+1}].$$

This is well-defined since $S \cap I_{n+1}$ is infinite. Then the sequence $\langle a_{f(n)} \mid n \in \mathbb{N}\rangle$ converges to β. To see this, let $\varepsilon > 0$. For any $n \in \mathbb{N}$ such that $\frac{1}{2^n} < \varepsilon$,

$$| \beta - a_{f(n)} | < c_n - b_n = \frac{1}{2^n} < \varepsilon.$$

Therefore $\langle a_{f(n)} \mid n \in \mathbb{N}\rangle$ is a convergent subsequence converging to β. □

DEFINITION. Cauchy sequence Let $\langle a_n \mid n \in \mathbb{N}\rangle$ be a sequence. The sequence $\langle a_n \rangle$ is a Cauchy sequence if

$$(\forall\, \varepsilon > 0)(\exists\, N \in \mathbb{N})(\forall\, m, n \in \mathbb{N})\ [m, n \geq N] \Rightarrow [|\ a_m - a_n\ | < \varepsilon].$$

THEOREM 8.7. *A real sequence converges iff it is a Cauchy sequence.*

PROOF. $\Rightarrow$

Let $\langle a_n \mid n \in \mathbb{N}\rangle$ be a sequence of real numbers that converges to $a \in \mathbb{R}$. Let $\varepsilon > 0$ and $N \in \mathbb{N}$ be such that

$$(\forall\, n \geq N) \;\; \mid a - a_n \mid < \frac{\varepsilon}{2}.$$

Then for any $m, n \geq N$,

$$\mid a_n - a_m \mid \leq \mid a_n - a \mid + \mid a - a_m \mid < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \;=\; \varepsilon.$$

Therefore $\langle a_n \mid n \in \mathbb{N}\rangle$ is a Cauchy sequence.

$\Leftarrow$

Let $\langle a_n \mid n \in \mathbb{N}\rangle$ be a Cauchy sequence. Then

$$(\exists\, N \in \mathbb{N})(\forall\, m, n > N) \;\; \mid a_n - a_m \mid < 1.$$

Every term in the sequence after the N^{th} term is in the ε-neighborhood of a_N. So

$$(\forall\, n \geq N)\; a_n \in [a_N - 1, a_N + 1].$$

The sequence $\langle a_n \mid n \geq N\rangle$ satisfies the hypotheses of the Bolzano-Weierstrass Theorem, and thus has a convergent subsequence.

Let $\langle a_{f(n)} \mid n \in \mathbb{N}\rangle$ be a convergent subsequence of $\langle a_n \mid n \in \mathbb{N}\rangle$ converging to $a \in \mathbb{R}$. Let $\varepsilon > 0$. Since $\langle a_n\rangle$ is Cauchy, there is N_1 such that

$$(\forall\, m, n \geq N_1) \;\; \mid a_m - a_n \mid < \frac{\varepsilon}{2}.$$

Furthermore, there is $N_2 \in \mathbb{N}$ such that

$$(\forall\, n \geq N_2) \;\; \mid a_{f(n)} - a \mid < \frac{\varepsilon}{2}.$$

Let $N_3 \geq N_1, f(N_2)$. Then $N_3 \geq N_2$ and

$$(\forall\, n \geq N_3) \;\; \mid a_n - a \mid \leq \mid a_n - a_{f(n)} \mid + \mid a_{f(n)} - a \mid < \varepsilon.$$

Therefore the sequence $\langle a_n \mid n \in \mathbb{N}\rangle$ converges to a. □

Cauchy sequences get at the essence of the order-completeness of the real numbers. A Cauchy sequence of rational numbers need not converge to a rational number. For instance, let a be any irrational number, and let a_n be the decimal approximation of a to the n^{th} digit.

The sequence $\langle a_n \rangle$ is a Cauchy sequence of rational numbers that converges to an irrational number. However if a Cauchy sequence fails to converge in a set of numbers, it is reasonable to say that there is a gap in the set of numbers. The real numbers are defined so that these gaps are filled.

8.7. Ratio Test

One of the uses of the order-completeness of the real numbers is proving that an infinite sequence converges, without having to know much about the number to which it converges. In Chapter 5 we allude to the ratio test in claiming that the Taylor polynomial for the exponential function evaluated at a real number a, $\sum_{k=0}^{\infty} \frac{a^k}{k!}$, converges. How do we prove that an infinite sum converges? If we have an idea of its limit, we might show that the sequence of partial sums approaches this value. This is how we prove that the geometric sum with ratio less than 1 converges. Many important mathematical functions are defined by infinite sums, and the limit of the sum defines the value of the function. In this case we need to show that the sum converges using properties of the real numbers.

DEFINITION. Absolute convergence Let $\langle a_n \rangle$ be an infinite sequence. If the infinite sum

$$\sum_{k=0}^{\infty} | a_k |$$

converges then the infinite sum $\sum_{k=0}^{\infty} a_k$ is said to converge absolutely.

LEMMA 8.8. *If an infinite sum converges absolutely, then it converges.*

PROOF. Assume $\sum_{k=0}^{\infty} a_k$ converges absolutely. We show that the sequence of partial sums of this series, $\langle s_n \mid n \in \mathbb{N} \rangle$, is a Cauchy sequence. For $n \in \mathbb{N}$, let

$$b_n = | a_n | .$$

Then $\sum_{k=0}^{\infty} b_k$ converges. Let $\langle t_n \mid n \in \mathbb{N} \rangle$ be the sequence of partial sums of $\sum_{k=0}^{\infty} b_k$. By Theorem 8.7, $\langle t_n \rangle$ is a Cauchy sequence. Let $\varepsilon > 0$. Then there is $N \in \mathbb{N}$ such that for any $n \geq m \geq N$,

$$\mid t_n - t_m \mid \leq \varepsilon.$$

By a generalization of the triangle inequality (see Exercise 8.24)

$$\mid s_n - s_m \mid = \left| \sum_{k=m+1}^{n} a_k \right| \leq \sum_{k=m+1}^{n} b_k = \mid t_n - t_m \mid < \varepsilon.$$

Hence $\langle s_n \rangle$ is a Cauchy sequence and converges. Therefore $\sum_{k=0}^{\infty} a_k$ converges. □

THEOREM 8.9. *Ratio test* *Suppose $\langle a_k \rangle$ is an infinite sequence of real numbers and that there is $N \in \mathbb{N}$ and a positive real number $r < 1$ such that for all $n \geq N$,*

$$\left| \frac{a_{n+1}}{a_n} \right| \leq r.$$

Then $\sum_{k=0}^{\infty} a_k$ converges.

PROOF. Let $\sum_{k=0}^{\infty} a_k$ be an infinite sum with terms satisfying the hypothesis. For $n \in \mathbb{N}$, let $b_n = \mid a_n \mid$. By assumption, there is $N \in \mathbb{N}$ and a positive real number $r < 1$ such that for all $n \geq N$,

$$\frac{b_{n+1}}{b_n} \leq r.$$

We may assume without loss of generality that $N = 0$, since the series $\sum_{k=0}^{\infty} b_k$ converges iff $\sum_{k=N}^{\infty} b_k$ converges, and if necessary we may ignore finitely many terms of the infinite sum. We claim that for all $n \in \mathbb{N}$,

$$b_n \leq b_0 r^n.$$

If $n = 0$ the claim is obvious. Assume the claim holds at n. By assumption,

$$\frac{b_{n+1}}{b_n} \leq r.$$

Therefore

$$b_{n+1} \leq r b_n \leq r b_0 r^n \leq b_0 r^{n+1}.$$

By Exercise 5.28, the geometric sum with radius $-1 < r < 1$ converges to $\frac{1}{1-r}$. Therefore, for any $n \in \mathbb{N}$,

$$s_n := \sum_{k=0}^{n} b_k \leq \sum_{k=0}^{n} b_0 r^k \;=\; b_0 \left(\sum_{k=0}^{\infty} r^k \right) \leq \frac{b_0}{1-r}.$$

The sequence of partial sums, $\langle s_n \rangle$, is a monotonic bounded sequence and by Lemma 8.5, converges. Therefore $\sum_{k=0}^{\infty} a_k$ converges absolutely. By Lemma 8.8 the sum converges. □

8.8. Real Functions

If you reread your calculus text, you will observe that many of the theorems of calculus are ultimately dependent on the Intermediate Value Theorem.

THEOREM 8.10. *Intermediate Value Theorem* *Let f be a continuous real function on a closed bounded interval $[a, b]$. If $f(a) < L < f(b)$ or $f(b) < L < f(a)$ then*

$$(\exists\ c \in (a, b)) \quad f(c) \;=\; L.$$

PROOF. Let f be a continuous real function on a closed bounded interval $[a, b]$, and $f(a) < L < f(b)$. We prove the special case $L = 0$. Given the result for $L = 0$, the theorem follows from application of the special case to the function $f(x) - L$.

Let

$$X \;=\; \{x \in [a, b] \mid (\forall\, y \in [a, x])\ f(y) \leq 0\}.$$

Then $X \neq \emptyset$ and X is bounded above by b. By the Least Upper Bound Property, X has a least upper bound, $m \leq b$. The function f is continuous, and hence $\lim_{x \to m} f(x) = f(m)$. If $f(m) = 0$, the theorem is proved.

(i) Assume that $f(m) > 0$. Let $0 < \varepsilon < f(m)$. For any $x \in [a, m)$, $f(x) \leq 0$ and

$$|\, f(x) - f(m) \,| \geq f(m) > \varepsilon.$$

Consequently for any $\delta > 0$, there is x in the punctured δ-neighborhood of m such that

$$| f(x) - f(m) | \geq \varepsilon.$$

This contradicts the assumption that $\lim_{x \to m} f(x) = f(m)$. Therefore $f(m) \leq 0$.

(ii) Assume that $f(m) < 0$. Let $0 < \varepsilon < | f(m) |$. For any $\delta > 0$, there is $x \in (m, m + \delta)$ such that $f(x) > 0$. Otherwise

$$[a, m + \delta) \subseteq X,$$

contradicting the assumption that m is the least upper bound for X. So for any $\delta > 0$ there is x in the punctured δ-neighborhood of m such that

$$| f(x) - f(m) | \geq | f(m) | > \varepsilon.$$

This contradicts the assumption that f is continuous at m. Therefore $f(m) = 0$. □

THEOREM 8.11. *Extreme Value Theorem If f is a continuous real function on a closed bounded interval $[a, b]$, then f achieves a maximum and a minimum on $[a, b]$.*

PROOF. We show first that the range of $f|_{[a,b]}$ is bounded above and below. By way of contradiction suppose that the range of f is not bounded above. For $n \in \mathbb{N}$, let $a_n \in [a, b]$ be such that $f(a_n) > n$. By the Bolzano-Weierstrass Theorem, the sequence $\langle a_n \rangle$ has a convergent subsequence, $\langle a_{g(n)} \rangle$, converging to some number $c \in [a, b]$. By the continuity of f, if $c \in (a, b)$ then

$$f(c) = \lim_{x \to c} f(x) = \lim_{n \to \infty} f(a_{g(n)}).$$

(See Exercise 8.25.) If c is an endpoint of $[a, b]$, we make the corresponding claim for the appropriate one-sided limit. However, for any $n \in \mathbb{N}$,

$$f(a_{g(n)}) > g(n) > n.$$

Hence, $\lim_{n \to \infty} f(a_{g(n)})$ does not exist. Therefore the range of f is bounded above. Similarly, the range of f is bounded below. By the

Least Upper Bound Property, the range of f has a least upper bound, M, and a greatest lower bound, L.

