

Set Theory & Logic, 2nd ed.

Richard P. Millspaugh

Preface

These notes are intended to be your guide through Set Theory & Logic (Math 330) at the University of North Dakota. We have three primary goals for this course. First, you should develop some facility in working with the basic properties of sets and the most common operations on sets. The second goal is that you should get over any fear you may yet harbor of the word *proof*. You will begin to write your own proofs and develop some understanding of what a proof is and what it is for. Finally, you will be introduced to some of the basic properties of the real numbers.

Contents

1	Elementary Logic	1
1.1	Basic Notions	2
1.1.1	Propositions	2
1.1.2	Truth values, negation, and truth tables	2
1.2	Quantifiers	3
1.2.1	Syllogisms	4
1.3	Logical Connectives	5
1.3.1	Conjunction and disjunction	5
1.3.2	Implication and equivalence	6
1.4	Tautology and contradiction	6
1.5	Negations Revisited	7
2	Elementary Proof Techniques	11
2.1	Direct Proofs	12
2.1.1	Modus ponens	12
2.1.2	Modus tolens	12
2.2	Proof by Contradiction	13
2.3	Cases	14
2.4	Mathematical Induction	16
2.5	Counterexamples	17
3	Set Theory Basics	19
3.1	What is a Set?	19
3.1.1	Naive set theory	19
3.1.2	Russell's Paradox	20
3.2	Elements and Subsets	20
3.2.1	Elements	20
3.2.2	Subsets	21
3.2.3	Universal sets	22

3.2.4	Equality of sets	22
3.3	Operations on Sets	23
3.3.1	Union and intersection	23
3.3.2	Complements	24
3.3.3	Cartesian products	25
3.4	Collections of Sets	26
3.4.1	The power set of a set	26
4	Relations and Functions	31
4.1	Relations	31
4.1.1	Properties of relations	32
4.2	Equivalence Relations	32
4.2.1	Equivalence classes	33
4.3	Functions	34
4.3.1	Binary Operations	37
4.4	Injective, Surjective, and Bijective Functions	38
4.5	Compositions of Functions	39
4.5.1	Inverses of functions	40
5	The Real Numbers	47
5.1	Field Axioms	47
5.2	Order Axioms	49
5.3	Completeness of \mathbb{R}	51
5.3.1	Upper and lower bounds	51
6	Introduction to Cardinality	57
6.1	The Cardinality of a Set	57
6.2	Finite Sets	60
6.3	Denumerable Sets	62
6.3.1	The set \mathbb{Q}	65
6.3.2	The set \mathbb{R}	66
A	The Cantor-Bernstein Theorem	71

Chapter 1

Elementary Logic

Introduction

In some sense, this text and this course are about proofs. Why do mathematicians require proofs? How are you supposed to read and understand a proof? If you have to prove something yourself, how do you know where to start? How do you know when you're finished? What is a proof anyway?

Mathematicians use proofs for many reasons, but a very simple answer to the last question above is that a proof is a logical explanation of why a statement is true. Unlike the experimental sciences, mathematicians do not accept a statement as true based on data or on statistical reasoning. We might collect data, but only to determine whether or not we believe that something is true. Consider the following:

Theorem. *If p is an even integer, then p^2 is even.*

This statement is true. How do we know that? We might be tempted to start by collecting some data. In other words, experiment. Some mathematicians refer to this as "playing." Let's try a few numbers: $2^2 = 4$ is even, $8^2 = 64$ is even, $(-12)^2 = 144$ is even. When have we tried enough numbers to be sure the theorem is true? Never! No matter how unlikely, it is possible that the first 3,000,012 examples we try will work, but the next one won't. Mathematicians do not accept a statement as true unless we can say with certainty that it is always true. In order to be convinced of the truth of the theorem we must show that the square of an arbitrary even integer is even, not just that the squares of all of the even integers we have tried so far are even. We will prove this theorem in the next chapter.

1.1 Basic Notions

1.1.1 Propositions

Mathematical results are statements that are either true or false. Such statements are called *propositions*. The following sentences are examples of propositions.

Example 1.1. *Every student in this class is under six feet tall.*

Two plus two equals four.

It will snow on March 17, 2525.

Note that a statement can be true or false without our knowing which, as is the case of the last proposition above. The following sentences are neither true nor false, so are not propositions.

Example 1.2. *Go clean your room!*

Is it raining outside?

This statement is false.

The last statement above is particularly interesting. It makes a claim that apparently should be either true or false, but cannot be either. If the claim is true, then the statement must be false. If on the other hand the claim is false, that would make the statement true. Statements which refer to themselves are interesting in their own right, but we will not consider them in this course.

We will use uppercase letters, frequently P or Q , to stand for variable propositions in much the same way that you might use x or y to represent variable numbers in algebra.

1.1.2 Truth values, negation, and truth tables

Since each proposition is either true or false, each proposition has a *truth value* associated with it. In the case of a true proposition this value is true (frequently abbreviated T) and in the case of a false proposition the value is false (or F). As noted previously, a statement may be a proposition even if we do not know what the truth value is.

The *negation* of a proposition P is the statement which is true if P is false and false if P is true. If P is a proposition, the negation of P is denoted $\neg P$. One way to represent this is with a table that lists the truth value of $\neg P$ for each possible truth value of P . This kind of table is called a *truth table*. The truth table for $\neg P$ is given below.

P	$\neg P$
T	F
F	T

1.2 Quantifiers

Many mathematical statements involve variables and are actually examples of quantified statements. We are interested in two types of quantifiers, *universal* and *existential*. A universally quantified statement is one that claims that every member of a particular universe satisfies some property. The following are examples of universally quantified statements.

Example 1.3. *All math majors are brilliant.*

For all real x , $x^2 \geq 0$.

Every eight foot tall student at UND is a Wookie.

An existentially quantified statement says that there is a member of a universe that satisfies a particular property. For example:

Example 1.4. *There is an odd integer which is also even.*

Some people in this room are professors.

There exists a completely regular space that is not normal.

We may extend our previous notation to include these kinds of statements in the following manner. We use $P(x)$ to indicate that P is a proposition involving the variable x . The symbol \forall is used to indicate “for every” and the symbol \exists is used to indicate “there exists.” For example, we may decide that our universe consists of math majors (in other words, x represents some unknown math major) and $P(x)$ might be the proposition “ x is brilliant.” Then the proposition “every math major is brilliant” can be written as $\forall x, P(x)$. The proposition “some math major is brilliant” would be denoted $\exists x, P(x)$.

The universally quantified statement $\forall x, P(x)$ is true if $P(x)$ is true for all possible choices of x in our universe. So while the first proposition in Example 1.3 might be subject to debate, the second is definitely true. The third part of this example is trickier. One might suspect that the statement is false since there are no Wookies at UND (a diversity issue perhaps?), but there are also no eight foot tall students at UND. In this case the universe is empty, so we aren’t really making a claim about anything. Should such a statement be given a truth value of true or false? Maybe we shouldn’t

even allow this kind of statement to be a proposition. To gain some insight, consider the negation of $\forall x, P(x)$.

The negation of $\forall x, P(x)$ should be true if $P(x)$ is not true for some x , which requires the existence of an x for which $P(x)$ is false. In other words, the negation of $\forall x, P(x)$ is the statement $\exists x, \neg P(x)$.

Now let's return to our Wookiee statement. The negation of that statement would be: "*There is an eight foot tall UND student who is not a Wookiee.*" This statement is clearly false since it posits the existence of someone (an eight foot tall UND student) who doesn't exist. Since the negation of the original statement is false, we conclude that the statement itself must be true. In general, a universally quantified statement about an empty universe is said to be vacuously true. Such statements are true simply because there is nothing to make them false.

Unlike a universal quantifier, an existential quantifier requires the existence of something satisfying property $P(x)$. For such a statement to be false, $P(x)$ would have to be false for all choices of x . In other words, the negation of $\exists x, P(x)$ is the statement $\forall x, \neg P(x)$. With this in mind, the truth values of the statements in Example 1.4 are false, true, and true, respectively. Note that the second statement is true simply because there is at least one person in the room who is a professor. Despite what you might think given the wording, this statement does not require more than one person in the room to be a professor.

1.2.1 Syllogisms

A (categorical) syllogism is a type of logical argument that dates at least to the time of Aristotle. A valid syllogism typically consists of three statements — a major premise, a minor premise, and a conclusion which must be valid if both of the premises are true. Here is a classic example.

Example 1.5. *All men are mortal. Socrates is a man. Therefore Socrates is mortal.*

The following example is not a valid syllogism, but demonstrates a fairly common type of syllogistic mistake.

Example 1.6. *Some tenured professors are lazy. Some of my teachers are tenured professors. Hence some of my teachers are lazy.*

Why isn't this argument valid? The major premise says that some tenured professors are lazy, but certainly allows for the possibility that some (or even most) tenured professors are not lazy. In fact, the major premise will

be satisfied even if there is only one tenured professor who is lazy. The minor premise says that some of my teachers are tenured professors. However, we have no valid reason for concluding that any of my teachers are in the group of lazy professors, they might all be in the group that is not lazy.

1.3 Logical Connectives

1.3.1 Conjunction and disjunction

We will use several ways of joining propositions (called connectives) to form more complex propositions. The first two of these are *conjunction* and *disjunction*. The conjunction of two statements P and Q is the statement “ P and Q ,” denoted $P \wedge Q$. If P and Q are propositions, then $P \wedge Q$ is true when both P and Q are true, and false otherwise. The disjunction of P and Q is the statement “ P or Q ,” which is true if at least one of P or Q is true, and false otherwise. We use the notation $P \vee Q$ to denote “ P or Q .” Here is a truth table summarizing the values of $P \wedge Q$ and $P \vee Q$.

P	Q	$P \wedge Q$	$P \vee Q$
T	T	T	T
T	F	F	T
F	T	F	T
F	F	F	F

One final note about our use of the word *or*. In standard English this connective is used in two distinct ways. Consider the following statements:

Example 1.7. *Either it's Monday or I'm having a bad day.*

Eat your vegetables or you won't get any ice cream.

In the first case, the connective means that it's Monday, or I'm having a bad day, but it might also be the case that both are true. In the second case, *or* is used differently. You must eat your vegetables, or you won't get any ice cream, but we presume that both will not happen here. The first use of the connective *or* is said to be inclusive (it includes the case that both statements are true) while the second is exclusive (it excludes the case that both statements are true). In mathematical usage, *or* is always inclusive. If we want an exclusive *or*, we will have to use a construction like *either P or Q , but not both*.

1.3.2 Implication and equivalence

Most mathematical propositions have the form “If P , then Q .” This kind of statement is called an *implication* and can be denoted $P \Rightarrow Q$. The elementary proposition P is called the *hypothesis* of the statement and Q is the *conclusion*. The implication $P \Rightarrow Q$ is true if whenever P is true, Q must also be true. Summarized in a truth table:

P	Q	$P \Rightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

As a student of mathematics, it is absolutely vital that you understand what an implication means and what it doesn't mean. To say that P implies Q is true does not mean that P is true or that Q is true, it only means that there is a relationship between the truth values of these two propositions. Note in particular that the only way to have $P \Rightarrow Q$ be false is when P is true and Q is false. In other words, the negation of $P \Rightarrow Q$ is $P \wedge \neg Q$.

On occasion we will be concerned with propositions P and Q for which $P \Rightarrow Q$ and $Q \Rightarrow P$ are both true. In this case we say that P and Q are (logically) equivalent and write $P \Leftrightarrow Q$.

1.4 Tautology and contradiction

A *tautology* is a proposition that is always true, regardless of the truth values of the elementary propositions that make it up. A *contradiction* is a proposition that is always false.

Example 1.8. The proposition $P \vee \neg P$ is an example of a tautology. One way to see this is to look at a truth table. Notice that all truth values in the column for $P \vee \neg P$ are true. We can also use this truth table to see that $P \wedge \neg P$ is a contradiction because every truth value in that column is false.

P	$\neg P$	$P \vee \neg P$	$P \wedge \neg P$
T	F	T	F
F	T	T	F

1.5 Negations Revisited

Negating complex propositions can seem difficult at first, but is really a matter of applying the rules we have already talked about step by step. The most important elementary negations for you to remember are listed in the following theorem, which summarizes the results of Exercises ?? and ??:

Theorem 1.1. *Given any propositions P and Q :*

(i) *(DeMorgan's Laws)*

$$(a) \neg(P \wedge Q) \Leftrightarrow (\neg P \vee \neg Q)$$

$$(b) \neg(P \vee Q) \Leftrightarrow (\neg P \wedge \neg Q)$$

(ii) $\neg(P \Rightarrow Q) \Leftrightarrow (P \wedge \neg Q)$

Example 1.9. *Find the negation of the proposition $(P \wedge (Q \Rightarrow \neg R)) \vee S$. This proposition is the disjunction of two other propositions, so we start by negating the disjunction, which gives us:*

$$\neg(P \wedge (Q \Rightarrow \neg R)) \wedge \neg S$$

There's nothing else to do to the proposition $\neg S$, but we still have to negate the proposition $P \wedge (Q \Rightarrow \neg R)$. To do so we start by negating the conjunction, which gives us

$$(\neg P \vee \neg(Q \Rightarrow \neg R)) \wedge \neg S$$

for the negation of our original proposition. We still have to negate the implication, which yields as our final negation the proposition

$$(P \vee (Q \wedge R)) \wedge \neg S$$

Chapter 1 Exercises

1.1. Determine which of the following are valid and which are invalid.

- (i) All the men in Lake Wobegone are good looking. Steve is a man who lives in Lake Wobegone. Hence Steve is good looking.
- (ii) No cats are dogs. Some cats are crazy. Hence some dogs are not crazy.
- (iii) All rats are rodents. No rats are bats. So some rodents are not bats.¹
- (iv) All criminals have shifty eyes. Steve has shifty eyes. So Steve is a criminal.
- (v) Some students live in North Dakota. Everyone who lives in North Dakota has seen snow. So some students have seen snow.
- (vi) Some students live in San Diego. Some people who live in San Diego have not seen snow. Hence some students have not seen snow.
- (vii) All students are smart. Some politicians are not students. Hence some politicians are not smart.

1.2. Show that $\neg(\neg P) \Leftrightarrow P$ is a tautology.

1.3. Determine whether each of the following is a tautology, a contradiction, or neither one.

- (i) $(P \Rightarrow Q) \Rightarrow Q$
- (ii) $(P \wedge (P \Rightarrow Q)) \Rightarrow Q$
- (iii) $(P \Rightarrow Q) \Leftrightarrow (Q \Rightarrow P)$
- (iv) $((P \Rightarrow Q) \wedge (Q \Rightarrow R)) \Rightarrow (P \Rightarrow R)$
- (v) $(P \Rightarrow Q) \wedge (P \Rightarrow \neg Q)$
- (vi) $P \wedge (P \Rightarrow Q) \wedge (P \Rightarrow \neg Q)$
- (vii) $(\neg Q \wedge (P \Rightarrow Q)) \Rightarrow \neg P$
- (viii) $((P \vee Q) \wedge \neg Q) \Rightarrow P$
- (ix) $((P \vee Q) \Rightarrow R) \Rightarrow (P \Rightarrow R)$

¹It is a common misconception that bats are rodents. They are not.

1.4. Use truth tables for each of the following:

- (i) Show that conjunction distributes over disjunction, i.e. that $P \wedge (Q \vee R)$ is equivalent to $(P \wedge Q) \vee (P \wedge R)$.
- (ii) Show that disjunction distributes over conjunction.

1.5. Show that $P \wedge (Q \vee R)$ is not equivalent to $(P \wedge Q) \vee R$.

1.6. Negate each of the following:

- (i) Every cloud has a silver lining.
- (ii) If it's Tuesday, this must be Belgium.
- (iii) There is a light at the end of every tunnel.

1.7. One way of defining what it means for a function f to be continuous at a point x_0 is as follows:

A function f is said to be continuous at a point $x_0 \in \mathbb{R}$ if for every $\epsilon > 0$ there is a $\delta > 0$ so that $|f(x_0) - f(t)| < \epsilon$ for all $t \in \mathbb{R}$ with $|x_0 - t| < \delta$.

What does it mean to say that f is not continuous at x_0 ?

