Introduction to Set Theory

Prison Mathematics Project

Introduction

Hello and welcome to the module on Introduction to Set Theory! What follows is a module intended to support the reader in learning this fascinating topic. The Prison Mathematics Project (PMP) realises that you may be practising mathematics in an environment that is highly restrictive, so this text can both be used independently and does not require a calculator.

What is Set Theory?

We begin by introducing some of the most basic, yet important, notions in mathematics. Time and again, we shall need to refer to these concepts so it is vital that we see them early on and get used to working with them.

Learning in this Module

The best way to learn mathematics is to do mathematics. Indeed, education isn't something that happens more than it is something we should all participate in. You will find various exercise questions and worked examples in these notes so that you may try to solve problems and deepen your understanding of this topic. Although the aim is for everything to only require the content of this module, you are encouraged to use any other sources you have at your disposal.

Acknowledgements

These notes are based on lecture courses by K. Houston at the University of Leeds.

Contents

1 Preliminaries

2 Basic Set Theory

Set theory is primarily contained within the area of mathematical logic. At a basic level, it concerns the study of collections of objects (they can be numbers, solutions to equations, arrays, etc.). Modern set theory was pioneered by Dedekind and Cantor in the 1870s. After the discovery of a number of paradoxes, a number of axiomatic systems were proposed in which to study set theory. The most common and well-understood of these systems is Zermelo–Fraenkel set theory.

Definition 2.1 A set is a collection of objects (elements) whose order doesn't matter.

Example 2.2 The following are all examples of sets.

(i) $\{2, -5, 6\}$ is a set containing three elements.

(ii) $\{2, -5, \{6\}\}$ is also a set containing three elements (where the set $\{6\}$ is here an element).

(iii) $\{7, 7, 7\}$ is a set containing one element (we don't count repeats).

(iv) $\{\}$ is a set containing zero elements.

(v) $\{1, 2, 3, 4, ...\}$ is a set containing an infinite-number of element.

The set $\{\}$ is called the empty set, which we denote $\emptyset$. What Example 2.2 shows is that sets can really be any size we want. We now develop the notion of sizes of sets and, when we come to looking at functions, we can take this one step further.

Remark 2.3 It is common to call sets capital letters, such as X, and call the elements of the set lowercase letters, such as x. This isn't always followed but it is helpful to introduce the idea in this way and, the vast majority of the time, this is what occurs in the 'real' world.

Notation 2.4 Let X be a set. Then, we will henceforth use the following notation.

- We write $x \in X$ to mean that x **is** an element of X.
- We write $x \notin X$ to mean that x **isn't** an element of X.

Exercise 1 State either true or false to the following, giving a brief reason for each.

(i) $32 \in \{2, \{32\}, \{-5\}, -5\}$.

(ii) $64 \in \{$the set of even numbers$\}$.

(iii) $\{4\} \in \{100, \pi, \{4\}, 7\}$.

(iv) $\emptyset = \{$the set of prime numbers which are even$\}$.

Although we didn't properly introduce this idea before Exercise 1, the idea that two sets X and Y are equal just means that they contain the same elements, no more and no less. We will now discuss a more mathematical way to define what it means for two sets to be equal.

Definition 2.5 Let X be a set. Then, the set Y is a subset of X if every $y \in Y$ is also an element of X, that is $y \in X$. In this case, we write $Y \subseteq X$. We say that Y is a proper subset if $Y \neq X$, which we write $Y \subsetneq X$.

Notation 2.6 Some mathematicians like to use $\subset$ to mean proper subset; this is because the symbols $\subseteq$ and $\subset$ then act like the set-versions of $\leq$ and $<$. However, many mathematicians use $\subseteq$ and $\subset$ interchangeably; this is the convention that we take. If we need to emphasise that a subset is proper, we will forever use $\subsetneq$.

Example 2.7 The following are all examples of subsets.

(i) $\{2, 3\} \subseteq \{2, 3, 4, 97\}$.

(ii) $\{2, 3\} \subsetneq \{2, 3, 4, 97\}$.

(iii) $\emptyset \subseteq X$, where X is any set.

(iv) $Y \subseteq Y$, where Y is any set.

Exercise 2 List all subsets of the following and note how many subsets there are of each.
(i) $\emptyset$.
(ii) $\{1\}$.
(iii) $\{1, 2\}$.
(iv) $\{1, 2, 3\}$.
Do you notice a pattern? How many subsets are there of $\{1, 2, ..., n\}$?

We can now make what it means for two sets to be equal into a proper definition.

Definition 2.8 Let X and Y be sets. Then, they are equal if both $X \subseteq Y$ and $Y \subseteq X$.

Before advancing, we will write down a number of important sets, some of which we define later but nevertheless it is useful to have notation for them now:

$$\begin{aligned}
\mathbb{N} &= \text{set of natural numbers } \{0, 1, 2, 3, ...\}. \\
\mathbb{Z}^+ &= \text{set of positive integers } \{1, 2, 3, ...\}. \\
\mathbb{Z} &= \text{set of integers } \{..., -2, -1, 0, 1, 2, ...\}. \\
\mathbb{Q} &= \text{set of rational numbers.} \\
\mathbb{R} &= \text{set of real numbers.}
\end{aligned}$$

$$\mathbb{C} \quad = \text{set of complex numbers.}$$

Looking back at Exercise 1, we defined some sets using words. This is fine, but it isn't very mathematical. We will now look at more efficient ways to define sets.

Definition 2.9 A restriction on a set is a rule which must be satisfied by the elements of that set. We write the restrictions endowed on a set like this:

$$\{\text{[elements here]} : \text{[restrictions here]}\}.$$

Notation 2.10 Some people prefer to use a vertical bar $\{\text{[elements here]} \mid \text{[restrictions here]}\}$ instead of a colon. This is, once again, just a preference thing.

Example 2.11 We can write the set $\{1, 2, 3, 4\}$ in this restriction notation in a variety of ways:

$$\{n \in \mathbb{Z}^+ : n < 5\} \quad \text{or} \quad \{n \in \mathbb{Z}^+ : n \leq 4\} \quad \text{or} \quad \{n \in \mathbb{N} : n \leq 4\} \quad \text{and so forth.}$$

Exercise 3 Write each of the following sets using the restriction form discussed above.

(i) $\{3, 4, 5, 6, 7, 8, 9\}$.

(ii) $\{1, -8, 27, -64, 125, -216, 343, ...\}$.

(iii) $\{\text{Ford, Carter, Reagan, Bush, Clinton, Bush, Obama, Trump, Biden}\}$.

Given some sets, we can construct new sets out of them using a number of operations. Some of these may be familiar to you if you can recall what a Venn diagram of is.

Definition 2.12 Let X and Y be sets. Then, we can define the following operations.

(i) Their union is $X \cup Y = \{z : z \in X \text{ or } z \in Y\}$.

(ii) Their intersection is $X \cap Y = \{z : z \in X \text{ and } z \in Y\}$.

(iii) The complement of Y in X is $X \setminus Y = \{z : z \in X \text{ and } z \notin Y\}$.

Let's put into words what we have got in Definition 2.12. The union contains everything from either set; the intersection contains only the things that appear in both sets; the complement effectively removes anything in X that appears in Y.

Example 2.13 Let $X = \{1, 2, 4, 9\}$ and $Y = \{2, 3, 4\}$. We now perform the above operations:

- $X \cup Y = \{1, 2, 3, 4, 9\}$.
- $X \cap Y = \{2, 4\}$.
- $X \setminus Y = \{1, 9\}$.

- $Y \setminus X = \{3\}$.

Lemma 2.14 *The union of sets is commutative, that is* $X \cup Y = Y \cup X$.

Proof: From our intuitive understanding of Definition 2.12, this seems clear. However, we now need to *prove* it, that is give a rigorous argument using only what we know is true to derive a new true statement. Well, both $X \cup Y$ and $Y \cup X$ are sets, so Lemma 2.14 is really asking us to prove that two sets are equal, so we need to show that they are each contained in the other.

Let $x \in X \cup Y$. We need to show that $x \in Y \cup X$ and this will establish that $X \cup Y \subseteq Y \cup X$. By definition, $x \in X$ or $x \in Y$, which is the same as saying $x \in Y$ or $x \in X$, meaning $x \in Y \cup X$.

Let $x \in Y \cup X$. We need to show that $x \in X \cup Y$ and this will establish that $Y \cup X \subseteq X \cup Y$. By definition, $x \in Y$ or $x \in X$, which is the same as saying $x \in X$ or $x \in Y$, meaning $x \in X \cup Y$.

Consequently, since $X \cup Y \subseteq Y \cup X$ and $Y \cup X \subseteq X \cup Y$, we can conclude $X \cup Y = Y \cup X$. □

Exercise 4 Let $A, B \subseteq X$. Prove that $A \setminus B \subseteq A \cap (X \setminus B)$. Is this a proper subset? Is this an equality? Justify your answer either with a proof or with a counterexample.

Definition 2.15 Let X and Y be sets. Their Cartesian product is the set

$$X \times Y = \{(x, y) : x \in X \text{ and } y \in Y\}.$$

In words, Definition 2.15 says that the Cartesian product of two sets is a set of pairs, where the first entry comes from the first set and the second entry comes from the second set.

Remark 2.16 In general, the Cartesian product is non-commutative, i.e. $X \times Y \neq Y \times X$.

Example 2.17 Let $X = \{1, 2, 4, 9\}$ and $Y = \{2, 3, 4\}$. We now write the Cartesian product:

$$X \times Y = \{(1,2), (1,3), (1,4), (2,2), (2,3), (2,4), (4,2), (4,3), (4,4), (9,2), (9,3), (9,4)\}.$$

Exercise 5 Let $X = \{2, 1\}$ and $Y = \{3, 6\}$ and $Z = \{8, 7\}$. Write out $X \times Y \times Z$ in full.

Finally, we formally introduce the sizes of sets and use it to prove an elementary case of the so-called Inclusion-Exclusion Principle.

Definition 2.18 The cardinality of a set X is the number of elements, denoted $|X|$.

By convention, if X is not a finite set, we say that $|X| = \infty$. We shall see, when looking at functions and countability, this may be a tad misleading as there are 'different sizes of infinity'.

Lemma 2.19 *Let X and Y be sets. Then,*

$$|X \cup Y| \leq |X| + |Y|.$$

Proof: In the case that X and Y are disjoint, meaning there is no element which appears in both of them, it is clearly true. More than that, it is an equality.

In the case that X and Y are not disjoint, we look to the complement $Y \setminus X$. It is first clear that $\big|Y \setminus X\big| < |Y|$ since we have removed some elements. But now, it is clear that X and $Y \setminus X$ are disjoint (by definition, the second set contains nothing that is in X so there will be no element appearing in both of them), so the first part of the proof tells us that

$$\big|X \cup (Y \setminus X)\big| = |X| + \big|Y \setminus X\big| < |X| + |Y|.$$

Nearly there, we need only note that $X \cup (Y \setminus X) = X \cup Y$ (the left-hand-side we chuck away X and then add it back in, the right-hand-side we just add it). Hence, either $|X \cup Y| = |X| + |Y|$ or $|X \cup Y| < |X| + |Y|$, which combine to give us the result: $|X \cup Y| \leq |X| + |Y|$. □

Proposition 2.20 (Simple Inclusion-Exclusion Principle) *Let X and Y be sets. Then,*

$$|X \cup Y| = |X| + |Y| - |X \cap Y|.$$

Proof: Similar to the proof of Lemma 2.19; we prove a stronger result later. □

Exercise 6 The power set of a set X is the set of all subsets, denoted $\mathcal{P}(X)$. Write out the power set of $X = \{a, \{b, c\}\}$ and determine its cardinality $\big|\mathcal{P}(X)\big|$. Does this agree with your answer from Exercise 2?

Definition 2.21 An interval is a subset of $\mathbb{R}$ of one of the following forms, where $a, b \in \mathbb{R}$:

- $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$.
- $[a, b) = \{x \in \mathbb{R} : a \leq x < b\}$.
- $(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$.
- $(a, b) = \{x \in \mathbb{R} : a < x < b\}$.
- $[a, \infty) = \{x \in \mathbb{R} : x \geq a\}$.
- $(a, \infty) = \{x \in \mathbb{R} : x > a\}$.
- $(-\infty, b] = \{x \in \mathbb{R} : x \leq b\}$.
- $(-\infty, b) = \{x \in \mathbb{R} : x < b\}$.

Note: There is a degenerate situation to consider. If $a = b$ in Definition 2.21, we see that $[a, a] = \{a\}$ is a *singleton* set, i.e. contains one element, and the other intervals are empty.

3 Mathematical Statements

We have seen three examples of statements: Lemma 2.14, Lemma 2.19 and Proposition 2.20. However, they were rather standard with unambiguous language. Usually, mathematicians use specific vocabulary when writing a mathematical statement to make clear what is being assumed and what comes as a result of a (collection of) assumption(s). We really wish to understand different methods of proof, but this first requires us to know about mathematical statements.

Definition 3.1 An if-then statement is one of the form "if P, then Q".

We dissect Definition 3.1 further; in this definition, P and Q are symbols which represent assumptions and conclusions, respectively. Thus, if we assume P, then we get Q as a result.

Example 3.2 The following are examples of if-then statements.

(i) If $x = 2$, then $x + 1 = 3$.

(ii) If a person is the US president, then they are a US citizen.

(iii) If we cut a square into two equally-sized pieces, then we get two pentagons.

Remark 3.3 Looking at the if-then statements in Example 3.2, it is clear that (iii) is absolute nonsense. This is fine; yes it is false, but it definitely is still an if-then statement.

Conjecture (Riemann Hypothesis) *If s is a non-trivial zero of ζ, then it has real part $\frac{1}{2}$.*

This conjecture, meaning a result which is yet to be shown true or false, is a famous if-then statement (if you can solve it, then you will be awarded $\$1,000,000$ and will probably be regarded as one of, if not **the**, most famous mathematicians to ever live). At this level, we don't need to understand the Riemann Hypothesis; the point is that even famous problems with big prizes are if-then statements.

Exercise 7 Write two examples of if-then statements, with one true and the other false.

Definition 3.4 An if and only if statement is one of the form "P if and only if Q".

Again, we will dissect Definition 3.4; an if and only if statement is really two statements rolled into one: "if P, then Q" and "if Q, then P". This means that P and Q can be concluded from each other, so we call these statements equivalent, denoted $P \Leftrightarrow Q$. We further define these:

- The statement "if P, then Q" is called the only if direction, denoted $P \Rightarrow Q$.

- The statement "if Q, then P" is called the if direction, denoted $P \Leftarrow Q$.

Example 3.5 The following are examples of if and only if statements.

(i) $x = 7$ if and only if $x - 4 = 3$.

(ii) A person is the UK prime minister if and only if they are a UK citizen.

Remark 3.6 Looking at the if and only if statements in Example 3.5, it is clear that (ii) is false. Why? Even though the only if direction is true (if you are the prime minister, then you are a UK citizen), it is absolutely false that the if direction is true (if you are a UK citizen, then you are the prime minister). Here's a counterexample: me (I'm a UK citizen but I'm certainly not the prime minister!).

4 Mathematical Proof

We already have some experience with proofs. However, there are a number of different strategies that one can take to prove a statement. The success of the method may well depend on the type of result we are asked to prove. This subsection will introduce these proofs, along with a few notions from number theory both to get the ball rolling in that field but also to give us a score of examples we can attempt. Note that the proofs we have done before are direct proofs, meaning we employed no trickery; we took what the assumptions were given in the statements and used them to produce a proof.

Definition 4.1 A proof by contradiction is achieved by assuming the opposite of a result's outcome and working your way to a mathematical fallacy.

Example 4.2 Suppose we are asked to prove the following statement:

There is no largest even number.

Assume to the contrary that $N \in \mathbb{Z}$ is the largest even number. It being even means it is divisible by two, so $N = 2k$ for some $k \in \mathbb{Z}$. But now, it is clear that $N + 2 = 2k + 2 = 2(k + 1)$ is also divisible by two, i.e. it is even, but $N + 2 > N$, contradicting maximality. □

Exercise 8 Prove, using contradiction, that if n^2 is even, then n is even, for all $n \in \mathbb{Z}$.

Remark 4.3 One of the most famous proofs by contradiction is Euclid's proof that there are infinitely-many prime numbers. But what is a prime number? Have we got enough theory to look at this? I think we could just about discuss it but we will wait; it deserves it's moment in the limelight; see Corollary 5.10.

Definition 4.4 A proof by contrapositive of the statement "if A, then B" is achieved by proving the statement "if **not** B, then **not** A" directly.

Example 4.5 Suppose we are asked to prove the following statement:

If $n^2 - 6n + 5$ is even, then n is odd, where $n \in \mathbb{N}$.

Firstly, the contrapositive statement is the following:

If n is even, then $n^2 - 6n + 5$ is odd, where $n \in \mathbb{N}$.