Since M is a least upper bound for the range of f, for any $\varepsilon > 0$, there is $x \in [a, b]$ such that

$$| f(x) - M | < \varepsilon.$$

For $n \in \mathbb{N}^+$, let $a_n \in [a, b]$ be such that

$$| f(a_n) - M | < \frac{1}{n}.$$

The sequence $\langle a_n \rangle$ has a convergent subsequence by the Bolzano-Weierstrass Theorem. Let $\langle c_n \rangle$ be a convergent subsequence of $\langle a_n \rangle$ with limit $c \in [a, b]$. Since $\langle c_n \rangle$ is a subsequence of $\langle a_n \rangle$, for any $n \in \mathbb{N}^+$,

$$| f(c_n) - M | < \frac{1}{n}.$$

Hence

$$\lim_{n \to \infty} f(c_n) \;=\; M.$$

By the continuity of f, if $c \in (a, b)$ then

$$\lim_{x \to c} f(x) \;=\; f(c) \;=\; \lim_{n \to \infty} f(c_n) \;=\; M.$$

If c is an endpoint of $[a, b]$ we have the analogous claim for the appropriate one-sided limit. Therefore f achieves a maximum value on $[a, b]$. By an analogous argument, f achieves a minimum value on $[a, b]$. $\square$

By the Extreme Value Theorem, a continuous function achieves extreme values on a closed bounded interval. It is easy to construct examples for which the theorem fails for open intervals. The extreme value theorem has in common with the least upper bound property that it guarantees the existence of a number satisfying a desirable condition without providing additional information about the number itself. Quite often it is enough to know abstractly that a function attains its extremum without having to further distinguish the object. What more can we conclude about the extreme values of a function?

THEOREM 8.12. *Let f be a real function defined on an interval (a, b). If $c \in (a, b)$ is such that $f(c)$ is an extreme value of f on (a, b) and f is differentiable at c, then $f'(c) = 0$.*

PROOF. Let f and c satisfy the hypotheses of the theorem. Suppose that $f(c)$ is the maximum value achieved by f on (a, b). For any $x \in (a, c)$, $f(x) \leq f(c)$ and

$$\frac{f(c) - f(x)}{c - x} \geq 0.$$

Therefore

$$\lim_{x \to c^-} \frac{f(c) - f(x)}{c - x} \geq 0.$$

Similarly,

$$\lim_{x \to c^+} \frac{f(c) - f(x)}{c - x} \leq 0.$$

However f is differentiable at c, so

$$0 \leq \lim_{x \to c^-} \frac{f(c) - f(x)}{c - x} = f'(c) = \lim_{x \to c^+} \frac{f(c) - f(x)}{c - x} \leq 0.$$

A similar argument proves the claim for $f(c)$ a minimum value of f on (a, b). □

COROLLARY 8.13. *Let f be a continuous real function on a closed bounded interval $[a, b]$. Then f achieves a maximum and minimum on $[a, b]$ and if $c \in [a, b]$ is a number at which f achieves an extreme value, then one of the following must be true of c:*
(i) $f'(c) = 0$
(ii) f is not differentiable at c
(iii) c is an endpoint of $[a, b]$.

THEOREM 8.14. *Mean Value Theorem Let f be a continuous real function on a closed bounded interval $[a, b]$ and differentiable on (a, b). Then there is $c \in (a, b)$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

PROOF. We first prove a special case of the Mean Value Theorem, known as Rolle's Theorem. Assume that $f(a) = f(b)$. We prove that there is $x \in (a, b)$ such that $f'(x) = 0$.

If f is constant then $f'(x) = 0$ for all $x \in (a, b)$. Assume that f is non-constant and that there is $x \in (a, b)$ such that $f(x) > f(a)$. By the Extreme Value Theorem f achieves a maximum value M on $[a, b]$. Thus,

$$M > f(a) = f(b).$$

Let $c \in (a, b)$ be such that $f(c) = M$. By Theorem 8.12, $f'(c) = 0$. If there is $x \in (a, b)$ such that $f(x) < f(a)$, the proof is similar.

To prove the Mean Value Theorem in general, we reduce it to Rolle's Theorem. We subtract from $f(x)$ the line segment formed by $(a, f(a))$ and $(b, f(b))$. Let

$$g(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a).$$

The function $g(x)$ satisfies the hypotheses of Rolle's Theorem. So there is $c \in (a, b)$ such that $g'(c) = 0$. Since

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}$$

we have

$$g'(c) = f'(c) - \frac{f(b) - f(a)}{b - a} = 0$$

and

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

□

The Mean Value Theorem has many practical consequences, one of which we state here.

COROLLARY 8.15. *Let f be a differentiable function on (a, b). If $f'(x) > 0$ (resp. $f'(x) < 0$) on (a, b) then f is increasing (resp. decreasing) on (a, b).*

8.9. Cardinality of the Real Numbers

We finished Chapter 6 with the unproved claim that the real numbers are uncountable. Now that we have a formal definition of the real numbers, we are ready to complete our investigation of the cardinality of $\mathbb{R}$. By Theorem 6.11 the set of infinite decimal sequences is uncountable, with cardinality $2^{\aleph_0}$. We went on to claim that this had consequences for the cardinality of $\mathbb{R}$. We consider the related question of the cardinality of the closed unit interval $[0, 1]$.

PROPOSITION 8.16. $| [0, 1] | = | \mathbb{R} |$.

PROOF. Define $f : [0, \infty) \to (1/2, 1]$ by

$$f(x) = \frac{1}{x + 2} + 1/2.$$

Then f is an injection. Let $\mathbb{R}^-$ be the negative real numbers, and define $g : \mathbb{R}^- \to [0, 1/2)$ by

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

Then g is an injection. Let $h : \mathbb{R} \to [0, 1]$ be the union of the functions f and g. Then h is clearly an injection. The identity function on $[0, 1]$ is an injection into $\mathbb{R}$. By the Schröder-Bernstein Theorem,

$$| [0, 1] | = | \mathbb{R} | .$$

□

We investigate the relationship between infinite decimal expansions (which are related to infinite decimal sequences) and the real numbers. We restrict our attention to infinite decimal expansions of numbers in the unit interval $[0, 1]$.

DEFINITION. Infinite decimal expansion For all $n \in \mathbb{N}^+$, let $a_n \in \ulcorner 10 \urcorner$. Then

$$.a_1 a_2 \ldots a_n \ldots$$

is an infinite decimal expansion.

Let s be an infinite decimal expansion $.a_1a_2\ldots$. For $n \in \mathbb{N}$, let

$$s_n := .a_1 \ldots a_n = \sum_{k=1}^{n} a_k 10^{-k}.$$

We want to associate infinite decimal expansions with real numbers (understood as Dedekind cuts). We interpret infinite decimal expansions as Cauchy sequences of rational numbers.

Let D be the set of infinite decimal expansions, and let $f : D \to \mathbb{R}$ be defined by

$$f(.a_1 \ldots) = \lim_{n\to\infty} s_n.$$

The sequence $\langle s_n \rangle$ is a Cauchy sequence so it converges to a real number. Let

$$L := \{x \in \mathbb{Q} \mid (\exists\, n \in \mathbb{N})\ x < s_n\}.$$

The set L is a Dedekind cut and $f(s) = L$. That is

$$\lim_{n\to\infty} s_n = L.$$

L is the least upper bound of the set $\{s_n \mid n \in \mathbb{N}\}$. We can associate with every infinite decimal expansion a real number in the unit interval, and can thereby define a function $f : D \to [0, 1]$. Is f a surjection? That is, can every real number in the unit interval be realized as an infinite decimal expansion? Let $x \in [0, 1]$. We define an increasing sequence of rational numbers converging to x. For $n \in N^+$, let s_n be the largest decimal expansion to n decimal places that is no greater than x. If $n < m$, then s_n is a truncation of s_m. Let

$$s = \lim_{n\to\infty} s_n.$$

Then $f(s) = x$. Therefore f is a surjection onto $[0, 1]$.

It would be ideal if f were an injection, for it would follow that Dedekind cuts are just the infinite decimal expansions. However this is not true. Suppose that

$$s = .a_1 \ldots a_n a_{n+1} \ldots$$

where $a_n \neq 9$ and for all $k > n$, $a_k = 9$. If

$$t = .a_1 \ldots a_{n-1}(a_n + 1)000 \ldots$$

then

$$f(s) \;=\; f(t).$$

If neither s nor t are infinite decimal expansions that terminate in repeating 9's, and $s < t$, then there is some n such that $s < t_n$. So the rational number $(s_n + t_n)/2$ is in the Dedekind cut $f(t)$ and not in $f(s)$, so $f(s) \neq f(t)$. Therefore we have proved the following theorem.

THEOREM 8.17. *Let D_0 be the set of infinite decimal expansions for numbers in the unit interval. Let $f : D_0 \to [0,1]$ be defined by*

$$f(.a_1a_2\ldots) \;=\; \lim_{n\to\infty} .a_1\ldots a_n \;=\; \sum_{k=1}^{\infty} a_k 10^{-k}.$$

Then f is a surjection. Moreover, two distinct decimal expansions are identified by f iff one of them is of the form $.a_1a_2\ldots a_n9999\ldots$ with $a_n \neq 9$ and the other is $.a_1a_2\ldots(a_n+1)000\ldots$.

COROLLARY 8.18. $|\,[0,1]\,| = \; 2^{\aleph_0}$.

PROOF. By Lemma 8.17, Proposition 6.15 and Theorem 6.11,

$$|\,[0,1]\,| \leq |\,D_0\,| = |\,{}^{\ulcorner}10^{\neg\mathbb{N}}\,| = \; 2^{\aleph_0}.$$

Let $g : {}^{\ulcorner}2^{\neg\mathbb{N}^+} \to D_0$ be defined by

$$g(\langle a_n\rangle) \;=\; .a_1a_2\ldots$$

and $h : D_0 \to [0,1]$ be defined as in the argument for Theorem 8.17. Then $h \circ g : {}^{\ulcorner}2^{\neg\mathbb{N}} \to [0,1]$ is an injection, and so

$$2^{\aleph_0} \leq |\,[0,1]\,|\,.$$

By the Schröder-Bernstein Theorem,

$$|\,[0,1]\,| = \; 2^{\aleph_0}.$$

□

COROLLARY 8.19. $|\,\mathbb{R}\,| = \; 2^{\aleph_0}$.

8.10. Order-Completeness

We give an argument for the uncountability of $\mathbb{R}$ depending only on its abstract order properties.

DEFINITION. Order-complete Let $(X, \leq)$ be a linearly ordered set. It is called order-complete if, whenever A and B are non-empty subsets of X with the property that

$$(\forall\ a\ \in\ A)\ (\forall\ b\ \in\ B) \quad a \leq b,$$

then there exists c in X such that

$$(\forall\ a\ \in\ A)\ (\forall\ b\ \in\ B) \quad a \leq c \leq b. \tag{8.20}$$

Note that any order-complete set must have the least upper bound property — if A is any non-empty bounded set, let B be the set of all upper bounds for A, and then c from (8.20) is the (unique) least upper bound for A.