Chapter 2

Elementary Proof Techniques

Introduction

There are a few common proof techniques that can be used to prove most mathematical propositions. In this chapter, we will discuss several of these techniques and look at some elementary examples. These techniques should serve you well throughout this course and in future mathematics courses. As the mathematics becomes more involved the proofs will also become more involved. You may even find it necessary to combine techniques in a single proof, but many of the proofs you are asked to construct, or to read and understand, will depend on the same few techniques.

Before we begin, note that when we say that a proposition is a *theorem* we mean that it is true and that it has been proved.¹ A *Lemma* or a *corollary* is also a true statement which has been proved, but these terms are usually used only in special cases. A lemma is typically a proposition which is only proved in order to use it to prove a more important, or at least more interesting, proposition. A corollary usually refers to a proposition which follows immediately, and frequently in an obvious way, from a prior theorem.

¹In 1930 Kurt Gödel proved that in any reasonably complex mathematical system there are true statements that cannot be proved within that system.

2.1 Direct Proofs

2.1.1 Modus ponens

Although mathematicians use proofs in many ways, the most obvious is that a proof is the way we determine that a statement is true. One (overly simplified) definition might be that a proof is a demonstration, using accepted rules of implication, that we can deduce the conclusion of a statement from the hypothesis. The simplest form a proof can take begins by assuming the hypothesis of the statement, deduces another statement, then uses that to deduce a third statement, etc., until we are able to conclude that the desired conclusion is true. This kind of proof, frequently called a direct proof, uses only the following rule of implication.

Modus ponens: From P and $P \Rightarrow Q$, we may conclude Q .

Here is an elementary example of a proof using only the rule of modus ponens.

Theorem 2.1. *The square of an even integer is even.*

Before proving this theorem we note that it is not, as currently stated, an implication. We could, however, restate the theorem as the following implication: *If n is an even integer, then n^2 is even.* You may find it useful to restate theorems as implications before trying to prove them. Whether or not you actually rewrite the theorem as an implication, it is absolutely imperative that you identify the hypothesis and conclusion of the theorem before you make any attempts to prove it.

Proof of Theorem 2.1. Let n be an even integer, then $n = 2k$ for some integer k . Squaring both sides yields $n^2 = 4k^2 = 2(2k^2)$. Since $2k^2$ is an integer, $2(2k^2)$ is even as desired. \square

Theorem 2.2. *The square of an odd integer is odd.*

2.1.2 Modus tollens

Modus tollens: From $\neg Q$ and $P \Rightarrow Q$, we may conclude $\neg P$.

Here is a simple example using the rule of modus tollens.

Theorem 2.3. *If n is an integer and n^2 is even, then n is even.*

We might try to begin this proof as we did the proof of Theorem 2.1. First assume that n^2 is even. By definition we know that $n^2 = 2k$ for some integer k . Now what??? In the previous example we were able to square $2k$ and see that the result was even, but taking the square root doesn't give us anything to work with. Here is a proof that does work.

Proof of Theorem 2.3. Assume that n^2 is even, then n^2 is not odd. Since the square of an odd integer is odd by Theorem 2.2, we may conclude that n is not odd. Every integer is either even or odd, so n must be even. \square

Theorem 2.4. *If n is an integer and n^2 is odd, then n is odd.*

2.2 Proof by Contradiction

Reductio ad absurdum: From $\neg P \Rightarrow (Q \wedge \neg Q)$, we may conclude P .

A proof by contradiction begins by assuming that the result we are trying to prove is false, then deriving a contradiction from that assumption. Here is a classic example.

Theorem 2.5. *There is no rational number x such that $x^2 = 2$.*

Proof. Suppose to the contrary that there is a rational number x with $x^2 = 2$. Since x is rational, we may express x as a fraction p/q in lowest terms, in other words p and q are integers and no integer evenly divides both of them. Since $x = p/q$ and $x^2 = 2$, we have $p^2/q^2 = 2$, or:

$$p^2 = 2q^2 \tag{2.1}$$

Since q^2 is an integer, $p^2 = 2q^2$ is even. So p is an integer and p^2 is even, so Theorem 2.3 implies that p is even. This allows us to write $p = 2k$ for some integer k . We plug this into equation 2.1 to see that $4k^2 = 2q^2$, so $q^2 = 2k^2$. Now q is an integer and q^2 is even, so another application of Theorem 2.3 implies that q is even. We now know that both p and q are even, so 2 divides both of them. This contradicts our choice of p and q , so our assumption that x is rational must be incorrect. \square

2.3 Cases

From $P \vee Q$ and $P \Rightarrow R$ and $Q \Rightarrow R$, we may conclude R .

On occasion, a proposition does not lend itself to a single proof. It may be necessary to consider several cases and show that each of them yields the desired result. This is especially true when the hypothesis of the statement to be proved involves a disjunction, either directly or as part of a definition.

Theorem 2.6. For any two real numbers a and b , $|ab| = |a||b|$.

Proof. We first recall that:

$$|a| = \begin{cases} a & \text{if } a \geq 0 \\ -a & \text{if } a < 0 \end{cases}$$

Case 1: a and b are both nonnegative.

Since a and b are both nonnegative, so is ab . Using the definition above yields:

$$|ab| = ab = |a||b|$$

Case 2: $a = 0$.

Since $ab = 0$, we have:

$$|ab| = |0| = 0 = 0(|b|) = |a||b|$$

Case 3: $b = 0$.

Since $ab = 0$, we have:

$$|ab| = |0| = 0 = (|a|)0 = |a||b|$$

Case 4: $a > 0$ and $b < 0$.

We have $ab < 0$, so applying the definition yields:

$$|ab| = -ab = a(-b) = |a||b|$$

Case 5: $a < 0$ and $b > 0$.

We have $ab < 0$, so applying the definition yields:

$$|ab| = -ab = (-a)b = |a||b|$$

Case 6: $a < 0$ and $b < 0$.

We have $ab > 0$, so applying the definition yields:

$$|ab| = ab = (-a)(-b) = |a||b|$$

Any real numbers a and b must fit into one of the cases above, and in every case we have $|ab| = |a||b|$, so the theorem is true. \square

Before moving on we want to note a couple of things about the previous proof. First, our cases must encompass all possibilities. Second, the first attempt to distinguish possibilities might not be very efficient, so once we have a proof it is worthwhile looking at our cases to see if some of them can be combined. In the proof above, cases 2 and 3 are certainly very similar, as are cases 4 and 5. Here is a somewhat more efficient proof of the same theorem, using the sign of ab rather than the signs of a and b to choose cases.

Alternate proof of Theorem 2.6. Either $ab > 0$, $ab = 0$, or $ab < 0$.

Case 1: $ab > 0$.

In this case we have $|ab| = ab$. It must also be the case that a and b are either both positive or both negative. If they are both positive, then $|a||b| = ab$. If they are both negative, then $|a||b| = (-a)(-b) = ab$. Hence $|ab| = |a||b|$ as desired.

Case 2: $ab = 0$.

In this case we have either $a = 0$ or $b = 0$. Either way, $|ab| = 0$ and $|a||b| = 0$, so $|ab| = |a||b|$.

Case 3: $ab < 0$.

In this case we have $|ab| = -ab$. Since one of a or b is positive and the other is negative, we also have $|a||b| = (-a)b = -ab$ or $|a||b| = a(-b) = -ab$. Hence $|ab| = |a||b|$ as desired.

In every case we have $|ab| = |a||b|$ as desired. \square

2.4 Mathematical Induction

The Principle of Mathematical Induction. Let $P(n)$ be a statement about the natural number n .

- (i) If $P(1)$ is true, and
- (ii) if $P(k) \implies P(k+1)$ for every natural number k ,

then $P(n)$ is true for every natural number n .

Proof by induction is frequently used to prove statements about the natural numbers. All proofs by induction use the following basic outline. To prove $P(n)$ for all natural numbers n ,

- (i) Base Case: Show that $P(1)$ is true.
- (ii) Inductive Step: Prove the implication $P(k) \implies P(k+1)$ for all $k \in \mathbb{N}$. Note: in proving that $P(k) \implies P(k+1)$, the assumption that $P(k)$ is true is often referred to as the inductive hypothesis.
- (iii) It is considered good form to clearly state that the statement $P(n)$ is now true for all natural numbers n by induction.

You might think of a proof by induction as similar to knocking over a row of dominoes. If you know that knocking over any domino will cause the next one to fall, then knocking over the first domino will cause all of the dominoes to fall. Keep in mind that *you must complete all of the steps above in an inductive proof*. Here are a couple of examples.

Theorem 2.7. For each natural number n , $\sum_{i=1}^n i = \frac{n(n+1)}{2}$.

Proof. If $n = 1$, then

$$\sum_{i=1}^n i = \sum_{i=1}^1 i = 1 = \frac{1(2)}{2} = \frac{n(n+1)}{2},$$

so the formula holds for $n = 1$.

Now suppose that the formula holds for some $k \in \mathbb{N}$. In other words, we are assuming that $\sum_{i=1}^k i = \frac{k(k+1)}{2}$. We want to show that the formula

must hold for $n = k + 1$. Using our assumption, we compute:

$$\begin{aligned} \sum_{i=1}^{k+1} i &= \left(\sum_{i=1}^k i \right) + (k + 1) \\ &= \frac{k(k + 1)}{2} + (k + 1) \\ &= \frac{k(k + 1)}{2} + \frac{2(k + 1)}{2} \\ &= \frac{(k + 1)(k + 2)}{2} \end{aligned}$$

Hence the formula holds for $n = k + 1$. By induction, the formula must hold for all $n \in \mathbb{N}$. \square

2.5 Counterexamples

Consider the following proposition:

Every integer is even.

It probably seems clear to you that this statement is false, but how could you prove that? To prove that a proposition is false we must show that its negation is true, so we first find the negation of the original statement:

Some integer is not even.

We have transformed our original problem (proving that a statement is false) into something more familiar (proving that a statement is true): How do we prove the second statement is true? One way to prove an existentially quantified statement is simply to find the object it says exists. In this case, we must find an integer that is not even. Of course, we know many such integers. Any odd integer will do. So our proof that the original statement is false consists of finding an example that makes it false: the integer $x = 1$ is not even. Note that you shouldn't just say that there are such examples, you should give a particular one that your reader can check. An example showing that a universally quantified statement is false is called a *counterexample* to that statement.

Chapter 2 Exercises

2.1. Using the proof of Theorem 2.1 as an example, prove Theorem 2.2.

2.2. Using the proof of Theorem 2.3 as an example, prove Theorem 2.4.

2.3. Find counterexamples to show that each of the following statements is false.

(i) Every student in this class has green hair.

(ii) For every real number x , $x^2 - 1 \geq 0$.

(iii) Every prime number is less than 1000.

2.4. Explain why a counterexample cannot be used to prove that an existentially quantified statement is false. You may use the following example: *There is an integer a with the property that a and $a + 1$ are both even.*

2.5. Use induction to prove that $\sum_{i=0}^n 2^i = (2^{n+1} - 1)$ for every natural number n .

2.6. Prove that $\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$ for every natural number n .

2.7. It is sometimes true that statements about the natural numbers are not true for small numbers, but are true for all natural numbers that are large enough. We can modify the Principle of Mathematical Induction so that it is applicable in these situations as follows:

Let $P(n)$ be a statement about the natural number n .

(i) If $P(a)$ is true, and

(ii) if $P(k) \implies P(k+1)$ for all $k \geq a$,

then $P(n)$ is true for all natural numbers $n \geq a$.

Use this version of induction to prove that $3^n < n!$ for every integer $n \geq 7$.

Chapter 3

Set Theory Basics

Introduction

The work of Cantor, Dedekind, Weierstrass, and others in the last half of the 19th century allowed all of mathematics to be based on a common foundation. Based on the ideas and insights of Georg Cantor, the new theory of sets was not immediately accepted by many outstanding mathematicians of the day. The most serious problem was that Cantor treated infinite sets as objects and defined operations on those objects. Poincaré felt that “Later mathematicians will regard set theory as a disease from which we have recovered.” Gauss and Kronecker were also among the mathematical greats who attacked set theory, sometimes in a particularly vicious and personal way. Hilbert, on the other hand, said that “no one will evict us from the paradise that Cantor has built for us.” Cantor himself was unable to obtain a position at any of Germany’s prestigious research universities and spent his entire 44 year academic career in a relatively minor position. Distraught over continuing resistance to his work, Cantor spent the last years of his life in a mental institution.

3.1 What is a Set?

3.1.1 Naive set theory

Early work in set theory assumed only that any collection that could be clearly specified (there is a rule for determining whether or not something is in the collection) could be considered a set, and that two sets were the same if they contained the same elements. These properties are sufficient

to allow most mathematics students, and most working mathematicians for that matter, to study the groups or fields or topological spaces they are interested in. The finer points of what actually is, or is not, a set just don't come up most of the time.

A student of mathematics should be aware, however, that there are in fact some restrictions. Perhaps the most important is that sets are not generally allowed to be elements of themselves. Students who find set theory of interest for its own sake are encouraged to delve further into this field of study.

3.1.2 Russell's Paradox

In 1902, philosopher and mathematician Bertrand Russell published his famous paradox. Closely related to the Liar's Paradox, Russell's Paradox exploits a form of self-reference. More specifically, the paradox works only if we allow the possibility that a set might be an element of itself. To avoid this kind of problem, we will not allow any set to be an element of itself. We will talk more about Russell's version of this paradox later, but for now let's deal with a popularization that doesn't depend so much on the language of sets.

Example 3.1. The Barber's Paradox. *A certain small town has only one barber. He shaves only those men in town who do not shave themselves. Who shaves the barber?*

3.2 Elements and Subsets

3.2.1 Elements

We will think of a set as a collection of objects, called its *elements*. These objects might be points, numbers, people, kitchen appliances, other sets, or any other objects we are interested in talking about. For a set to be well-defined, we must have a way to determine whether or not a given object is in the set. We consider two sets to be equal if they contain exactly the same elements. Note that an object is either an element of a set or not, the elements of a set do not occur in a particular order and the same object cannot be an element of the set more than once.

One way to indicate a set is to simply list all of the elements between set braces $\{$ and $\}$. The set of positive integers less than 3 is $\{1, 2\}$. We will

find it useful to have a special notation for the set which has no elements. To this end, let $\emptyset = \{\}$.

Example 3.2. Let $A = \{1, 2, 4, 8\}$, $B = \{8, 2, 1, 4\}$, $C = \{1, 2, 1, 4, 8\}$, and $D = \{1, 2, 3, 4, 8\}$. Of these sets, $A = B = C$ because they each contain the same elements. It doesn't matter that we listed the elements in a different order when we defined A than when we defined B , or that we listed 1 as an element of C more than once. The set D is different from the others since it contains the element 3 and the other sets do not.

We indicate that the object x is an element of the set A by writing $x \in A$. We write $x \notin A$ if x is not an element of A . Listing all of the elements of a set works well when the set contains only a few elements, but what about sets with many elements? In this case, we use *set builder* notation to denote the set. The important part of this notation is the use of the vertical bar $|$ (some texts use a colon in place of the vertical bar), read "such that." So the set $A = \{x \mid x^2 - 2 = 0\}$ is read "the set of all elements x such that $x^2 - 2 = 0$." An object will be an element of this set if and only if it satisfies the equation $x^2 - 2 = 0$.

There are also a few sets that are useful enough to deserve their own special notation. In particular, \mathbb{N} denotes the set of natural numbers, \mathbb{Z} the set of integers, \mathbb{Q} the set of rational numbers, \mathbb{R} the set of real numbers, and \mathbb{C} the set of complex numbers.

3.2.2 Subsets

Given two sets A and B we say that A is a *subset* of B , or that A is contained in B , if every element of A is also an element of B . In this case we write $A \subset B$ (sometimes $A \subseteq B$). Since every natural number is also an integer and every integer is a real number, $\mathbb{N} \subset \mathbb{Z}$ and $\mathbb{Z} \subset \mathbb{R}$.

Given any set A , it is certainly true that every element of A is an element of A . In other words, every set is a subset of itself. Note also that the empty set is a subset of every set A since there are no elements of the empty set which are not in A . A subset B of A is said to be a *proper subset* of A if $B \neq \emptyset$ and $B \neq A$.

Theorem 3.1. If $A \subset B$ and $B \subset C$, then $A \subset C$.