Thus, we now begin a direct proof; let n be even, meaning it is of the form $n = 2k$ for some $k \in \mathbb{N}$. Therefore $n^2 - 6n + 5 = 4k^2 - 12k + 5 = 2(2k^2 - 12k + 2) + 1$, which is one above an even number, meaning it is odd. □

Exercise 9 Prove, by contrapositive, that if $7y + 9$ is even, then y is odd, where $y \in \mathbb{Z}$.

Remark 4.6 The difference between a proof by contradiction and a proof by contrapositive is lost on many a mathematician. Let's fix a statement to prove: "if P, then Q".

- A proof by contradiction is to show that "if P and **not** Q, then nonsense".
- A proof by contrapositive is to show that "if **not** Q, then **not** P".

When many people do a proof by contradiction, they may end up showing that "if P and **not** Q, then **not** P", which is perfectly fine for a contradiction because it means that both P and **not** P are true (which is clearly nonsense). In situations like this, they can actually be rephrased as proofs by contrapositive.

Definition 4.7 A proof by induction is applied when a statement is indexed by some $n \in \mathbb{Z}^+$ or an infinite subset thereof. If $A(n)$ is an infinite collection of statements, we can show they are all true (i.e. for all $n \in \mathbb{Z}^+$) as follows:

- Show that the base case $A(1)$ is true.
- Show that the inductive step $A(k)$ implies that $A(k+1)$ is true for some $k \in \mathbb{Z}^+$.

Example 4.8 Suppose we are asked to prove the following formula:

The sum of the squares of 1 *to* n *satisfies* $\sum_{t=1}^{n} t^2 = \frac{1}{6}n(n+1)(2n+1)$ *for all* $n \in \mathbb{Z}^+$.

We proceed by induction. First, one should check the base case, which is where $n \in \mathbb{Z}^+$ is smallest (i.e. $n = 1$). This is what the base case tells us:

$$\sum_{t=1}^{1} t^2 = \frac{1}{6}1(1+1)(2(1)+1).$$

Is this true? Well, the left-hand-side is $1^2 = 1$ and the right-hand-side is $6/6 = 1$, so the base case holds true. The inductive hypothesis is to assume that the following is true for some $k \in \mathbb{Z}^+$:

$$\sum_{t=1}^{k} t^2 = \frac{1}{6}k(k+1)(2k+1).$$

We will now use this to show that a similar expression for $n = k+1$ is true. Indeed,

$$\begin{aligned}\sum_{t=1}^{k+1} t^2 &= \sum_{t=1}^{k} t^2 + (k+1)^2 \\ &= \frac{1}{6}k(k+1)(2k+1) + (k+1)^2, \text{ by the inductive hypothesis,}\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{6}(2k^3 + 2k^2 + k) + (k^2 + 2k + 1) \\
&= \frac{1}{6}(2k^3 + 9k^2 + 13k + 6) \\
&= \frac{1}{6}(k+1)(k+2)(2(k+1)+1), \text{ by factorising,}
\end{aligned}
$$

which is precisely the formula we have to prove with $n = k+1$. By the principal of mathematical induction, the statement holds for all $n \in \mathbb{Z}^+$. □

Exercise 10 Prove, by induction, that $\sum_{t=1}^{n} t = \frac{1}{2}n(n+1)$ for all $n \in \mathbb{Z}^+$.

Sometimes, we need to stop and ask ourselves this question: why does this work? At a lower level, assuming proof by induction is taught, this is swept under the carpet (mathematics is almost like magic before one studies it at undergraduate-level, and it still is a bit magical beyond that). We can use proof by contradiction to show why proof by induction works!

Theorem 4.9 *Let $A(n)$ be an infinite collection of statements indexed by $n \in \mathbb{Z}^+$. If $A(1)$ is true, and $A(k)$ implies $A(k+1)$ for any $k \in \mathbb{Z}^+$, then all $A(n)$ are true.*

Proof: Suppose to the contrary at least one of the statements is false and let $j \in \mathbb{Z}^+$ be the smallest positive integer where $A(j)$ is false. Since $A(1)$ is assumed true, it must be that $j > 1$. Next, since j is the smallest index where a statement fails, it must be that $A(j-1)$ is true or else it contradicts minimality. But now, the second assertion tells us that $A(j-1)$ being true implies $A(j-1+1) = A(j)$ is true, a contradiction. □

5 Basic Number Theory

Number theory is an area of mathematics primarily concerned with looking at integers (and integer-valued functions, but we delay this until we have defined what a function is), amongst generalisations of these things. The oldest-known fragment of number theory dates back to 1800BC, where a clay tablet contains a list of Pythagorean triples (again, something we can study at length later). At a higher level, people can encode number-theoretic questions in other branches of mathematics (e.g. the famous Riemann hypothesis has many links to number theory), but we will first define the basics and provide Euclid's proof of infinite primes.

Definition 5.1 Let $a, b \in \mathbb{Z}$. It is said that a divides b if there exists $k \in \mathbb{Z}$ with $b = ka$. This is denoted by $a \mid b$, whereas $a \nmid b$ denotes that a does not divide b.

Example 5.2 The following are examples of divisors and non-divisors.

(i) $2 \mid 6$ because we can take $k = 3$ in Definition 5.1, that is $6 = 3(2)$.

(ii) $2 \nmid 5$ because there is no $k \in \mathbb{Z}$ where $5 = 2k$.

(iii) $a \mid a$ for every integer $a \in \mathbb{Z}$; simply take $k = 1$.

(iv) $a \mid 0$ for every integer $a \in \mathbb{Z}$; simply take $k = 0$.

Proposition 5.3 *Let $a, b, c \in \mathbb{Z}$ such that $a \mid b$ and $a \mid c$. Then, for all $m, n \in \mathbb{Z}$, we have*

$$a \mid (mb + nc).$$

Proof: By the definition of divides, there exists $k_1 \in \mathbb{Z}$ such that $b = k_1 a$ and there exists $k_2 \in \mathbb{Z}$ such that $c = k_2 a$. But now, we can quickly see

$$mb + nc = mk_1 a + nk_2 a = (mk_1 + nk_2)a,$$

which is, again by Definition 5.1, telling us that $a \mid (mb + nc)$. □

Corollary 5.4 *Let $a, b \in \mathbb{Z}$ such that $a \mid b$. Then, $a \mid b^2$.*

Proof: Simply take $m = b$ and $n = 0$ in Proposition 5.3. □

Exercise 11 The converse of "if P, then Q" is defined as "if Q, then P". Determine the converse of Corollary 5.4. Is it true? If so, prove it. If not, give a counterexample.

Definition 5.5 The absolute value $|a|$ of an integer $a \in \mathbb{Z}$ is defined as follows:

$$|a| = a \text{ if } a \geq 0 \qquad \text{and} \qquad |a| = -a \text{ if } -a < 0.$$

Example 5.6 The absolute value $|-4| = -(-4) = 4$ by Definition 5.5. We can easily convince ourselves that the absolute value changes any negative sign to a positive sign. Thus, $|a| \geq 0$ for all $a \in \mathbb{Z}$. In fact, this definition holds for any real number $a \in \mathbb{R}$; this is important later on.

Proposition 5.7 *Let $a, b, c \in \mathbb{Z}$. Then, the following hold true.*

(i) *If $a \mid b$ and $b \mid c$, then $a \mid c$.*

(ii) *If $a \mid b$ and $b \mid a$, then $a = |b|$ (meaning either $a = b$ or $a = -b$).*

Proof: (i) By the definition of divides, there exist $k_1, k_2 \in \mathbb{Z}$ such that $b = k_1 a$ and $c = k_2 b$. Combining these equations, we get $c = k_2 k_1 a$, which is precisely to say $a \mid c$, as required.

(ii) By the definition of divides, there exist $k_1, k_2 \in \mathbb{Z}$ such that $b = k_1 a$ and $a = k_2 b$. Combining these equations, we get $b = k_1 k_2 b$, which is only true when $k_1 k_2 = 1$. Because these are integers, the only possibilities are $k_1 = 1 = k_2$ and $k_1 = -1 = k_2$, so $a = b$ or $a = -b$, respectively. □

Definition 5.8 We say $p \in \mathbb{Z}^+$ is prime if its only distinct positive divisors are 1 and p.

Exercise 12 Carefully use Definition 5.8 to justify whether or not 1 is prime.

Theorem 5.9 (Fundamental Theorem of Arithmetic) *Let $n > 1$ be an integer. Then, it can be written as a unique (up to order) product of primes $n = p_1^{a_1} \cdots p_r^{a_r}$, where $p_1, ..., p_r$ are distinct primes and $a_1, ..., a_r \in \mathbb{Z}^+$.*

Proof: Let $A(n)$ be the statement "n is either a prime or a product of primes". We proceed by a special type of induction called strong induction. This is near-identical to what we have seen in Definition 4.7. Indeed, the base case here is $A(2)$, because $n > 1$, and 2 is clearly prime, so the base case is true. As for the inductive hypothesis, we assume that $A(2), A(3), ..., A(k)$ all hold true for some integer $k > 1$ and we use this to prove that $A(k+1)$ is true (it is clearly similar to usual induction, just that we ask that more things are true in this inductive step).

(i) If $k+1$ is prime, then $A(k+1)$ is true. By the principal of strong mathematical induction, the result follows.

(ii) If $k+1$ is not prime, then it can be written as a product $k+1 = xy$ of two integers $1 < x, y < k+1$. By the inductive hypothesis, we know that $A(x)$ and $A(y)$ are true, meaning that either they are prime or can be written as a product of primes. This means that they have the form

$$x = p_1^{a_1} \cdots p_r^{a_r} \quad \text{and} \quad y = q_1^{b_1} \cdots q_s^{b_s},$$

where the $p_1, ..., p_r, q_1, ..., q_s$ are primes and the powers are positive integers. Thus, $k+1$ is also a product of primes, where we still have distinctness of primes since if $p_i = q_i$, then the product will contain $p_i^{a_i+b_i}$. Hence, $A(k+1)$ is true. By the principal of strong mathematical induction, the result follows. □

Corollary 5.10 (Euclid's Theorem) *There are infinitely-many primes.*

Proof: Suppose to the contrary that there are finitely-many primes: $p_1, ..., p_k$. Then, we consider the number $N = p_1 \cdots p_k + 1$; we multiply together all of our finitely-many primes and add one.

(i) If N is prime, then it is a prime larger than all that came before, so we missed it out on our list of primes, contradicting the finiteness of primes we assumed.

(ii) If N is not prime, then the Fundamental Theorem of Arithmetic implies that it can be written as a product of primes, but we clearly see from Proposition 5.3 that $p_1 \not| N, ..., p_k \not| N$. Hence, the prime divisors of N are also missed out on our list of primes, again a contradiction. □

Definition 5.11 Let $a, b \in \mathbb{Z}$ be non-zero. Their greatest common divisor is the largest positive integer that divides both a and b, which we denote $\gcd(a, b)$.

Example 5.12 To compute $\gcd(16, 36)$, an acceptable strategy for small numbers is to write out all positive factors of each and look at the largest one appearing in both lists:

$$\{\text{+ve factors of } 16\} = \{1, 2, 4, 8, 16\},$$
$$\{\text{+ve factors of } 36\} = \{1, 2, 3, 4, 6, 9, 12, 18, 36\}.$$

Consequently, we can see that the greatest common divisor is $\gcd(16, 36) = 4$.

Exercise 13 Justify the following statements where $a, b \in \mathbb{Z}^+$.

(i) The greatest common divisor is symmetric, that is $\gcd(a, b) = \gcd(b, a)$.

(ii) The greatest common divisor is bounded below, namely $\gcd(a, b) \geq 1$.

Compute $\gcd(18, -57)$ and suggest how it compares to both $\gcd(-18, 57)$ and $\gcd(18, 57)$.

As mentioned in Example 5.12, computing the greatest common divisor is easy to do exhaustively when the numbers are small; exhaustion in mathematics basically means that we write down every possible case and just look for the answer. However, if we were asked to compute $\gcd(14\,441, 3563)$, it would be very tiresome writing down all factors. Hence, we will now develop some theory which will give us an algorithm to do this in a much more efficient way.

Lemma 5.13 (Division Lemma) *Let $a, b \in \mathbb{Z}^+$ be non-zero with $b > 0$. There exist unique $q, r \in \mathbb{Z}$ called the* quotient *and* remainder*, respectively, with $0 \leq r < b$ where $a = qb + r$.*

Proof: Consider the subset $S = \{a - sb : s \in \mathbb{Z} \text{ and } a - sb \geq 0\} \subseteq \mathbb{N}$. We now show that $S \neq \emptyset$. If $a < 0$, then $a - ab = a(1 - b) \geq 0$, since $1 - b \leq 0$ for all $b > 0$ Thus, $a(1 - b) \in S$. On the other hand, if $a > 0$, then $a - 0b = a \geq 0$. Thus, $a \in S$. Either way, S contains something. This means it has a smallest element, which we call r. We now define $q \in \mathbb{Z}$ to be the integer such that $r = a - qb$. As such, we at least know that we can find $q, r \in \mathbb{Z}$ such that $a = qb + r$.

However, there is a bit more to prove. Because $r \in S$, by definition, it must be non-negative. This establishes that $r \geq 0$. Furthermore, it is clear that

$$r - b = (a - qb) - b = a - (q + 1)b.$$

It now follows that $r - b \in S$ is equivalent to $a - (q + 1)b \geq 0$. But remember that r is the smallest element of S, so $r - b \notin S$ and so $a - (q + 1)b < 0$, which is the same as saying $r - b < 0$, that is $r < b$. The final thing to prove is uniqueness. Suppose that there exist two options for each of the quotient and remainder, meaning that

$$a = qb + r \qquad \text{and} \qquad a = q'b + r', \tag{5.1}$$

where $q, q', r, r' \in \mathbb{Z}$ with $0 \leq r < b$ and $0 \leq r' < b$. Assume to the contrary that $q \neq q'$ and $r \neq r'$. Without loss of generality, let $q > q'$ (we could equally argue that $q < q'$, it really doesn't matter; this is why we will not lose out on how general this argument is by choosing one of these options). From this inequality, it follows that $(q - q')b = r' - r$, by equating the expressions in equation (5.1) and rearranging. This means that $q - q' \geq 1$ and so $b \leq r' - r$, but this contradicts $r, r' < b$. Hence, $q = q'$ and $r = r'$, giving uniqueness and completing the proof. □

The next result gives us the algorithm that we want; it is really just repeated use of the Division Lemma we had above, so we don't even need to prove it!

Theorem 5.14 (Euclidean Algorithm) *Let $a, b \in \mathbb{Z}$ with $b \neq 0$. Then, repeated application of the Division Lemma gives the following:*

$$\begin{aligned} a &= q_1 b + r_1, && \text{for } 0 \leq r_1 < |b_1|, \\ b &= q_2 r_1 + r_2, && \text{for } 0 \leq r_2 < r_1, \\ r_1 &= q_3 r_2 + r_3, && \text{for } 0 \leq r_3 < r_2, \\ &\vdots \\ r_{k-2} &= q_k r_{k-1} + r_k, && \text{for } 0 \leq r_k < r_{k-1}. \end{aligned}$$

The algorithm terminates when $r_k = 0$ for some $k \in \mathbb{Z}^+$ and we get $\gcd(a, b) = r_{k-1}$.

Although this looks complicated, basically all we are doing is dividing so that the remainders get smaller and smaller (until we end up with remainder zero) and the greatest common divisor is then the smallest non-zero remainder. Let's give it a go.

Example 5.15 We will apply the Euclidean Algorithm to compute $\gcd(14\,441, 3563)$. Indeed, suppose that $a = 14\,441$ and $b = 3563$. Then,

$$\begin{aligned} 14\,441 &= 4(3563) + 189, \\ 3563 &= 18(189) + 161, \\ 189 &= 1(161) + 28, \\ 161 &= 5(28) + 21, \\ 28 &= 1(21) + 7, \\ 21 &= 3(7) + 0. \end{aligned}$$

Thus, by the Euclidean Algorithm, we see that $\gcd(14\,441, 3563) = 7$.

Corollary 5.16 (Bézout's Lemma) *Let $a, b \in \mathbb{Z}$ with $b \neq 0$. There exist $s, t \in \mathbb{Z}$ such that*

$$\gcd(a, b) = sa + tb.$$

Sketch of Proof: Rearrange the second-to-last equation in the Euclidean Algorithm for the greatest common divisor and substitute into it the other equations. □

Example 5.17 We wish to express $\gcd(14\,441, 3563)$ in the form $14\,441s + 3563t$ for some $s, t \in \mathbb{Z}$. We will use Example 5.15, rearranging the second-to-last equation for the greatest common divisor (which we computed was 7) and substituting the other equations into it:

$$\begin{aligned}
7 &= 28 - 1(21), && \text{by rearranging the fifth equation,}\\
&= 28 - 1(161 - 5(28)), && \text{by rearranging the fourth equation,}\\
&= 6(28) - 1(161) && \\
&= 6(189 - 1(161)) - 1(161), && \text{by rearranging the third equation,}\\
&= 6(189) - 7(161) && \\
&= 6(189) - 7(3563 - 18(189)), && \text{by rearranging the second equation,}\\
&= 132(189) - 7(3563) && \\
&= 132(14\,441 - 4(3563)) - 7(3563), && \text{by rearranging the first equation,}\\
&= 14\,441(132) + 3563(-535). &&
\end{aligned}$$

Consequently, we can see that $s = 132$ and $t = -535$ will give us the result we want.