DEFINITION. Dense Let $(X, \leq)$ be a linearly ordered set, and $Y \subseteq X$. We say Y is dense in X if

$$(\forall\ a < b\ \in\ X)\ (\exists\ y\ \in\ Y)\ a < y < b.$$

DEFINITION. Extension Let $(X, \leq_X)$ and $(Y, \leq_Y)$ be linearly ordered sets. We say $(Y, \leq_Y)$ is an extension of $(X, \leq_X)$ if $X \subseteq Y$ and, for all x_1, x_2 in X,

$$x_1\ \leq_X\ x_2 \quad \text{iff} \quad x_1\ \leq_Y\ x_2.$$

THEOREM 8.21. *Let $(X, \leq)$ be an extension of $(\mathbb{Q}, \leq)$. If $(X, \leq)$ is order-complete and $\mathbb{Q}$ is dense in X, then X is uncountable.*

PROOF. Suppose that X is a countable order-complete extension of $\mathbb{Q}$ and that $\mathbb{Q}$ is dense in X.

Let the sequence $\langle a_n \mid n \in \mathbb{N} \rangle$ be a bijection from $\mathbb{N}$ to X. Observe that the sequence imposes an ordering on X. Let $\preceq$ be defined on X by

$$(\forall\, m, n \in \mathbb{N})\ a_m \preceq a_n \iff m \leq n.$$

That is, for any $x, y \in X$, $x \preceq y$ if x appears in the sequence $\langle a_n \rangle$ before y. Then $\preceq$ is a well-ordering of X.

Given $Y \subseteq X$ and $y_0 \in Y$, we say that y_0 is the $\preceq$-minimal element of Y if

$$(\forall\, x \in Y)\ y_0 \preceq x.$$

So every subset of X has a $\preceq$-minimal element.

We shall define two subsequences of $\langle a_n \rangle$, called $\langle a_{f(n)} \rangle$ and $\langle a_{g(n)} \rangle$, so that for any $n \in \mathbb{N}$

(1) $f(n+1) > g(n)$
(2) $g(n) > f(n)$
(3) $a_{f(n+1)}$ is the $\preceq$-minimal element of the set

$$\{y \in X \mid a_{f(n)} < y < a_{g(n)}\}$$

(4) $a_{g(n+1)}$ is the $\preceq$-minimal element of the set

$$\{y \in X \mid a_{f(n+1)} < y < a_{g(n)}\}.$$

We define the subsequences by recursion using the sequence $\langle a_n \rangle$ to carefully control the construction. This argument is called a back-and-forth argument. Given finite sequences of length N satisfying the properties enumerated above, we define $a_{f(N+1)}$ subject to constraints imposed by $a_{f(N)}$ and $a_{g(N)}$. We then define $a_{g(N+1)}$ subject to constraints imposed by $a_{f(N+1)}$ and $a_{g(N)}$. We then define $a_{f(N+2)}$, $a_{g(N+2)}$, and so on.

Let $f(0) = 0$. So $a_{f(0)} = a_0$. Let $g(0)$ be the smallest integer n such that $a_0 < a_n$. Note that this is equivalent to defining $g(0)$ so that $a_{g(0)}$ is the $\preceq$-minimal element of X greater than a_0. Assume we have defined finite subsequences $\langle a_{f(n)} \mid n \leq N \rangle$, $\langle a_{g(n)} \mid n \leq N \rangle$ satisfying the order properties listed above. We shall define $a_{f(N+1)}$ and $a_{g(N+1)}$ satisfying the ordering properties listed above. The set X contains the rational numbers and since $\mathbb{Q}$ is dense in X, there is an element of X, x, such that

$$a_{f(N+1)} < x < g_{(N+1)}.$$

Let $a_{f(N+1)}$ be the $\preceq$-minimal element of X such that

$$a_{f(N)} < a_{f(N+1)} < a_{g(N)}.$$

Since $\preceq$ is a well-ordering of X, $f(N+1)$ is well-defined. We let $a_{g(N+1)}$ be the $\preceq$-minimal element of X such that

$$a_{f(N+1)} < a_{g(N+1)} < a_{g(N)}.$$

By our previous discussion, $g(N+1)$ is well-defined. Observe that for any $m, n \in \mathbb{N}$,

$$a_{f(m)} < a_{g(n)}.$$

Therefore the increasing sequence $\langle a_{f(n)} \mid n \in \mathbb{N}\rangle$ is bounded above, and by Lemma 8.5, the sequence converges to its least upper bound, a.

For any $n \in \mathbb{N}$,

$$a_{f(n)} < a < a_{g(n)}.$$

So a is not a term of either subsequence. We show that a is not a term in the sequence $\langle a_n\rangle$. Suppose by way of contradiction that $a = a_n$ for some $n \in \mathbb{N}$. Since $f(0) = 0$, $n \neq 0$. Let

$$Y = (f[\mathbb{N}] \cup g[\mathbb{N}]) \cap \ulcorner n \urcorner.$$

Then $Y \neq \emptyset$ is finite, and has a maximal element.

If the maximal element of Y is $f(0)$, then for every $1 \leq k < n$, we must have $a_k < a_0$. But then $g(0)$ would be n, which contradicts the fact that n is not in the range of g.

If the maximal element of Y is $f(m+1)$ for some m, then $g(m+1) > n$, and

$$f(m+1) < n < g(m+1).$$

However

$$a_{f(m+1)} < a_n < a_{g(m+1)} < a_{g(m)}.$$

This is impossible since $a_{g(m+1)}$ is the $\preceq$-minimal element of X in the open interval $(a_{f(m+1)}, a_{g(m)})$.

If the maximal element of Y is $g(m)$ for some m, then $f(m+1) > n$ and

$$g(m) < n < f(m+1).$$

However

$$a_{f(m)} < a_{f(m+1)} < a_n < a_{g(m)}.$$

This is impossible since $a_{f(m+1)}$ is the $\preceq$-minimal element of X in the open interval $(a_{f(m)}, a_{g(m)})$. So a is not a term in the sequence $\langle a_n \rangle$. Therefore there is no bijection from $\mathbb{N}$ to X, and X is uncountable. □

By Exercise 8.20, $\mathbb{Q}$ is dense in $\mathbb{R}$. As the set of real numbers is order-complete by the least upper bound theorem, we get:

COROLLARY 8.22. *The set of real numbers is uncountable.*

THEOREM 8.23. *Let $(X, \leq_X)$ be an order-complete extension of $\mathbb{Q}$ in which $\mathbb{Q}$ is dense, and such that X has no maximal or minimal element. Then there is an order-preserving bijection from $\mathbb{R}$ onto X that is the identity on $\mathbb{Q}$.*

PROOF. Let us define a map $f : \mathbb{R} \to X$. If $q \in \mathbb{Q}$, define $f(q) = q$. If $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, define $f(\alpha)$ to be the least upper bound in X of $\{q \in \mathbb{Q} \mid q \leq \alpha\}$. The function f is well-defined, because X has the Least Upper Bound Property. It is injective, because if $\alpha \neq \beta$, there are rational numbers between α and β.

To show f is onto, suppose $x \in X$. Define $\alpha \in \mathbb{R}$ to be the least upper bound in $\mathbb{R}$ of $\{q \in \mathbb{Q} \mid q \leq_X x\}$. Then $f(\alpha) = x$.

Finally, f is order-preserving because if $\alpha \leq \beta$, then $f(\beta)$ is defined as the least upper bound of a superset of the set whose least upper bound is $f(\alpha)$, and so $f(\alpha) \leq_X f(\beta)$. □

REMARK. What happens if we drop the requirement that X have no maximal or minimal element?

8.11. Exercises

EXERCISE 8.1. Let S be the successor function in Definition 8.1. Prove that

$$S(\emptyset) \neq \emptyset.$$

Prove that for any set X,

$$S(X) \neq X.$$

EXERCISE 8.2. Prove that no proper subset of $\mathbf{N}$ (see equation 8.1) is inductive.

EXERCISE 8.3. Let $\mathcal{F} = \{X_\alpha \mid \alpha \in Y\}$ be a family of inductive sets indexed by Y. Prove that

$$\bigcap_{\alpha \in Y} X_\alpha$$

is inductive.

EXERCISE 8.4. Prove that addition and multiplication in $\mathbb{N}$ (as formally defined in Section 8.1) are associative, commutative and distributive.

EXERCISE 8.5. Prove that the relation $\leq$ defined on $\mathbb{N}$ in Section 8.1 is a linear ordering of $\mathbb{N}$.

EXERCISE 8.6. Prove that addition and multiplication in $\mathbb{Z}$ (as formally defined in Section 8.2) are associative, commutative and distributive.

EXERCISE 8.7. Prove that the relation $\leq$ defined on $\mathbb{Z}$ in Section 8.2 is a linear ordering of $\mathbb{Z}$.

EXERCISE 8.8. Prove that $\leq$ is a well ordering of $\mathbb{N}$ but not of $\mathbb{Z}$ (using the formal definition of the relation).

EXERCISE 8.9. Prove that addition and multiplication in $\mathbb{Z}$ and the relation $\leq$ on $\mathbb{Z}$ extends the operations and relation on $\mathbb{N}$. Let $I : \mathbb{N} \to \mathbb{Z}$ be defined by

$$I(n) = [\langle n, 0 \rangle].$$

Prove that I is an injection and that for all $m, n \in \mathbb{N}$,

$$I(m+n) = I(m) + I(n), \tag{8.24}$$

$$I(m \cdot n) \;=\; I(m) \cdot I(n) \tag{8.25}$$

and

$$m \leq n \;\Rightarrow\; I(m) \leq I(n). \tag{8.26}$$

Note that the operations on the left hand sides of equations 8.24 and 8.25 are defined in $\mathbb{N}$ and on the right hand side are defined in $\mathbb{Z}$. Similarly, the antecedent of statement 8.26 is defined in $\mathbb{N}$ and the consequence is defined in $\mathbb{Z}$.

EXERCISE 8.10. Prove that addition and multiplication $\mathbb{Q}$ (as formally defined in Section 8.3) are associative, commutative and distributive.

EXERCISE 8.11. Prove that the relation $\leq$ defined on $\mathbb{Q}$ in Section 8.3 is a linear ordering of $\mathbb{Q}$.

EXERCISE 8.12. Prove that addition and multiplication on $\mathbb{Q}$ and the relation $\leq$ on $\mathbb{Q}$ extends the operations and relation on $\mathbb{Q}$. Let $I : \mathbb{Z} \to \mathbb{Q}$ be defined by

$$I(a) \;=\; [\langle a, 1\rangle].$$

Prove that I is an injection and that for all $a, b \in \mathbb{Z}$,

$$I(a + b) \;=\; I(a) + I(b), \tag{8.27}$$

$$I(a \cdot b) \;=\; I(a) \cdot I(b) \tag{8.28}$$

and

$$a \leq b \;\Rightarrow\; I(a) \leq I(b). \tag{8.29}$$

Note that the operations on the left hand sides of equations 8.27 and 8.28 are defined in $\mathbb{Z}$ and on the right hand side are defined in $\mathbb{Q}$. Similarly, the antecedent of statement 8.29 is defined in $\mathbb{Z}$ and the consequence is defined in $\mathbb{Q}$.