Proof. Suppose that $a \in A$. Since $A \subset B$ it follows that $a \in B$. Now $B \subset C$, so we may say that $a \in C$ as desired. \square

The preceding proof is a simple example of a direct proof. Let's pause for a moment and analyze this proof. The statement we want to prove is a universally quantified statement about elements of A : every element of A is also an element of C . We begin by considering an arbitrary element of A , which we have named a . The goal is to use our hypotheses ($A \subset B$ and $B \subset C$) to arrive at the conclusion that $a \in C$. First we note that every element of A is also an element of B since $A \subset B$. This allows us to state that $a \in B$. Once we have $a \in B$ we may use the second hypothesis to see that $a \in C$ since $B \subset C$. So starting with any element at all of the set A we have shown that it must also be an element of C , which is the definition of $A \subset C$. We conclude that $A \subset C$ as desired.

Note that the preceding proof is a syllogism: *All elements of A are elements of B . All elements of B are elements of C . Hence all elements of A are elements of C .* This is certainly not true of all direct proofs, but will be true on occasion.

3.2.3 Universal sets

In many instances, all of the sets we may be interested in are subsets of some particular set U . In this case we say that U is a *universal set*, or that U is the *universe*. For example, all of the functions we study in calculus have domains and ranges that are subsets of the universal set \mathbb{R} .

3.2.4 Equality of sets

We say that two sets are equal if they contain exactly the same elements. In other words, $A = B$ means that every element of A is an element of B and every element of B is an element of A . In other words, we have the following:

Theorem 3.2. *Given any two sets A and B , $A = B$ if and only if $A \subset B$ and $B \subset A$.*

This theorem will become our most valuable tool for showing that two sets are equal.

3.3 Operations on Sets

3.3.1 Union and intersection

The *union* of two sets A and B is defined as follows:

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\}$$

The *intersection* of A and B is defined as:

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}$$

Example 3.3. Define $A = \{1, 2, 3, 4, 5\}$ and $B = \{2, 4, 6, 8, 10\}$. Then

$$A \cup B = \{1, 2, 3, 4, 5, 6, 8, 10\} \text{ and } A \cap B = \{2, 4\}.$$

The following theorem summarizes several useful algebraic properties.

Theorem 3.3. Let A , B , and C be any sets in the universal set U . Then:

(i) (*Commutative Laws*)

(a) $A \cup B = B \cup A$

(b) $A \cap B = B \cap A$

(ii) (*Associative Laws*)

(a) $A \cup (B \cup C) = (A \cup B) \cup C$

(b) $A \cap (B \cap C) = (A \cap B) \cap C$

(iii) (*Distributive Laws*)

(a) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

(b) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

(iv) (*Idempotence*)

(a) $A \cup A = A$

(b) $A \cap A = A$

(v) (*Identities*)

(a) $A \cup \emptyset = A$

(b) $A \cap U = A$

Proof of i(a). Suppose that $x \in A \cup B$, then $x \in A$ or $x \in B$. By a tautology $((P \vee Q) \Leftrightarrow (Q \vee P))$ ¹ we have $x \in B$ or $x \in A$. Therefore $x \in B \cup A$ and we have $(A \cup B) \subset (B \cup A)$. Conversely, if $y \in B \cup A$, then $y \in B$ or $y \in A$. Once again we may reverse this to get $y \in A$ or $y \in B$, which implies that $y \in A \cup B$. Hence $(B \cup A) \subset (A \cup B)$. Finally, we apply Theorem 3.2 to see that $A \cup B = B \cup A$ as desired. \square

Proof of iii(a). Let $x \in A \cup (B \cap C)$, then $x \in A$ or $x \in B \cap C$. By definition, we have $x \in A$ or $(x \in B \text{ and } x \in C)$. Applying the tautology from Exercise 1.3(d) we may now say that $(x \in A \vee x \in B) \wedge (x \in A \vee x \in C)$. Hence $x \in A \cup B$ and $x \in A \cup C$, so $x \in (A \cup B) \cap (A \cup C)$. Therefore $A \cup (B \cap C) \subset (A \cup B) \cap (A \cup C)$.

To see the converse, start by assuming that $y \in (A \cup B) \cap (A \cup C)$. Then $y \in A \cup B$ and $y \in A \cup C$. By definition this yields $(y \in A \vee y \in B) \wedge (y \in A \vee y \in C)$, which is equivalent to $y \in A \vee (y \in B \wedge y \in C)$ by the same tautology as we used above. Hence $y \in A$ or $y \in B \cap C$, so $y \in A \cup (B \cap C)$ and we have shown that $(A \cup B) \cap (A \cup C) \subset A \cup (B \cap C)$. Finally, we apply Theorem 3.2 to achieve the desired result. \square

Proof of v(a). Assume first that $x \in A \cup \emptyset$, then $x \in A$ or $x \in \emptyset$. Since $x \in \emptyset$ would be a contradiction, we may apply the tautology from Exercise 1.3(i) to say that $x \in A$. Therefore $A \cup \emptyset \subset A$. Letting $B = \emptyset$, we may apply part (vii) to see that $A \subset A \cup \emptyset$. Now Theorem 3.2 implies that $A \cup \emptyset = A$ as desired. \square

3.3.2 Complements

The *relative complement* of B in A (also called the set difference) is the set:

$$A \setminus B = \{a \mid a \in A \text{ and } a \notin B\}$$

In a given universe U , we may also define the *complement* of a set A to be the set:

$$A' = U \setminus A$$

Theorem 3.4 (DeMorgan). *Let A , B , and C be sets. Then:*

- (i) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$, and
- (ii) $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$.

¹In general, tautologies may be used without specifically mentioning them, but we will mention the tautologies we use in proofs of the various portions of this theorem.

Proof of Theorem 3.4(i). Rather than using Theorem 3.2 we use Theorem 1.1 to show that an element is in $A \setminus (B \cup C)$ if and only if it is in $(A \setminus B) \cup (A \setminus C)$.

$$\begin{aligned}
 x \in A \setminus (B \cup C) &\Leftrightarrow (x \in A) \wedge (x \notin B \cup C) \\
 &\Leftrightarrow (x \in A) \wedge \neg(x \in B \vee x \in C) \\
 &\Leftrightarrow (x \in A) \wedge (\neg(x \in B) \wedge \neg(x \in C)) \\
 &\Leftrightarrow (x \in A) \wedge (x \notin B) \wedge (x \notin C) \\
 &\Leftrightarrow ((x \in A) \wedge (x \notin B)) \wedge ((x \in A) \wedge (x \notin C)) \\
 &\Leftrightarrow (x \in A \setminus B) \wedge (x \in A \setminus C) \\
 &\Leftrightarrow x \in (A \setminus B) \cap (A \setminus C)
 \end{aligned}$$

□

3.3.3 Cartesian products

The (*Cartesian*) *product* of two sets A and B is the set:

$$A \times B = \{(a, b) \mid a \in A \text{ and } b \in B\}.$$

In this context, (a, b) denotes an ordered pair, not an interval, so the product $A \times B$ is simply the set of all ordered pairs with first coordinates in A and second coordinates in B . For example, the Cartesian plane used in Calculus is the set $\mathbb{R} \times \mathbb{R}$.

Theorem 3.5. *For any sets A , B , and C :*

$$(i) \quad (A \cup B) \times C = (A \times C) \cup (B \times C)$$

$$(ii) \quad (A \cap B) \times C = (A \times C) \cap (B \times C)$$

$$(iii) \quad (A \setminus B) \times C = (A \times C) \setminus (B \times C)$$

Proof of Theorem 3.5(i). First suppose that $(x, y) \in (A \cup B) \times C$. By definition we have $x \in A \cup B$ and $y \in C$. Since $x \in A \cup B$, either $x \in A$ or $x \in B$. If $x \in A$, then we have $x \in A$ and $y \in C$, so $(x, y) \in A \times C$. If $x \in B$, then we have $x \in B$ and $y \in C$, so $(x, y) \in B \times C$. We can now say that $(x, y) \in A \times C$ or $(x, y) \in B \times C$, so $(x, y) \in (A \times C) \cup (B \times C)$. Hence $(A \cup B) \times C \subset (A \times C) \cup (B \times C)$.

To see that the converse is true, suppose that $(x, y) \in (A \times C) \cup (B \times C)$. Either $(x, y) \in A \times C$ or $(x, y) \in B \times C$. If $(x, y) \in A \times C$, then $x \in A$ and $y \in C$, so $(x, y) \in (A \cup B) \times C$. Similarly, if $(x, y) \in B \times C$, then $x \in B$ and $y \in C$, so $(x, y) \in (A \cup B) \times C$. Hence $(A \times C) \cup (B \times C) \subset (A \cup B) \times C$.

$(A \cup B)$ and $y \in C$, so $(x, y) \in (A \cup B) \times C$. If $(x, y) \in B \times C$, then $x \in B \subset (A \cup B)$ and $y \in C$, so $(x, y) \in (A \cup B) \times C$. Hence $(A \times C) \cup (B \times C) \subset (A \cup B) \times C$. Therefore, $(A \cup B) \times C = (A \times C) \cup (B \times C)$ as desired. \square

Since $A \times B$ is a set of ordered pairs, $A \times B \neq B \times A$. You should make sure that you look at an example to understand why this is true. This requires us to state another theorem that seems very similar to the last one. The proof of the following theorem involves making obvious changes to the previous proof and checking that all of the details still work.

Theorem 3.6. *For any sets A , B , and C :*

$$(i) \quad A \times (B \cup C) = (A \times B) \cup (A \times C)$$

$$(ii) \quad A \times (B \cap C) = (A \times B) \cap (A \times C)$$

$$(iii) \quad A \times (B \setminus C) = (A \times B) \setminus (A \times C)$$

It may be tempting to assume that we could combine these two theorems somehow and obtain statements like $(A \times B) \setminus (C \times D) = (A \setminus C) \times (B \setminus D)$. This statement is generally false, however, as shown by the following.

Example 3.4. *Let $A = \{1, 2, 3\}$, $B = \{5, 6\}$, $C = \{1, 2\}$, and $D = \{6\}$. Then:*

$$(A \times B) \setminus (C \times D) = \{(1, 5), (2, 5), (3, 5), (3, 6)\}$$

but

$$(A \setminus C) \times (B \setminus D) = \{(3, 5)\}$$

3.4 Collections of Sets

Some of the structures used in pure mathematics require the use of sets whose elements are other sets. It is customary, though not necessary, to refer to these kinds of sets as collections of sets. When working with sets and collections of sets, it can be particularly confusing to keep track of which set is an element of which other set, as opposed to being a subset.

3.4.1 The power set of a set

Let A be any set. The *power set* of A is the set $\mathcal{P}(A) = \{B \mid B \subset A\}$. In words, the power set of A is the set whose elements are the subsets of A . Note that for any set A we have $\emptyset \in \mathcal{P}(A)$ and $A \in \mathcal{P}(A)$, so $\mathcal{P}(A)$ is nonempty for every set A . The power set of A is frequently denoted 2^A .

Theorem 3.7. *For any sets A and B , $A \subset B$ if and only if $\mathcal{P}(A) \subset \mathcal{P}(B)$.*

Proof. First assume that $A \subset B$ and let $X \in \mathcal{P}(A)$. By definition $X \subset A$, so Theorem 3.1 implies that $X \subset B$. Hence $X \in \mathcal{P}(B)$ and $\mathcal{P}(A) \subset \mathcal{P}(B)$. Conversely, if we assume that $\mathcal{P}(A) \subset \mathcal{P}(B)$, then $A \in \mathcal{P}(A) \subset \mathcal{P}(B)$ and $A \subset B$ by definition. \square

Theorem 3.8. *For any sets A and B , $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cap \mathcal{P}(B)$.*

Proof. First note that $A \cap B \subset A$ by Exercise 3.6, so Theorem 3.7 implies that $\mathcal{P}(A \cap B) \subset \mathcal{P}(A)$. The same reasoning shows that $\mathcal{P}(A \cap B) \subset \mathcal{P}(B)$, so we have $\mathcal{P}(A \cap B) \subset \mathcal{P}(A) \cap \mathcal{P}(B)$ by Exercise 3.7.

Now suppose that $X \in \mathcal{P}(A) \cap \mathcal{P}(B)$, then $X \in \mathcal{P}(A)$ and $X \in \mathcal{P}(B)$. By definition $X \subset A$ and $X \subset B$, so $X \subset A \cap B$ by Exercise 3.7. It follows that $\mathcal{P}(A \cap B) \supset \mathcal{P}(A) \cap \mathcal{P}(B)$, so $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cap \mathcal{P}(B)$ as desired. \square

Chapter 3 Exercises

3.1. Explain why there is no consistent answer to the question in Example 3.1.

3.2. Here are some common infinite sets:

$P = \{n \mid n \text{ is prime}\}$ is the set of all prime numbers.

$E = \{n \mid n = 2k \text{ for some } k \in \mathbb{Z}\}$ is the set of all even integers.

How would you write the set of all odd integers in set builder notation? What about the set of all integral powers of 2? The set of all UND students taking Math 330 this semester?

3.3. Consider the sets \mathbb{N} , \mathbb{Z} , \mathbb{Q} , \mathbb{R} , and \mathbb{C} . Which of these sets are subsets of \mathbb{Q} ? Of \mathbb{C} ?

3.4. Define the following sets: $A = \{1, 3, 9, 27\}$, $B = \{1, 2, 4, 8\}$, $P = \{n \mid n \text{ is a prime integer}\}$, and $E = \{n \mid n = 2k \text{ for some } k \in \mathbb{N}\}$. Find each of the following:

(i) $A \cup B$

(ii) $A \cap B$

(iii) $P \cup E$

(iv) $P \cap E$

3.5. Prove the remaining parts of Theorem 3.3. In each case, identify any tautologies you use in your proofs.

3.6. For any sets A and B , prove that:

(i) $A \subset A \cup B$

(ii) $A \cap B \subset A$

(iii) $A \cap \emptyset = \emptyset$

3.7. Let A , B , and C be sets. Prove that $A \subset B \cap C$ if and only if $A \subset B$ and $A \subset C$.

3.8. Let A , B , and C be sets. Prove that if $A \subset B \cup C$ and $A \cap B = \emptyset$, then $A \subset C$.

3.9. Prove part (ii) of Theorem 3.4.

3.10. Let A and B be sets in a universal set U . Prove the following:

- (i) $A \setminus B = A \cap B'$
- (ii) $A \setminus B = A$ if and only if $A \cap B = \emptyset$.
- (iii) $A \setminus B = \emptyset$ if and only if $A \subset B$.

3.11. Prove the remaining parts of Theorem 3.5.

3.12. Determine whether or not each of the following is true. If so, provide a proof. If not, provide a counterexample.

- (i) $(A \cup B) \times (C \cup D) = (A \times C) \cup (B \times D)$
- (ii) $(A \cap B) \times (C \cap D) = (A \times C) \cap (B \times D)$

3.13. Let A and B be sets.

- (i) Show that $\mathcal{P}(A) \cup \mathcal{P}(B) \subset \mathcal{P}(A \cup B)$.
- (ii) Show that it is not necessarily true that $\mathcal{P}(A) \cup \mathcal{P}(B) = \mathcal{P}(A \cup B)$.

Chapter 4

Relations and Functions

Introduction

Anybody who has taken a calculus class is aware that functions are important objects in mathematics. It wasn't until the eighteenth and nineteenth centuries that mathematicians developed a rigorous definition of functions. In this chapter we will learn how to define a function in terms of sets, which first requires us to define a relation. We will also consider some special kinds of relations and functions.

4.1 Relations

Given two sets A and B , a *relation* from A to B is a subset of $A \times B$. The relations we will be most interested in are usually from a set A to itself, in which case we say that the relation is a relation on A . If S is a relation, we frequently use the notation xSy to indicate that the ordered pair (x, y) is in S . Let's look at some examples of relations.

Example 4.1. *The usual ordering $<$ on \mathbb{R} is a relation on \mathbb{R} . We don't usually think of this relation as a set of ordered pairs, but we could. Define a subset L of $\mathbb{R} \times \mathbb{R}$ by $L = \{(a, b) \mid b - a \text{ is positive}\}$. In other words, an ordered pair (a, b) is in the relation if $a < b$.*

Example 4.2. *The set of points on a circle is another example of a relation on \mathbb{R} . For example, we could let $C = \{(x, y) \mid x^2 + y^2 = 1\}$.*

Example 4.3. *A relation need not be something you've encountered before or that you necessarily have a use for. Let A denote the set of people in this room, and*

B the set of possible hair colors. One might define a relation H from A to B by:
 $H = \{(x, y) \mid y \text{ is } x\text{'s hair color}\}$.