Exercise 14 Use the Euclidean Algorithm to compute $\gcd(4635, 873)$. Consequently, apply Bézout's Lemma to express your answer in the form $4635s + 873t$, for some $s, t \in \mathbb{Z}$. Are these integers that we get in Bézout's Lemma unique? Briefly justify your answer.

Definition 5.18 Two integers $a, b \in \mathbb{Z}$ are coprime (or relatively prime) if $\gcd(a, b) = 1$.

Example 5.19 Clearly, $\gcd(n, 1) = 1$ for all $n \in \mathbb{Z}$ so 1 is coprime with everything. Similarly, -1 is coprime with everything. In fact, 1 and -1 are the only integers that are coprime with 0.

Lemma 5.20 (Euclid's Lemma) *Let $a, b, n \in \mathbb{Z}^+$. If $n \mid ab$ and $\gcd(a, n) = 1$, then $n \mid b$.*

Proof: By Bézout's Lemma, there exist $s, t \in \mathbb{Z}$ such that $sa + nt = 1$. Multiplying this equation by b gives us $sab + ntb = b$. Because $n \mid ab$, it clearly follows that $n \mid sab$ (and it is trivial that $n \mid ntb$). Thus, n divides the left-hand-side so it divides the right-hand-side, that is $n \mid b$. □

Corollary 5.21 *Let $a, b, p \in \mathbb{Z}^+$ with p prime. If $p \mid ab$, then $p \mid a$ or $p \mid b$.*

Exercise 15 Prove Corollary 5.21.

6 Equivalence Relations

We will now return to the more abstract setting of set theory, but we will soon be able to apply this directly to the number theory we have discussed. The discussion now amounts to seeing how elements of a set may be 'equivalent' to other elements but perhaps not to others. This will allow us to split our set into parts, where each part contains 'equivalent' elements.

> **Definition 6.1** Let X be a set. A relation on X is a subset $R \subseteq X \times X$ of the Cartesian product of X with itself. We say that x is related to y via R if the pair $(x, y) \in R$. We denote this by $x \sim y$, or by $x \sim_R y$ to really make explicit what the relation is.

Example 6.2 We can define the following relation on the set $X = \{1, 2, 3\}$:

$$R = \{(1,1), (1,2), (1,3), (2,2), (2,3), (3,3)\}.$$

This tells us that $1 \sim_R 1$ and $1 \sim_R 2$ and so forth. In fact, if we stare long enough at R, we can convince ourselves that $x \sim_R y$ precisely when $x \leq y$, so this a so-called ordering relation.

Remark 6.3 In practice, we call $\sim_R$ the relation and we don't write out the set R explicitly.

We are interested in so-called 'equivalence'; if we look back at Example 6.2, maybe we can convince ourselves that R cannot represent an 'equivalence' because, although $1 \sim_R 2$, we can see that $2 \not\sim_R 1$ (simply because $(2, 1) \notin R$). Surely for two things to be 'equivalent', we need this symmetry in our relation. What other rules must we impose?

> **Definition 6.4** A relation $\sim$ on a set X is an equivalence relation if the following hold.
>
> (i) $x \sim x$ for all $x \in X$. (Reflexivity)
>
> (ii) If $x \sim y$, then $y \sim x$ for all $x, y \in X$. (Symmetry)
>
> (iii) If $x \sim y$ and $y \sim z$, then $x \sim z$ for all $x, y, z \in X$. (Transitivity)

Example 6.5 We have already seen a number of relations on $\mathbb{Z}$; let's consider two of them.

(i) The relation $=$ is an equivalence relation on $\mathbb{Z}$. Indeed, it is clear that $x = x$ for every $x \in \mathbb{Z}$. Furthermore, $=$ is definitely symmetric and transitive.

(ii) The relation $<$ is **not** an equivalence relation on $\mathbb{Z}$. Indeed, $3 \in \mathbb{Z}$ but the following is false: $3 < 3$. It is not reflexive, so it is not an equivalence relation.

> **Exercise 16** Is $\leq$ an equivalence relation on $\mathbb{Z}$? Justify your answer.

We now have a notion of equivalence on the elements of a set. If two elements are related by an equivalence relation, it would be neat to collect them together into a ‘class’ where all elements in that class are related by that equivalence. This is precisely the next definition and it leads to the main result of this section.

Definition 6.6 Let $\sim$ be an equivalence relation on a set X. The equivalence class of an element $x \in X$ is the set $[x] = \{y \in X : y \sim x\}$, that is the set of all elements in X that are related to x by $\sim$.

Example 6.7 Consider the set $X = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$, that is X consists of pairs of integers (a, b) where the second integer $b \neq 0$. We can define an equivalence relation as follows:

$$(a, b) \sim (c, d) \quad \text{if and only if} \quad ad = bc.$$

This looks quite abstract, but this essentially tells us that two pairs are equivalent if the fractions $a/b = d/c$ are equal (simply rearrange $ad = bc$). For instance, we see that $(6, 2) \sim (12, 4)$, yes because $6 \times 4 = 2 \times 12$ but this is made clearer by the fact that $6/2 = 3 = 12/4$ are equal as fractions. The equivalence class $[(a, b)]$ under this relation consists of fractions which are the same as a/b. In fact, the set of all equivalence classes is called $\mathbb{Q}$; this is one way to define the rational numbers. A more traditional approach is given and used in Definition 9.1.

Exercise 17 Describe the equivalence classes in $\mathbb{N}$ under the relation $\sim$ defined as follows:

$$m \sim n \quad \text{if and only if} \quad 2 \mid (m - n).$$

You do not need to prove that $\sim$ is an equivalence relation.

We can now provide the set-up for, and the proving of, the main result of this section.

Definition 6.8 A partition of X is a collection of subsets $A_1, ..., A_k \subseteq X$ satisfying these:

(i) $A_i \neq \emptyset$ for all $i = 1, ..., n$.

(ii) $A_i \cap A_j = \emptyset$ for all $i \neq j$.

(iii) $X = A_1 \cup \cdots \cup A_k$.

Theorem 6.9 *For $\sim$ an equivalence relation on X, the equivalence classes partition X.*

Proof: It suffices to show each of the properties in Definition 6.8.

(i) Let $[x]$ be an equivalence class. Then, $x \in [x]$ by reflexivity (i.e. $x \sim x$), and so $[x] \neq \emptyset$.

(ii) Take two equivalence classes $[x]$ and $[y]$ such that $[x] \cap [y] \neq \emptyset$. Then, there exists an element $z \in [x] \cap [y]$, meaning that $z \sim x$ and $z \sim y$, by definition of equivalence class and of intersection. Using symmetry, we see that $x \sim z$ and, using transitivity, we get that $x \sim y$. Therefore, $[x] = [y]$ (note that this is the contrapositive of (ii) in Definition 6.8, so we have proven what we need).

(iii) To show that $X = \bigcup_{x \in X} [x]$, that is X is the same as the union of all the equivalence classes, we simply need to show that each of them is contained in the other. Firstly, let $x \in X$. Then, it is clear that $x \in [x]$ by reflexivity, meaning that $x \in \bigcup_{x \in X} [x]$. Thus, we have that $X \subseteq \bigcup_{x \in X} [x]$. Secondly, it is clear that $[x] \subseteq X$ for every $x \in X$, by Definition 6.6, so it follows trivially that $\bigcup_{x \in X} [x] \subseteq X$. Hence, both inclusions are shown, so we have equality. $\square$

7 Modular Arithmetic

We again switch back to the number theory setting, now weaponised with the knowledge of equivalence relations. We will define a relation which we can prove is an equivalence relation, and we can then determine its equivalence classes. This section then provides us with a neat real-world application of number theory.

Definition 7.1 Let $x, y, n \in \mathbb{Z}$ with $n > 1$. We say that x is congruent to y modulo n if $x = y + kn$ for some $k \in \mathbb{Z}$. We denote this by $x \equiv y \pmod{n}$.

Remark 7.2 Thus, $x \equiv y \pmod{n}$ if x and y have the same remainder on division by n.

Example 7.3 Here are some examples and non-examples of the modulo congruence.

(i) We have $15 \equiv 0 \pmod{3}$ because $15 = 0 + 5(3)$.

(ii) We have $16 \equiv 4 \pmod{12}$ because $16 = 4 + 1(12)$.

(iii) We have $3 \not\equiv 2 \pmod{5}$ because $3 = 2 + 5k$ has no solutions for $k \in \mathbb{Z}$

Theorem 7.4 (Arithmetic of Modulo Congruence) *Let $x, y, s, t, n \in \mathbb{Z}$ with $n > 1$ where $x \equiv s \pmod{n}$ and $y \equiv t \pmod{n}$. Then, the following properties hold true.*

(i) $x + y \equiv s + t \pmod{n}$.

(ii) $xy \equiv st \pmod{n}$.

Exercise 18 Prove Theorem 7.4 by appealing directly to Definition 7.1.

To get used to this modular arithmetic, we will prove the following number-theoretic result.

Proposition 7.5 *Let $n \in \mathbb{Z}$ be a square number. Then, $n \equiv 0 \pmod{4}$ or $n \equiv 1 \pmod{4}$.*

Proof: Let $n = k^2$ for some $k \in \mathbb{Z}$ (this is the definition of a square number). We need only proceed case-by-case on what the value of $k \pmod{4}$ is. Throughout, $m \in \mathbb{Z}$.

(i) If $k \equiv 0 \pmod{4}$, then $k = 4m$, so $n = (4m)^2 = 16m^2 \equiv 0 \pmod{4}$.

(ii) If $k \equiv 1 \pmod{4}$, then $k = 4m + 1$, so $n = (4m + 1)^2 = 16m^2 + 8m + 1 \equiv 1 \pmod{4}$.

(iii) If $k \equiv 2 \pmod{4}$, then $k = 4m + 2$, so $n = (4m + 2)^2 = 16m^2 + 16m + 4 \equiv 0 \pmod{4}$.

(iv) If $k \equiv 3 \pmod{4}$, then $k = 4m + 3$, so $n = (4m + 3)^2 = 16m^2 + 24m + 9 \equiv 1 \pmod{4}$. □

Proposition 7.6 *Let $x, y, n \in \mathbb{Z}$ with $n > 1$. Then, $x \equiv y \pmod{n}$ if and only if $n \mid x - y$.*

Proof: If $x \equiv y \pmod{n}$, then $x = y + kn$ for some $k \in \mathbb{Z}$, by definition. Hence, $x - y = kn$ is clearly divisible by n. Conversely, if $n \mid x - y$, then $x - y = kn$ for some $k \in \mathbb{Z}$, by definition. Hence, $x = y + kn$, which is precisely to say $x \equiv y \pmod{n}$. Thus, the two are equivalent. □

Remark 7.7 We can write the proof to Proposition 7.6 much quicker by combining the two arguments and using the symbol ⇔ (which we recall means 'equivalent to'). Indeed then,

$$x \equiv y \pmod{n} \quad \Leftrightarrow \quad x = y + kn \quad \Leftrightarrow \quad x - y = kn \quad \Leftrightarrow \quad n \mid x - y.$$

We now state Fermat's Little Theorem, from which we can get a better handle on a way to determine remainders upon division by primes.

Theorem 7.8 (Fermat's Little Theorem) *For all $a \in \mathbb{Z}$ and p prime, $a^p \equiv a \pmod{p}$.*

Proof: Omitted; we prove this result in Corollary **??**. □

Corollary 7.9 *Let p be prime.*

(i) *For all $a \in \mathbb{Z}$, we have $a^p - a \equiv 0 \pmod{p}$.*

(ii) *If $a \in \mathbb{Z}$ where $p \nmid a$, then $a^{p-1} \equiv 1 \pmod{p}$.*

Exercise 19 Prove Corollary 7.9 by using Fermat's Little Theorem.

Example 7.10 Suppose we wish to compute the remainder of 3^{293} upon division by 97; this can be done by exploiting Fermat's Little Theorem. The first task is to express the exponent in terms of the divisor:

$$293 = 3(97) + 2.$$

Thus, we know from the rules of indices that $3^{293} = 3^{3^{97}} \times 3^2 = 27^{97} \times 9$. We can use Fermat's Little Theorem with $p = 97$ and $a = 27$ to get a handle on the large power. Indeed, the theorem tell us that $27^{97} \equiv 27 \pmod{97}$. Hence, we can compute the remainder we are asked to find:

$$3^{293} = 27^{97} \times 9 \equiv 27 \times 9 = 243 \equiv 49 \pmod{97}.$$

Exercise 20 Use Fermat's Little Theorem to prove that 39 is **not** prime.

Coming from Section 6, we can now prove the following.

Proposition 7.11 *The modulo congruence is an equivalence relation.*

Proof: Throughout, suppose that $x, y, n \in \mathbb{Z}$ with $n > 1$. We clearly have that $x \equiv x \pmod{n}$ because $x = x + 0n$; this proves reflexivity. Next, assume that $x \equiv y \pmod{n}$, which means $x = y + kn$ for some $k \in \mathbb{Z}$. But this is equivalent to $y = x - kn$, but $-k \in \mathbb{Z}$ so this satisfies Definition 7.1, i.e. $y \equiv x \pmod{n}$; this proves symmetry. Finally, suppose $x \equiv y \pmod{n}$ and $y \equiv z \pmod{n}$. Then, there exist $k_1, k_2 \in \mathbb{Z}$ such that $x = y + k_1 n$ and $y = z + k_2 n$. If we substitute the first into the second, we get $x = z + (k_1 + k_2)n$ and $k_1 + k_2 \in \mathbb{Z}$ so this satisfies Definition 7.1, i.e. $x \equiv z \pmod{n}$; this proves transitivity. □

Remark 7.12 If we consider the set of equivalence classes $\mathbb{Z}/\equiv$ under the modulo congruence equivalence relation (modulo n), we can convince ourselves that it is precisely

$$\mathbb{Z}/\equiv = \{[0], [1], ..., [n-1]\}.$$

We would denote this by either $\mathbb{Z}_n$ or $\mathbb{Z}/n\mathbb{Z}$; we will see later that this notation is most often used to mean the set $\{0, 1, ..., n-1\}$ but we can see a similarity between this and $\mathbb{Z}/\equiv$ above.

Definition 7.13 An operation on equivalence classes is called well-defined if it doesn't depend on the representative taken from that class.

Example 7.14 Once again, consider the modulo congruence equivalence relation (modulo n) $\equiv$. We can define how to multiply two equivalence classes under this relation as follows:

$$[x] \cdot [y] := [xy].$$

This is well-defined. To prove this, we need to show that if we choose different representatives for each class $[x]$ and $[y]$, replacing each of x and y, then the multiplication will still work as defined above. Indeed, suppose that $[x] = [u]$ and $[y] = [v]$, that is our representatives are now u and v. By definition of equivalence classes, it means that $x \equiv u \pmod{n}$ and $y \equiv v \pmod{n}$. By definition, there exist $k_1, k_2 \in \mathbb{Z}$ such that $x = u + k_1 n$ and $y = v + k_2 n$. Therefore, we have

$$xy = (u + k_1 n)(v + k_2 n) = uv + (k_1 v + k_2 u + k_1 k_2 n)n \qquad \Rightarrow \qquad xy \equiv uv \pmod{n}.$$

Thus, we have that $[x] \cdot [y] = [xy] = [uv] = [u] \cdot [v]$, so the operation is well-defined.

Exercise 21 Show $[x] + [y] := [x + y]$ is well-defined under the equivalence relation $\equiv$.

Example 7.15 Consider the congruence modulo 3 equivalence relation and define this operation:

$$[x]^{[y]} := [x^y].$$

This is actually **not** well-defined. Indeed, notice that $[1] = [4]$ are the same class but

$$[2]^{[1]} = [2^1] = [2] \qquad \text{whereas} \qquad [2]^{[4]} = [2^4] = [16] = [1].$$

Since we get different classes depending on the representative of the class we have in the exponent, the operation is not well-defined (we say it is *ill-defined* or *ambiguous*).

Now that we are used to the idea of modular congruence and arithmetic, we can develop the theory further to dealing with solving so-called 'congruence equations'. We will first define what we mean by this, before applying it both mathematically and in a quasi-real-world context.

Definition 7.16 A congruence equation is an expression of the form $ax \equiv b \pmod{n}$, which is to be solved for $x \in \mathbb{Z}$ such that $0 < x < n$.