EXERCISE 8.13. Prove that every non-zero element of $\mathbb{Q}$ has a multiplicative inverse in $\mathbb{Q}$.

EXERCISE 8.14. Prove statements (1), (2) and (3) in Section 8.4.

EXERCISE 8.15. Prove Theorem 8.2.

EXERCISE 8.16. Let $X \subseteq \mathbb{R}$, $Y \subseteq \mathbb{R}$ and let every element of X be less than every element of Y. Prove that there is $a \in \mathbb{R}$ satisfying

$$(\forall\, x \in X)(\forall\, y \in Y)\ x \leq a \leq y.$$

EXERCISE 8.17. Let $X \subseteq \mathbb{R}$ be bounded above. Prove that the least upper bound of X is unique.

EXERCISE 8.18. Let $X \subseteq \mathbb{R}$ be bounded below. Prove that X has a greatest lower bound.

EXERCISE 8.19. Only the special case of the Bolzano-Weierstrass Theorem (Theorem 8.6) was proved (where $[b, c]$ is the closed unit interval, $[0, 1]$). Generalize the proof to arbitrary $b, c \in \mathbb{R}$ where $b \leq c$.

EXERCISE 8.20. Let $X \subseteq \mathbb{R}$. We say that X is dense in $\mathbb{R}$ if given any $a, b \in \mathbb{R}$ with $a < b$, there is $x \in X$ such that

$$a \leq x \leq b.$$

a) Prove that $\mathbb{Q}$ is dense in $\mathbb{R}$.
b) Prove that $\mathbb{R} \setminus \mathbb{Q}$ is dense in $\mathbb{R}$.

EXERCISE 8.21. Let $\langle a_n \rangle$ be an injective sequence. What is the cardinality of the set of all subsequences of $\langle a_n \rangle$? What can you say about the set of subsequences of a non-injective sequence?

EXERCISE 8.22. Let s be an infinite decimal expansion, and for any $n \in \mathbb{N}^+$, let s_n be the truncation of s to the n^{th} decimal place. Prove that the sequence $\langle s_n \rangle$ is a Cauchy sequence.

EXERCISE 8.23. Let $\langle a_n \rangle$ be a convergent sequence and $\langle a_{f(n)} \rangle$ be a subsequence of $\langle a_n \rangle$. Prove that

$$\lim_{n \to \infty} a_n \ = \ \lim_{n \to \infty} a_{f(n)}.$$

EXERCISE 8.24. Prove the following generalization of the triangle inequality: if the series $\sum_{n=0}^{\infty} a_n$ converges, then

$$\left| \sum_{n=0}^{\infty} a_n \right| \leq \sum_{n=0}^{\infty} | \, a_n \, | \, .$$

EXERCISE 8.25. Let f be a real function continuous at a, and let $\langle a_n \rangle$ be a sequence converging to a. Prove that

$$\lim_{n \to \infty} \; f(a_n) \; = \; f(a).$$

EXERCISE 8.26. Give an example of a continuous function on an open interval that achieves its extreme values on the interval. Give an example of a continuous function defined on an open interval that does not achieve its extreme values on the interval.

EXERCISE 8.27. Complete the proof of Theorem 8.12 — that is, prove the result for $f(c)$ a minimum value of f on (a, b).

EXERCISE 8.28. Prove Corollary 8.15.

EXERCISE 8.29. Prove that any continuous injective real function on an interval is monotonic on that interval.

EXERCISE 8.30. Prove that there is no continuous bijection from $(0, 1)$ to $[0, 1]$.

EXERCISE 8.31. Prove that every polynomial in $\mathbb{R}[x]$ of odd degree has at least one real root.

EXERCISE 8.32. Prove that if you have a square table, with legs of equal length, and a continuous floor, you can always rotate the table so that all 4 legs are simultaneously in contact with the floor. (Hint: Apply the Intermediate value theorem to an appropriately chosen function). This is one of the earliest applications of mathematics to coffee-houses.

EXERCISE 8.33. The proof of Proposition 8.16 requires that non-zero real numbers have reciprocals (and hence quotients of real numbers are well-defined). Prove that non-zero real numbers have reciprocals.

EXERCISE 8.34. Show that there are exactly 4 order-complete extensions of $\mathbb{Q}$ in which $\mathbb{Q}$ is dense.

CHAPTER 9

Complex Numbers

9.1. Cubics

How does one find the roots of a cubic polynomial? The Babylonians knew the quadratic formula in the second millennium BC, but a formula for the cubic was only found in the 16^{th} century. The history of the discovery is complicated, but most of the credit should go to Nicolo Tartaglia. The solution was published in 1545 in Girolomo Cardano's very influential book *Artis magnae sive de regulis algebraicis liber unus*. Formula 9.2 is known today as the Tartaglia-Cardano formula. For a historical account, see *e.g.* [**6**].

Consider a cubic polynomial in $\mathbb{R}[x]$

$$p(x) \;=\; a_3x^3 + a_2x^2 + a_1x + a_0. \tag{9.1}$$

If we want to find the roots, there is no loss of generality in assuming that $a_3 = 1$, since the zeroes of p are the same as the zeroes of $\frac{1}{a_3}p$.

The second simplification is that we can assume $a_2 = 0$. Indeed, make the change of variable

$$x \;=\; y - \beta,$$

for some β to be chosen later. Then

$$\begin{aligned} p(x) &= x^3 + a_2x^2 + a_1x + a_0 \\ &= (y-\beta)^3 + a_2(y-\beta)^2 + a_1(y-\beta) + a_0 \\ &= y^3 + [a_2 - 3\beta]y^2 + [a_1 - 2a_2\beta + 3\beta^2]y + [a_0 - a_1\beta + a_2\beta^2 - \beta^3] \\ &=: q(y). \end{aligned}$$

Choose $\beta = a_2/3$. Then the coefficient of y^2 in $q(y)$ vanishes. Suppose you can find the roots of q, call them $\alpha_1, \alpha_2, \alpha_3$. Then the roots of the original polynomial p are $\alpha_1 - \beta, \alpha_2 - \beta, \alpha_3 - \beta$.

Therefore it is sufficient to find a formula for the roots of a cubic in which the quadratic term vanishes. This is called a *reduced cubic*. As there are now only two coefficients left, we shall drop the subscripts and write our reduced cubic as

$$q(x) \;=\; x^3 + ax + b. \tag{9.2}$$

The key idea is to make another, more ingenious, substitution. Let us introduce a new variable w, related to x by

$$x \;=\; w + \frac{c}{w}, \tag{9.3}$$

where c is a constant we shall choose later. Then

$$\begin{aligned} q(x) &= \left(w + \frac{c}{w}\right)^3 + a\left(w + \frac{c}{w}\right) + b \\ &= w^3 + [3c + a]w + [3c^2 + ac]\frac{1}{w} + c^3\frac{1}{w^3} + b. \end{aligned} \tag{9.4}$$

Choose

$$c \;=\; -\frac{a}{3},$$

so both the coefficient of w and $1/w$ in (9.4) vanish. Then finding x so that $q(x) = 0$ is the same as finding w so that

$$\begin{aligned} w^3 + \frac{c^3}{w^3} + b \;&=\; 0 \\ \iff\; w^6 + bw^3 + c^3 \;&=\; 0. \end{aligned} \tag{9.5}$$

Equation (9.5) is of degree 6, which seems worse than the original cubic; but so many terms vanish that it is actually a quadratic equation in w^3. Therefore it can solved by the quadratic formula:

$$w^3 \;=\; \frac{-b \pm \sqrt{b^2 - 4c^3}}{2}. \tag{9.6}$$

Knowing w, we can recover x by

$$x \;=\; w + \frac{c}{w} \;=\; w - \frac{a}{3w}.$$

So we arrive at the Tartaglia-Cardano formula for the roots of the reduced cubic (9.2):

$$x \;=\; \left[\frac{-b \pm \sqrt{b^2 + \frac{4a^3}{27}}}{2}\right]^{1/3} \;-\; \frac{a}{3\left[\frac{-b \pm \sqrt{b^2 + \frac{4a^3}{27}}}{2}\right]^{1/3}}. \tag{9.7}$$

How does the formula work in practice?

EXAMPLE 9.8. Let $p(x) = x^3 - 3x + 2$. Then $c = 1$, and (9.6) says $w^3 \;=\; -1$. Therefore $w = -1$, and so $x = -2$ is a root. Therefore, by Lemma 4.13, $(x + 2)$ is a factor of p. Factoring, we get

$$x^3 - 3x + 2 \;=\; (x + 2)(x^2 - 2x + 1).$$

The last term factors as $(x - 1)^2$, so we conclude that the roots are $-2, 1, 1$.

In Example 9.8, the formula worked, but only gave us one of the roots. Consider the next example:

EXAMPLE 9.9. Let

$$p(x) \;=\; x^3 - 3x + 1. \tag{9.10}$$

Then $c = 1$, and

$$w^3 \;=\; \frac{-1 \pm \sqrt{-3}}{2}. \tag{9.11}$$

Now we have a worse problem: w^3 involves the square root of a negative number, and even if we make sense of that, we then have to extract a cube root. Is this analagous to trying to solve the quadratic equation

$$q(x) \;:=\; x^2 + x + 1 \;=\; 0?$$

The quadratic formula again gives the right-hand side of (9.11), and we explain this by saying that in fact q has no real roots. Indeed, graphing shows that q looks like Figure 9.12.

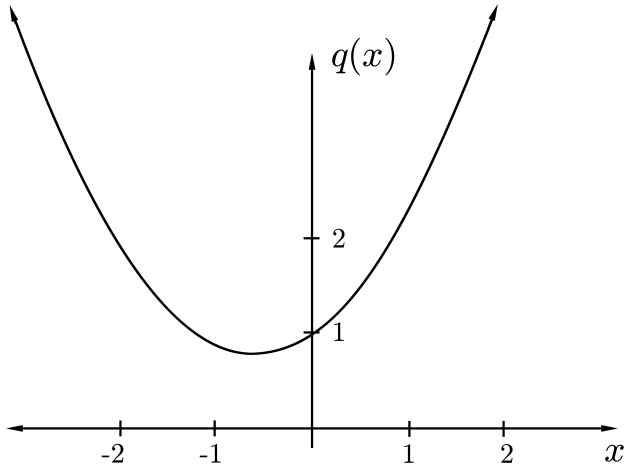

FIGURE 9.12. Plot of $q(x) = x^2 + x + 1$

But this cannot be the case for p. Indeed,

$$\begin{aligned} p(-2) &= -1 < 0 \\ p(0) &= 1 > 0 \\ p(1) &= -1 < 0 \\ p(2) &= 3 > 0. \end{aligned}$$

Therefore, by the Intermediate Value Theorem 8.10, p must have a root in each of the intervals $(-2, 0)$, $(0, 1)$ and $(1, 2)$. As p can have at most 3 roots by Theorem 4.10, it must therefore have exactly three roots. A graph of p looks like Figure 9.13.