4.1.1 Properties of relations

Let A be a nonempty set and let \sim be a relation on A . We say that \sim is:

- (i) *reflexive* if $a \sim a$ for every $a \in A$;
- (ii) *symmetric* if $a \sim b$ implies $b \sim a$ for every $a, b \in A$;
- (iii) *antisymmetric* if $a \sim b$ and $b \sim a$ imply $a = b$ for every $a, b \in A$.
- (iv) *transitive* if $a \sim b$ and $b \sim c$ imply $a \sim c$ for every $a, b, c \in A$.

If \sim is reflexive, symmetric, and transitive, then we say that \sim is an *equivalence relation* on A . We will investigate equivalence relations more deeply later in this chapter.

Let $<$ denote the relation “less than” on \mathbb{R} . Since $1 \not< 1$, $<$ is not reflexive. This relation is also not symmetric since $1 < 7$, but $7 \not< 1$. The relation is antisymmetric since the hypothesis $a < b$ and $b < a$ is never satisfied. Finally, this relation is transitive because if $a < b$ and $b < c$, then $a < c$.

Note that symmetry and antisymmetry are not opposites. There are relations that satisfy both of these properties and relations that satisfy neither.

4.2 Equivalence Relations

We begin with a question: Is $1/2 = 2/4$?

After years of working with fractions, most people would answer yes. In that case, it should certainly be true that $(-1)^{1/2} = (-1)^{2/4}$, shouldn't it? But

$$(-1)^{1/2} = \sqrt{-1} = i$$

and

$$(-1)^{2/4} = \sqrt[4]{(-1)^2} = \sqrt[4]{1} = 1.$$

Maybe it's not so clear whether $1/2 = 2/4$ after all. What's actually happening here? When we work with rational numbers, there is an equivalence relation at play (see the next example) and our standard operations of multiplication and addition work well in conjunction with that equivalence relation. We should be more careful when using rational numbers as exponents.

Example 4.4. Let $\mathbb{F} = \{\frac{a}{b} \mid a, b \in \mathbb{Z} \text{ and } b \neq 0\}$, in other words \mathbb{F} is the set of fractions of integers. We define a relation \cong on \mathbb{F} by: $\frac{a}{b} \cong \frac{c}{d}$ if $ad = bc$. We claim that this is an equivalence relation on \mathbb{F} .

If $\frac{a}{b} \in \mathbb{F}$, clearly $ab = ab$. Hence $\frac{a}{b} \cong \frac{a}{b}$ and \cong is reflexive.

If $\frac{a}{b} \cong \frac{c}{d}$, then $ad = bc$. Using standard properties of multiplication, it follows that $cb = da$. Thus $\frac{c}{d} \cong \frac{a}{b}$ and \cong must be symmetric.

Finally, suppose that $\frac{a}{b} \cong \frac{c}{d}$ and $\frac{c}{d} \cong \frac{e}{f}$. Then $ad = bc$ and $cf = de$. Since we know $d \neq 0$, we can rewrite these as $a = \frac{bc}{d}$ and $e = \frac{cf}{d}$. Now we have

$$af = \frac{bc}{d}f = b\frac{cf}{d} = be,$$

so $\frac{a}{b} \cong \frac{e}{f}$ and \cong is transitive.

4.2.1 Equivalence classes

Let A be a set and let \sim be an equivalence relation on A . For any $a \in A$ the set $[a] = \{b \in A \mid b \sim a\}$ is called the *equivalence class* of a . The set of equivalence classes always satisfy the following:

- (i) For each $a \in A$, $a \in [a]$.
- (ii) If $[a] \neq [b]$, then $[a] \cap [b] = \emptyset$

Proof of (i). Since \sim is reflexive, $a \sim a$ for each $a \in A$. By definition this implies that $a \in [a]$ and (i) is satisfied. \square

Proof of (ii). Suppose that $[a] \cap [b] \neq \emptyset$ and let $c \in [a] \cap [b]$. Then $c \sim a$ and $c \sim b$.

Since $c \sim a$ and \sim is symmetric, it follows that $a \sim c$. Now we have $a \sim c$ and $c \sim b$, so $a \sim b$ by transitivity. Let $x \in [a]$, then $x \sim a$ by definition. From transitivity it follows that $x \sim b$, so $x \in [b]$. Hence $[a] \subset [b]$.

Since $a \sim b$, it must also be true that $b \sim a$ by symmetry. Let $y \in [b]$, then $y \sim b$ by definition. From transitivity it follows that $y \sim a$, so $y \in [a]$. Hence $[b] \subset [a]$.

Since $[a] \subset [b]$ and $[b] \subset [a]$, $[a] = [b]$ by Theorem 3.2. \square

Example 4.5. Consider the equivalence relation \equiv_5 of Exercise 4.2 there are five

equivalence classes associated with this equivalence relation:

$$\begin{aligned} [0] &= \{\dots, -10, -5, 0, 5, 10, \dots\} \\ [1] &= \{\dots, -9, -4, 1, 6, 11, \dots\} \\ [2] &= \{\dots, -8, -3, 2, 7, 12, \dots\} \\ [3] &= \{\dots, -7, -2, 3, 8, 13, \dots\} \\ [4] &= \{\dots, -6, -1, 4, 9, 14, \dots\} \end{aligned}$$

The equivalence classes are sets of integers that have the same remainder when divided by 5.

In the previous example, note that we had several ways to refer to each equivalence class. For example $[1] = [16]$, so we may as well have used $[16]$ as a name for this equivalence class. In this context the numbers 1 and 16 (or any other member of the class) are called *representatives* of this equivalence class, and any representative can be used to name the equivalence class.

Now let's consider the situation in Example 4.4. In this case there are infinitely many equivalence classes. It is not hard to show that the equivalence classes are sets of fractions that can all be reduced to the same fraction. One way of defining the rational numbers precisely is to first define the integers, then let the rational numbers be the set of equivalence classes of fractions of integers under this equivalence relation. Now $1/2$ and $2/4$ are different fractions, but represent the same rational number since they are in the same equivalence class.

4.3 Functions

You probably think of functions in several ways based on what you have seen in previous courses. Functions can be rules for computing things, some people think of them as machines (plug in one number and get out another), and of course we all know that a graph is only the graph of a function if it passes the vertical line test. Most of these ideas match the way that mathematicians thought of functions (and the way they worked with them) at some time or another. It wasn't until the late nineteenth century that the concept of a function was defined in terms of sets. We present such a definition here:

Let A and B be sets. A *function* f from A to B is a relation from A to B with the additional property that for each $a \in A$ there is exactly one ordered pair (a, b) in f having a as a first coordinate. In this case, the set A

is called the *domain* of f and the set B is called the *codomain* of f . If f is a function from A to B , we denote this by writing $f : A \rightarrow B$.

Do not confuse the codomain of a function with its range. The codomain of a function is the set that second coordinates must come from. The range is the subset of the codomain containing all of those second coordinates.

Note that our definition requires that every element of the domain be the first coordinate of one, and only one, ordered pair in a function. It does not, however, require that each element of the codomain be a second coordinate, or that it be the second coordinate of only one ordered pair. Satisfying these requirements would make our function surjective or injective, respectively. We discuss these kinds of functions in Section 4.4.

Before going any further, let's consider a couple of examples.

Example 4.6. Consider the sets $A = \{1, 2, 3, 4\}$ and $B = \{1, 3, 5, 7\}$. Define the following relations from A to B :

$$(i) f = \{(1, 3), (2, 1), (3, 5), (4, 7)\}$$

$$(ii) g = \{(1, 1), (3, 3)\}$$

$$(iii) h = \{(1, 1), (2, 3), (3, 5), (4, 1)\}$$

$$(iv) j = \{(1, 1), (2, 3), (3, 5), (4, 7), (2, 7)\}$$

$$(v) Pat = \{(1, 1), (2, 3), (3, 5), (4, 7), (1, 1)\}$$

The relation f is certainly a function. Each element of A is a first coordinate of one and only one ordered pair.

The relation g is not a function because 2 and 4 are elements of A , but are not first coordinates of ordered pairs.

The relation h is a function. The fact that 1 is a second coordinate of two ordered pairs, or that 7 is not the second coordinate of any, do not violate our definition.

The relation j is not a function because 2 is the first coordinate of two distinct ordered pairs.

Finally, Pat is also a function. We make two comments here. First, the fact that the ordered pair $(1, 1)$ is listed twice does not violate our definition of a function. Each ordered pair is either in the relation or not, it cannot be two distinct elements of the relation. Second, while we will usually use letters like f or g to denote functions, and adhering to this convention makes life easier for us, there is no real requirement that we do so. We can name a relation, and thus also a function, in some other manner if there is a reason to do so.

The functions in the previous example are obviously constructed to fit the definition, but may not strike you as the kinds of things you think of as functions. What about the kinds of functions we are used to?

Example 4.7. Define the following relations from \mathbb{R} to \mathbb{R} .

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

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

The relation f is a function (make sure you understand why). In fact, it is a function that should be familiar to you. The second coordinates are squares of the first coordinates, so this is the usual squaring function on \mathbb{R} .

The relation g is not a function from \mathbb{R} to \mathbb{R} . Despite appearances, there are elements of \mathbb{R} which are not first coordinates: -1 is one such element, but any negative number will do. This relation also violates our definition in another way. Both $(4, 2)$ and $(4, -2)$ satisfy the definition of g , so 4 is the first coordinate of more than one ordered pair in g . In fact, every positive real number is the first coordinate of two ordered pairs in g .

It probably seems unnatural to you to write the squaring function $f : \mathbb{R} \rightarrow \mathbb{R}$ as we did in Example 4.7. Wouldn't it be easier to write this function as $f(x) = x^2$, the same way we did in algebra or calculus classes? Once we know that f is a function, we can use this notation.

Suppose that $f : A \rightarrow B$ is a function. If $a \in A$ and $(a, b) \in f$, we say that b is the *value* of the function f at a and denote this by writing $b = f(a)$.

Now the notation $f(x) = x^2$ indicates that for each $x \in \mathbb{R}$, x^2 is the value of the function at x . Please be careful to distinguish between the name of the function f and an arbitrary value of the function $f(x)$.

Example 4.8. Here are some examples of important types of functions. Assume each of the sets is nonempty.

(i) Let A be any set. The identity function $i : A \rightarrow A$ is defined by $i(a) = a$ for each $a \in A$.

(ii) Let A and B be sets and let $b_0 \in B$. The function $k : A \rightarrow B$ defined by $k(a) = b_0$ for all $a \in A$ is called a constant function.

(iii) Let A and B be any sets. The coordinate projections $\pi_A : A \times B \rightarrow A$ and $\pi_B : A \times B \rightarrow B$ are defined by $\pi_A(a, b) = a$ and $\pi_B(a, b) = b$ for each $(a, b) \in A \times B$.

(iv) Let A be any subset of \mathbb{R} . The characteristic function $\chi_A : \mathbb{R} \rightarrow \{0,1\}$ of A is defined by

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

Note: try and sketch a graph of the characteristic function of \mathbb{Q} .

4.3.1 Binary Operations

Standard addition and multiplication on \mathbb{R} are both examples of functions used to compute a single real number from two given real numbers. A *binary operation* on a set X is a function $*$: $X \times X \rightarrow X$. If $*$ is a binary operation on X , we usually use the notation $x * y$ to denote the value $*(x, y)$.

Example 4.9. Standard addition and multiplication are binary operations on the set \mathbb{N} of natural numbers.

Example 4.10. Subtraction is not a binary operation on the set \mathbb{N} because it is not defined for all ordered pairs in $\mathbb{N} \times \mathbb{N}$. For example, $4 - 10$ is not a natural number. Note, however, that subtraction is a binary operation on the larger set \mathbb{Z} of integers.

Associative and Commutative operations

We say that a binary operation $*$ on a set X is *commutative* if $x * y = y * x$ for all $x, y \in X$, and *associative* if $(x * y) * z = x * (y * z)$ for all $x, y, z \in X$.

Addition and multiplication on \mathbb{Z} (or on \mathbb{N} or \mathbb{R} for that matter) are commutative and associative.

Subtraction on \mathbb{Z} is not commutative since, for example, $4 - 2 \neq 2 - 4$. Subtraction on \mathbb{Z} also fails to be associative since, for example, $(5 - 1) - 3 = 1 \neq 7 = 5 - (1 - 3)$.

Example 4.11. Define the operation $*$ on \mathbb{Z} by $a * b = (ab)^2$. We claim that $*$ is commutative. To see this, let a and b be arbitrary integers. Then:

$$\begin{aligned} a * b &= (ab)^2 \\ &= a^2b^2 \\ &= b^2a^2 \\ &= (ba)^2 \\ &= b * a \end{aligned}$$

We leave it as an exercise to determine whether or not $*$ is associative.

4.4 Injective, Surjective, and Bijective Functions

A function $f : X \rightarrow Y$ is said to be *injective* if for all $x, y \in X$, if $f(x) = f(y)$, then $x = y$. Another way of saying this is that each element of Y can be the second coordinate of at most one ordered pair in the function. Injective functions are also said to be *one-to-one*.

A function $f : X \rightarrow Y$ is said to be *surjective* if for each $y \in Y$, there is an $x \in X$ such that $f(x) = y$. In other words, every element of Y is the second coordinate of at least one ordered pair in the function. Surjective functions are also said to be *onto*.

A function $f : X \rightarrow Y$ that is both injective and surjective is said to be *bijective*. In this case, every element of Y is the second coordinate of exactly one ordered pair in the function. Bijective functions are also called *one-to-one correspondences*.

Example 4.12. Let's consider the functions we defined in Example 4.8.

For any nonempty set A , the identity function $i : A \rightarrow A$ is bijective. Assume first that $a, b \in A$ with $i(a) = i(b)$. Since $i(a) = a$ and $i(b) = b$, this implies that $a = b$ and i is injective. To see that i is surjective, let c be any element of A . Then $i(c) = c$, so i is surjective.

In general, constant functions are neither injective nor surjective. Assume for the moment that A and B each have more than one element and let $k : A \rightarrow B$ be the constant function $k(a) = b_0$. Choose two elements $a_1 \neq a_2$ in A , then $k(a_1) = b_0 = k(a_2)$ and k is not injective. If $b \neq b_0$ and $b \in B$, then $b \neq k(a)$ for any $a \in A$ and k is not surjective.

Next we consider the coordinate projection $\pi_A : A \times B \rightarrow A$. Once again, we assume that A and B have more than one element each. The function π_A is surjective. To see this suppose that $a \in A$ and choose some $b_0 \in B$, then $\pi_A(a, b_0) = a$. Choose $a \in A$ and $b_1 \neq b_2$ both in B . Then $(a, b_1) \neq (a, b_2)$ but $\pi_A(a, b_1) = a = \pi_A(a, b_2)$, so π_A is not injective. The coordinate projection $\pi_B : A \times B \rightarrow B$ is also surjective but not injective. The proofs are similar.

Finally, consider the characteristic function $\chi_A : \mathbb{R} \rightarrow \{0, 1\}$ of some subset A of \mathbb{R} . Once again we assume that A has more than one element. For any two elements $a \neq b$ of A we have $\chi_A(a) = 1 = \chi_A(b)$, so χ_A is not injective. If $a \in A$ and $c \in \mathbb{R} \setminus A$, then $\chi_A(a) = 1$ and $\chi_A(c) = 0$. Since 0 and 1 are the only elements of the codomain, χ_A is surjective. Note however that the proof that χ_A is surjective requires that we be able to find points in A and points in $\mathbb{R} \setminus A$. If $A = \mathbb{R}$, then χ_A is not surjective because $\mathbb{R} \setminus A = \emptyset$.

4.5 Compositions of Functions

Let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ be functions. The *composition* $g \circ f : X \rightarrow Z$ is the function from X to Z defined by $g \circ f(x) = g(f(x))$. In other words to find $g \circ f(x)$, first find $f(x)$, then plug the result into the function g . In terms of ordered pairs, the composition is

$$\{(x, z) \mid (x, y) \in f \text{ and } (y, z) \in g \text{ for some } y \in Y\}.$$

Note that the composition only makes sense if the codomain of f is contained in the domain of g .