Theorem 7.17 *Let $a, b, n \in \mathbb{Z}$ with $n \geq 1$ and define $h := \gcd(a, n)$.*

(i) *If $h \nmid b$, then $ax \equiv b \pmod{n}$ has* **no** *solutions.*

(ii) *If $h \mid b$, then $ax \equiv b \pmod{n}$ has h solutions given by $x = (sb + kn)/h$, where $k \in \mathbb{Z}$ such that $0 \leq k < h$ and $s \in \mathbb{Z}$ is from Bézout's Lemma (i.e. whereby $h = sa + tn$).*

Proof: (i) Suppose that $ax \equiv b \pmod{n}$ has a solution, meaning that $n \mid (ax - b)$. By definition, we have $h \mid n$, from which it follows that $h \mid (ax - b)$. By definition, we also have $h \mid a$, so it is true also that $h \mid b$. This proves the contrapositive of, and therein, the statement.

(ii) Suppose that $h \mid b$, meaning $sb/h \in \mathbb{Z}$ for **any** $s \in \mathbb{Z}$. Multiplying by a gives the following:

$$a\frac{sb}{h} = sa\frac{b}{h} = (h - tn)\frac{b}{h} = b - \frac{tn}{h} \equiv b \pmod{n},$$

where Bézout's Lemma is used to establish $h = sa + tn$. It is therefore clear that $x = sb/h$ is a solution of $ax \equiv b \pmod{n}$. Now, let x be **any** solution to the congruence equation. We can subtract the above equation from the congruence equation, giving

$$ax - a\frac{sb}{h} = a\left(x - \frac{sb}{h}\right) \equiv 0 \pmod{n}.$$

Thus, $n \mid a(x - sb/h)$ which is to say that $(n/h) \mid (a/h)(x - sb/h)$. If we consider the prime factorisations of these integers, we must have that n/h and a/h have **no** common factors (because $h = \gcd(a, n)$ by definition). Therefore, for this division to make sense, it must be that $(n/h) \mid$

$(x - sb/h)$; this is equivalent to

$$x - \frac{sb}{h} = k\frac{n}{h}$$

for some $k \in \mathbb{Z}$. Rearranging this gives us $x = (sb+kn)/h$ as needed. Finally, as for the restriction on k, notice that if we have two congruent solutions $x_1 = (sb + k_1 n)/h$ and $x_2 = (sb + k_2 n)/h$, it follows that $h \mid (k_1 - k_2)$, so it suffices to restrict $0 \leq k < h$. □

Example 7.18 Consider the congruence equation $759x \equiv 115 \pmod{12\,167}$. We can apply the Euclidean Algorithm to see that $h = \gcd(759, 12\,167) = 23$ and that $23 \mid 115$. By Theorem 7.17, we know that there are exactly 23 solutions. One can apply Bézout's Lemma to see that $23 = 759(-16) + 12\,167(1)$, so we know now that $s = -16$. Explicitly, the solutions are

$$x = \frac{-1840 + 12\,167k}{23}, \qquad \text{where } 0 \leq k < 23.$$

Remark 7.19 We are able to shift the interval $0 \leq k < h$ to a different place, say $1 \leq k < h+1$. In fact, we can actually move it anywhere, just so long as the *width* of the interval is still h.

Exercise 22 Find all solutions to the congruence equation $759x \equiv 100 \pmod{12\,167}$.

Solving congruence equations is the backbone of RSA encryption, that and the fact it is 'difficult' to factorise large numbers into primes; this means computationally, so the numbers worked with in practice will be ludicrously large but we will demonstrate the idea behind this encryption for relatively small numbers.

Definition 7.20 Let p and q be prime numbers and consider their product $n = pq$.

(i) The private key is the number k such that $\gcd(k, (p-1)(q-1)) = 1$.

(ii) The public key is the pair (s, n) such that $sk + t(p-1)(q-1) = 1$.

Remark 7.21 The public key is obtained from the private key by applying Bézout's Lemma.

Example 7.22 Let $p = 53$ and $q = 61$, which are both prime. Give the private key $k = 1013$, we can compute the public key. First, it is easy to see that $n = pq = 3233$. Furthermore, $(p-1)(q-1) = 3120$. If we apply the Euclidean Algorithm and reverse the process, then we can see that $77(1013) + (-25)(3120) = 1$. Hence, the public key is $(77, 3233)$.

Definition 7.23 Let n, s, k be as in Definition 7.20.

(i) A message is $m \in \mathbb{Z}$ such that $0 \leq m < n$.

(ii) An encrypted message is $r \equiv m^s \pmod{n}$.

(iii) A decrypted message is $m \equiv r^k \pmod{n}$.

Example 7.24 Consider the same information from Example 7.22, that is the private key is $k = 1013$ and the public key is $(s, n) = (77, 3233)$. We will encrypt the message $m = 10$. By definition, the encrypted message is the number $r \equiv 10^{77} \pmod{3233}$ so we will build up powers of 10 as follows, where all congruences are modulo 3233:

$$\begin{aligned}
10^2 &= 100 \equiv 100,\\
10^4 &= 10\,000 \equiv 301,\\
10^8 &\equiv 301^2 = 90\,601 \equiv 77,\\
10^{16} &\equiv 77^2 = 5929 \equiv 2696,\\
10^{32} &\equiv 2696^2 = 7\,268\,416 \equiv 632,\\
10^{64} &\equiv 632^2 = 399\,424 \equiv 1765,\\
10^{72} &= 10^{64} \cdot 10^8 \equiv 1765 \cdot 77 = 135\,905 \equiv 119,\\
10^{77} &= 10^{72} \cdot 10^4 \cdot 10 \equiv 119 \cdot 301 \cdot 10 = 358\,190 \equiv 2560.
\end{aligned}$$

Therefore, the encrypted message is $r = 2560$.

Exercise 23 Decrypt $r = 8363$ using the encryption $k = 11\,787$ and $(s, n) = (3, 17\,947)$.

Proposition 7.25 *Let (s, n) be a public key, k be a private key and m a message. Then,*

$$m^{sk} \equiv m \pmod{n}.$$

Proof: By rearranging Bézout's Lemma, we get this from Definition 7.20: $sk = 1 - t(p-1)(q-1)$. Thus, we can work with the left-hand-side of the congruence in the statement of the result:

$$m^{sk} = m^{1-t(p-1)(q-1)} = m \cdot (m^{p-1})^{t(q-1)} \equiv m \pmod{p},$$

where we use Fermat's Little Theorem on the factor $(m^{p-1})^{t(q-1)}$. This looks promising, but we want to conclude it is true modulo n, not just modulo p as we have above. However, what we do know is that $p \mid (m^{sk} - m)$. Similarly, we could have used Fermat's Little Theorem as follows:

$$m^{sk} = m^{1-t(p-1)(q-1)} = m \cdot (m^{q-1})^{t(p-1)} \equiv m \pmod{q},$$

and so $q \mid (m^{sk} - m)$. Hence, $n \mid (m^{sk} - m)$, which is precisely to say $m^{sk} \equiv m \pmod{n}$. □

8 Functions

We are finally ready to do justice to the other half of this chapter's title. Informally, we may already know what a function 'looks like', that is it tells us an output from a given input. This is really the right definition, we just need to make it a bit more rigorous.

Definition 8.1 Let X and Y be sets. A function is a subset $f \subseteq X \times Y$ where, for every $x \in X$, there is a unique pair $(x, y) \in f$.

(i) The set X is called the domain of the function f.

(ii) The set Y is called the co-domain of the function f.

Remark 8.2 It is very rare that a mathematician would refer to a function directly in the sense of Definition 8.1. Instead, we think of a function f as a rule which assigns a unique $y \in Y$ to each input $x \in X$. In this way, we write $f : X \to Y$ to mean a function which assigns to elements of X and element of Y. We also write $f(x) \in Y$ to mean the element of Y that is assigned to x.

Example 8.3 Here are some examples of functions.

(i) The map $f : \mathbb{N} \to \mathbb{N}$ given by $f(n) = n^3$ is a function.

(ii) The map $g : \mathbb{R} \to \mathbb{R}$ given by $g(x) = x^3$ is also a function.

(iii) The map $h : \mathbb{R} \to \mathbb{R}$ given by $h(x) = x^2 + 5x + 6$ is a function.

There is a point to be made about the difference between (i) and (ii) in Example 8.3. Indeed, the formulae for the functions are identical (albeit with different letters used); the main differences are the domain and co-domain. We can see that there is **no** way to get the number 2 coming out of f, yet we **can** get 2 coming out of g simply by inputting $\sqrt[3]{2}$.

Definition 8.4 The image (or range) of a function $f : X \to Y$ is the subset $\mathrm{im}(f) \subseteq Y$ which contains precisely the elements of the co-domain that are obtained from X.

Example 8.5 In this language, $\mathrm{im}(f) \subsetneq \mathbb{N}$ whereas $\mathrm{im}(g) = \mathbb{R}$, in Example 8.3.

Exercise 24 Consider the function $p : \mathbb{N} \to \mathbb{N} \cup \{0\}$ where $p(n)$ is the number of distinct prime factors of n. Why do we need 0 in the co-domain?

Definition 8.6 Let $f : X \to Y$ be a function and consider subsets $U \subseteq X$ and $V \subseteq Y$.

(i) The image of U under f is the set $f(U) = \{f(u) : u \in U\}$.

(ii) The pre-image of V under f is the set $f^{-1}(V) = \{x \in X : f(x) \in V\}$.

Remark 8.7 In other words, the image of $U \subseteq X$ is found by applying f to everything inside U and the pre-image of $V \subseteq Y$ is the set of things in X that get sent into V by f.

Example 8.8 Let's consider again the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^3$, that is it cubes our input. Suppose we want to consider the subset $W = \{\ldots, -4, -2, 0, 2, 4, \ldots\} \subseteq \mathbb{R}$. Then, we see that $f^{-1}(W) = \{\ldots, -\sqrt[3]{4}, -\sqrt[3]{2}, 0, \sqrt[3]{2}, \sqrt[3]{4}, \ldots\}$ is the pre-image of W under f.

Note: Strictly speaking, we don't know that the cube root of a real number exists; we haven't yet even started to look at rational numbers, let alone real numbers. You will have to suspend disbelief for now, but Chapter **??** will fill in these gaps!

Definition 8.9 Let $f : X \to Y$ be a function.

(i) It is injective if $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$ for all $x_1, x_2 \in X$.

(ii) It is surjective if for every $y \in Y$, there exists $x \in X$ such that $f(x) = y$.

(iii) It is bijective if it is both injective and surjective.

Example 8.10 We consider a number of functions and determine if they are injective/surjective.

(i) Let the function $f : \mathbb{N} \to \mathbb{N}$ be given by $f(n) = n^3$. This is injective but **not** surjective. Indeed, it is injective because $f(n_1) = f(n_2)$ is equivalent to $n_1^3 = n_2^3$, which is true if and only if $n_1 = n_2$. It is not surjective because for $2 \in \mathbb{N}$, say, there is **no** $n \in \mathbb{N}$ with $f(n) = 2$.

(ii) Let the function $g : \mathbb{Z} \to \mathbb{Z}_7$ be given by $g(n) = n \pmod 7$. This is **not** injective but is surjective. Indeed, it is not injective because $0 \neq 7$ yet $f(0) = 0 = f(7)$ are equal. It is surjective because, for $m \in \mathbb{Z}_7$, we can see that $f(m + 7) = m$, where $m + 7 \in \mathbb{Z}$.

(iii) Let the function $h : \mathbb{R} \to \mathbb{R}$ be given by $h(x) = x + 1$. This is bijective. Indeed, it is injective because $h(x_1) = h(x_2)$ is equivalent to $x_1 + 1 = x_2 + 1$, which clearly implies that $x_1 = x_2$. It is surjective because any $y \in \mathbb{R}$ is obtained by $f(y - 1)$, where $y - 1 \in \mathbb{R}$.

Remark 8.11 There is something special to be said about a bijection. In a certain sense, giving a bijection between two sets means that they are the 'same'. In Example 8.10, the only bijection we saw was between $\mathbb{R}$ and itself (and, of course, every set is the same as itself!) but there are situations where we can construct bijections between different sets. As a quick example, the interval $[0, 1)$, meaning all numbers $x \in \mathbb{R}$ where $0 \leq x < 1$, and the unit circle S^1, meaning the pairs $(x, y) \in \mathbb{R}^2$ where $x^2 + y^2 = 1$, are in bijection. More will be said on this in Chapter **??**.

Exercise 25 Determine if $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ where $\varphi(x, y) = (x - y, xy)$ is injective/surjective.

Notation 8.12 Let $f : X \to Y$ be a function. We know now that $f(x)$ is one way to denote the image of $x \in X$ under the map f. However, there is another notation commonly used:

$$\begin{array}{rccc} f & : & X & \longrightarrow & Y \\ & & x & \longmapsto & f(x) \end{array}.$$

This arrow ↦ says that the object at the stick-end "maps to" to the object at the arrow end. For example, we have seen this function a number of times already:

$$\begin{array}{rccc} g & : & \mathbb{N} & \longrightarrow & \mathbb{N} \\ & & n & \longmapsto & n^3 \end{array}.$$

Exercise 26 Determine which of these functions is injective:

$$\begin{array}{rcccc} f & : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & & x & \longmapsto & x^2 \end{array}, \qquad \begin{array}{rcccc} g & : & \mathbb{Q} & \longrightarrow & \mathbb{Q} \\ & & x & \longmapsto & x^2 \end{array},$$

$$\begin{array}{rcccc} h & : & \mathbb{Z} & \longrightarrow & \mathbb{Z} \\ & & x & \longmapsto & x^2 \end{array}, \qquad \begin{array}{rcccc} k & : & \mathbb{N} & \longrightarrow & \mathbb{N} \\ & & x & \longmapsto & x^2 \end{array},$$

$$\begin{array}{rcccc} s & : & \mathbb{N} & \longrightarrow & \mathbb{Z} \\ & & x & \longmapsto & x^2 \end{array}, \qquad \begin{array}{rcccc} t & : & \mathbb{Z} & \longrightarrow & \mathbb{N} \\ & & x & \longmapsto & x^2 \end{array}.$$

[**Hint:** The injectivity of a function is determined by both its rule $x \mapsto \cdots$ and its domain.]

Definition 8.13 The composition of $f : X \to Y$ and $g : Y \to Z$ is $g \circ f : X \to Z$.

Remark 8.14 Informally, this means we hit our element $x \in X$ by the function f and then we hit the resulting element $f(x) \in Y$ by the function g. Formally, this means $(g \circ f)(x) = g(f(x))$. Thus, the notation $g \circ f$ means "first f, then g"; we read the order in which the functions are applied right-to-left.

Note: Even if the composition $g \circ f$ exists, there is **no** guarantee that $f \circ g$ will exist. Even if it does, there is **no** guarantee that $g \circ f$ is equal to $f \circ g$. Mathematicians would say that "function composition is non-commutative", meaning it matters which order we do it in (unlike multiplication of integers, which can be done in any order, for instance).

Lemma 8.15 *If $f : X \to Y$ and $g : Y \to Z$ are injective, $g \circ f : X \to Z$ is injective.*

Proof: We need to show that $(g \circ f)(x_1) = (g \circ f)(x_2) \Rightarrow x_1 = x_2$ for any $x_1, x_2 \in X$. Well, $(g \circ f)(x_1) = (g \circ f)(x_2)$ is precisely the same as saying $g(f(x_1)) = g(f(x_2))$ by Remark 8.14.

Because g is injective, we know that the things on the inside are equal, that is $f(x_1) = f(x_2)$, but because f is also injective, we know that the things on the inside of this are equal, that is $x_1 = x_2$, as required. □

Lemma 8.16 *If $f : X \to Y$ and $g : Y \to Z$ are surjective, $g \circ f : X \to Z$ is surjective.*

Exercise 27 Prove Lemma 8.16.

Corollary 8.17 *If $f : X \to Y$ and $g : Y \to Z$ are bijective, $g \circ f : X \to Z$ is bijective.*

Proof: This is a consequence of Lemmata (traditional plural of lemma) 8.15 and 8.16. □

The aim is to now discuss the concept of a bijection in a bit more detail in the context of so-called *inverse functions*. To do this, we must first show that, given a bijection from X to Y, we can 'flip the arrows' to get a bijection from Y to X.

Proposition 8.18 *Let $f : X \to Y$ be a bijection. Then, there is a bijection $f^{-1} : Y \to X$.*

Proof: In terms of Definition 8.1, which I said wasn't used much directly but we nevertheless use here, we can define $f^{-1} : Y \to X$ as follows:

$$f^{-1} = \{(y, x) \in Y \times X : (x, y) \in X \times Y\}.$$

Essentially, if $f : x \mapsto y$, then $f^{-1} : y \mapsto x$. There a number of things we need to show about this, namely that (i) it is a function, (ii) it is injective and (iii) it is surjective.

(i) To show that f^{-1} is a function, we must show that for every point in its domain $y \in Y$, there is one and only one point in its co-domain $x \in X$ such that $(y, x) \in f^{-1}$. Well, because $f : X \to Y$ is surjective, we know that for every $y \in Y$, there is at least one $x \in X$ such that $(x, y) \in f$, which means that $(y, x) \in f^{-1}$. As f is injective, there is precisely one such $x \in X$.