It turns out that one can find the roots of p in Example 9.9 by correctly interpreting the Tartaglia-Cardano formula. We shall come back to this example in Section 9.3, after we develop the necessary ideas. The big idea is to introduce the notion of a *complex number*.

9.2. Complex Numbers

DEFINITION. Complex number A complex number is an expression of the form $a + ib$, where a and b are real numbers.

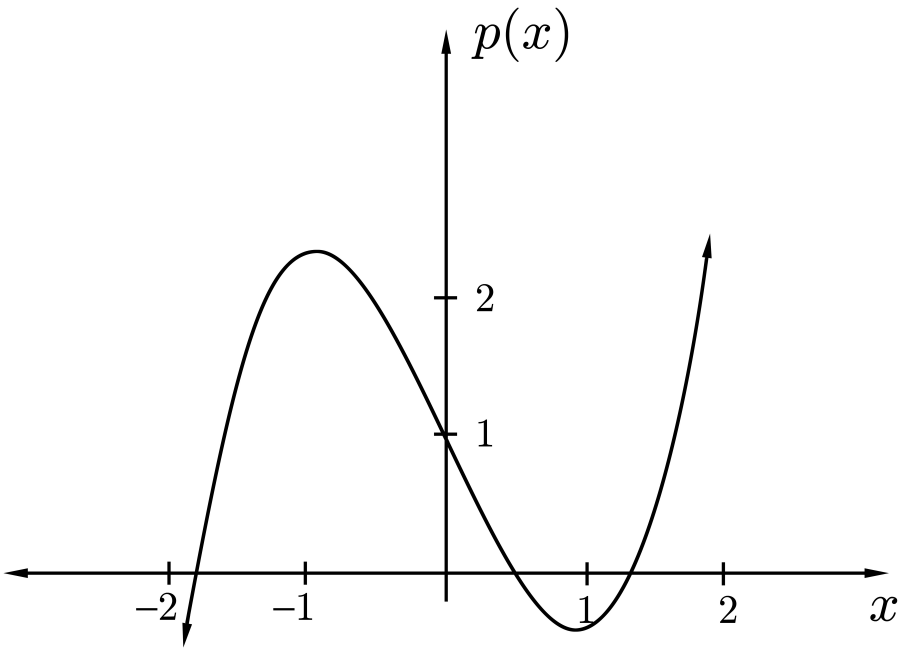

FIGURE 9.13. Plot of $p(x) = x^3 - 3x + 1$

For the moment, you can think of the i in $a+ib$ as a formal symbol, or a place-holder. Later, we shall see that it has another interpretation.

NOTATION. $\mathbb{C}$ *We shall let* $\mathbb{C}$ *denote the set of all complex numbers:*

$$\mathbb{C} \;=\; \{a + ib \;:\; a, b \in \mathbb{R}\}.$$

As a set, one can identify $\mathbb{C}$ with $\mathbb{R}^2$ in the obvious way. This allows us to define addition; what is not so obvious is that there is also a good definition for multiplication.

DEFINITION. Let $a+ib$ and $c+id$ be complex numbers. Then their sum and product are defined by

$$(a + ib) + (c + id) \quad = \quad (a + c) + i(b + d) \tag{9.14}$$

$$(a + ib) \times (c + id) \quad = \quad (ac - bd) + i(ad + bc). \tag{9.15}$$

The formula for the sum (9.14) is just what you would get if you identified the complex number $a + ib$ with the vector (a, b) in $\mathbb{R}^2$ and

used vector addition. The product is more subtle. If you multiply out the left-hand side of (9.15), you get

$$ac + i(ad + bc) + i^2 bd.$$

One arrives at the right-hand side of (9.15) by *defining*

$$i^2 = -1. \tag{9.16}$$

So i is the square root of -1; that is, it is an algebraic quantity we introduce that is defined to have the property that its square is -1. Obviously this precludes i from being a real number.

In essence we have continued the program of defining number systems that we began in Chapter 8. Addition and multiplication of complex numbers have been defined by algebraic operations on $\mathbb{R}\times\mathbb{R}$. Since algebraic operations on the real numbers were defined set-theoretically, we have thereby defined algebraic operations on $\mathbb{C}$ by set operations. Unlike the other numbers systems we have defined, we do not define a linear ordering of $\mathbb{C}$. It is not generally useful to think of complex numbers on a number line. However it is very useful to think of complex numbers as points in the plane $\mathbb{R}^2$, and to describe them in polar coordinates.

As usual, the point with Cartesian coordinates (x, y) has polar coordinates (r, θ), where they are related by

$$\begin{aligned} r &= \sqrt{x^2 + y^2} & \tan(\theta) &= y/x \\ x &= r\cos\theta & y &= r\sin\theta. \end{aligned}$$

So the complex number $z = x + iy$ can also be written as

$$z = r(\cos\theta + i\sin\theta). \tag{9.18}$$

The form (9.18) is so widely used that there is a special notation for it.

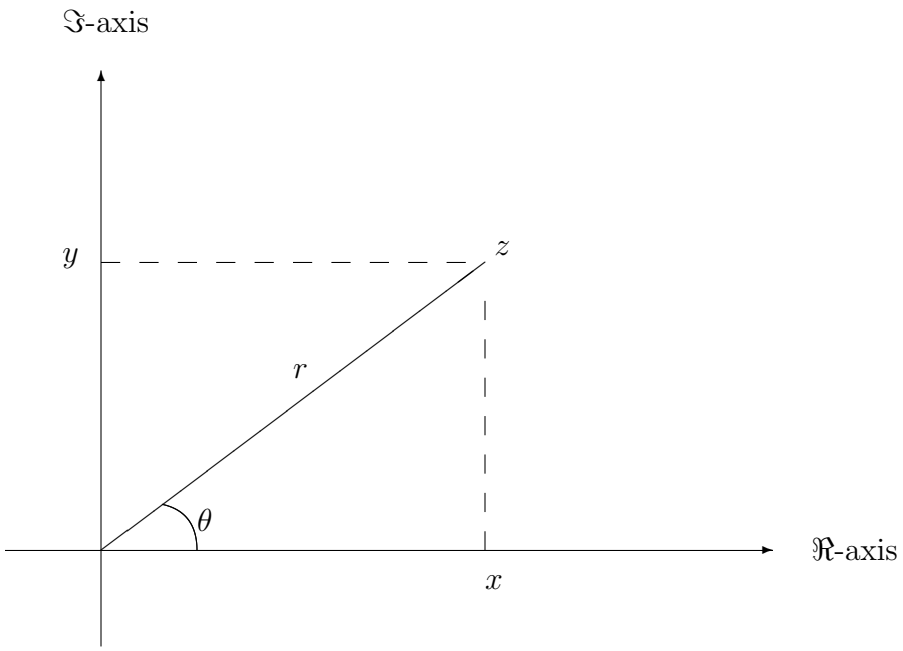

FIGURE 9.17. Polar Coordinates

NOTATION. Cis

$$\text{Cis}(\theta) \;:=\; \cos\theta + i\sin\theta.$$

DEFINITION. For the complex number $z = x + iy = r\text{Cis}(\theta)$, we have the following:
$\Re(z)$ x is called the real part of z, written $\Re(z)$;
$\Im(z)$ y is called the imaginary part of z, written $\Im(z)$;
$|z|$ r is called the modulus of z, or absolute value of z, written $|z|$;
$\arg(z)$ θ is called the argument of z, written $\arg(z)$.
$\bar{z}$ The number $x - iy$ is called the conjugate of z, written $\bar{z}$.

REMARK. There is an important point to bear in mind about the argument: it is only unique up to addition of multiples of 2π. In other words, if θ_0 is an argument of the complex number z, then so are all the numbers $\{\theta_0 + 2k\pi \;:\; k \in \mathbb{Z}\}$.

Addition is easiest in Cartesian coordinates: add the real and imaginary parts. Multiplication is easiest in polar coordinates: multiply the moduli and add the arguments.

PROPOSITION 9.19. *Let* $z_1 = r_1\text{Cis}(\theta_1)$ *and* $z_2 = r_2\text{Cis}(\theta_2)$. *Then*

$$z_1 z_2 \;=\; r_1 r_2 \text{Cis}(\theta_1 + \theta_2).$$

PROOF. Multiplying out, we get

$$\begin{aligned} z_1 z_2 \;=\; & r_1 r_2 [\cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2 \\ & +\; i\,(\cos\theta_1 \sin\theta_2 + \cos\theta_2 \sin\theta_1)]. \end{aligned}$$

The result follows by the trigonometric identities for the cosine and sine of the sum of two angles. □

A consequence of Proposition 9.19 is the following formula for raising a complex number to a power, called De Moivre's theorem.

THEOREM 9.20. *De Moivre's Theorem Let* $z = r\text{Cis}(\theta)$ *be a non-zero complex number, and let* $n \in \mathbb{Z}$. *Then*

$$z^n \;=\; r^n \text{Cis}(n\theta). \tag{9.21}$$

PROOF. If $n \geq 0$, then (9.21) can be proved by induction from Proposition 9.19. For n negative, it is enough to observe that by Proposition 9.19

$$[r\text{Cis}(\theta)]\left[r^{-1}\text{Cis}(-\theta)\right] \;=\; 1\text{Cis}(0) \;=\; 1.$$

□

We can now prove that every non-zero complex number has *exactly* n distinct n^{th} roots.

THEOREM 9.22. *Let* $z = r\text{Cis}(\theta)$ *be a non-zero complex number, and let* n *be an integer greater than* 1. *Then there are exactly* n *complex numbers* w *satisfying the equation* $w^n = z$. *They are*

$$\left\{ r^{1/n}\text{Cis}\left(\frac{\theta}{n} + \frac{2k\pi}{n}\right) \;:\; k = 0, 1, \ldots, n-1 \right\}. \tag{9.23}$$

PROOF. Suppose $w = \rho\text{Cis}(\phi)$ is an n^{th} root of z. Then by De Moivre's theorem, $\rho^n = r$ and $n\phi$ is an argument of z. As ρ must be a positive real number, it is the unique positive n^{th} root of r. The number $n\phi$ can be *any* argument of z, so we have

$$n\phi \;=\; \theta + 2k\pi, \quad k \in \mathbb{Z}.$$

So ϕ can have the form

$$\frac{\theta}{n} + \frac{2k\pi}{n}$$

for any integer k. However, different ϕ's will give rise to the same complex number w if they differ by a multiple of 2π. So there are exactly n different w's that are n^{th} roots of z. □

EXAMPLE 9.24. What does Theorem 9.22 tell us are the square roots of -1? We let $r = 1$ and $\theta = \pi$, and we get the square roots are $\text{Cis}(\pi/2) = i$ and $\text{Cis}(-\pi/2) = -i$.

EXAMPLE 9.25. Find the cube roots of 1.