Theorem 4.1. *Let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ be functions.*

(i) *If f and g are injective, then $g \circ f : X \rightarrow Z$ is injective.*

(ii) *If f and g are surjective, then $g \circ f : X \rightarrow Z$ is surjective.*

Proof of (i). Assume that $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ are injective. Let x_1 and x_2 be distinct elements of X . Since f is injective, $f(x_1)$ and $f(x_2)$ are distinct elements of the set Y . Now since g is injective, $g(f(x_1))$ and $g(f(x_2))$ are distinct elements of Z . Therefore $g \circ f$ is injective as desired. \square

Proof of (ii). Assume that $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ are surjective. Let z be an arbitrary element of Z . Since g is surjective, there must be some element $y \in Y$ such that $g(y) = z$. Now since f is surjective there must be an element $x \in X$ with $f(x) = y$, so $g(f(x)) = g(y) = z$ and $g \circ f : X \rightarrow Z$ is surjective as desired. \square

Combining the two parts of Theorem 4.1, we have the following:

Corollary 4.1.1. *If $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ are bijective functions, then $g \circ f : X \rightarrow Z$ is bijective.*

It is natural to ask whether or not the converses of the statements in Theorem 4.1 are true. We consider the converse of 4.1 (i) in the following example and theorem.

Example 4.13. *Define $f : \mathbb{N} \rightarrow \mathbb{R}$ by $f(n) = n^2$. Since no two natural numbers (which are all positive) have the same square, f is injective. Define $g : \mathbb{R} \rightarrow \mathbb{R}$ by $g(x) = x^2$. Since $g(-2) = 4 = g(2)$, g is not injective. Now we consider the composition $g \circ f : \mathbb{N} \rightarrow \mathbb{R}$ of these functions. For each $n \in \mathbb{N}$ we have $g(f(n)) = g(n^2) = (n^2)^2 = n^4$. Let $m, n \in \mathbb{N}$ with $m^4 = n^4$, then*

$$0 = m^4 - n^4 = (m^2 + n^2)(m - n)(m + n),$$

so $m = n$ or $m = -n$. But m and n are both positive, so $m \neq -n$ and it must be true that $m = n$. Hence $g \circ f : \mathbb{N} \rightarrow \mathbb{R}$ is injective.

This example shows that it is possible for $g \circ f$ to be injective when g is not. The next result says that if $g \circ f$ is injective, then it does follow that f is injective.

Theorem 4.2. *Let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ be functions such that $g \circ f : X \rightarrow Z$ is injective. Then $f : X \rightarrow Y$ must be injective.*

Proof. We prove that if $f : X \rightarrow Y$ is not injective, then $g \circ f : X \rightarrow Z$ is not injective, which is the contrapositive of the desired proposition. Suppose that f is not injective, then there must be elements $x_1 \neq x_2$ in X with $f(x_1) = f(x_2)$. Since g is a function, it must be true that $g(f(x_1)) = g(f(x_2))$. Since $x_1 \neq x_2$ but $g \circ f(x_1) = g \circ f(x_2)$, $g \circ f$ is not injective. \square

4.5.1 Inverses of functions

If $f : X \rightarrow Y$ is a function, the *inverse* of f is the relation $\{(y, x) \mid (x, y) \in f\}$ from Y to X . In general, the inverse of a function is not a function, for example:

Example 4.14. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be the function defined by $f(x) = x^2$, then the inverse of f is the relation $g = \{(x^2, x) \mid x \in \mathbb{R}\}$. Note that g is not a function since both $(4, 2)$ and $(4, -2)$ are in g .*

If $f : X \rightarrow Y$ is a function and the inverse of f is also a function, we use f^{-1} to denote the inverse.

Question. What conditions must a function $f : X \rightarrow Y$ satisfy in order to insure that the inverse of f is a function from Y to X ? Such a function is said to be *invertible*.

Thinking back to what you've seen in previous courses for a moment, you are likely to have been taught that a function from \mathbb{R} to \mathbb{R} will have an inverse if it passes the "horizontal line test." Do you remember why? One way to think of this is that the graph of the inverse is the reflection through the line $y = x$ of the graph of the function. Since we want the reflection (the graph of the inverse) to pass the vertical line test, the graph of the original function should pass the horizontal line test. There should be some relationship between this condition and the answer to our question.

For the graph of a function to pass the horizontal line test, i.e. no horizontal line intersects the graph more than once, means that no two points

of the graph have the same height. In other words, no two distinct points on the graph have the same y -coordinate. Rephrasing, if (x_1, y) and (x_2, y) are both on the graph, then $x_1 = x_2$. But this is just our definition of what it means for a function to be injective. Perhaps injective functions and invertible functions are the same thing? We can show that every invertible function is injective.

Theorem 4.3. *If $f : X \rightarrow Y$ is an invertible function, then f is injective.*

Proof. Suppose that f is not injective, then there are two elements $x_1 \neq x_2$ of X such that $f(x_1) = f(x_2)$. Let $y = f(x_1)$. By definition both (y, x_1) and (y, x_2) must be elements of the inverse of f . Since $x_1 \neq x_2$, this implies that the inverse of f is not a function. Therefore, f is not invertible. \square

Unfortunately, this is not a complete answer to our question because the converse of Theorem 4.3 is not true. We have found a condition that all invertible functions must satisfy, but not all functions that satisfy this condition are invertible. The following example illustrates the difficulty.

Example 4.15. *Define $f : \mathbb{N} \rightarrow \mathbb{N}$ by $f(n) = n + 1$. This function is injective (if $m + 1 = n + 1$, then $m = n$), but not invertible according to our definition. The inverse relation g contains all ordered pairs of the form $(n + 1, n)$ where $n \in \mathbb{N}$. For g to be a function from \mathbb{N} to \mathbb{N} , every element of \mathbb{N} must be the first coordinate of exactly one ordered pair in g . The problem here is that 1 is in \mathbb{N} , but $1 \neq n + 1$ for any $n \in \mathbb{N}$, so 1 is not the first coordinate of any of the ordered pairs in g . Therefore g is not a function from \mathbb{N} to \mathbb{N} .*

For a function $f : X \rightarrow Y$ to be invertible, every element of the codomain Y must be a first coordinate of some ordered pair in the inverse. That in turn means that every element of the codomain must be the second coordinate of some ordered pair in f . This is exactly what it means to say that f is surjective. This leads us to believe the following:

Theorem 4.4. *If $f : X \rightarrow Y$ is invertible, then f is surjective.*

Proof. Assume that $f : X \rightarrow Y$ is not surjective, then there is some element $y \in Y$ so that $y \neq f(x)$ for any $x \in X$. In other words, y is not the second coordinate of an ordered pair in f . By definition then, y will not be the first coordinate of any ordered pair in the inverse of f . Hence the inverse of f is not a function from Y to X and f is not invertible. \square

We are now ready to give a complete answer to our question in the form of the following theorem.

Theorem 4.5. *A function $f : X \rightarrow Y$ is invertible if and only if it is bijective.*

Proof. If f is invertible, then we may use Theorems 4.3 and 4.4 to establish the fact that f is bijective.

To see that the converse is true, suppose that $f : X \rightarrow Y$ is a bijective function. Let $g = \{(y, x) \mid (x, y) \in f\}$ be the inverse of f . We must show that g is a function from Y to X , i.e. that every $y \in Y$ is the first coordinate of exactly one ordered pair in g . Let $y \in Y$. Since f is surjective, $y = f(x)$ for some $x \in X$. By definition $(x, y) \in f$ implies that $(y, x) \in g$, so y is the first coordinate of at least one ordered pair in g . Now suppose that (y, x_1) and (y, x_2) are both in g . By definition this means that $f(x_1) = y$ and $f(x_2) = y$. Since f is injective this implies that $x_1 = x_2$. It follows that y cannot be the first coordinate of more than one ordered pair in g , so g is a function from Y to X and f is invertible. \square

Chapter 4 Exercises

4.1. Determine whether or not each of the following relations is reflexive, symmetric, antisymmetric, and/or transitive. Are any of these equivalence relations?

- (i) The relation \leq on \mathbb{R} .
- (ii) Equality on \mathbb{R} , i.e. the set of ordered pairs of real numbers whose first and second coordinates are equal.
- (iii) The relation $|$ on \mathbb{N} , where $a | b$ means that $b = an$ for some $n \in \mathbb{N}$.
- (iv) The relation \sim on \mathbb{N} defined by $a \sim b$ if there is an integer $n > 1$ that evenly divides both a and b .

4.2. Define the relation \equiv_5 on \mathbb{Z} by: $a \equiv_5 b$ if there is an integer n so that $b - a = 5n$. Show that \equiv_5 is an equivalence relation on \mathbb{Z} . Note: this is a fairly common equivalence relation. The phrase $a \equiv_5 b$ is usually read “ a is equivalent to b modulo 5.”

4.3. Define a relation \sim on $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ by $(x, y) \sim (w, z)$ if $x^2 + y^2 = w^2 + z^2$. Show that \sim is an equivalence relation on \mathbb{R}^2 .

4.4. Consider the equivalence relation defined in Exercise 4.3. How many equivalence classes are there? What are the equivalence classes? In other words, for a typical $(x, y) \in \mathbb{R}^2$, what is $[(x, y)]$?

4.5. Let A be a nonempty set. Explain why there are no functions from A to \emptyset .

4.6. Is standard division a binary operation on \mathbb{N} ? on \mathbb{R} ? Justify your answer in each case.

4.7. Is the dot product a binary operation on \mathbb{R}^3 ? Recall that the dot product is defined by $(a, b, c) \cdot (d, e, f) = ad + be + cf$.

4.8. Determine whether or not the operation $*$ defined in Example 4.11 is an associative operation on \mathbb{Z} .

4.9. Determine whether or not each of the following binary operations on \mathbb{R} is (a) commutative, (b) associative.

(i) $a * b = |a - b|$.

(ii) $a * b = \frac{a}{b^2+1}$.

(iii) $a * b = \max\{a, b\}$.

4.10. Let $f : X \rightarrow Y$ be a function; let A and B be subsets of X . Determine which of the following are true. If a statement is true, prove it. If a statement is false, find a counterexample.

(i) $f(A \cup B) \subset f(A) \cup f(B)$

(ii) $f(A \cup B) \supset f(A) \cup f(B)$

(iii) $f(A \cap B) \subset f(A) \cap f(B)$

(iv) $f(A \cap B) \supset f(A) \cap f(B)$

(v) $f(X \setminus A) \subset Y \setminus f(A)$

(vi) $f(X \setminus A) \supset Y \setminus f(A)$

Note: the notation $f(A)$ is used to indicate the set $\{f(a) \mid a \in A\}$.

4.11. For each of the false statements in the previous exercise, determine whether or not they are true under the following conditions. Prove or give a counterexample in each case.

(i) $f : X \rightarrow Y$ is an injective function.

(ii) $f : X \rightarrow Y$ is a surjective function.

4.12. Let $A = \{1, 2, 3\}$. For each of the following, either find a function satisfying the indicated properties or prove that no such function exists.

(i) A bijective function $f : A \rightarrow A$ other than the identity function.

(ii) An injective function $g : A \rightarrow A$ that is not surjective.

(iii) A surjective function $h : A \rightarrow A$ that is not injective.

(iv) A function $j : A \rightarrow A$ that is neither injective nor surjective.

4.13. Let $A = \{1, 2, 3\}$ and $B = \{4, 5\}$. For each of the following, either find a function satisfying the indicated properties or prove that no such function exists.

- (i) A bijective function $f : A \rightarrow B$.
- (ii) An injective function $g : A \rightarrow B$ that is not surjective.
- (iii) A surjective function $h : A \rightarrow B$ that is not injective.
- (iv) A function $j : A \rightarrow B$ that is neither injective nor surjective.

4.14. For each of the following, either find a function satisfying the indicated properties or prove that no such function exists.

- (i) A bijective function $f : \mathbb{N} \rightarrow \mathbb{N}$ other than the identity.
- (ii) An injective function $g : \mathbb{N} \rightarrow \mathbb{N}$ that is not surjective.
- (iii) A surjective function $h : \mathbb{N} \rightarrow \mathbb{N}$ that is not injective.
- (iv) A function $j : \mathbb{N} \rightarrow \mathbb{N}$ that is neither injective nor surjective.

4.15. Find examples of sets X , Y , and Z and functions $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ so that $g \circ f : X \rightarrow Z$ is surjective but $f : X \rightarrow Y$ is not surjective.

4.16. Prove the following:

Theorem 4.6. *Let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ be functions such that $g \circ f : X \rightarrow Z$ is surjective. Then $g : Y \rightarrow Z$ must be surjective.*

4.17. The standard integer addition operation $+$ is a function from $\mathbb{Z}^2 \rightarrow \mathbb{Z}$. Show that if $a \equiv_5 b$ (see Example 4.2) and $c \equiv_5 d$, then $a + c \equiv_5 b + d$. This shows that addition is *well-defined* with respect to this equivalence relation.

4.18. Let $f : X \rightarrow Y$ be an invertible function and let $f^{-1} : Y \rightarrow X$ denote the inverse of f . Show that f^{-1} is bijective and that f is the inverse of f^{-1} .

4.19. In an algebra or calculus class, we usually say that two functions $f : X \rightarrow Y$ and $g : Y \rightarrow X$ are inverses if $f \circ g : Y \rightarrow Y$ and $g \circ f : X \rightarrow X$ are identity functions, i.e. $f(g(y)) = y$ for all $y \in Y$ and $g(f(x)) = x$ for all $x \in X$. Is this definition equivalent to the definition given in this chapter? Explain.

Chapter 5

The Real Numbers

Introduction

We now turn our attention to developing a mathematical description of the real numbers. All of the results of calculus can be derived from these basic properties of the real number system, which is usually done in a first course in real analysis. Our basic assumptions about the real numbers are called axioms, and they are divided into several groups. Many of these axioms will be familiar to you from previous courses in algebra.

5.1 Field Axioms

Taken as a group, the field axioms tell us that the real numbers together with the operations of addition and multiplication form what is known as a *field*. Very informally, you might think of a field as something that satisfies the usual rules you encountered in high school algebra. Fields and related objects are studied in more depth in a course in abstract algebra.

Field Axioms. The set of real numbers \mathbb{R} is a collection of objects together with the two binary operations addition and multiplication that satisfy the following properties:

- (i) Commutative Laws: For every $a, b \in \mathbb{R}$, $a + b = b + a$ and $ab = ba$.
- (ii) Associative Laws: For every $a, b, c \in \mathbb{R}$, $(a + b) + c = a + (b + c)$ and $(ab)c = a(bc)$.
- (iii) Distributive Law: For every $a, b, c \in \mathbb{R}$, $a(b + c) = ab + ac$.

- (iv) Identities: There are distinct elements 0 and 1 in \mathbb{R} such that $a \cdot 1 = a$ and $a + 0 = a$ for every $a \in \mathbb{R}$.
- (v) Additive Inverses: For every $a \in \mathbb{R}$, there is an element $-a \in \mathbb{R}$ such that $a + (-a) = 0$.
- (vi) Multiplicative Inverses: For every $a \in \mathbb{R}$ with $a \neq 0$, there is an element $a^{-1} \in \mathbb{R}$ such that $aa^{-1} = 1$.

Any set of objects satisfying all of these properties is a field. Other examples of fields include the fields of rational numbers \mathbb{Q} , real numbers \mathbb{R} , and complex numbers \mathbb{C} . As we develop further axioms for the real numbers, we will also narrow the list of fields that satisfy all of the axioms.

A number of the familiar algebraic properties of the real numbers are immediate consequences of the field axioms. A few of them are collected in the following theorem, though there are certainly many others.

Theorem 5.1. *For any real numbers a, b, c , the following are true:*

- (i) If $a + b = a + c$, then $b = c$.
- (ii) $-(-a) = a$.
- (iii) If $a \neq 0$ and $ab = ac$, then $b = c$.
- (iv) If $a \neq 0$, then $(a^{-1})^{-1} = a$.
- (v) $a \cdot 0 = 0$.
- (vi) $-a = (-1)a$.
- (vii) If $ab = 0$, then $a = 0$ or $b = 0$.

Proof. We prove parts (ii), (iii), and (vi). The remaining proofs are exercises.