(ii) To show that f^{-1} is injective, we must show $(y_1, x), (y_2, x) \in f^{-1}$ implies that $y_1 = y_2$. Well, this is precisely to say that $(x, y_1), (x, y_2) \in f$ which implies that $y_1 = y_2$, since f is a function.

(iii) To show that f^{-1} is surjective, we must show for all $x \in X$, there exists at least one $y \in Y$ such that $(y, x) \in f^{-1}$. Because f is a function, for every $x \in X$, there is some $y \in Y$ with $(x, y) \in f$, which implies that $(y, x) \in f^{-1}$. □

Definition 8.19 The bijection $f^{-1} : Y \to X$ in Proposition 8.18 is the inverse of f.

Remark 8.20 More practically than Definition 8.19, we introduce the identity function, just assigns any element to itself. This is denoted $\mathrm{id}_X : X \to X$ where $\mathrm{id}_X(x) = x$. Then, we say that $f : X \to Y$ has an inverse $g : Y \to X$ if the following compositions hold:

$$g \circ f = \mathrm{id}_X \qquad \text{and} \qquad f \circ g = \mathrm{id}_Y \, .$$

We then re-label the inverse function $g = f^{-1}$ to make clear that it is indeed the inverse to f.

Lemma 8.21 *The inverse of a bijective map $f : X \to Y$ is unique.*

Proof: Assume $f : X \to Y$ has two inverse functions, $f_1^{-1}, f_2^{-1} : Y \to X$. By Remark 8.20, we know that $f_1^{-1} \circ f = \mathrm{id}_X = f_2^{-1} \circ f$. But this equality implies that

$$f_1^{-1} \circ (f \circ f_1^{-1}) = f_2^{-1} \circ (f \circ f_1^{-1}) \qquad \Rightarrow \qquad f_1^{-1} = f_2^{-1},$$

by cancelling the first two terms in each of the compositions. □

Note: Suppose we wish to solve $\sqrt{x^2 + x + 1} = x$ for $x \in \mathbb{R}$. It is tempting to think that by squaring both sides and rearranging, we can get an answer. If we do this, we would get $x^2 + x + 1 = x^2$ which means that $x + 1 = 0$ and so $x = -1$. However, if we substitute this into the original equation, we would get $\sqrt{1} = -1$, which is utterly incorrect. The problem is with this chain of implications:

$$\sqrt{x^2 + x + 1} = x \quad \Rightarrow \quad x^2 + x + 1 = x^2 \quad \Leftrightarrow \quad x + 1 = 0 \quad \Leftrightarrow \quad x = -1.$$

The first implication is **not** two-sided; this boils down to the fact that $f(x) = x^2$ is not injective. Hence, we can only conclude that $x = -1$ is a **candidate** for a solution, not that it is a solution.

9 Countability

We can now 'properly' introduce the rational numbers $\mathbb{Q}$ and use what we know of functions from Section 8 to develop some theory on rational numbers.

Definition 9.1 A number $r \in \mathbb{R}$ is rational if $r = m/n$ for some $m, n \in \mathbb{Z}$ with $n \neq 0$.

Remark 9.2 There is a slight flaw in Definition 9.1, as helpful as it is to visualise what a rational is: we don't really know what $\mathbb{R}$ is. We will forgo this issue for now but note that there is a nice way to construct the rational numbers from the integers by defining $\sim$ to be the following equivalence relation on $\mathbb{Z} \times \mathbb{Z} \setminus \{0\}$:

$$(a, b) \sim (c, d) \qquad \Leftrightarrow \qquad ad = bc.$$

Under this relation, the equivalence class $[(a, b)]$ is then identified with a/b. Why? Because the defining relation $ad = bc$ is equivalent to $a/b = c/d$, so the equivalence classes really consist of pairs which are integer multiples of each other, i.e. fractions which are equivalent to one another.

Exercise 28 Consider a rational $r = m/n \in \mathbb{Q}$. How many representations are there of r as a ratio of two integers, that is how many elements are there in the equivalence class $[(m, n)]$ under the relation $\sim$ defined in Remark 9.2?

Definition 9.3 A set is called closed under an operation if applying said operation to elements of that set result in another element of that set.

Example 9.4 The following are examples and non-examples of closure under operations.

(i) $\mathbb{Z}$ is closed under the operation $+$.

(ii) $\mathbb{N}$ is **not** closed under the operation $-$ because $2 - 5 \notin \mathbb{N}$.

(iii) $\mathbb{R}$ is **not** closed under the operation $\boldsymbol{\nabla}\cdot$ because $3\boldsymbol{\nabla} \cdot 0 \notin \mathbb{R}$ (it isn't even defined).

(iv) $\mathbb{R} \setminus \{0\}$ is closed under the operation $\boldsymbol{\nabla}\cdot$.

Proposition 9.5 *The rationals $\mathbb{Q}$ are closed under addition, subtraction, multiplication and division by non-zeros.*

Proof: Throughout, suppose that $a = m/n$ and $b = x/y$ where $m, n, x, y \in \mathbb{Z}$ with $n, y \neq 0$.

(i) We have $a + b = (my + nx)/ny$, with $my + nx \in \mathbb{Z}$ and $ny \in \mathbb{Z} \setminus \{0\}$ by closure, so $a + b \in \mathbb{Q}$.

(ii) We have $a - b = (my - nx)/ny$, with $my - nx \in \mathbb{Z}$ and $ny \in \mathbb{Z} \setminus \{0\}$ by closure, so $a - b \in \mathbb{Q}$.

(iii) We have $ab = mx/ny$, with $mx \in \mathbb{Z}$ and $ny \in \mathbb{Z} \setminus \{0\}$ by closure, so $ab \in \mathbb{Q}$.

(iv) We have $a/b = my/nx$, with $my \in \mathbb{Z}$ and $nx \in \mathbb{Z} \setminus \{0\}$ if $b \neq 0$ by closure, so $a/b \in \mathbb{Q}$. □

Definition 9.6 A number $r \in \mathbb{R}$ is irrational if $r \neq m/n$ for any $m, n \in \mathbb{Z}$ with $n \neq 0$.

Example 9.7 These are examples of irrational numbers (it is non-trivial to prove this).

(i) The k^{th} root of any non-k^{th} power is irrational, e.g. $\sqrt[3]{2}$ and $\sqrt[5]{67}$ and $\sqrt[654]{9}$.

(ii) The number $\pi = 3.1415926\ldots$ is irrational.

(iii) The number $e = 2.7182818\ldots$ is irrational.

Proposition 9.8 *The number $\sqrt{2}$ is irrational.*

Proof (w/o Fundamental Theorem of Arithmetic): Assume to the contrary $\sqrt{2} \in \mathbb{Q}$, so it can be expressed as $\sqrt{2} = m/n$ for some $m, n \in \mathbb{Z}$ with $n \neq 0$. In particular, take these integers so that they are in lowest terms, meaning $\gcd(m, n) = 1$ and so the fraction cannot be cancelled down further. Then, one can square the expression and multiply by n^2 to get

$$2n^2 = m^2.$$

Since the left-hand-side is divisible by two, it follows that so is the right-hand-side, i.e. $2 \mid m^2$ which implies that $2 \mid m$. Thus, there exists $k \in \mathbb{Z}$ such that $m = 2k$. Substituting this into the expression gives

$$2n^2 = (2k)^2 = 4k^2 \qquad \Rightarrow \qquad n^2 = 2k^2.$$

By a similar argument, $2 \mid n^2$ which implies that $2 \mid n$, which contradicts $\gcd(m, n) = 1$. □

Proof (w/ Fundamental Theorem of Arithmetic): Assume to the contrary $\sqrt{2} \in \mathbb{Q}$, so it can be expressed as $\sqrt{2} = m/n$ for some $m, n \in \mathbb{Z}$ with $n \neq 0$. By the Fundamental Theorem of Arithmetic, we can find prime factorisations of the form $m = p_1 \cdots p_s$ and $n = q_1 \cdots q_t$. One can square the expression for $\sqrt{2}$, multiply by n^2 and substitute in the prime factorisations to get

$$2q_1^2 \cdots q_t^2 = p_1^2 \cdots p_s^2.$$

The left-hand-side contains an odd number of factors of two whereas the right-hand-side contains an even number of factors of two, a contradiction to the Fundamental Theorem of Arithmetic. □

Note: This is the general method for proving irrationality of a square root.

(i) Assume to the contrary that the root is rational.

(ii) (a) If one **doesn't** use the Fundamental Theorem of Arithmetic, contradict that the rational is in lowest terms by finding a common divisor.

(b) If one **does** use the Fundamental Theorem of Arithmetic, contradict by getting a different number of prime factors on each side of an equality.

Remark 9.9 The above method works for **all** roots, not just square roots, as now shown.

Example 9.10 We will prove that $\sqrt[7]{25}$ is irrational **without** the Fundamental Theorem of Arithmetic. Assume to the contrary $\sqrt[7]{25} \in \mathbb{Q}$, so it can be expressed as $\sqrt[7]{25} = m/n$ for some $m, n \in \mathbb{Z}$ in lowest terms, that is $\gcd(m, n) = 1$, with $n \neq 0$. Then, we can see that $25n^7 = m^7$. Hence, $5 \mid m^7$ which means that $5 \mid m$, so there exists $k \in \mathbb{Z}$ such that $m = 5k$. Substituting this into our expression gives $25n^7 = (5k)^7 = 5^7k^7 \Rightarrow n^7 = 5^5k^7$. Similarly, $5 \mid n^7$ which means that $5 \mid n$, contradicting $\gcd(m, n) = 1$. □

Exercise 29 Prove that $\sqrt[7]{25}$ is irrational **with** the Fundamental Theorem of Arithmetic.

Conjecture *The numbers* $\pi + e$, $\pi - e$, πe, π/e, π^e, π^π, e^e *are irrational.*

This conjecture always baffles me; the amount of mathematics that is known to us grows by the day, yet a problem as 'simple' as showing that $\pi + e \neq m/n$ for any $m, n \in \mathbb{Z}$ with $n \neq 0$ still escapes us. It seems so likely, as π and e are 'very' (this is explained in Chapter **??**) irrational.

The next step is to study the way in which we usually denote real numbers, that is the so-called *decimal number system*. In this way, we will develop theory to see the type of decimal expression which is assigned to a rational number. This will lead to the important property known as *density* and then, we can extend the discussion by introducing countability.

Definition 9.11 A number has a terminating decimal expression if it has a finite number of non-zero decimal places. Otherwise, it is a non-terminating decimal expression.

Example 9.12 Here are some examples of terminating and non-terminating decimals.

(i) The number 0.2857 is a terminating decimal expression.

(ii) The number 3/8 has a terminating decimal expression, namely 0.375.

(iii) The number 1/3 does **not** have a terminating decimal expression, as it is 0.3333....

Lemma 9.13 *Every real number with a terminating decimal expression is rational.*

Proof: Let $x \in \mathbb{R}$ have a terminating decimal expression, so it has the form $x = n.a_1 a_2 \dots a_k$, where $n \in \mathbb{Z}$ and the $a_i \in \{0, 1, 2, \dots, 9\}$ are the digits after the decimal point. Thus, it is clear that multiplying this number by a large enough power of ten gives an integer. In particular, $10^k x \in \mathbb{Z}$. Therefore, $x = 10^k x/10^k$ is a ratio of integers, which means that $x \in \mathbb{Q}$. □

Definition 9.14 A number $x \in \mathbb{R}$ has a recurring decimal expression if it can be written in the form $x = n.a_1 \dots a_k b_1 \dots b_p b_1 \dots b_p \dots$, where $n \in \mathbb{Z}$ and $a_i, b_j \in \{0, 1, 2, \dots, 9\}$.

Example 9.15 Here are some examples of recurring decimals.

(i) The number 1/3 has a recurring decimal expression, namely 0.3333....

(ii) The number 2/7 has a recurring decimal expression, namely 0.285714285714....

(iii) The number 3/8 **has** a recurring decimal expression, namely 0.374999....

Proposition 9.16 *Let $x \in \mathbb{R}$. Then, it is true that*

$$1 + x + x^2 + x^3 + \cdots + x^n = \sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1 - x}.$$

Sketch of Proof: Simply proceed inductively on the power n, with trivial base case $n = 0$. □

Corollary 9.17 (Geometric Series) *Let $x \in \mathbb{R}$ such that $|x| < 1$. Then, it is true that*

$$1 + x + x^2 + x^3 + \cdots = \sum_{k=0}^{\infty} x^k = \frac{1}{1 - x}.$$

Proof: Take the limit as $n \to \infty$ in Proposition 9.16. However, we haven't yet defined the notion of a limit so all further details are deferred; see Proposition **??**. □

The aim is to show that any number with a recurring decimal expression can be written as a ratio of integers, i.e. it is rational. We first, by way of an example, demonstrate how one can convert a recurring decimal into a fraction by using the Geometric Series.

Example 9.18 Suppose we wish to convert 1.53555555... to a fraction; we proceed as follows:

$$1.53555555... = \frac{1}{100} 153.55555...$$

$$
\begin{aligned}
&= \frac{1}{100}\left(153 + \frac{5}{10}\left(1 + \frac{1}{10} + \frac{1}{100} + \cdots\right)\right)\\
&= \frac{1}{100}\left(153 + \frac{1}{2}\sum_{k=0}^{\infty}\frac{1}{10^k}\right)\\
&= \frac{1}{100}\left(153 + \frac{1}{2}\frac{1}{1-1/10}\right)\\
&= \frac{691}{450}.
\end{aligned}
$$

Exercise 30 Prove that $0.dddddd... = d/9$ for any digit $d \in \{0, 1, 2, \ldots, 9\}$.

Exercise 31 Prove that $0.dedede... = (10d + e)/99$ for any digits $d, e \in \{0, 1, 2, \ldots, 9\}$.

Corollary 9.19 *The recurring decimal* $0.99999... = 1$.

Proof: Simply take $d = 9$ in Exercise 30. □

This may look unsettling but, mathematically, there is nothing wrong with Corollary 9.19. This suggests that if we have a 'tail' of repeated nines, we may get some strange behaviour. In fact, the next result guarantees that this is the only strange behaviour that occurs in our decimals.

Theorem 9.20 *If two different decimal expressions represent the same number, one expression terminates and the other expression ends in recurring nines.*

Proof: Let $x \in \mathbb{R}$ have two different decimal expressions. Then, we will be able to move the decimal place far enough to the left to ensure the following is true, that is there exists $k \in \mathbb{N}$ such that multiplying each expression by 10^{-k} gives us these expressions:

$$0.a_1a_2a_3... \qquad \text{and} \qquad 0.b_1b_2b_3...,$$

again where the $a_i, b_j \in \{0, 1, 2, \ldots, 9\}$. It suffices to work only with these expressions to prove the result. Let n be the first position where the two decimals differ, that is they have the form

$$0.a_1 \ldots a_{n-1}a_na_{n+1}... \qquad \text{and} \qquad 0.a_1 \ldots a_{n-1}b_nb_{n+1}....$$

Without loss of generality, assume that $a_n < b_n$, which means that $a_n + 1 \leq b_n$ because the digits are non-negative integers. By definition, both expressions are precisely equal to x, which means

$$0.a_1 \dots a_{n-1} b_n 000... \leq x \leq 0.a_1 \dots a_{n-1}(a_n + 1)000...,$$

which implies that $b_n \leq a_n + 1$. Since both directions of the inequality hold true, it must be that it is a straight-up equality, that is $a_n + 1 = b_n$. It follows from this that $a_k = 9$ and $b_k = 0$ for all $k \geq n+1$. Indeed, if this was **not** the case, then we would achieve this contradiction:

$$x - x = 0.a_1 \dots a_{n-1}(a_n + 1)b_{n+1}... - 0.a_1 \dots a_{n-1} a_n b_{n+1}... > 0. \qquad \square$$

Proposition 9.21 *The recurring decimal $x = 0.a_1 \dots a_k a_1 \dots a_k...$ can be expressed as*

$$\frac{10^{k-1}a_1 + 10^{k-2}a_2 + \cdots + a_k}{10^k - 1}.$$

Sketch of Proof: Use the Geometric Series on $x = (10^{k-1}a_1 + 10^{k-2}a_2 + \cdots + a_k) + (10^k - 1)x$. □

Theorem 9.22 *A real number is rational if and only if it has a recurring or terminating decimal expression.*

Proof: ($\Rightarrow$) Suppose that $x = m/n \in \mathbb{R}$ is rational. Then, one can find a decimal expression by carrying out long division, dividing n into m, noting that the remainder r at each stage is an integer satisfying $0 \leq r < n$. After exhausting all digits of m and finitely-many steps after that, either the remainder will be zero (terminating expression) or it will be a remainder we had before (recurring expression).

($\Leftarrow$) Suppose that $x \in \mathbb{R}$ has a decimal expression of the form $x = k.a_1 \dots a_n b_1 \dots b_m b_1 \dots b_m....$ In the case it is terminating, $m = 1$ and $b_1 = 0$, which is clearly rational. If this is not the case, and the decimal recurs, then Proposition 9.21 implies that

$$x = k + \frac{10^{n-1}a_1 + 10^{n-2}a_2 + \cdots + a_n}{10^n - 1} + \frac{1}{10^n}\frac{10^{m-1}b_m + 10^{m-2}b_m + \cdots + b_m}{10^m - 1},$$

which is clearly rational. □

We finish with an important property on the rational numbers, which is covered in more generality in Chapter **??**, that informs the discussion on cardinality promised by this subsection's title.