In the notation of Theorem 9.22, $r = 1$ and $\theta = 0$. Therefore the cube roots are

$$\begin{aligned} 1 & & & \\ \omega &= \text{Cis}(2\pi/3) &= -\frac{1}{2} + i\frac{\sqrt{3}}{2} \\ \omega^2 &= \text{Cis}(4\pi/3) &= -\frac{1}{2} - i\frac{\sqrt{3}}{2}. \end{aligned}$$

The number ω is called a *primitive cube root of unity*, because all the cube roots are obtained as $\omega, \omega^2, \omega^3$.

DEFINITION. Primitive root of unity A primitive n^{th} root of unity is a number ω such that $\{1, \omega, \omega^2, \ldots, \omega^{n-1}\}$ constitute all the n^{th} roots of 1.

PROPOSITION 9.26. *Let z be a complex number, and w_0 be some n^{th} root of z. Let ω be a primitive n^{th} root of unity. Then all the n^{th} roots of z are $\{w_0, \omega w_0, \omega^2 w_0, \ldots, \omega^{n-1} w_0\}$.*

9.3. Tartaglia-Cardano Revisited

Let us consider again Example 9.9. We wanted to find the cube roots of

$$\zeta_\pm \;=\; \frac{-1 \pm \sqrt{-3}}{2}.$$

If we take the + sign, we get

$$\zeta_+ \;=\; \mathrm{Cis}(2\pi/3),$$

and if we take the − sign, we get

$$\zeta_- \;=\; \mathrm{Cis}(4\pi/3).$$

So ζ_+ has 3 roots, namely

$$\{\mathrm{Cis}\left(\frac{2\pi}{9} + \frac{2k\pi}{3}\right) \;:\; k = 0, 1, 2\},$$

and ζ_- has 3 roots, namely

$$\{\mathrm{Cis}\left(\frac{4\pi}{9} + \frac{2k\pi}{3}\right) \;:\; k = 0, 1, 2\},$$

Knowing w, we want to find x, which for Example 9.9 is given by $w + 1/w$. For any number w that can be written as $\mathrm{Cis}(\theta)$ (*i.e.* any complex number of modulus 1), we have

$$\begin{aligned} w + \frac{1}{w} &= \cos\theta + i\sin\theta + \cos(-\theta) + i\sin(-\theta) \\ &= 2\cos\theta. \end{aligned}$$

Therefore the roots of the polynomial given in (9.10) are

$$\left\{2\cos\frac{2\pi}{9}, 2\cos\frac{8\pi}{9}, 2\cos\frac{14\pi}{9}, 2\cos\frac{4\pi}{9}, 2\cos\frac{10\pi}{9}, 2\cos\frac{16\pi}{9}\right\}. \quad (9.27)$$

Are these 6 different roots? Theorem 4.10 says that p can have at most 3 different roots. As $\cos(\theta) = \cos(2\pi - \theta)$, we see our set (9.27) may be written as

$$\{2\cos\frac{2\pi}{9}, 2\cos\frac{4\pi}{9}, 2\cos\frac{8\pi}{9}\}. \quad (9.28)$$

It turns out that the Tartaglia-Cardano formula (9.7) does give all three roots of the cubic, and moreover it does not matter whether one chooses the + or − sign, as long as one calculates all 3 cube roots of

(9.6) for some choice of sign. We shall use $\mathbb{C}[z]$ to denote the set of polynomials in z with coefficients from $\mathbb{C}$.

THEOREM 9.29. *Consider the polynomial*

$$p(z) = z^3 + az + b \tag{9.30}$$

in $\mathbb{C}[z]$*, and assume* $a \neq 0$*. Let* $c = -a/3$*, and let* ζ *be*

$$\zeta \;=\; \frac{-b + \sqrt{b^2 - 4c^3}}{2}. \tag{9.31}$$

Let w_1, w_2, w_3 *be the three distinct cube roots of* ζ*. For each* w_i*, define* z_i *by*

$$z_i \;=\; w_i + \frac{c}{w_i}.$$

Then

$$p(z) \;=\; (z - z_1)(z - z_2)(z - z_3). \tag{9.32}$$

REMARK. It will follow from the proof that it doesn't matter which square root of $b^2 - 4c^3$ one chooses in (9.31).

PROOF. If p is given by (9.32), then

$$p(z) \;=\; z^3 - (z_1+z_2+z_3)z^2 + (z_1z_2+z_2z_3+z_3z_1)z - (z_1z_2z_3). \tag{9.33}$$

We must show that the coefficients in (9.33) match those in (9.30). By Proposition 9.26, we can assume

$$w_1 \;=\; \omega w_3, \qquad w_2 \;=\; \omega^2 w_3$$

where $\omega = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$ is a primitive cube root of unity. In the following calculations, we use the facts that $\omega^2 = 1/\omega$ and $1 + \omega + \omega^2 = 0$. (Why are these true?) Notice that $w_3 \neq 0$, as that would force $c = 0$.

The coefficient of z^2 in (9.33) is

$$\begin{aligned} -(z_1 + z_2 + z_3) &= -w_3(\omega + \omega^2 + 1) \;-\; \frac{1}{w_3}(\omega^2 + \omega + 1) \\ &= 0. \end{aligned}$$

The coefficient of z is

$$\begin{aligned} z_1z_2 + z_2z_3 + z_3z_1 &= \left(\omega w_3 + c\omega^2\frac{1}{w_3}\right)\left(\omega^2 w_3 + c\omega\frac{1}{w_3}\right) \\ &\quad + \left(\omega^2 w_3 + c\omega\frac{1}{w_3}\right)\left(w_3 + c\frac{1}{w_3}\right) \\ &\quad + \left(w_3 + c\frac{1}{w_3}\right)\left(\omega w_3 + c\omega^2\frac{1}{w_3}\right) \\ &= w_3^2(1+\omega^2+\omega) \ + \ 3c(\omega+\omega^2) \ + \ \frac{c^2}{w_3^2}(1+\omega+\omega^2) \\ &= -3c \\ &= a. \end{aligned}$$

The constant term in (9.33) is

$$\begin{aligned} -z_1z_2z_3 &= -\left(\omega w_3 + c\omega^2\frac{1}{w_3}\right)\left(\omega^2 w_3 + c\omega\frac{1}{w_3}\right)\left(w_3 + \frac{1}{w_3}\right) \\ &= -w_3^3 \ - \ cw_3(1+\omega^2+\omega) \ - \ \frac{c^2}{\frac{1}{w_3}}(\omega+1+\omega^2) \ - \ \frac{c^3}{w_3^3} \\ &= -\zeta - \frac{c^3}{\zeta} \\ &= -\frac{-b+\sqrt{b^2-4c^3}}{2} \ - \ \frac{2c^3}{-b+\sqrt{b^2-4c^3}} \\ &= \frac{-b^2+2b\sqrt{b^2-4c^3}-(b^2-4c^3)-4c^3}{2(-b+\sqrt{b^2-4c^3})} \\ &= \frac{b(-b+\sqrt{b^2-4c^3})}{-b+\sqrt{b^2-4c^3}} \\ &= b. \end{aligned}$$

Therefore all the coefficients of (9.30) and (9.32) match, so they are the same polynomial. □

The Tartaglia-Cardano formula therefore gives all three roots to a reduced cubic polynomial p with complex coefficients (repeated roots can occur). If the coefficients a and b are real, we know from the Intermediate Value Theorem that at least one of the three roots of p will be real (See Exercise 8.31). As Example 9.9 shows, however, it

may still be necessary to take the cube root of a complex ζ to obtain the real roots of a real cubic. This realization was what led to the acceptance of complex numbers as useful objects rather than a bizarre fantasy.

9.4. Fundamental Theorem of Algebra

Algebra over the complex numbers is in many ways easier than over the real numbers. The reason is that a polynomial of degree N in $\mathbb{C}[z]$ has *exactly* N zeroes, counting multiplicity. This is called the Fundamental Theorem of Algebra. To prove it, we must establish some preliminary results.

9.4.1. Some Analysis.

DEFINITION. We say that a sequence $\langle z_n = x_n + iy_n \rangle$ of complex numbers *converges* to the number $z = x + iy$ iff $\langle x_n \rangle$ converges to x and $\langle y_n \rangle$ converges to y. We say the sequence is Cauchy iff both $\langle x_n \rangle$ and $\langle y_n \rangle$ are Cauchy.

REMARK. This is the same as saying that $\langle z_n \rangle$ converges to z iff $|z - z_n|$ tends to zero, and that $\langle z_n \rangle$ is Cauchy iff

$$(\forall\, \varepsilon > 0)\ (\exists\, N)\ (\forall\, m, n > N)\ |z_m - z_n| < \varepsilon.$$

DEFINITION. Let $G \subseteq \mathbb{C}$. We say a function $f : G \to \mathbb{C}$ is continuous on G if, whenever $\langle z_n \rangle$ is a sequence in G that converges to some value z_∞ in G, then $\langle f(z_n) \rangle$ converges to $f(z_\infty)$.

PROPOSITION 9.34. *Polynomials are continuous functions on* $\mathbb{C}$.

PROOF. Repeat the proof of Proposition 5.23 with complex numbers instead of real numbers. □

DEFINITION. A *closed rectangle* is a set of the form $\{z \in \mathbb{C} \mid a \leq \Re(z) \leq b,\ c \leq \Im(z) \leq d\}$ for some real numbers $a \leq b$ and $c \leq d$.

We would like a version of the Extreme Value Theorem, but it is not clear how the minimum and maximum values of a complex valued function should be defined. However, our definition of continuity makes sense even if the range of f is contained in $\mathbb{R}$, and every complex valued continuous function g has three naturally associated real-valued continuous functions, *viz.* $\Re(g), \Im(g)$ and $|g|$.

THEOREM 9.35. *Let R be a closed rectangle in $\mathbb{C}$, and $f : R \to \mathbb{R}$ a continuous function. Then f attains its maximum and its minimum.*

PROOF. Let $R = \{z \in \mathbb{C} \mid a \leq \Re(z) \leq b,\ c \leq \Im(z) \leq d\}$. Let $\langle z_n = x_n + iy_n \rangle$ be a sequence of points such that $f(z_n)$ tends to either the least upper bound of the range of f, if this exists, or let $f(z_n) > n$ for all n, if the range is not bounded above. By the Bolzano-Weierstrass Theorem 8.6, there is some subsequence for which the real parts converge to some number x_∞ in $[a, b]$. By Bolzano-Weierstrass again, some subsequence of this subsequence has the property that the imaginary parts also converge, to some point y_∞ in $[c, d]$. So, replacing the original sequence by this subsequence of the subsequence, we can assume that z_n converges to the point $z_\infty = x_\infty + iy_\infty \in R$. By continuity, $f(z_\infty) = \lim_{n\to\infty} f(z_n)$. If the original sequence were unbounded then $f(z_n) > n$ in the subsequence. This is impossible since the sequence $\langle f(z_n) \rangle$ converges to $f(z_\infty)$. Therefore the subsequence is bounded and $f(z_\infty)$ must be the least upper bound of the range of f. Therefore $f(z_\infty)$ is the maximum of f over R.

A similar argument shows that the minimum is also attained. □

REMARK. The previous theorem can be improved to show that a continuous real-valued function on a closed bounded set in $\mathbb{C}$ attains its extrema. A set F is *closed* if whenever a sequence of points $\langle z_n \rangle$ converges to some complex number z_∞, then z_∞ is in F. A set is *bounded* if it is contained in some rectangle.