We first show that (ii) is true. To see this, note that $-(-a)$ is the additive inverse of $-a$. The reader should justify each of the following steps:

$$\begin{aligned} a + (-a) &= 0 \\ &= -a + (-(-a)) \\ &= -(-a) + (-a) \end{aligned}$$

so $a + (-a) = -(-a) + (-a)$. We apply part (i) to see that $a = -(-a)$ as desired.

To prove (iii), suppose that $a \neq 0$ and $ab = ac$. Since $a \neq 0$, a has a multiplicative inverse a^{-1} . The reader should give a justification for each step of the following:

$$\begin{aligned} b &= 1 \cdot b \\ &= (a^{-1}a)b \\ &= a^{-1}(ab) \\ &= a^{-1}(ac) \\ &= (a^{-1}a)c \\ &= 1 \cdot c \\ &= c \end{aligned}$$

To prove (vi), let a be any real number. Then:

$$\begin{aligned} a + (-1)a &= 1 \cdot a + (-1)a \\ &= a(1 + (-1)) \\ &= a \cdot 0 \\ &= 0 \end{aligned}$$

Hence $a + (-1)a = 0 = a + (-a)$ and we may apply part (i) to see that $(-1)a = -a$. \square

5.2 Order Axioms

As noted previously, there are a number of common fields. One of the differences between \mathbb{R} and \mathbb{C} is that we think of the real numbers as lying in order along a line. There is no natural way to organize the complex numbers in this fashion. The next group of axioms make precise what we mean by saying that the real numbers form an ordered field.

Order Axioms. There is an order $<$ defined on the real numbers \mathbb{R} satisfying:

- (i) Transitivity: If $a < b$ and $b < c$, then $a < c$.
- (ii) Trichotomy: For every two real numbers a and b , exactly one of the following holds:

$$a < b \text{ or } a = b \text{ or } b < a$$

(iii) If $a < b$, then $a + c < b + c$ for every $c \in \mathbb{R}$.

(iv) If $a < b$ and $c > 0$, then $ac < bc$.

We derive several consequences of the fact that \mathbb{R} is an ordered field.

Theorem 5.2. *The following statements are true in \mathbb{R} :*

(i) $a > 0$ iff $-a < 0$.

(ii) If $a < b$, then $-a > -b$.

(iii) If $a \neq 0$, then $a^2 > 0$.

(iv) $1 > 0$.

Proof. To prove part (i), assume first that $a > 0$. By Trichotomy we have either $-a < 0$, $-a = 0$, or $-a > 0$. We will show that two of these possibilities lead to contradictions, forcing the remaining option to be true. If $-a = 0$ then we have $a = a + 0 = a + (-a) = 0$, which contradicts the fact that $a > 0$. If $-a > 0$ then we have $0 = a + (-a) > a + 0 = a$, once again contradicting the fact that $a > 0$. We have shown that neither $-a = 0$ nor $-a > 0$ can be true, so it must be the case that $-a < 0$ as desired. The remaining direction of the proof of part (i) is left as a homework exercise.

We next prove part (ii). The reader should determine which axiom or theorem justifies each step of the following:

$$\begin{aligned} a &< b \\ a + (-b) &< b + (-b) \\ a + (-b) &< 0 \\ ((-a) + a) + (-b) &< (-a) + 0 \\ -b &< -a \end{aligned}$$

The proofs of the remaining parts of the theorem are left to the reader. \square

Theorem 5.3. *Let $a, b \in \mathbb{R}$.*

(i) If $a > 0$ and $b > 0$, then $ab > 0$.

(ii) If $a < 0$ and $b < 0$, then $ab > 0$.

(iii) If $a > 0$ and $b < 0$, then $ab < 0$.

Theorem 5.4. Let $a \in \mathbb{R}$.

(i) If $a > 0$, then $a^{-1} > 0$.

(ii) If $a < 0$, then $a^{-1} < 0$.

Proof. To prove part (i) we assume $a > 0$ and apply Trichotomy. If $a^{-1} < 0$, then $1 = aa^{-1} < 0$ by Theorem 5.3, but this contradicts Theorem 5.2(iv). If $a^{-1} = 0$, then we have $1 = aa^{-1} = a \cdot 0 = 0$, which again contradicts Theorem 5.2(iv). The only remaining possibility is that $a^{-1} > 0$ as desired.

The proof of (ii) is similar and is left for the reader. \square

Theorem 5.5. Let $a \geq 0$ and $b \geq 0$. Then $a < b$ iff $a^2 < b^2$.

Proof. We assume throughout that $a \geq 0$ and $b \geq 0$. By Trichotomy we have exactly one of $a < b$, $a = b$, or $a > b$; we also have exactly one of $a^2 < b^2$, $a^2 = b^2$, or $a^2 > b^2$.

If $a < b$, then $a^2 < b^2$ by an application of exercise 5.8.

If $a = b$, then $a^2 = b^2$.

If $a > b$, then an $a^2 > b^2$ by application of exercise 5.8.

It now follows that $a < b$ if and only if $a^2 < b^2$. \square

5.3 Completeness of \mathbb{R}

Our axioms so far insure that \mathbb{R} is an ordered field. Unfortunately, \mathbb{R} is not the only ordered field. The set \mathbb{Q} of rationals also forms an ordered field. We require one more assumption to insure that we are talking about the real numbers as we usually think of them. Intuitively, we need to know that there are no holes in the real number line.

5.3.1 Upper and lower bounds

Definition. Let A be a nonempty set of real numbers. We say that a number u is an *upper bound* for A if $x \leq u$ for every $x \in A$. We say that a number m is a *lower bound* for A if $x \geq m$ for every $x \in A$. We say that A is *bounded above* if A has an upper bound, that A is *bounded below* if A has a lower bound, and that A is *bounded* if A is bounded above and below.

Example 5.1. Let $A = \{a \mid a^2 < 5\}$, $B = [0, 7)$, $C = \mathbb{N}$, and $D = \mathbb{Z}$.

The number 5 is an upper bound for A and -3 is a lower bound for A .

The set B has an upper bound at 7 and a lower bound at 0.

The set C has a lower bound at 1, but no upper bound.

The set D has no upper or lower bounds.

The previous example illustrates several things. It is possible for a nonempty set to have both upper and lower bounds, just one bound, or no bounds at all. Upper and lower bounds may or may not be elements of the set. Finally, upper and lower bounds are not unique. Transitivity implies that if M is an upper bound for a set A , then every number larger than M is also an upper bound for A . Similarly, if m is a lower bound for A then every number smaller than m is a lower bound for A .

Consider the set $A = \{a \mid a^2 \leq 5\}$ from the previous example again. We claimed that 5 is an upper bound, and that is certainly true, but note that 3 is also an upper bound. In some sense this smaller upper bound is a “better” bound for A because it puts a tighter restriction on the size of the elements of A . This is approximately like saying that Grand Forks is a better way to describe the location of UND than North Dakota. In this sense the best possible upper bound would be the smallest one, if there is one. In this particular case, A has a smallest upper bound in \mathbb{R} ($\sqrt{5}$) but not in \mathbb{Q} . This is the basic difference between \mathbb{Q} and \mathbb{R} that we wish to capture in an axiom. First, we need another couple of definitions.

Definition. Let A be a nonempty subset of \mathbb{R} . We say that a number u is a *least upper bound* (LUB) for A if both of the following are true:

- (i) For every $a \in A$, $a \leq u$.
- (ii) If b is an upper bound for A , then $u \leq b$.

We say that a number m is a *greatest lower bound* (GLB) for A if both of the following are true:

- (i) For every $a \in A$, $a \geq m$.
- (ii) If b is a lower bound for A , then $m \geq b$.

Note that a LUB is sometimes called a *supremum* and a GLB is sometimes called an *infimum*.

It is not hard to prove that, unlike upper bounds, a set can have only one least upper bound. This fact is made explicit in the following theorem, whose proof is an exercise.

Theorem 5.6. *Let A be a nonempty subset of \mathbb{R} . If u and v are least upper bounds for A , then $u = v$.*

The following theorem says that the least upper bound of a set must in some sense be “close to” the set.

Theorem 5.7. *Let A be a nonempty subset of \mathbb{R} and let u be the least upper bound for A . If $x < u$, then there is an element $a \in A$ such that $x < a$.*

Proof. Suppose that u is the least upper bound for A and that $x < u$. Assume that there is no $a \in A$ such that $x < a$, then $a \leq x$ for every $a \in A$ by Trichotomy. It follows by definition that x is an upper bound for A , but this contradicts the fact that u is the least upper bound for A since $x < u$. \square

We are now ready to state our final axiom, which says that \mathbb{R} is *complete*.

The Completeness Axiom. Let A be a nonempty subset of \mathbb{R} . If A has an upper bound, then A has a least upper bound.

There are a number of consequences of the Completeness Axiom, most of which are beyond the scope of this text. We will look at only a couple of them. We begin with the fairly intuitive seeming fact that for any real number x , there is a natural number larger than x .

Theorem 5.8. (*Archimedean Property*) *If $x \in \mathbb{R}$, then there is a number $n_x \in \mathbb{N}$ such that $x \leq n_x$.*

Proof. Assume to the contrary that there is a real number x so that $n < x$ for every natural number n . In this case x is an upper bound for the set \mathbb{N} . Applying the Completeness Axiom, \mathbb{N} must have a least upper bound m in \mathbb{R} . Since $m - 1 < m$, Theorem 5.7 implies that there is a natural number n such that $m - 1 < n$. Now $n + 1$ is also a natural number and $m < n + 1$, which contradicts the fact that m is an upper bound for \mathbb{N} . \square

We previously commented, without proof, that the Completeness Axiom would allow us to distinguish \mathbb{R} from \mathbb{Q} . We will now make this explicit by proving that there is at least one real number that is not rational.¹ We proved previously (Theorem 2.5) that there is no rational number whose square is 2. We now show that the Completeness Axiom implies that there must be a real number x with $x^2 = 2$.

Theorem 5.9. *There is a number $x \in \mathbb{R}$ such that $x^2 = 2$.*

Proof. Let $A = \{a \in \mathbb{R} \mid a^2 \leq 2\}$. Note that A is not empty since $1 \in A$. We claim that 2 is an upper bound for A . To see this note that if $t > 2$, then $t^2 > 2 \cdot 2 = 4 > 2$ (see exercise 5.8), so $t \notin A$. Since A is a nonempty subset

¹In fact there are more irrational numbers than there are rational numbers, as we shall see in the final chapter.

of \mathbb{R} that is bounded above, A must have a least upper bound x . We will use Trichotomy to show that $x^2 = 2$.

Assume first that $x^2 < 2$. Note that $\frac{2x+1}{2-x^2}$ is a positive real number. By the Archimedean Property there must be a natural number n such that $n > \frac{2x+1}{2-x^2}$. Since $n > \frac{2x+1}{2-x^2} > 0$, it follows that $\frac{1}{n} < \frac{2-x^2}{2x+1}$. We will show that $x + \frac{1}{n} \in A$, which will contradict the fact that x is an upper bound for A . To see that $x + \frac{1}{n} \in A$, we compute:

$$\begin{aligned} \left(x + \frac{1}{n}\right)^2 &= x^2 + \frac{2x}{n} + \frac{1}{n^2} \\ &= x^2 + \frac{1}{n} \left(2x + \frac{1}{n}\right) \\ &\leq x^2 + \frac{1}{n}(2x+1) \\ &< x^2 + (2-x^2) \\ &= 2 \end{aligned}$$

Now $x + \frac{1}{n} \in A$, which leads to the desired contradiction. It follows that $c^2 < 2$ cannot hold.

Next suppose that $x^2 > 2$. In this case $\frac{2x}{x^2-2}$ is a positive number and we may find $m \in \mathbb{N}$ such that $m > \frac{2x}{x^2-2}$. This in turn implies that $\frac{1}{m} < \frac{x^2-2}{2x}$. We show that $(x - \frac{1}{m})^2 > 2$:

$$\begin{aligned} \left(x - \frac{1}{m}\right)^2 &= x^2 - \frac{2x}{m} + \frac{1}{m^2} \\ &> x^2 - \frac{2x}{m} \\ &> x^2 - (2x) \frac{x^2-2}{2x} \\ &= 2 \end{aligned}$$

Now Theorem 5.5 implies that if $s > x - \frac{1}{m}$, then $s^2 > (x - \frac{1}{m})^2 > 2$, so $x - \frac{1}{m}$ is an upper bound for A . This would contradict the fact that x is the least upper bound for A , so $x^2 > 2$ cannot hold.

The only possibility left is that $x^2 = 2$ as desired. \square

Chapter 5 Exercises

- 5.1. Prove part (i) of Theorem 5.1.
- 5.2. Prove part (iv) of Theorem 5.1.
- 5.3. Prove part (v) of Theorem 5.1.
- 5.4. Prove part (vii) of Theorem 5.1.
- 5.5. Use the field axioms to show that $-0 = 0$ and $1^{-1} = 1$.
- 5.6. Prove that $(-a)b = -(ab) = a(-b)$ for all real numbers a and b .
- 5.7. Use the field axioms and Theorem 5.1 to show that for any $a \in \mathbb{R}$, $(-a)(-a) = a^2$.
- 5.8. If $a > b \geq 0$ and $c > d \geq 0$, prove that $ac > bd$.
- 5.9. Complete the proof of part (i) of Theorem 5.2 by showing that if $a < 0$, then $-a > 0$.
- 5.10. Prove part (iii) of Theorem 5.2.
- 5.11. Prove part (iv) of Theorem 5.2.
- 5.12. Prove Theorem 5.3.
- 5.13. Prove part (ii) of Theorem 5.4.
- 5.14. Prove Theorem 5.6.
- 5.15. Prove that if u is an upper bound for A and $u \in A$, then u is the least upper bound for A .
- 5.16. Show that every nonempty subset of \mathbb{R} with a lower bound has a greatest lower bound.
- 5.17. Let A and B be two nonempty subsets of \mathbb{R} such that $A \cup B = \mathbb{R}$. If A and B satisfy the further property that $a < b$ for every $a \in A$ and $b \in B$, then A and B form a *Dedekind cut* of \mathbb{R} . The Completeness Axiom is sometimes replaced with Dedekind's Axiom, which says that given any Dedekind cut of \mathbb{R} , either A has a largest element or B has a smallest element. Prove that Dedekind's Axiom is logically equivalent to the Completeness Axiom. In other words, show that if we assume only the field and order axioms, then the Completeness Axiom is true iff Dedekind's Axiom is true.

Chapter 6

Introduction to Cardinality

Introduction

What do we mean when we say that there are *four* suits in a standard deck of cards? More generally, what does it mean to count any collection of objects? Since you may no longer actually think about counting, it may help to watch a child who is just learning to count. Given four objects to count, the child is likely to point to each object in turn and count “one, two, three, four.” In other words, the child is explicitly constructing a bijective function between the objects she is counting and the elements of the set $\{1, 2, 3, 4\}$. Our goal in this chapter is to extend this idea to infinite sets.

6.1 The Cardinality of a Set

Let A and B be two sets. We define the relation \equiv by $A \equiv B$ if there is a bijective function $f : A \rightarrow B$. In this case we say that A and B are *equinumerous* or that they have the same *cardinality*. We define $A \preceq B$ to mean that there is an injective function $f : A \rightarrow B$. We may also write this as $B \succeq A$. We write $A \prec B$ to indicate that $A \preceq B$ and $A \not\equiv B$. Note: it is fairly common to use the notation $|A| = |B|$ rather than $A \equiv B$.

Theorem 6.1. *The relation \equiv defined above is an equivalence relation.*

Proof. Let A , B , and C be arbitrary sets.

Since the identity function on any set is a bijection, $A \equiv A$ and \equiv is reflexive.

If $A \equiv B$, then there is a bijection $f : A \rightarrow B$. Applying Theorem 4.5, the function f is invertible. By Exercise 4.18 the inverse function $f^{-1} : B \rightarrow A$ is a bijection. Hence $B \equiv A$ and \equiv is symmetric.

To see that \equiv is transitive, suppose that $A \equiv B$ and $B \equiv C$. By definition there are bijective functions $f : A \rightarrow B$ and $g : B \rightarrow C$. Now apply Corollary 4.1.1 to see that $g \circ f : A \rightarrow C$ is a bijective function. Hence $A \equiv C$ and \equiv is transitive. Therefore, \equiv is an equivalence relation as desired. \square

Note that Theorem 6.1 allows us to say that two sets A and B have the same cardinality if we are able to find a bijection from A to B or a bijection from B to A .