Definition 9.23 A set X is densely-ordered if there is an element of X between any two elements of X.

Theorem 9.24 *The set of rationals $\mathbb{Q}$ is densely-ordered.*

Proof: Let $a, b \in \mathbb{Q}$ and assume $a < b$ without loss of generality. Note that $(a+b)/2 \in \mathbb{Q}$ by the closure established in Proposition 9.5. Since $a < b$, it follows that

$$a + a < a + b < b + b \quad \Leftrightarrow \quad 2a < a + b < 2b \quad \Leftrightarrow \quad a < \frac{a+b}{2} < b,$$

so there is a rational number between two arbitrary rationals, as required. □

Exercise 32 Prove that the set of irrationals $\mathbb{R} \setminus \mathbb{Q}$ is densely-ordered.

We can see that $\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}$. However, all of these sets have infinite cardinality. What is going on here? This is where countability can be used to provide some more insight.

Definition 9.25 Sets X and Y have equal cardinality if there is a bijection $f : X \to Y$.

We can see Definition 9.25 as an extension to Definition 2.8; the idea of a bijection is independent of whether or not our sets have infinite size. Indeed, we have seen examples of bijections from/to infinitely-large sets such as $\mathbb{N}$.

Remark 9.26 If a set X has cardinality $|X| = n$, this means there is a bijection $f : X \to \mathbb{Z}_n$.

Exercise 33 Prove rigorously that the set $X = \{2, 4, 6, 8, 10\}$ has cardinality 5.

Definition 9.27 A set is countably infinite if it is in bijection with $\mathbb{Z}^+$. A set is then called countable if it is either finite or countably infinite.

Example 9.28 The set of positive even numbers $2\mathbb{Z}^+ := \{2, 4, 6, \ldots\}$ is countable. Well, we can define a bijection $f : \mathbb{Z}^+ \to 2\mathbb{Z}^+$ simply by $f(n) = 2n$. Moreover, this tells us that the positive integers and the positive even numbers have the same cardinality; this does seem counter-intuitive given that, informally, ‘there are twice as many numbers as there are evens’.

Lemma 9.29 *Let X be countably infinite and $y \notin X$. Then, $X \cup \{y\}$ is countably infinite.*

Proof: By assumption, we have a bijection $f : \mathbb{Z}^+ \to X$. Hence, we define $g : \mathbb{Z}^+ \to X \cup \{y\}$ as follows, using so-called *cases notation*:

$$g(n) = \begin{cases} y, & \text{if } n = 1 \\ f(n-1), & \text{if } n > 1 \end{cases}.$$

Essentially, a bijection $f : \mathbb{Z}^+ \to X$ gives us a labelling of each element of X with a positive integer, so g just shifts that labelling along so that we free up the first label $n = 1$ to give to our new element y. From this, we conclude that g is a bijection. □

Proposition 9.30 *Let X be countable. Then, any subset $U \subseteq X$ is countable.*

Proof: (i) If U is finite, the result is trivial.

(ii) If U is infinite, then it must be that $|X| = \infty$ also. Then, by assumption, we have a bijection $f : \mathbb{Z}^+ \to X$. In other words, we label each element of X so that the set looks like $X = \{x_1, x_2, x_3, \ldots\}$. We define $g : \mathbb{Z}^+ \to U$ *inductively*:

$$\begin{aligned} g(1) &= x_i, \text{ where } x_i \in U \text{ has the smallest possible } i, \\ g(n+1) &= x_i, \text{ where } x_i \in U \setminus \{g(1), \ldots, g(n)\} \text{ has the smallest possible } i. \end{aligned}$$

It is true that g is a bijection, giving the result. □

Proposition 9.31 *Let X and Y be countable. Then, $X \cup Y$ is countable.*

Proof: (i) If X and Y are both finite, the result is trivial.

(ii) If one of X and Y is finite and the other infinite, the result follows from Proposition 9.30.

(iii) If X and Y are both infinite, we first assume we have bijections $f : \mathbb{Z}^+ \to X$ and $g : \mathbb{Z}^+ \to Y$. Without loss of generality, we can assume that X and Y are disjoint (recall this means $X \cap Y = \emptyset$). Then, we define $h : \mathbb{Z}^+ \to X \cup Y$ as follows:

$$h(n) = \begin{cases} g(n/2), & \text{if } n \text{ is even,} \\ f\left(\dfrac{n+1}{2}\right), & \text{if } n \text{ is odd} \end{cases}.$$

It is true that h is a bijection, giving the result. □

Exercise 34 Prove that $\mathbb{Z}$ is countable by finding a bijection $f : \mathbb{Z}^+ \to \mathbb{Z}$.

The end goal is to say something on the countability of both $\mathbb{Q}$ and $\mathbb{R}$. To this end, we consider the next result, noting $\mathbb{Z}^+$ is trivially countable (the bijection is the identity map $n \mapsto n$).

Lemma 9.32 *The Cartesian product $\mathbb{Z}^+ \times \mathbb{Z}^+$ is countable.*

Proof: We define $\varphi : \mathbb{Z}^+ \times \mathbb{Z}^+ \to \mathbb{Z}^+$ by $\varphi(n, m) = 2^{n-1}(2m - 1)$. This is bijective. Indeed, for injectivity, assume that $\varphi(a, b) = \varphi(c, d)$, which is equivalent to $2^{a-1}(2b - 1) = 2^{c-1}(2d - 1)$. Without loss of generality, let $a \geq c$ and divide both sides by 2^{c-1} to get $2^{a-c}(2b - 1) = 2d - 1$. Because the right-hand-side is odd, it must be that the left-hand-side is also odd, so $a = c$ and thus $b = d$. Therefore, as pairs, $(a, b) = (c, d)$. For surjectivity, for any $x \in \mathbb{Z}^+$, the Fundamental Theorem of Arithmetic implies it can be written as $x = 2^k y$, where $k \in \mathbb{N}$ (which means $k + 1 \in \mathbb{Z}^+$) and y is a product of odd prime factors. Since y is therefore odd, it will have the form $2t - 1$ for some $t \in \mathbb{Z}^+$. As such, we see that $x = \varphi(k + 1, t)$. □

Corollary 9.33 *Let X and Y be countable. Then, $X \times Y$ is countable.*

Proof: Let $f : \mathbb{Z}^+ \to X$ and $g : \mathbb{Z}^+ \to Y$ be bijective. We define $h : X \times Y \to \mathbb{Z}^+$ as follows:

$$h(x, y) = \varphi(f(x), g(y)),$$

where φ is the map in the proof of Lemma 9.32. as f, g, φ are bijective, and Corollary 8.17 implies says that composing bijections preserves the bijective property, we know h is bijective. □

Note: We now want to show that we can relax what it means for a set to be countable. The next result is of the form 'statements A, B, C are equivalent'. If we say that three (or more) statements are equivalent, that means each statement is related by an 'if and only if' to each other statement. In the case of three statements, we have

$$A \Leftrightarrow B, \quad B \Leftrightarrow C, \quad A \Leftrightarrow C.$$

Proposition 9.34 *Let $X \neq \emptyset$. The following statements are equivalent.*

(i) *X is countable.*

(ii) *There is a surjection $\mathbb{Z}^+ \to X$.*

(iii) *There is a surjection $S \to X$, where S is any countable subset.*

Proof: $\big((\text{i}) \Rightarrow (\text{ii})\big)$ This is trivial, by the definition of countability.

$\big((\text{ii}) \Rightarrow (\text{iii})\big)$ This is trivial, since if S is countable, there is a bijection $S \to \mathbb{Z}^+$. In particular, this is surjective and Lemma 8.16 implies the composition $S \to \mathbb{Z}^+ \to X$ is surjective.

$\big((\text{iii}) \Rightarrow (\text{i})\big)$ Let $f : S \to X$ be surjective. As S is countable, it has the form $S = \{s_1, s_2, s_3, \ldots\}$. Define a new function $g : X \to S$ by $g(x) = s_i$ where $s_i \in S$ has the smallest index i such that $f(s_i) = x$. By surjectivity, there will always exist some s_i. Therefore, g is bijective onto its image $\text{im}(g)$, but because $\text{im}(g) \subseteq S$, Proposition 9.30 implies that $\text{im}(g)$ is countable. Therefore, X is countable. □

Theorem 9.35 *The set of rationals $\mathbb{Q}$ is countable.*

Exercise 35 Prove Theorem 9.35.

[**Hint:** Consider the function $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Q}$ which sends the pair (a, b) either to the rational number $a/2b$ or to the number 0 in the case that $b = 0$. Argue that this is surjective and use Corollary 9.33 and Proposition 9.34 to complete your proof.]

Definition 9.36 A set X is uncountable if there exist no bijections with $\mathbb{Z}^+$.

Lemma 9.37 *The power set $\mathcal{P}(\mathbb{Z}^+)$ is uncountable.*

Proof: Assume to the contrary $\mathcal{P}(\mathbb{Z}^+)$ **is** countable, i.e. there exists a bijection $f : \mathbb{Z}^+ \to \mathcal{P}(\mathbb{Z}^+)$. Define the set $X = \{n \in \mathbb{Z}^+ : n \notin f(n)\}$. It is clear that $X \subseteq \mathbb{Z}^+$, but because f is a bijection, it follows that there exists $m \in \mathbb{Z}^+$ such that $f(m) = X$. There are two cases to consider.

(i) We have $m \in X \Leftrightarrow m \in f(m)$ since $f(m) = X$, but by definition of X, $m \notin f(m)$.

(ii) We have $m \notin X \Leftrightarrow m \notin f(m)$ since $f(m) = X$, but by definition of X, $m \in f(m)$.

Each yields a contradiction, so the only option is $\mathcal{P}(\mathbb{Z}^+)$ is uncountable. □

Theorem 9.38 *The set of real numbers $\mathbb{R}$ is uncountable.*

Proof: Assume to the contrary $\mathbb{R}$ **is** countable. In particular, Proposition 9.30 implies that the interval $[0, 1)$ is countable (recall this is the interval containing all numbers between zero and one, *including* zero but *excluding* one). Hence, there exists a bijection $f : \mathbb{Z}^+ \to [0, 1)$. Suppose this function is of the form

$$\begin{aligned} f(1) &= 0.d_{11}d_{12}d_{13}d_{14}... \\ f(2) &= 0.d_{21}d_{22}d_{23}d_{24}... \\ f(3) &= 0.d_{31}d_{32}d_{33}d_{34}... \\ &\vdots \end{aligned}$$

where d_{ij} represents the j^{th} decimal place in the number $f(i)$ and these are decimal expressions **without** repeated nines. Then, it is possible to define the number $x = 0.e_1e_2e_3e_4...$ where

$$e_i = \begin{cases} 4, & \text{if } d_{ii} \neq 4 \\ 5, & \text{if } d_{ii} = 4 \end{cases}.$$

In this way, we ensure that $x \in [0,1)$ but that it is completely different to every other number on the list above, so $x \notin \text{im}(f)$, contradicting f being bijective. □

> **Exercise 36** Prove that $\mathbb{R} \setminus \mathbb{Z}$ is uncountable. Can your proof also work with $\mathbb{R} \setminus \mathbb{Q}$?

10 Boundedness

The final thing we discuss here is to really prepare for the more rigorous Chapter **??**. This begins with a discussion on what are called *ordered fields*, something which is generalised by Chapter **??**, and will then concern with finding upper and lower bounds on subsets of $\mathbb{R}$.

Definition 10.1 An ordered field is a set X on which we have the operations $+, -, \times, \boldsymbol{\nabla\cdot}$ (where we don't allow division by zero) which have the usual properties such as commutativity and distributivity (again, this is treated in more depth later), as well as the structure of an ordering relation $<$, which means that for distinct $x, y \in X$, exactly one of $x < y$ and $y < x$ is true.

Example 10.2 Here are some examples and non-examples of ordered fields.

(i) $\mathbb{R}$ is an ordered field with the usual relation $<$.

(ii) $\mathbb{Q}$ is an ordered field with the same relation $<$ as $\mathbb{R}$.

(iii) $\mathbb{Z}$ is **not** an ordered field as it isn't even a field ($\boldsymbol{\nabla\cdot}$ isn't well-defined).

(iv) $\mathbb{C}$ is **not** an ordered field as there is no well-defined ordering relation.

Definition 10.3 Let $A \subseteq \mathbb{R}$ be a subset.

(i) An upper bound on A is a number $K \in \mathbb{R}$ such that $a \leq K$ for all $a \in A$.

(ii) A lower bound on A is a number $L \in \mathbb{R}$ such that $a \geq L$ for all $a \in A$.

We say that A is unbounded (above/below) if these numbers do not exist.

Example 10.4 Consider the interval $A = [1, 2)$; it is clear that A is bounded below by -100 and above by 23. Can we *tighten* these bounds? Yes we can. In fact, the tightest bound we can get is the following: A is bounded below by 1 and above by 2. Although we can't prove it yet, this is the tightest possible bound on A.

Note: It is clear from Example 10.4 that a tight bound may or may not be in the set itself. Indeed, the greatest lower bound $1 \in A$ whereas the least upper bound $2 \notin A$.

Exercise 37 Suggest the tightest possible bounds on the interval $(-2, 4]$. Considering this and Example 10.4, conjecture what the tightest bounds are on the following intervals, where $x, y \in \mathbb{R}$ are some real numbers.

(i) $[x, y]$.
(ii) $[x, y)$.
(iii) $(x, y]$.
(iv) (x, y).

Although this may seem 'obvious' given our intuition on real numbers and integers, the following result is fundamentally important; it will be used time and again when dealing rigorously with limits.

Theorem 10.5 (Archimedean Property of $\mathbb{R}$) *Given any real number $K \in \mathbb{R}$, there exists a positive integer $n \in \mathbb{Z}^+$ such that $n > K$.*

Proof: Deferred; see just after Exercise 39. □

Definition 10.6 Let $A \subseteq \mathbb{R}$ be a subset.

(i) The supremum of A is the least upper bound of A, denoted $\sup(A)$.
(ii) The infimum of A is the greatest lower bound of A, denoted $\inf(A)$.

Example 10.7 Consider the set $A = \{1/n : n \in \mathbb{Z}^+\}$. We will now discuss its supremum and its infimum (if they exist). We claim that $\sup(A) = 1$. Well, suppose we have a real number $K < 1$. Because $1 \in A$, it is clear that K is **not** an upper bound on A. Therefore, 1 is the smallest upper bound and that is the definition of supremum. We claim that $\inf(A) = 0$. Well, suppose we have a real number $L > 0$. Then, it is clear that $1/L \in \mathbb{R}$. By the Archimedean Property of $\mathbb{R}$, there exists a positive integer $n \in \mathbb{Z}^+$ such that $n > 1/L$, which we can rearrange to $1/n < L$, but $1/n \in A$ by definition, so L is **not** a lower bound on A. Therefore, 0 is the greatest lower bound and that is the definition of infimum.

Note: This is the general method for proving what the supremum/infimum of a set is.

(i) Have an 'educated guess' as to what the supremum/infimum appears to be.
(ii) Show that any number K **less than** your supremum is **not** an upper bound.
(iii) Show that any number L **greater than** your infimum is **not** a lower bound.

Exercise 38 Determine the supremum and infimum of $A = \{1/p - 2/q : p, q \in \mathbb{Z}^+\}$.

Axiom 10.8 (Axiom of Completeness) *Any non-empty subset of $\mathbb{R}$ bounded above has a supremum in $\mathbb{R}$.*

Exercise 39 Determine if there is an equivalent Axiom of Completeness for $\mathbb{Q}$, that is does every subset of rational numbers with an upper bound have a supremum in the rationals?

[**Hint:** The key part of the statement is that the supremum is in the **rationals**.]

Proof (of the Archimedean Property of $\mathbb{R}$): Assume to the contrary that $\mathbb{Z}^+$ is bounded above. By the Axiom of Completeness, there exists a supremum $K \in \mathbb{R}$, say. Since $K-1 < K$, it is clear that $K-1 \in \mathbb{R}$ is **not** an upper bound. Therefore, there exists $n \in \mathbb{Z}^+$ such that $n > K-1$, which is equivalent to $n+1 > K$, but $n+1 \in \mathbb{Z}^+$ so K is **not** an upper bound, a contradiction. □

Proposition 10.9 *Any non-empty subset of $\mathbb{R}$ bounded below has an infimum in $\mathbb{R}$.*

Proof: Let $A \subseteq \mathbb{R}$ be non-empty and bounded below. Then, define the set $B = \{-x : x \in A\}$. It is clear that $B \neq \emptyset$ because $A \neq \emptyset$. By assumption, A is bounded below by L, say, which means that B is bounded above by $-L$. Hence, the Axiom of Completeness implies that B has a supremum in $\mathbb{R}$. For notation, suppose that $\sup(B) = K$ and take some $x \in A$ (which means $-x \in B$). Since K is an upper bound, we have $-x \leq K$ which is equivalent to $x \geq -K$. This shows that $-K$ is a lower bound on A. Next, assume that $M > -K$; this is equivalent to $-M < K$, but because K is the least upper bound on B, it follows that $_M$ is **not** an upper bound on B. Hence, there exists some $-y \in B$ such that $-y > -M$, which is equivalent to $y < M$. By definition, $y \in A$ so this means that M is **not** a lower bound on A. Thus, we have shown that $\inf(A) = -K \in \mathbb{R}$, so the infimum certainly exists and is a real number. □

Exercise 40 A subset $A \subseteq \mathbb{R}$ is said to have the *predecessor property* if $a \in A$ implies that $a - 1 \in A$. Show that any subset with the predecessor property is unbounded below.