We need one more geometric fact.

LEMMA 9.36. *Triangle inequality* *Let z_1, z_2 be complex numbers. Then*

$$|z_1 + z_2| \ \leq \ |z_1| + |z_2|$$

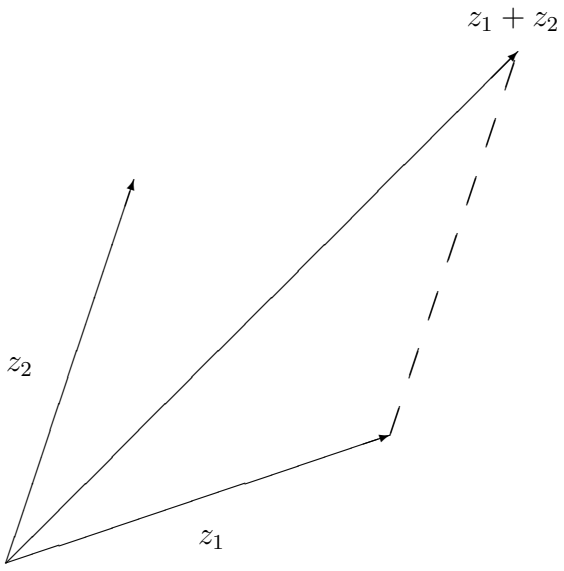

FIGURE 9.18. Triangle inequality

PROOF. Write $z_1 = r_1\text{Cis}(\theta_1)$ and $z_2 = r_2\text{Cis}(\theta_2)$. Then

$$\begin{aligned}
&|r_1\text{Cis}(\theta_1) + r_2\text{Cis}(\theta_2)| \\
&\quad = \quad \left[(r_1\cos\theta_1 + r_2\cos\theta_2)^2 + (r_1\sin\theta_1 + r_2\sin\theta_2)^2\right]^{1/2} \\
&\quad = \quad \left[r_1^2 + r_2^2 + 2r_1r_2(\cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2)\right]^{1/2} \\
&\quad = \quad \left[r_1^2 + r_2^2 + 2r_1r_2\cos(\theta_1 - \theta_2)\right]^{1/2} \\
&\quad \leq \quad \left[r_1^2 + r_2^2 + 2r_1r_2\right]^{1/2} \\
&\quad = \quad r_1 + r_2.
\end{aligned}$$

□

COROLLARY 9.38. *Let* $z_1, \ldots, z_n \in \mathbb{C}$. *Then*

$$|z_1 + \cdots + z_n| \leq |z_1| + \cdots + |z_n|.$$

9.4.2. The Proof of the Fundamental Theorem of Algebra. First we observe that finding roots and finding factors are closely related.

LEMMA 9.39. *Let* p *be a polynomial of degree* $N \geq 1$ *in* $\mathbb{C}[z]$. *A complex number,* c, *is a root of* p *iff*

$$p(z) = (z - c)q(z),$$

where q *is a polynomial of degree* $N - 1$.

PROOF. Repeat the proof of Lemma 4.13 with real numbers replaced by complex numbers. □

Now we prove D'Alembert's lemma, which states that the modulus of a polynomial cannot have a local minimum except at a root.

LEMMA 9.40. *D'Alembert's Lemma Let* $p \in \mathbb{C}[z]$ *and* $\alpha \in \mathbb{C}$. *If* $p(\alpha) \neq 0$, *then*

$$(\forall\, \varepsilon > 0)\ (\exists\, \zeta)\ [\, |\zeta - \alpha| < \varepsilon\,] \ \wedge\ [\, |p(\zeta)| < |p(\alpha)|\,]. \tag{9.41}$$

PROOF. Fix α, not a root of p. Write p as

$$p(z) = \sum_{k=0}^{N} a_k (z - \alpha)^k,$$

where neither a_0 nor a_N are 0. Let

$$m = \min\{j \in \mathbb{N}^+ \mid a_j \neq 0\}.$$

So

$$p(z) = a_0 + a_m(z - \alpha)^m + \cdots + a_N(z - \alpha)^N. \tag{9.42}$$

Let $a_0 = r_0 \mathrm{Cis}(\theta_0)$ and $a_m = r_m \mathrm{Cis}(\theta_m)$. We will choose ζ of the form

$$\zeta = \alpha + \rho \mathrm{Cis}(\phi)$$

in such a way as to get some cancellation in the first two terms of (9.42). So, let

$$\phi \;=\; \frac{\theta_0 + \pi - \theta_m}{m}.$$

Then

$$a_0 + a_m(\zeta - \alpha)^m \;=\; r_0\mathrm{Cis}(\theta_0) - r_m\rho^m\mathrm{Cis}(\theta_0).$$

It remains to show that, for ρ small enough, we can ignore all the higher order terms. Note that if $\rho < 1$, we have

$$\begin{aligned}
&|a_{m+1}(\zeta-\alpha)^{m+1} + \cdots + a_N(\zeta-\alpha)^N| \\
&\quad\leq\quad |a_{m+1}(\zeta-\alpha)^{m+1}| + \cdots + |a_N(\zeta-\alpha)^N| \\
&\quad=\quad |a_{m+1}|\rho^{m+1} + \cdots + |a_N|\rho^N \\
&\quad\leq\quad \rho^{m+1}[|a_{m+1}| + \cdots + |a_N|\,] \\
&\quad=:\quad C\rho^{m+1}.
\end{aligned}$$

Choose ρ so that $r_m\rho^m < r_0$. Then

$$p(\zeta) \;=\; (r_0 - r_m\rho^m)\mathrm{Cis}(\theta_0) + a_{m+1}(\zeta-\alpha)^{m+1} + \cdots + a_N(\zeta-\alpha)^N,$$

so

$$|p(\zeta)| \;\leq\; r_0 - r_m\rho^m + C\rho^{m+1}. \tag{9.43}$$

If $\rho < r_m/C$, the right-hand side of (9.43) is smaller than r_0.

So we conclude that by taking

$$\rho \;=\; \frac{1}{2}\min\left(1,\, \frac{r_m}{C},\, \left[\frac{r_0}{r_m}\right]^{1/m},\, \varepsilon\right)$$

then

$$\zeta \;=\; \rho\mathrm{Cis}\left(\frac{\theta_0 + \pi - \theta_m}{m}\right)$$

satisfies the conclusion of the lemma. □

THEOREM 9.44. *Fundamental Theorem of Algebra* *Let $p \in \mathbb{C}[z]$ be a polynomial of degree $N \geq 1$. Then p can be factored as*

$$p(z) \;=\; c(z-\alpha_1)\ldots(z-\alpha_N) \tag{9.45}$$

for complex numbers $c, \alpha_1, \ldots, \alpha_N$. Moreover the factoring is unique up to order.

PROOF. (i) Show that p has at least one root.

Let $p(z) = \sum_{k=0}^{N} a_k z^k$, with $a_N \neq 0$. Let S be the closed square $\{z \in \mathbb{C} \mid -L \leq \Re(z) \leq L,\ -L \leq \Im(z) \leq L\}$, where L is some (large) number to be chosen later.

If $|z| = R$ then

$$\left|\sum_{k=0}^{N-1} a_k z^k\right| \ \leq \ \sum_{k=0}^{N-1} |a_k| R^k.$$

Choose L_0 so that if $R \geq L_0$, then

$$\sum_{k=0}^{N-1} |a_k| R^k \ \leq \ \frac{1}{2}|a_N| R^N.$$

Then if $L \geq L_0$ and z is outside S, we have

$$\begin{aligned} |a_N z^N| &= \left|p(z) - \sum_{k=0}^{N-1} a_k z^k\right| \\ &\leq |p(z)| + \left|\sum_{k=0}^{N-1} a_k z^k\right| \\ &\leq |p(z)| + \frac{1}{2}|a_N| L^N, \end{aligned}$$

where the first inequality is the triangle inequality, and the second because $|z| > L$. Choose L_1 such that

$$\frac{1}{2}|a_N| L_1^N \ > \ |a_0|.$$

Let $L = \max(L_0, L_1)$, and let S be the corresponding closed square. The function $|p|$ is continuous on S, so it attains its minimum at some point, α_1 say, by Theorem 9.35. On the boundary of S, we know

$$|p(z)| \ \geq \ \frac{1}{2}|a_N| L^N \ > \ |a_0| \ = \ |p(0)|.$$

Therefore α_1 must be in the interior of S. By D'Alembert's lemma, we must have $p(\alpha_1) = 0$, or else there would be a nearby point ζ, also in S, where $|p(\zeta)|$ was smaller than $|p(\alpha_1)|$. So α_1 is a root of p.

(ii) Now we apply Lemma 9.39 to conclude that we can factor p as

$$p(z) \ = \ (z - \alpha_1) q(z)$$

where q is a polynomial of degree $N-1$. By a straightforward induction argument, we can factor p into N linear factors.

(iii) Uniqueness is obvious. The number c is the coefficient a_N. The numbers a_k are precisely the points at which the function p vanishes, as it follows from Proposition 9.19 that the product of finitely many complex numbers can be 0 if and only if one of the numbers is itself 0. □

9.5. Application to Real Polynomials

If p is a polynomial in $\mathbb{R}[x]$, it follows from the Fundamental Theorem of Algebra that it does have roots, but they may be complex. If it has complex roots, they must occur in complex conjugate pairs.

THEOREM 9.46. *Let $p \in \mathbb{R}[x]$. Let α be a root of p. Then so is $\bar{\alpha}$.*

PROOF. Let $p(x) = \sum_{k=0}^{N} a_k x^k$. Then

$$p(\alpha) \;=\; \sum_{k=0}^{N} a_k \alpha^k \;=\; 0,$$

so

$$p(\bar{\alpha}) \;=\; \sum_{k=0}^{N} a_k \bar{\alpha}^k \;=\; \overline{p(\alpha)} \;=\; 0. \qquad \square$$

Let $\alpha = a + ib$. Then

$$\begin{aligned}(x-\alpha)(x-\bar{\alpha}) &= (x-(a+ib))(x-(a-ib)) \\ &= x^2 - 2ax + a^2 + b^2 \\ &= (x-a)^2 + b^2. \end{aligned} \tag{9.47}$$

So applying the Fundamental Theorem of Algebra to the real polynomial p, we first factor out the real roots, and for each pair of complex conjugate roots we get a factor as in (9.47). Thus we get:

THEOREM 9.48. *Let $p \in \mathbb{R}[x]$ be a polynomial of degree N. Then p can be factored into a product of linear factors $(x-c_k)$ and quadratic*

factors $((x-a_k)^2+b_k^2)$*:*

$$p(x) \;=\; c\left(\prod_{k=1}^{N_1}(x-c_k)\right)\left(\prod_{j=1}^{N_2}((x-a_j)^2+b_j^2)\right)$$

for some (not necessarily distinct) real numbers c, c_j, a_j, b_j*. We have* $N_1+2N_2=N$*, and the factoring is unique, up to ordering and replacing any* b_j *by* $-b_j$*.*

9.6. Further remarks

In Chapter 5 we defined cosine and sine in terms of power series. In Section 9.2, we interpreted them geometrically and used trigonometric identities. Showing that the power series and the trigonometric interpretation are really describing the same function is part of a course in Complex Analysis.