Theorem 6.2. *The relation \preceq is reflexive and transitive.*

Proof. Let A , B , and C be any sets. Since the identity function on any set is injective, $A \preceq A$ and \preceq is reflexive. To see that \preceq is transitive, suppose that $A \preceq B$ and $B \preceq C$. By definition there are injections $f : A \rightarrow B$ and $g : B \rightarrow C$. We apply Theorem 4.1 to see that the composition $g \circ f : A \rightarrow C$ is also injective, hence $A \preceq C$ as desired. \square

While we would probably not expect \preceq to be symmetric, it wouldn't be too surprising if it was antisymmetric. If $A \preceq B$ and $B \preceq A$, does it follow that $A \equiv B$? Georg Cantor (1845-1918) was interested in just this question. One of his doctoral students, Felix Bernstein (1878-1956) was able to prove that the answer is yes. The resulting theorem is usually known as the Cantor-Bernstein Theorem.

Theorem 6.3 (Cantor-Bernstein). *If $A \preceq B$ and $B \preceq A$, then $A \equiv B$.*

We defer the proof of this theorem to the Appendix. The following example uses the same ideas as the proof in order to construct a bijection between the open interval $(-1, 1)$ and the closed interval $[-1, 1]$.

Example 6.1. *Define the functions $f : (-1, 1) \rightarrow [-1, 1]$ and $g : [-1, 1] \rightarrow (-1, 1)$ by $f(x) = x$ and $g(x) = x/2$, respectively. It is easy to show that both of these functions are injective, so $(-1, 1) \preceq [-1, 1]$ and $[-1, 1] \preceq (-1, 1)$. The Cantor-Bernstein Theorem implies that there must be a bijection between the open interval $(-1, 1)$ and the closed interval $[-1, 1]$, but the theorem itself doesn't tell us how to find such a bijection. We construct such a bijection here.*

We wish to find a bijective function $h : (-1, 1) \rightarrow [-1, 1]$. For most elements $x \in (-1, 1)$ we want $h(x) = f(x) = x$. Unfortunately, if we let $h(x) = f(x)$ for all x , then the function is not bijective because $-1 \neq f(x)$ and $1 \neq f(x)$ for

all $x \in (-1, 1)$. We use the function g to help us fix this problem. Consider first the element $1 \in [-1, 1]$. While $1 \neq f(x)$ for any x , there is an element of the open interval associated with 1 by g . In particular $g(1) = 1/2$. We could define h so that $h(1/2) = 1$ and $h(x) = f(x)$ for other $x \in (-1, 1)$, but that creates a new problem. Now 1 is in the image of the function, but $1/2$ is not. To fix this we let $h(1/4) = 1/2$, because $g(1/2) = 1/4$. Of course, now $1/4$ is not in the image of h . We continue fixing one problem at a time using g , each time creating a new problem. Naturally, we will need to do the same with -1 .

These particular functions are simple enough that we can write down a formula for the function h that we end up with. We end up defining $h : (-1, 1) \rightarrow [-1, 1]$ by

$$h(x) = \begin{cases} 2^{-(n-1)} & \text{if } x = 2^{-n} \text{ for some } n \in \mathbb{N} \\ -2^{-(n-1)} & \text{if } x = -2^{-n} \text{ for some } n \in \mathbb{N} \\ x & \text{otherwise} \end{cases}$$

We claim that this function is the desired bijection.

We first show that h is surjective. To see this, let $t \in [-1, 1]$. If $t = 2^{-(n-1)}$ for some $n \in \mathbb{N}$, then $h(2^{-n}) = t$. If $t = -2^{-(n-1)}$ for some $n \in \mathbb{N}$, then $h(-2^{-n}) = t$. For any other $t \in [-1, 1]$ we have $h(t) = t$. In any case $t = h(x)$ for some $x \in (-1, 1)$, so h is surjective.

To see that h is injective, suppose that $x \neq y$ are both elements of $(-1, 1)$. We consider several cases.

Case 1. $x = 2^{-n}$, $y = 2^{-k}$ for some $n, k \in \mathbb{N}$. Since $x \neq y$, it follows that $n \neq k$. Hence $n - 1 \neq k - 1$ and we have $h(x) = 2^{-(n-1)} \neq 2^{-(k-1)} = h(y)$.

Case 2. $x = -2^{-n}$, $y = -2^{-k}$ for some $n, k \in \mathbb{N}$. Since $x \neq y$, it follows that $n \neq k$. Hence $n - 1 \neq k - 1$ and we have $h(x) = -2^{-(n-1)} \neq -2^{-(k-1)} = h(y)$.

Case 3. $x = 2^{-n}$, $y = -2^{-k}$ for some $n, k \in \mathbb{N}$. In this case we have $h(x) = 2^{-(n-1)} \neq -2^{-(k-1)} = h(y)$.

Case 4. $x = -2^{-n}$, $y = 2^{-k}$ for some $n, k \in \mathbb{N}$. In this case we have $h(x) = -2^{-(n-1)} \neq 2^{-(k-1)} = h(y)$.

Case 5. $x = \pm 2^{-n}$ for some $n \in \mathbb{N}$ and $y \neq \pm 2^{-k}$ for any $k \in \mathbb{N}$. In this case we have $h(x) = \pm 2^{-(n-1)} \neq y = h(y)$.

Case 6. $y = \pm 2^{-n}$ for some $n \in \mathbb{N}$ and $x \neq \pm 2^{-k}$ for any $k \in \mathbb{N}$. In this case we have $h(x) = x \neq \pm 2^{-(n-1)} = h(y)$.

Case 7. $x \neq \pm 2^{-n}$ and $y \neq \pm 2^{-n}$ for any $n \in \mathbb{N}$. In this case we have $h(x) = x \neq y = h(y)$.

In all cases we have $h(x) \neq h(y)$, so h is injective.

6.2 Finite Sets

Question. What does it mean to say that a set A is finite?

At first glance, this may seem like something you've known for a long time. Don't we just mean that we can count the elements of A ? If so, is the number of cells in your body finite? Can you count them? Maybe the answer to our question isn't quite so obvious after all.

Let's try to make our answer a bit more precise. First, for any natural number n we define the set $\mathbb{N}_n = \{k \in \mathbb{N} \mid k \leq n\}$. So $\mathbb{N}_4 = \{1, 2, 3, 4\}$, for example.

Next, we use the sets \mathbb{N}_n to formalize the idea of counting introduced in the introduction to this chapter. We say that a set A has n elements if $\mathbb{N}_n \equiv A$.

Now it seems that we can say the set A is finite if A has n elements for some $n \in \mathbb{N}$. Almost, but we're still forgetting something. Is the empty set finite? Clearly we would like to say that the empty set has zero elements, making it finite. This doesn't quite fit our scheme, so we must treat the empty set as a special case. Keeping our previous definitions, here is one way to do so.

We say that A has 0 elements if $A = \emptyset$. Now we say that a set is *finite* if it has n elements for some $n \in \mathbb{N} \cup \{0\}$.

Applying Theorem 4.1 and Exercise 6.1, we obtain the following:

Theorem 6.4. *If A is finite, then $A \preceq \mathbb{N}$.*

Now that we have a definition of finite, let's reconsider the set C of cells in your body. Is C finite? If so, for which n is $C \equiv \mathbb{N}_n$? We still can't really count them, so perhaps we've just obscured the question rather than answering it. Scientists estimate that there are about 10,000,000,000,000 cells in the average adult human body. That's not an actual count of the number of cells in any individual human body, though. These kinds of estimates are based on the sizes of various kinds of cells and the approximate proportion of each kind of cell in the body. In fact, we could determine the maximum number of cells that might be in a person's body by figuring out how many of the smallest kinds of cells would be required to build a body of a particular volume, or weight, etc. It seems reasonable to think that we could say a set was finite if we were sure it had at most n elements for some natural number n . That is the intent of the next result.

Theorem 6.5. *If $A \preceq \mathbb{N}_n$ for some natural number n , then A is finite.*

Before attempting to prove this result, let's make sure we understand what we are trying to prove. We are assuming that there is an injective function $f : A \rightarrow \mathbb{N}_n$. Unfortunately, our definition of finite requires us to produce a bijective function to some \mathbb{N}_k and the function f is probably not bijective. Since f is injective, the potential difficulty is that there are extra elements of \mathbb{N}_n (i.e. f is not surjective). This seems like something we should be able to overcome without much difficulty since a set with fewer elements than some finite set should certainly be finite. How do we construct the required bijection though? Let's first consider a simpler result which will prove useful.

Lemma. 6.2.1. *Let $f : A \rightarrow \mathbb{N}_n$ be an injective function that is not surjective. Then there is an injective function $g : A \rightarrow \mathbb{N}_{n-1}$.*

Proof of Lemma. Since $f : A \rightarrow \mathbb{N}_n$ is not surjective, the set $B = \mathbb{N}_n \setminus f(A)$ is nonempty. Choose $b \in B$. If $b = n$, then $f(a) \neq n$ for any element $a \in A$ and we may define $g : A \rightarrow \mathbb{N}_{n-1}$ by $g(a) = f(a)$ for each $a \in A$. If $b \neq n$, we define g by:

$$g(a) = \begin{cases} f(a) & \text{if } f(a) \neq n \\ b & \text{if } f(a) = n \end{cases}$$

In either case $g : A \rightarrow \mathbb{N}_{n-1}$ as desired since for all $a \in A$, $g(a) \neq n$.

It remains to be shown that g is injective. Suppose that $a_1, a_2 \in A$ and that $g(a_1) = g(a_2) = m$. If $m = b$, then by definition of g we have $f(a_1) = n = f(a_2)$. If $m \neq b$, then our definition of g implies that $f(a_1) = m = f(a_2)$. In either case we have $f(a_1) = f(a_2)$, so $a_1 = a_2$ because the function f is injective. Therefore g is injective as desired. \square

We will now prove the theorem.

Proof of Theorem 6.5. First note that if $A = \emptyset$, then A is finite by definition. We assume for the remainder of the proof that $A \neq \emptyset$. By hypothesis, there is an injective function $f_0 : A \rightarrow \mathbb{N}_n$ for some natural number n . If f_0 is also surjective, then we have the desired bijection. If f_0 is not surjective, then we may apply Lemma 6.2.1 to find an injection $f_1 : A \rightarrow \mathbb{N}_{n-1}$. We now consider the function f_1 .

If $f_1 : A \rightarrow \mathbb{N}_{n-1}$ is surjective, then f_1 is a bijection. If f_1 is not surjective, then we again apply Lemma 6.2.1 to find an injective function $f_2 : A \rightarrow \mathbb{N}_{n-2}$.

We continue this process recursively. If any of the injective functions $f_i : A \rightarrow \mathbb{N}_{n-i}$ are surjective, then we have the desired bijection and the

proof is complete. We claim that this must occur for some $0 \leq i \leq n - 1$. To see this, suppose that $f_{n-2} : A \rightarrow \mathbb{N}_2$ is not surjective. Applying Lemma 6.2.1 we find an injection $f_{n-1} : A \rightarrow \mathbb{N}_1$. Since $A \neq \emptyset$, we may choose $a \in A$. Now $f_{n-1}(a)$ must be an element of $\mathbb{N}_1 = \{1\}$, so $f_{n-1}(a) = 1$. It follows that f_{n-1} is surjective as desired.

We have shown that there is a bijective function $f_i : A \rightarrow \mathbb{N}_{n-i}$ for some $0 \leq i \leq n - 1$, so $A \equiv \mathbb{N}_{n-i}$ and A is finite. \square

We conclude this section with a question, the answer to which may seem obvious to you.

Question. If $m, n \in \mathbb{N}$ and $m \neq n$, can you prove that $\mathbb{N}_m \not\equiv \mathbb{N}_n$?

6.3 Denumerable Sets

We say that a set A is *infinite* if it is not finite. Georg Cantor was able to define a complete system of infinite numbers and of arithmetic on those numbers. We will not discuss his system here, but we will look at one particular kind of infinite number that is important in many areas of mathematics.

Theorem 6.6. *If A is an infinite set and B is a finite subset of A , then the set $A \setminus B$ is infinite.*

Proof. Suppose to the contrary that $A \setminus B$ is finite. It is easy to show that for any subset $B \subset A$, $A = B \cup (A \setminus B)$. Since both B and $A \setminus B$ are finite, it follows from Exercise 6.5 that A is finite. This contradicts our hypothesis that A is infinite, so it must be true that $A \setminus B$ is infinite. \square

For any set A , we say that A is *denumerable* if $A \equiv \mathbb{N}$. We say that A is *countable* if A is either finite or denumerable. The set A is said to be *uncountable* if it is infinite and not denumerable.

Example 6.2. *The set $A = \{2k \mid k \in \mathbb{N}\}$ of even natural numbers is denumerable. To see this, define the function $f : \mathbb{N} \rightarrow A$ by $f(n) = 2n$. It is routine to check that f is a bijection, so $\mathbb{N} \equiv A$ as desired.*

The preceding example points out a very important difference between finite and infinite sets. The set A is a proper subset of \mathbb{N} , but has the same cardinality as \mathbb{N} . Compare this to Exercise 6.6. All infinite sets have proper subsets of the same cardinality. In fact, this property is sometimes used to define what it means for a set to be infinite.

We would like to determine which of our results about finite sets are also true for denumerable sets. If A and B are denumerable, must $A \cup B$ also be denumerable? Are subsets of denumerable sets denumerable? Is there an analog of Theorem 6.5? Are there other ways to tell that a set is denumerable?

Theorem 6.7. *If A is a denumerable set and $B \equiv A$, then B is denumerable.*

Proof. By definition we have $A \equiv \mathbb{N}$. Since \equiv is an equivalence relation, it follows that $B \equiv \mathbb{N}$ as desired. \square

Theorem 6.8. *If $A \subset \mathbb{N}$, then A is countable.*

Proof. If A is finite, then A is countable by definition.

Assume that A is infinite. We define a function $f : \mathbb{N} \rightarrow A$ as follows. Recall that every nonempty subset of \mathbb{N} has a smallest element. Let $f(1)$ be the smallest element of $A_0 = A$. Since A is infinite the set $A_1 = A \setminus \{f(1)\}$ is nonempty, so we may define $f(2)$ to be the smallest element of A_1 . Continuing recursively, suppose that we have defined $f(1), \dots, f(n)$ for some $n \in \mathbb{N}$. Let $A_n = A \setminus \{f(1), \dots, f(n)\}$. Since $\{f(1), \dots, f(n)\}$ is finite, A_n is infinite by Theorem 6.6. In particular, A_n is nonempty and we may define $f(n+1)$ to be the smallest element of this set. Before showing that f is bijective we note the following facts that follow immediately from our construction:

- (i) For every natural number n , $A \setminus A_n = \{f(1), \dots, f(n)\}$.
- (ii) For every natural number n , $f(n) \geq n$.
- (iii) For all natural numbers m, n , if $f(n) \in A_m$ then $m < n$.
- (iv) For all natural numbers $m < n$, $f(n) \in A_m$.
- (v) For all natural numbers $m \leq n$, $f(m) \notin A_n$.

For any two natural numbers $m < n$ we shown that $f(n) \in A_m$ and $f(m) \notin A_m$, so $f(n) \neq f(m)$ and f is injective.