[**Hint:** Perform a proof by contradiction using Proposition 10.9.]

We now give a slightly alternate formulation to Section 9 using intervals. We can then use the final part of Exercise 36 to notice a (possibly surprising) phenomenon the irrationals exhibit.

Definition 10.10 A collection of intervals $I_1, I_2, ... \subseteq \mathbb{R}$ is called nested if $I_1 \supseteq I_2 \supseteq \cdots$.

Example 10.11 The collection of intervals $I_n = (0, 1/n]$ is nested. Indeed, we can see that

$$I_1 = (0,1] \quad \supseteq \quad I_2 = (0,1/2] \quad \supseteq \quad I_3 = (0,1/3] \quad \supseteq \quad \cdots .$$

Lemma 10.12 (Nested Intervals Lemma) *Let $I_n = [a_n, b_n]$ be a nested collection of closed intervals. Then, there exists $x \in \mathbb{R}$ such that $x \in I_n$ for all $n \in \mathbb{Z}^+$, that is there is at least one element in every closed interval in the collection.*

Proof: Consider the set $A = \{a_n : n \in \mathbb{Z}^+\}$ of lower endpoints of our intervals; it is non-empty. Consider now the upper endpoint b_m of the interval I_m, for some $m \in \mathbb{Z}^+$. There are two cases to consider.

(i) If $n \geq m$, then $[a_n, b_n] \subseteq [a_m, b_m]$, which means $a_n \in [a_m, b_m]$ and thus $a_n \leq b_m$.

(ii) If $n < m$, then $[a_m, b_m] \subseteq [a_n, b_n]$ which means $b_m \in [a_n, b_n]$ and thus $a_n \leq b_m$.

Either way, we see that $a_n \leq b_m$ for every $n \in \mathbb{Z}^+$, so it follows that b_m is an upper bound on A. By the Axiom of Completeness, we know that $\sup(A) = K$ exists. Because K itself is an upper bound, we have that $a_n \leq K$ for all $n \in \mathbb{Z}^+$. Since m was arbitrary, it is true that every b_n is an upper bound on A, which implies that $K \leq b_n$. Thus, $K \in [a_n, b_n]$ for every $n \in \mathbb{Z}^+$, so we have found our element in all intervals; take $x = K$. □

Proposition 10.13 *The set $[0, 1]$ is uncountable.*

Remark 10.14 We essentially showed Proposition 10.13 when proving Theorem 9.38. However, there is a slick proof which makes use of the Nested Intervals Lemma.

Proof: Assume to the contrary $[0, 1]$ is countable, that is it has the form $[0, 1] = \{x_1, x_2, x_3, \ldots\}$. It is now possible to define a collection of closed intervals as follows:

I_1 is **any** closed interval in $[0, 1]$ **not** containing x_1,
I_2 is **any** closed interval in I_1 **not** containing x_2,
I_3 is **any** closed interval in I_2 **not** containing x_3,
$\vdots$

In this way, we get a nested collection of closed intervals $I_1 \supseteq I_2 \supseteq \cdots$, so the Nested Intervals Lemma implies the existence of an element $x \in [0, 1]$ such that $x \in I_n$ for all $n \in \mathbb{Z}^+$. However, by definition, $x_n \notin I_n$, so the number $x \neq x_n$ for any $n \in \mathbb{Z}^+$. Consequently, our enumeration $\{x_1, x_2, x_3, \ldots\}$ misses out the element x, a contradiction to being countable. □

Since we showed in Exercise 32 that $\mathbb{R} \setminus \mathbb{Q}$ is densely-ordered in $\mathbb{R}$, it follows that the irrationals are also densely-ordered in $[0, 1]$. Thus, combined with the fact that (spoiler alert) the irrationals are uncountable, as showed in Exercise 36, we know that there are more irrationals in $[0, 1]$ than there are rationals in all of $\mathbb{R}$.

Infinity is weird.

11 Exercise Solutions

We provide detailed solutions to the exercises interwoven within each section of the module. Hopefully you have given these questions a try whilst on your learning journey with the module. But mathematics is difficult, so don't feel disheartened if you had to look up an answer before you knew where to begin (we have all done it)!

Solutions to Exercises in Section 2

> **Exercise 1** State either true or false to the following, giving a brief reason for each.
>
> (i) $32 \in \{2, \{32\}, \{-5\}, -5\}$.
>
> (ii) $64 \in \{\text{the set of even numbers}\}$.
>
> (iii) $\{4\} \in \{100, \pi, \{4\}, 7\}$.
>
> (iv) $\emptyset = \{\text{the set of prime numbers which are even}\}$.

Solution:

(i) False, because inside the set is another *set* $\{32\}$, not just the number 32.

(ii) True, because $64 = 2 \times 32$ which means it is even.

(iii) True, because the set $\{4\}$ is an element of the set $\{100, \pi, \{4\}, 7\}$.

(iv) False, because 2 is an even prime; the set of even primes is certainly non-empty. □

> **Exercise 2** List all subsets of the following and note how many subsets there are of each.
>
> (i) $\emptyset$.
>
> (ii) $\{1\}$.
>
> (iii) $\{1, 2\}$.
>
> (iv) $\{1, 2, 3\}$.
>
> Do you notice a pattern? How many subsets are there of $\{1, 2, ..., n\}$?

Solution:

(i) There is one subset: $\emptyset$.

(ii) There are two subsets: $\emptyset, \{1\}$.

(iii) There are four subsets: $\emptyset, \{1\}, \{2\}, \{1, 2\}$.

(iv) There are eight subsets: $\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}$.

In general, there are 2^n subsets of the set $\{1, 2, ..., n\}$. □

> **Exercise 3** Write each of the following sets using the restriction form discussed above.
> (i) $\{3, 4, 5, 6, 7, 8, 9\}$.
> (ii) $\{1, -8, 27, -64, 125, -216, 343, ...\}$.
> (iii) {Ford, Carter, Reagan, Bush, Clinton, Bush, Obama, Trump, Biden}.

Solution:

(i) $\{n \in \mathbb{Z} : 3 \leq n \leq 9\}$.

(ii) $\{(-1)^{n+1}n^3 : n \in \mathbb{Z}^+\}$.

(iii) {US presidents : they became president on or after August 9th 1974}. □

> **Exercise 4** Let $A, B \subseteq X$. Prove that $A \setminus B \subseteq A \cap (X \setminus B)$. Is this a proper subset? Is this an equality? Justify your answer either with a proof or with a counterexample.

Solution: Let $x \in A \setminus B$, meaning that $x \in A$ and $x \notin B$. Because $B \subseteq X$ is a subset, the fact x isn't in B can be re-written as $x \in X \setminus B$. Thus, $x \in A$ and $x \in X \setminus B$. Finally, we recall the intersection symbol means 'and', giving us $x \in A \cap (X \setminus B)$. This proves the first part. Why? Because whenever we take an element of $A \setminus B$, we have shown that it also means it lives inside $A \cap (X \setminus B)$.

In fact, if we take some element $y \in A \cap (X \setminus B)$, we will always have that $y \in A \setminus B$ basically by reversing the above argument. Therefore, we also have that $A \cap (X \setminus B) \subseteq A \setminus B$. Since we have the subsets both ways, it must mean that they are equal. □

> **Exercise 5** Let $X = \{2, 1\}$ and $Y = \{3, 6\}$ and $Z = \{8, 7\}$. Write out $X \times Y \times Z$ in full.

Solution: $X \times Y \times Z = \{(2, 3, 8), (2, 3, 7), (2, 6, 8), (2, 6, 7), (1, 3, 8), (1, 3, 7), (1, 6, 8), (1, 6, 7)\}$. □

> **Exercise 6** The power set of a set X is the set of all subsets, denoted $\mathcal{P}(X)$. Write out the power set of $X = \{a, \{b, c\}\}$ and determine its cardinality $\big|\mathcal{P}(X)\big|$. Does this agree with your answer from Exercise 2?

Solution: The power set $\mathcal{P}(\{a, \{b, c\}\}) = \{\emptyset, \{a\}, \{\{b, c\}\}, \{a, \{b, c\}\}\}$. Be aware that the number of curly brackets in the correct places is vitally important. As such, we can see that $\big|\mathcal{P}(X)\big| = 4$, which agrees with our answer to Exercise 2, since $|X| = 2$ so there will be $2^2 = 4$ subsets. □

Note: Because the power set is the set of subsets, we essentially list out all the subsets (which are themselves sets) and put the whole list in curly brackets!

Exercise 7 Write two examples of if-then statements, with one true and the other false.

Solution: There are many examples; here are two which popped into my head: "if $x^2 = 9$ for $x \in \mathbb{Z}$, then $x = 3$ or $x = -3$" (true) and "if $x^2 = 9$ for $x \in \mathbb{N}$, then $x = 3$ or $x = -3$" (false). □

Exercise 8 Prove, using contradiction, that if n^2 is even, then n is even, for all $n \in \mathbb{Z}$.

Solution: Assume to the contrary that n^2 is even but that n is odd. Then, $n = 2k + 1$ for some $k \in \mathbb{Z}$, that is it is always one more than an even number (that's just the definition of being odd). Therefore, we see that

$$n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1,$$

so n^2 is one more than an even number, making it odd. This contradicts n^2 being even. □

Exercise 9 Prove, by contrapositive, that if $7y + 9$ is even, then y is odd, where $y \in \mathbb{Z}$.

Solution: Suppose that y is not odd, meaning that it is even. Thus, we can write it as $y = 2k$ for some $k \in \mathbb{Z}$. Therefore, $7y + 9 = 7(2k) + 9 = 14k + 9 = 2(7k + 4) + 1$, meaning that $7y + 9$ is odd. By the contrapositive, the result follows. □

Exercise 10 Prove, by induction, that $\sum_{t=1}^{n} t = \frac{1}{2}n(n + 1)$ for all $n \in \mathbb{Z}^+$.

Solution: The base case $n = 1$ is clear, because the left-hand-side is just 1 and the right-hand-side is $2/2 = 1$. As for the inductive hypothesis, we assume the formula holds for some general $n = k \in \mathbb{Z}^+$. We now look to the $n = k + 1$ case:

$$\begin{aligned}\sum_{t=1}^{k+1} t &= \sum_{t=1}^{k} t + (k + 1) \\ &= \frac{1}{2}k(k + 1) + (k + 1), \text{ by the inductive hypothesis,} \\ &= \frac{1}{2}(k(k + 1) + 2(k + 1))\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2}(k^2 + 3k + 2) \\
&= \frac{1}{2}(k+1)(k+2), \text{ by factorising,}
\end{aligned}
$$

which is precisely the formula we have to prove with $n = k+1$. By the principal of mathematical induction, the statement holds for all $n \in \mathbb{Z}^+$. □

> **Exercise 11** The converse of "if P, then Q" is defined as "if Q, then P". Determine the converse of Corollary 5.4. Is it true? If so, prove it. If not, give a counterexample.

Solution: The converse of Corollary 5.4 is this: "let $a, b \in \mathbb{Z}$ such that $a \mid b^2$. Then, $a \mid b$". We can see that the converse statement is false: $8 \mid 4^2$ but $8 \nmid 4$. □

> **Exercise 12** Carefully use Definition 5.8 to justify whether or not 1 is prime.

Solution: Per Definition 5.8, 1 is not prime as it doesn't have two distinct positive factors. □

> **Exercise 13** Justify the following statements where $a, b \in \mathbb{Z}^+$.
> (i) The greatest common divisor is symmetric, that is $\gcd(a, b) = \gcd(b, a)$.
> (ii) The greatest common divisor is bounded below, namely $\gcd(a, b) \geq 1$.
> Compute $\gcd(18, -57)$ and suggest how it compares to both $\gcd(-18, 57)$ and $\gcd(18, 57)$.

Solution:

(i) By definition, $\gcd(a, b)$ is the largest integer k such that $k \mid a$ and $k \mid b$. This is precisely the same as the largest integer where $k \mid b$ and $k \mid a$, the definition of $\gcd(b, a)$.

(ii) Given any pair of integers a and b, we always have that $1 \mid a$ and $1 \mid b$, so the value of $\gcd(a, b)$ will always be at least this.

We can proceed as in Example 5.12 and write out the positive factors of each:

$$
\begin{aligned}
\{\text{+ve factors of } 18\} &= \{1, 2, 3, 6, 9, 18\}, \\
\{\text{+ve factors of } -57\} &= \{1, 3, 19, 57\}.
\end{aligned}
$$

Therefore, $\gcd(18, -57) = 3$. In fact, the greatest common divisor is invariant under the sign of the integers we are discussing. In other words, the following is true:

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

As such, we can also conclude that $\gcd(-18, 57) = 3 = \gcd(18, 57)$. □

Exercise 14 Use the Euclidean Algorithm to compute $\gcd(4635, 873)$. Consequently, apply Bézout's Lemma to express your answer in the form $4635s + 873t$, for some $s, t \in \mathbb{Z}$. Are these integers that we get in Bézout's Lemma unique? Briefly justify your answer.

Solution: We apply the Euclidean Algorithm as in Example 5.15. Indeed,

$$\begin{aligned} 4635 &= 5(873) + 270, \\ 873 &= 3(270) + 63, \\ 270 &= 4(63) + 18, \\ 63 &= 3(18) + 9, \\ 18 &= 2(9) + 0. \end{aligned}$$

Therefore, we conclude $\gcd(4635, 873) = 9$. The next job amounts to following Example 5.17:

$$\begin{aligned} 7 &= 63 - 3(18), && \text{by rearranging the fourth equation,} \\ &= 63 - 3(270 - 4(63))), && \text{by rearranging the third equation,} \\ &= 13(63) - 3(270) && \\ &= 13(873 - 3(270)) - 3(270), && \text{by rearranging the second equation,} \\ &= 13(873) - 42(270) && \\ &= 13(873) - 42(4635 - 5(875)), && \text{by rearranging the first equation,} \\ &= 4635(-42) + 873(223). && \end{aligned}$$

Consequently, we see that $s = -42$ and $t = 223$ gives us the result we want. As for uniqueness, they are not unique. In fact, we can see that

$$4635(-42 + 873) + 873(223 - 4635) = 9$$

still works. Why? Well, if we expand out the brackets above, we get precisely the same equation as at the end of the Bézout calculation except with $4635(873) + 873(-4625)$ added on, but clearly this is zero. Hence, we can also take $s = 831$ and $t = -4412$. □

Exercise 15 Prove Corollary 5.21.

Solution: Let $p \mid ab$ where $p \nmid a$. Then, $\gcd(p, a) = 1$ and Euclid's Lemma applies: $p \mid b$. □

Note: These types of arguments are quite slick. If ever you need to prove that something satisfies condition P or condition Q, it is sufficient to assume that if it does not satisfy condition P, then it must satisfy condition Q because the only other option is that it does satisfy condition P, and so the statement that it satisfies P or Q is still very much valid.

Exercise 16 Is $\leq$ an equivalence relation on $\mathbb{Z}$? Justify your answer.

Solution: No. This is because it isn't symmetric: $3 \leq 4$ doesn't imply that $4 \leq 3$. □

Exercise 17 Describe the equivalence classes in $\mathbb{N}$ under the relation $\sim$ defined as follows:

$$m \sim n \quad \text{if and only if} \quad 2 \mid (m-n).$$

You do not need to prove that $\sim$ is an equivalence relation.

Solution: Under this equivalence relation, the natural numbers are related if and only if they are either both odd or they are both even. Hence, there are two equivalence classes: the even naturals $2\mathbb{N}$ and the odd naturals $2\mathbb{N}+1$ (this is non-standard notation but hopefully you can see why I have denoted the odd naturals in this way). □

Exercise 18 Prove Theorem 7.4 by appealing directly to Definition 7.1.

Solution: By definition, $x \equiv s \pmod{n}$ means there exists $k_1 \in \mathbb{Z}$ such that $x = s + k_1 n$ and $y \equiv t \pmod{n}$ means there exists $k_2 \in \mathbb{Z}$ such that $y = t + k_2 n$. Therefore, we see the following:

(i) $x + y = s + k_1 n + t + k_2 n = s + t + (k_1 + k_2)n$.