There are two main ingredients to a first course in Complex Analysis. The first is to show that if a function f has a derivative everywhere on some open disk, in the sense that

$$\lim_{z\to z_0}\frac{f(z_0)-f(z)}{z_0-z}$$

exists, then the function is automatically analytic, *i.e.* expressible by a convergent power series. This is not true for real functions, and explains much of the special nature of complex differentiable functions.

The second part of the course concerns evaluating contour integrals of complex differentiable functions. This is useful not only in its own right, but in applications to real analysis, such as inverting the Laplace transform, or evaluating definite integrals.

A good introduction to Complex analysis is the book by Donald Sarason [**7**].

9.7. Exercises

EXERCISE 9.1. What are the primitive fourth roots of unity?

EXERCISE 9.2. Show that if ω is any n^{th} root of unity other than 1, then $1+\omega+\omega^2+\cdots+\omega^{n-1}=0$.

EXERCISE 9.3. How many primitive cube roots of unity are there? How many primitive sixth roots? How many primitive n^{th} roots for a general n?

EXERCISE 9.4. Redo Example 9.8 to get all three roots from the Tartaglia-Cardano formula.

EXERCISE 9.5. Let $p(x) = x^3 + 3x + \sqrt{2}$. Show without using the Cardano-Tartaglia formula that p has exactly one real root. Find it. What are the complex roots?

EXERCISE 9.6. Fill in the proof of Proposition 9.34.

EXERCISE 9.7. Let $g : G \to \mathbb{C}$ be a continuous function on $G \subseteq \mathbb{C}$. Show that $\Re(g)$, $\Im(g)$ and $|g|$ are continuous. Conversely, show that the continuity of $\Re(g)$ and $\Im(g)$ imply the continuity of g.

EXERCISE 9.8. Show that every continuous real-valued function on a closed, bounded subset of $\mathbb{C}$ attains its extrema.

APPENDIX A

The Greek Alphabet

Lower-case	Upper-case	Name
α	A	alpha
β	B	beta
γ	Γ	gamma
δ	Δ	delta
ε	E	epsilon
ζ	Z	zeta
η	H	eta
θ	Θ	theta
ι	I	iota
κ	K	kappa
λ	Λ	lambda
μ	M	mu
ν	N	nu
ξ	Ξ	xi
o	O	omicron
π	Π	pi
ρ	R	rho
σ	Σ	sigma
τ	T	tau
υ	Υ	upsilon
ϕ	Φ	phi
χ	X	chi
ψ	Ψ	psi
ω	Ω	omega

APPENDIX B

Axioms of Zermelo-Fraenkel with the Axiom of Choice

Russell's Paradox (Section 1.7) demonstrates that the General Comprehension Principle is false, as it gives rise to a contradiction. So how do we decide whether a definable collection is a set? This question engendered a program to *axiomatize* set theory with the objective of producing uniform assumptions about sets that satisfied numerous constraints:

- The axioms are understandable and intuitively sound. We must be able to recognize when a statement about sets is an axiom.
- The axioms are sufficient to derive the standard theorems of mathematics.
- The axioms are not redundant. That is, no axiom can be derived from the remaining axioms.
- Every mathematical statement about sets is either provable or refutable from the axioms.
- The axioms are logically consistent and hence do not give rise to a contradiction.

As we will discuss later, no collection of axioms can simultaneously achieve these objectives. First we give the axioms on which mathematicians ultimately settled.

Axioms of Zermelo-Fraenkel (with the Axiom of Choice):

(1) **Extensionality** If sets X and Y have the same elements, then $X = Y$.
(2) **Pairing** For any sets X and Y, there is a set $Z = \{X, Y\}$.

(3) **Union** Let X be a set of sets. Then there is a set

$$\{x \mid (\exists Y \in X)\, x \in Y\}.$$

(4) **Power Set** If X is a set then the collection of all subsets of X is a set.
(5) **Infinity** There is an inductive set.
(6) **Schema of Separation** If $P(x, y_1, \ldots, y_n)$ is a formula with $n+1$ variables, and $X, X_1, \ldots, X_n$ are sets, then there is a set

$$\{x \in X \mid P(x, X_1, \ldots, X_n\}.$$

(7) **Schema of Replacement** If F is a function on arbitrary collections, X is a set and $f = F|_X$, then the range of f is a set.
(8) **Regularity** Let X be a set. Then there is no infinite sequence of elements of X, $\langle x_i \rangle$, such that for all $n \in \mathbb{N}$, $x_{n+1} \in x_n$.
(9) **Choice** Let X be a set of non-empty sets. Then there is a function f with domain X such that for all $x \in X$, $f(x) \in x$.

The axioms of Zermelo-Fraenkel with the Axiom of Choice are referred to as ZFC. There are seven axioms and two axiom *schemata*. The schemata give infinitely many axioms. The Extensionality Axiom characterizes set identity. It says that a set is defined by its members. The Axioms of Pairing, Union and Power Set guarantee that collections built from sets with these set operations will be sets. The Axiom of Infinity implies that the natural numbers are a set. The Schema of Separation says that any subset of a given set defined by a formula is a set. It is a weakened version of the General Comprehension Principle. The Schema of Replacement says that given a function, F, on arbitrary collections (not necessarily sets) and a set X, the range of $F|_X$ is a set. The Axiom of Regularity is a technical axiom that implies that no set may be a member of itself.

The Axiom of Choice is different than the other axioms in that it does not claim that a definable object in the universe of sets is also

a set. Rather, it implies the existence of a function without specifying the function. If X is a set and f is the function with domain X whose existence is guaranteed by the Axiom of Choice, then f is called a choice function for X. The Axiom of Choice is logically equivalent to axioms that are frequently used in arguments in many branches of mathematics. For instance, the Axiom of Choice is equivalent to the claim that every set may be well-ordered (the Well-ordering Principle). The axiom gives rise to interesting paradoxes that caused some mathematicians to question its validity. It was proved by Kurt Gödel that if the axioms of Zermelo-Fraenkel without Choice were logically consistent, then the axioms of Zermelo-Fraenkel with Choice were logically consistent. There were a few occasions in Chapter 6 when we invoked the Axiom of Choice. There were occasions (*e.g.* Cantor's Theorem) in which the axiom is actually necessary, but discussing it would have been unacceptably confusing. The axiom is considered necessary by most mathematicians. For instance without it, or some logically equivalent axiom, we cannot even conclude that any pair of sets can be compared (*i.e.* for any sets X and Y, either $X \preceq Y$ or $Y \preceq X$). The Axiom of Choice is referred to as AC, and the Zermelo-Fraenkel axioms without the Axiom of Choice is referred to as ZF.

Does ZFC achieve the objectives of an axiomatization of set theory? The axioms are generally intuitive with the possible exceptions of AC and the Regularity Axiom. It is also known that if ZFC without the Regularity Axiom is logically consistent, then ZFC with the Regularity Axiom is logically consistent. Mathematicians assume ZFC almost universally without giving it too much consideration. The axioms of ZFC have been sufficient for proving the theorems of standard mathematics.

We say that a set is decidable (or recursive) if membership in the set can be determined by rote computation. For instance, the set of even integers is decidable — you can use the division algorithm to check whether an integer is divisible by 2. ZFC is a recursive set

of axioms. Indeed it is necessary that a set of axioms be recursive to be of any practical use. It is a theorem of mathematics, Gödel's First Incompleteness Theorem, that any decidable set of axioms in which one can do arithmetic will be logically incomplete. That is, there are statements in the language of the axioms that are neither provable nor refutable from the axioms. It is not known, nor can it be known by a mathematical proof (using ZFC) whether ZFC is logically consistent. The consistency of a decidable set of axioms in which one can do arithmetic cannot be a logical consequence of those axioms. This result is known as Gödel's Second Incompleteness Theorem, and is one of the great results of the twentieth century mathematics.

For a good treatment of Set Theory at an undergraduate level, see Y. Moschovakis's book [**5**].

APPENDIX C

Hints to get started on early exercises

Exercise 1.2. You could do this with a Venn diagram. However, once there are more than three sets (see Exercise 1.13), this approach will be difficult. An algebraic proof will generalize more easily, so try to find one here. Argue for the two inclusions

$$\begin{aligned} (X \cup Y)^c &\subseteq X^c \cap Y^c \\ X^c \cap Y^c &\subseteq (X \cup Y)^c \end{aligned}$$

separately. In the first one, for example, assume that $x \in (X \cup Y)^c$ and show that it must be in both X^c and Y^c.

Exercise 1.13. Part of the problem here is notation — what if you have more sets than letters? Start with a finite number of sets contained in U, and call them $X_1, \ldots, X_n$. What do you think the complement of their union is? Prove it as you did when $n = 2$ in Exercise 1.2. (See the advantage of having a proof in Exercise 1.2 that did not use Venn diagrams? One of the reasons mathematicians like to have multiple proofs of the same theorem is that each proof is likely to generalize in a different way).

Can you make the same argument work if your sets are indexed by some infinite index set?

Now do the same thing with the complement of the intersection.

Exercise 1.14. Again there is a notational problem, but while Y and Z play the same rôle in Exercise 1.3, X plays a different rôle. So

rewrite the equations as

$$\begin{aligned} X \cap (Y_1 \cup Y_2) &= (X \cap Y_1) \cup (X \cap Y_2) \\ X \cup (Y_1 \cap Y_2) &= (X \cup Y_1) \cap (X \cup Y_2), \end{aligned}$$

and see if you can generalize these.

Exercise 1.35. (i) Again, this reduces to proving two containments. If y is in the left-hand side, then there must be some x_0 in some U_{α_0} such that $f(x) = y$. But then y is in $f(U_{\alpha_0})$, so y is in the right-hand side.

Conversely, if y is in the right-hand side, then it must be in $f(U_{\alpha_0})$ for some $\alpha_0 \in A$. But then y is in $f\left(\cup_{\alpha \in A} U_\alpha\right)$, and so is in the left-hand side.

Exercise 3.1 There are four possible assignments of truth values 0 and 1 to the two statements P and Q. For each such assignment, evaluate the truth values of the left-hand and right-hand sides of (3.3) and show they are always the same.

Bibliography

[1] M. Aigner and G.M. Ziegler. *Proofs from the Book*. Springer, Berlin, 2003.

[2] J.B. Fraleigh. *A first course in abstract algebra*. Addison-Wesley, Reading, 1982.

[3] I.N. Herstein. *Abstract Algebra*. J. Wiley, New York, 1999.

[4] I. Lakatos. *Proofs and Refutations: The Logic of mathematical discovery*. Cambridge University Press, Cambridge, 1976.

[5] Y.N. Moschovakis. *Notes on Set Theory*. Springer, Berlin, 1994.

[6] The MacTutor History of Mathematics Archives. http://www-groups.dcs.st-and.ac.uk/ history/index.html.

[7] D. Sarason. *Notes on Complex Function Theory*. Hindustan Book Agency, New Delhi, 1998.

[8] E.M. Stein and R. Shakarchi. *Fourier Analysis*. Princeton University Press, Princeton, 2003.

Index

Page numbers appear in bold for pages where the term is defined.