To see that f is surjective, let $k \in A$. We must show that $k = f(m)$ for some $m \in \mathbb{N}$. If $k = f(k)$, we are done. Otherwise we have $f(k) > k$. Now $f(k)$ is the smallest element of A_{k-1} and $k < f(k)$, so $k \notin A_{k-1}$. We also know that $k \in A$, so it follows that $k \in A \setminus A_{k-1} = \{f(1), \dots, f(k-1)\}$. Therefore $k = f(n)$ for some $1 \leq n < k$ and f is surjective as desired. \square

Applying Theorem 6.8, Theorem 4.1, and Corollary 4.1.1 we have:

Corollary 6.8.1. *A subset of a denumerable set is countable.*

Corollary 6.8.2. *A set A is countable if and only if $A \preceq \mathbb{N}$.*

Note that this Corollary allows us to say that a set A is countable if we can find an injection from A to \mathbb{N} . This is equivalent to finding a surjection from \mathbb{N} to A . (Can you prove this?) This allows us to conclude the following:

Corollary 6.8.3. *A set A is countable if either of the following is true:*

(i) *There is an injection $f : A \rightarrow \mathbb{N}$.*

(ii) *There is a surjection $f : \mathbb{N} \rightarrow A$.*

Theorem 6.9. *The union of two denumerable sets is denumerable.*

Proof. Suppose that we are given any two denumerable sets A and B . By definition there are bijective functions $f : A \rightarrow \mathbb{N}$ and $g : B \rightarrow \mathbb{N}$. We define a function $h : A \cup B \rightarrow \mathbb{N}$ by the following rule:

$$h(x) = \begin{cases} 2f(x) & \text{if } x \in A \\ 2g(x) + 1 & \text{if } x \notin A \end{cases}$$

We claim that h is an injection. To see this, let $x \neq y$ be two elements of $A \cup B$. If $x \in A$ and $y \notin A$, then $h(x) \neq h(y)$ since $h(x)$ is even and $h(y)$ is odd. Similarly if $x \notin A$ and $y \in A$, then $h(x) \neq h(y)$. If x and y are both in A , then $h(x) = 2f(x)$ and $h(y) = 2f(y)$, so $h(x) \neq h(y)$ because f is injective. Finally, if neither x nor y are in A , then $h(x) = 2g(x) + 1$ and $h(y) = 2g(y) + 1$, so $h(x) \neq h(y)$ because g is injective. In any case, we have shown that $h(x) \neq h(y)$ so $h : A \cup B \rightarrow \mathbb{N}$ is injective.

We have shown that $A \cup B \preceq \mathbb{N}$, so $A \cup B$ must be countable. Note however that the function h constructed above is not necessarily a bijection. (Do you see why?) To see that $A \cup B$ is actually denumerable, note that $A \subset A \cup B$. Applying Exercise 6.1 it follows that $A \preceq A \cup B$. Since $\mathbb{N} \preceq A$ we may apply Theorem 6.2 to obtain $\mathbb{N} \preceq A \cup B$. It now follows from the Cantor-Bernstein Theorem that $A \cup B \equiv \mathbb{N}$ as desired. \square

Using Mathematical Induction on the number of sets, we obtain:

Corollary 6.9.1. *The union of finitely many denumerable sets is denumerable.*

You have already shown that the set \mathbb{Z} is denumerable as part of Exercise 6.7. Are there any similarities between your solution to that exercise and the proof of Theorem 6.9?

Theorem 6.10. *The set $\mathbb{N} \times \mathbb{N}$ is denumerable.*

Proof. Since $\mathbb{N} \times \mathbb{N}$ is infinite, we need only show that it is countable. Define the function $f : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ by $f(a, b) = 2^{a-1}(2b - 1)$. We will show that f is injective, then apply Corollary 6.8.3 to obtain the desired result.

To see that f is injective, suppose that

$$2^{a-1}(2b - 1) = 2^{c-1}(2d - 1) \quad (6.1)$$

for $(a, b), (c, d) \in \mathbb{N} \times \mathbb{N}$. We show first that $a = c$. If not, then we may assume without loss of generality that $a < c$. Dividing both sides of equation 6.1 by 2^{a-1} yields $(2b - 1) = 2^{c-a}(2d - 1)$. But this is a contradiction since the quantity on the left side of the equation is odd and the quantity on the right is even. Since $a = c$, we may reduce equation 6.1 to $(2b - 1) = (2d - 1)$, from which it follows that $b = d$. Now $(a, b) = (c, d)$ and f is injective as desired.¹ \square

6.3.1 The set \mathbb{Q}

At first glance it seems that there are far more rational numbers than there are natural numbers. After all, there are infinitely many rational numbers between any two natural numbers. One of Cantor's accomplishments was to show that the set of rational numbers is actually denumerable. Our intuition developed from years of working with finite sets just doesn't serve us very well when working with infinite sets.

Lemma. 6.3.1. *The set \mathbb{Q}^+ of positive rational numbers is denumerable.*

Proof. Let $\mathbb{F} = \{\frac{a}{b} \mid a, b \in \mathbb{N}\}$ be the set of fractions whose numerators and denominators are natural numbers. We claim that $\mathbb{F} \equiv \mathbb{N} \times \mathbb{N}$. To see this, define $f : \mathbb{F} \rightarrow \mathbb{N} \times \mathbb{N}$ by $f(\frac{a}{b}) = (a, b)$. It is routine to show that f is a bijection, so $\mathbb{F} \equiv \mathbb{N} \times \mathbb{N}$. We may now apply Theorems 6.10 and 6.7 to see that \mathbb{F} is denumerable.

We now note that \mathbb{Q}^+ is the subset of \mathbb{F} consisting of all fractions $\frac{a}{b}$ for which a and b have no common divisors, so we may apply Corollary 6.8.1 to see that \mathbb{Q}^+ is countable. Since $\mathbb{N} \subset \mathbb{Q}^+$, \mathbb{Q}^+ is infinite by Exercise 6.4. It follows that \mathbb{Q}^+ is denumerable. \square

Since the function taking every positive rational to its additive inverse is bijective, the following Corollary follows immediately from our Lemma.

¹While it is not necessary for the purposes of this example, it is not too difficult to show that the function f defined here is actually bijective.

Corollary 6.10.1. *The set \mathbb{Q}^- of negative rational numbers is denumerable.*

We know that \mathbb{Q}^+ , \mathbb{Q}^- , and $\{0\}$ are all countable sets. Applying Exercise 6.8 we may conclude that:

Theorem 6.11. *The set of \mathbb{Q} of rational numbers is denumerable.*

6.3.2 The set \mathbb{R}

By this point it may seem possible that all sets are countable, making denumerability a useless distinction. We will show that this is not true by demonstrating that the set of real numbers is uncountable. First recall that every real number can be written as an infinite decimal. There is one danger in using such representations, it is possible to have different decimal representations that represent the same real number: e.g. $1.0\bar{0} = 0.9\bar{9}$. There is only one way that this can happen, though. A real number with a decimal expansion ending in an infinite string of 0's also has an expansion ending in an infinite string of 9's. If we disallow expansions ending in a string of 0's (the decimal expansions we normally think of as terminating), the infinite decimal representation of each real number is unique. This is important in the following proof since we will want to know that real numbers with different infinite decimal expansions are distinct.

Theorem 6.12. *The set \mathbb{R} of real numbers is uncountable.*

Proof. Suppose to the contrary that the set \mathbb{R} is countable, then there is a surjection $f : \mathbb{N} \rightarrow \mathbb{R}$. For convenience we use the notation x_n to denote the n th digit to the right of the decimal place in the unique infinite decimal expansion of x . In particular, the n th digit to the right of the decimal place in the expansion of $f(m)$ will be denoted $f(m)_n$. For example if $f(2) = \pi$, then $f(2)_4 = 5$. We will construct a real number y such that $y \neq f(n)$ for any $n \in \mathbb{N}$, which contradicts the fact that f is surjective.

For each $n \in \mathbb{N}$, define:

$$y_n = \begin{cases} 1 & \text{if } f(n)_n = 9 \\ 9 & \text{if } f(n)_n \neq 9 \end{cases}$$

Next we define

$$y = \sum_{i=1}^{\infty} \frac{y_i}{10^i}$$

so $y = 0.y_1y_2y_3\dots$. In other words, y is the unique real number in $(0, 1]$ such that the n th term to the right of the decimal place in the infinite decimal

expansion of y is y_n . By definition, $y_n \neq f(n)_n$ for every $n \in \mathbb{N}$, so $y \neq f(n)$ for any $n \in \mathbb{N}$ as desired. \square

Chapter 6 Exercises

- 6.1. Show that if $A \subset B$, then $A \preceq B$.
- 6.2. Prove that the relation \prec is transitive. Hint: there is a short proof that makes use of Theorem 6.2 and the Cantor-Bernstein Theorem.
- 6.3. It may seem easier to try and include the empty set in our original scheme simply by including 0 in the set of natural numbers, which in fact is sometimes done. This causes other difficulties, and doesn't really allow us to deal with the empty set as easily as it may seem. The following exercises investigate some of the difficulties. In the following, let $\mathbb{M} = \mathbb{N} \cup \{0\}$ and for each $n \in \mathbb{M}$ define $\mathbb{M}_n = \{k \in \mathbb{M} \mid k \leq n\}$.
- (i) For $n \geq 1$, explain why it is not a good idea to say that a set A has n elements if $A \equiv \mathbb{M}_n$. How would you define what it means for A to have n elements in terms of the sets \mathbb{M}_n ?
 - (ii) Do the sets \mathbb{M}_n make it easier to define what it means for a set to have 0 elements? Explain your answer.
- 6.4. Prove that a subset of a finite set is finite.
- 6.5. Let A and B be finite sets. Prove the following:
- (i) $A \cap B$ is finite.
 - (ii) $A \cup B$ is finite.
- 6.6. Let A be a finite set and let B be a proper subset of A .
- (i) Prove that $B \prec A$.
 - (ii) Prove that A and B are not equinumerous.
- 6.7. Prove that each of the following sets is denumerable.
- (i) The set of nonnegative integers $\mathbb{N} \cup \{0\}$.
 - (ii) The set of integers \mathbb{Z} .
- 6.8. Prove the following.
- (i) The union of two countable sets is countable.
 - (ii) The union of finitely many countable sets is countable.

6.9. Let A and B be denumerable sets. Prove that the set $A \times B$ is denumerable.

6.10. Let A , B , and C be sets. Define $A \times B \times C = \{(a, b, c) \mid a \in A \text{ and } b \in B \text{ and } c \in C\}$. Prove that $A \times B \times C \equiv (A \times B) \times C$.

6.11. For each natural number n , let \mathbb{N}^n denote the set of ordered n -tuples of natural numbers, so $\mathbb{N}^3 = \mathbb{N} \times \mathbb{N} \times \mathbb{N}$, etc. Prove that \mathbb{N}^n is countable.

6.12. Let S denote the set of sequences of 0's and 1's, so a typical element of S looks like (x_1, x_2, x_3, \dots) where each x_i is either 0 or 1. Prove that S is uncountable.

Appendix A

The Cantor-Bernstein Theorem

In this appendix we present a proof of the Cantor-Bernstein Theorem. The proof uses the same idea we used to construct the bijection between an open interval and a closed interval in Example 6.1. First we introduce some convenient notation.

Let $f : X \rightarrow Y$ be any function. For $A \subset X$, the *image* of A is the set $f(A) = \{f(x) \mid x \in A\}$. For $B \subset Y$ we use $f^{-1}(B)$ to denote the set $f^{-1}(B) = \{x \in X \mid f(x) \in B\}$. The set $f^{-1}(B)$ is called the *preimage* of B . We also use the notation $f^{-1}(x)$ to denote $f^{-1}(\{x\})$. You should not assume from our use of this notation that the function f is invertible! If $f : X \rightarrow X$ we use the notation f^2 to denote the function $f \circ f : X \rightarrow X$. Recursively, for a natural number $n \geq 2$, f^{n+1} is used to denote the function $f \circ f^n : X \rightarrow X$. Finally, we define f^0 to be the identity function on X .

Theorem (Cantor-Bernstein). *If $X \preceq Y$ and $Y \preceq X$, then $X \equiv Y$.*

Proof. By hypothesis there exist injections $f : X \rightarrow Y$ and $g : Y \rightarrow X$. We will find a bijective function $h : X \rightarrow Y$ with the property that $h(x) = f(x)$ for most $x \in X$ and $h(x) = g^{-1}(x)$ (since g is injective, $g^{-1}(x)$ is always a single point) for the remaining $x \in X$. As in Example 6.1, we use $g^{-1}(x)$ only as necessary to insure that the final function is bijective.

For every point $y \in Y \setminus f(X)$, we define a sequence of points in X as follows: $x_1 = g(y)$, $x_2 = g(f(x_1)) = g \circ f(g(y))$, $x_3 = g(f(x_2)) = (g \circ f)^2(g(y))$, and in general $x_{n+1} = g(f(x_n)) = (g \circ f)^{n-1}(g(y))$ for each $n \in \mathbb{N}$. Note that since g is injective $g^{-1}(x_1)$ contains only the point y and that for $k \geq 2$, $g^{-1}(x_k)$ contains only the point x_{k-1} . Define the set

$$S = \{x \mid x = (g \circ f)^n(g(y)) \text{ for some } y \in Y \setminus f(X) \text{ and for some } n \in \mathbb{N} \cup \{0\}\}.$$

In other words, the set S contains every point of each of the sequences we created above. This set will be the set of points on which h is not the same as f .

STEP 1: We prove the following facts about the set S .

- (i) If $y \in Y \setminus f(X)$, then $g(y) \in S$.
- (ii) If $x \in S$, then $g \circ f(x) \in S$.
- (iii) If $g(f(x)) \in S$, then $x \in S$.

For $y \in Y \setminus f(X)$ we have $g(y) = (g \circ f)^0(g(y))$, so (i) is true.

If $x \in S$, then $x = (g \circ f)^n(g(y))$ for some $n \in \mathbb{N} \cup \{0\}$ and some $y \in Y \setminus f(X)$. Now $g \circ f(x) = (g \circ f)^{n+1}(g(y))$, so $g \circ f(x) \in S$ and (ii) holds.

To see that (iii) is true, suppose that $g(f(x)) = (g \circ f)^n(g(y))$ for some $y \in Y \setminus f(X)$ and some $n \in \mathbb{N} \cup \{0\}$. If $n = 0$ then we have $g(f(x)) = g(y)$. Since g is injective this would imply $y = f(x)$, which contradicts our choice of y . Hence $n \geq 1$ and the point $w = (g \circ f)^{n-1}(g(y))$ is an element of S by definition. Now $g \circ f(w) = (g \circ f)^n(g(y)) = g \circ f(x)$. But $g \circ f$ is injective by Theorem 4.1, so we have $w = x$ and $x \in S$ as desired.

STEP 2: We now define the function $h : X \rightarrow Y$.

For $x \in X$ we define:

$$h(x) = \begin{cases} g^{-1}(x) & \text{if } x \in S \\ f(x) & \text{otherwise.} \end{cases}$$

Clearly $h(x)$ is defined for every $x \in X \setminus S$. If $x \in S$ then by definition $x = (g \circ f)^n(g(y))$ for some n and y , but this implies that $x \in g(Y)$ and $g^{-1}(x)$ is nonempty. We must also be sure that we have actually defined a function, i.e. that there is only one point in $g^{-1}(x)$ for each $x \in S$. To this end, suppose that y_1 and y_2 are each in $g^{-1}(x)$ for some $x \in S$. By definition we have $g(y_1) = g(y_2)$. Since g is injective this implies that $y_1 = y_2$ as desired. It remains to be shown that h is bijective.

STEP 3: We show that h is surjective.

Let $z \in Y$. Either $g(z) \in S$ or not.

CASE 1: If $g(z) \in S$, then $h(z) = g^{-1}(g(z)) = z$.

CASE 2: Suppose $g(z) \notin S$. By (i) $z \notin Y \setminus f(X)$, so there is an $x \in X$ such that $f(x) = z$. Since $g(f(x)) = g(z) \notin S$, $x \notin S$ by (ii). Therefore $h(x) = f(x) = z$.

In either case we have shown that $z \in h(X)$, so h is surjective as desired.

STEP 4: We show that h is injective, which will complete the proof of the theorem.

Suppose that $h(x_1) = h(x_2)$ for some $x_1, x_2 \in X$. We consider several cases.

CASE 1: Suppose that neither of x_1 or x_2 are in S . Then $f(x_1) = h(x_1) = h(x_2) = f(x_2)$ and $x_1 = x_2$ because f is injective.

CASE 2: Suppose both x_1 and x_2 are in S . Then $g^{-1}(x_1) = h(x_1) = h(x_2) = g^{-1}(x_2)$, so there is an element $y \in S$ so that $g(y) = x_1$ and $g(y) = x_2$. This implies that $x_1 = x_2$ because g is a function.

CASE 3: Finally, suppose that $x_1 \in S$ and $x_2 \notin S$. In this case we have $g^{-1}(x_1) = h(x_1) = h(x_2) = f(x_2)$, so $x_1 = g(g^{-1}(x_1) = g(f(x_2)))$ and $g(f(x_2)) \in S$. Applying (iii) it follows that $x_2 \in S$, which is a contradiction. Therefore this case is impossible. Clearly it is also impossible to have $x_1 \notin S$ and $x_2 \in S$. \square