(ii) $xy = (s + k_1 n)(t + k_2 n) = st + (k_1 t + k_2 s + k_1 k_2 n)n$.

Because the sums/products of two integers is an integer, the above are precisely the definition of congruence modulo n between (i) $x + y$ and $s + t$ and (ii) xy and st. □

Exercise 19 Prove Corollary 7.9 by using Fermat's Little Theorem.

Solution: (i) From Fermat's Little Theorem, we can apply the Arithmetic of Modulo Congruence (that we proved in Exercise 18) to see that $a^p \equiv a \pmod{p}$ is equivalent to $a^p - a \equiv 0 \pmod{p}$.

(ii) Given that $p \not| a$, another thing we can do is divide through in the formula of Fermat's Little Theorem by a, giving us $a^{p-1} \equiv 1 \pmod{p}$. □

Note: It is crucial in (ii) that p doesn't divide a, else we have $a \equiv 0 \pmod{p}$, so we would essentially be dividing by zero (which is absolutely not allowed!)

Exercise 20 Use Fermat's Little Theorem to prove that 39 is **not** prime.

Solution: This is obviously overkill (since we can see that $39 = 13 \cdot 3$) but let's go: assume to the contrary that $p = 39$ is prime and let $a = 2$. We have that $39 \not| 2$, so we should be able to apply Corollary 7.9(ii) (that we proved in Exercise 19), namely $2^{39-1} \equiv 1 \pmod{39}$. Let's check if this is true. Because the left-hand-side is quite large, it is best to re-write it so that we can hit it with some congruences (using the Arithmetic of Modulo Congruence). Indeed,

$$2^{38} = 2^{32} \cdot 2^6 = (2^8)^4 \cdot 2^6.$$

Now, $2^8 = 256 \equiv 22 \pmod{39}$ which means that $2^{32} \equiv 22^4 = 234\,256 \equiv 22 \pmod{39}$. As such,

$$2^{38} \equiv 22 \times 2^6 = 1408 \equiv 4 \not\equiv 1 \pmod{39},$$

a contradiction to Fermat's Little Theorem. Hence, 39 is not prime. □

Exercise 21 Show $[x] + [y] := [x + y]$ is well-defined under the equivalence relation $\equiv$.

Solution: Suppose that $[x] = [u]$ and $[y] = [v]$, that is x and u both represent/live inside the same equivalence class, and the same for y and v. By definition then, $x \equiv u$ and $y \equiv v$, which is to say there exist $k_1, k_2 \in \mathbb{Z}$ such that

$$x = u + k_1 n \quad \text{and} \quad y \cong v + k_2 n.$$

Now then, the Arithmetic of Modulo Congruence implies that $x + y \equiv u + v$, from which it follows that $[x] + [y] := [x + y] = [u + v] =: [u] + [v]$, so the operation is independent of the representatives chosen and thus it is well-defined. □

Exercise 22 Find all solutions to the congruence equation $759x \equiv 100 \pmod{12\,167}$.

Solution: We refer to Theorem 7.17 to check how many solutions we should expect. Indeed, we can calculate (via the Euclidean Algorithm, say) that $\gcd(759, 12\,167) = 23$. We notice that $23 \nmid 100$, so according to the aforementioned result, there are no solutions. □

Exercise 23 Decrypt $r = 8363$ using the encryption $k = 11\,787$ and $(s, n) = (3, 17\,947)$.

Solution: By definition, the decryption is the integer $m \equiv 8363^{11\,787} \pmod{17\,947}$. The power here is quite large, so it is advisable to break down 11 787 into sums of smaller powers, say powers of two. Indeed, it is true that $11\,787 = 2^{13} + 2^{11} + 2^{10} + 2^9 + 2^3 + 2^1 + 2^0$, which implies

$$8363^{11\,787} = (8363)^{2^{13}} \cdot (8363)^{2^{11}} \cdot (8363)^{2^{10}} \cdot (8363)^{2^{9}} \cdot (8363)^{2^{3}} \cdot (8363)^{2^{1}} \cdot (8363)^{2^{0}}.$$

Consequently, we can reduce each of these modulo 17 947 to get the result:

$$\begin{aligned}
(8363)^{2^0} &\equiv 8363 \pmod{17\,947},\\
(8363)^{2^1} &\equiv 310 \pmod{17\,947},\\
(8363)^{2^3} &\equiv 310^{2^2} = 6846 \pmod{17\,947},\\
(8363)^{2^9} &\equiv (6846)^{2^6} \equiv 2778 \pmod{17\,947},\\
(8363)^{2^{10}} &\equiv 2778^2 \equiv 74 \pmod{17\,947},\\
(8363)^{2^{11}} &\equiv 74^2 = 5476 \pmod{17\,947},\\
(8363)^{2^{13}} &\equiv (5476)^{2^2} \equiv 1489 \pmod{17\,947}.
\end{aligned}$$

Thus, if we multiply each of the above numbers together modulo 11 947, we get our decryption:

$$\begin{aligned}
m &= 1489 \cdot 5476 \cdot 74 \cdot 2778 \cdot 6846 \cdot 310 \cdot 8363\\
&\equiv 1489 \cdot 5476 \cdot 74 \cdot 2778 \cdot 6846 \cdot 8162 \pmod{17\,947},\\
&\equiv 1489 \cdot 5476 \cdot 74 \cdot 2778 \cdot 8041 \pmod{17\,947},\\
&\equiv 1489 \cdot 5476 \cdot 74 \cdot 11\,830 \pmod{17\,947},\\
&\equiv 1489 \cdot 5476 \cdot 13\,964 \pmod{17\,947},\\
&\equiv 1489 \cdot 12\,644 \pmod{17\,947},\\
&\equiv 513 \pmod{17\,947}.
\end{aligned}$$

□

Note: In the final calculation above, going line-to-line, we reduce modulo 17 947 the two things been multiplied at the end of the line; this way, it is still manageable (a computer should have no issue jumping from the first line to the last but we are mere humans!).

Exercise 24 Consider the function $p : \mathbb{N} \to \mathbb{N} \cup \{0\}$ where $p(n)$ is the number of distinct prime factors of n. Why do we need 0 in the co-domain?

Solution: Well, $1 \in \mathbb{N}$ is in the domain and the number of distinct prime factors of 1 is zero, that is we need to have $p(1) = 0$. This is why we include zero in the co-domain. □

Exercise 25 Determine if $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ where $\varphi(x, y) = (x - y, xy)$ is injective/surjective.

Solution: It is certainly not injective. Indeed, $\varphi(1, 1) = (0, 1) = \varphi(-1, -1)$ but the inputs are different: $(1, 1) \neq (-1, -1)$. It is also not surjective. Indeed, $(0, -1)$ is not in the image of φ for this reason: assume to the contrary that it was, meaning there exists a pair $(a, b) \in \mathbb{R}^2$ such that $\varphi(a, b) = (0, -1)$. This means $a - b = 0$ and $ab = -1$; in other words, $a = b$ and $ab = a^2 \geq 0$, contradicting the fact that $ab = -1$. □

Exercise 26 Determine which of these functions is injective:

$$\begin{array}{ccccc} f & : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & & x & \longmapsto & x^2 \end{array}, \qquad \begin{array}{ccccc} g & : & \mathbb{Q} & \longrightarrow & \mathbb{Q} \\ & & x & \longmapsto & x^2 \end{array},$$

$$\begin{array}{ccccc} h & : & \mathbb{Z} & \longrightarrow & \mathbb{Z} \\ & & x & \longmapsto & x^2 \end{array}, \qquad \begin{array}{ccccc} k & : & \mathbb{N} & \longrightarrow & \mathbb{N} \\ & & x & \longmapsto & x^2 \end{array},$$

$$\begin{array}{ccccc} s & : & \mathbb{N} & \longrightarrow & \mathbb{Z} \\ & & x & \longmapsto & x^2 \end{array}, \qquad \begin{array}{ccccc} t & : & \mathbb{Z} & \longrightarrow & \mathbb{N} \\ & & x & \longmapsto & x^2 \end{array}.$$

[**Hint:** The injectivity of a function is determined by both its rule $x \mapsto \cdots$ and its domain.]

Solution:

(i) The function f is not injective: $f(2) = 4 = f(-2)$ but $2 \neq -2$.

(ii) The function g is not injective for the same reason.

(iii) The function h is not injective for the same reason.

(iv) The function k is injective: $f(x) = f(y)$ means $x^2 = y^2$, so $x = y$ since they are positive.

(v) The function s is injective for the same reason.

(vi) The function t is not injective: $f(-3) = 9 = f(3)$ but $-3 \neq 3$. □

Exercise 27 Prove Lemma 8.16.

Solution: Let $z \in Z$. Then, because g is surjective, there exists $y \in Y$ such that $z = g(y)$. But now, because f is surjective, there exists $x \in X$ such that $y = f(x)$. Combining these gives that $z = g(f(x)) = (g \circ f)(x)$, but this is the definition of surjective for $g \circ f$. □

Exercise 28 Consider a rational $r = m/n \in \mathbb{Q}$. How many representations are there of r as a ratio of two integers, that is how many elements are there in the equivalence class $[(m, n)]$ under the relation $\sim$ defined in Remark 9.2?

Solution: There are infinitely-many ways to represent a rational as a ratio of integers. This is because we can always multiply the class representative by a non-zero integer, that is to say $[(m, n)] = [(km, kn)]$ for some $k \in \mathbb{Z} \setminus \{0\}$. This is precisely the property we learn at high school that allows us to cancel common factors on the top and bottom of a fraction. □

Exercise 29 Prove that $\sqrt[7]{25}$ is irrational **with** the Fundamental Theorem of Arithmetic.

Solution: Assume to the contrary that $\sqrt[7]{25}$ is rational. Then, there exist $m, n \in \mathbb{Z}$ with $n \neq 0$ such that $\sqrt[7]{25} = m/n$ in lowest terms. By the Fundamental Theorem of Arithmetic, it is possible to find prime factorisations of these integers, namely $m = p_1 \cdots p_s$ and $n = q_1 \cdots q_t$. Therefore,

$$\sqrt[7]{25} = \frac{p_1 \cdots p_s}{q_1 \cdots q_t} \quad \Leftrightarrow \quad 25q_1^7 \cdots q_t^7 = p_1^7 \cdots p_s^7.$$

The left-hand-side contains two factors of five more than the right-hand-side, a contradiction to the Fundamental Theorem of Arithmetic (because the above equality just tells us that one integer is the same as another, but the theorem says that any prime factorisation is unique up to the order we write the primes in). □

Exercise 30 Prove that $0.dddddd... = d/9$ for any digit $d \in \{0, 1, 2, \ldots, 9\}$.

Solution: We have $0.ddddd... = 0.d + 0.0d + 0.00d + \cdots = d10^{-1} + d10^{-2} + d10^{-3} + \cdots$, so

$$0.ddddd... = \frac{d}{10}\sum_{k=0}^{\infty}\frac{1}{10^k}.$$

Using the Geometric Series formula from Corollary 9.17, we see that

$$0.ddddd... = \frac{d}{10}\frac{1}{1-1/10} = \frac{1}{10}\frac{d}{9/10} = \frac{d}{9}.$$ □

Exercise 31 Prove that $0.dedede... = (10d + e)/99$ for any digits $d, e \in \{0, 1, 2, \ldots, 9\}$.

Solution: This is very much in the same vein as the solution to Exercise 30. Indeed, we have $0.dedede... = 0.de + 0.00de + 0.0000de + \cdots = (d10^{-1} + e10^{-2}) + (d10^{-3} + e10^{-4}) + \cdots$, so

$$0.dedede... = \frac{10d+e}{100}\sum_{k=0}^{\infty}\frac{1}{100^k}.$$

Using the Geometric Series formula from Corollary 9.17, we see that

$$0.dedede... = \frac{10d+e}{100}\frac{1}{1-1/100} = \frac{1}{100}\frac{10d+e}{99/100} = \frac{10d+e}{99}.$$ □

Exercise 32 Prove that the set of irrationals $\mathbb{R} \setminus \mathbb{Q}$ is densely-ordered.

Solution: Let $a, b \in \mathbb{R} \setminus \mathbb{Q}$ where $a < b$ without loss of generality. Then, we can choose some $n \in \mathbb{Z}$ such that $n > \sqrt{2}/(b-a)$. Said integer will be positive and, by definition, it means that

$$a < a + \frac{\sqrt{2}}{n} < b.$$

Finally, it remains to note that because $\sqrt{2}$ is irrational, it follows that the number $a + \sqrt{2}/n$ is irrational. Because we have found an irrational between two arbitrary irrationals, it follows that $\mathbb{R} \setminus \mathbb{Q}$ is densely-ordered (with respect to the usual ordering relation $\leq$ on $\mathbb{R}$). □

Exercise 33 Prove rigorously that the set $X = \{2, 4, 6, 8, 10\}$ has cardinality 5.

Solution: Again, this is an overkill exercise but the point is to construct an explicit bijection between X and $\mathbb{Z}_5$ and prove that it is a bijection. Indeed, consider the map $f : X \to \mathbb{Z}_5$ given by

$f(x) = x \pmod 5$. This is a bijection. Indeed, we can just compute the images of the elements of X and conclude there is a one-to-one correspondence between the inputs and outputs:

$$\begin{aligned} f(2) &= 2, \\ f(4) &= 4, \\ f(6) &= 1, \\ f(8) &= 3, \\ f(10) &= 0. \end{aligned}$$

Clearly, it is both one-to-one and onto, so it is a bijection. □

Exercise 34 Prove that $\mathbb{Z}$ is countable by finding a bijection $f : \mathbb{Z}^+ \to \mathbb{Z}$.

Solution: Let's define the map $f : \mathbb{Z}^+ \to \mathbb{Z}$ by

$$f(n) = \begin{cases} \dfrac{n}{2} & \text{if } n \text{ is even} \\ \dfrac{-n+1}{2} & \text{if } n \text{ is odd} \end{cases}.$$

It is clear the first case sends every positive **even** integer to the positive integers in a bijective way and it is clear the second case sends every positive **odd** integer to the non-positive (i.e. negative and zero) integers in a bijective way. Moreover, there is no overlap, so the resulting map f is certainly a bijection. □

Note: This is actually a special case of Proposition 9.31 where we are in part (iii) of the proof. Indeed, $\mathbb{Z} = \mathbb{Z}^+ \cup -\mathbb{N}$, where we have used the notation $-\mathbb{N} := \{0, -1, -2, ...\}$. This is a disjoint union so we get a bijection that looks like h in the aforementioned proof.

Exercise 35 Prove Theorem 9.35.

[**Hint:** Consider the function $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Q}$ which sends the pair (a, b) either to the rational number $a/2b$ or to the number 0 in the case that $b = 0$. Argue that this is surjective and use Corollary 9.33 and Proposition 9.34 to complete your proof.]

Solution: Per the hint, the map $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Q}$ is surjective. Indeed, any rational $m/n \in \mathbb{Q}$ can be written as $f(m, n/2)$. Note that it is certainly **not** injective because $f(x, 0) = 0 = f(y, 0)$ for all $x, y \in \mathbb{Z}$. Because $\mathbb{Z}$ is countable (proved in Exercise 34), it follows from Corollary 9.33 that

$\mathbb{Z} \times \mathbb{Z}$ is countable. Then, Proposition 9.34(iii) tells us that $\mathbb{Q}$ being countable is equivalent to having a surjection from a countable set ($\mathbb{Z} \times \mathbb{Z}$ in this case) to it, which we just proved we have. Consequently, $\mathbb{Q}$ is countable. □

> **Exercise 36** Prove that $\mathbb{R} \setminus \mathbb{Z}$ is uncountable. Can your proof also work with $\mathbb{R} \setminus \mathbb{Q}$?

Solution: Assume to the contrary that $\mathbb{R} \setminus \mathbb{Z}$ is countable. We know that the union of countable sets is countable (Proposition 9.31), meaning that $(\mathbb{R} \setminus \mathbb{Z}) \cup \mathbb{Z} = \mathbb{R}$ is countable, a contradiction to Theorem 9.38. The proof is practically identical for $\mathbb{R} \setminus \mathbb{Q}$. □

> **Exercise 37** Suggest the tightest possible bounds on the interval $(-2, 4]$. Considering this and Example 10.4, conjecture what the tightest bounds are on the following intervals, where $x, y \in \mathbb{R}$ are some real numbers.
>
> (i) $[x, y]$.
> (ii) $[x, y)$.
> (iii) $(x, y]$.
> (iv) (x, y).

Solution: The tightest possible bounds on the interval $(-2, 4]$ are -2 (lower bound) and 4 (upper bound). In general, it doesn't matter if the interval is open or closed or half-open/half-closed; the tightest bounds on the intervals (i)–(iv) are always x (lower bound) and y (upper bound). □

> **Exercise 38** Determine the supremum and infimum of $A = \{1/p - 2/q : p, q \in \mathbb{Z}^+\}$.

Solution: We will show that $\sup(A) = 1$. Indeed, recall we need to show that it is both an upper bound and no other upper bound is smaller than it. Let $x \in A$, meaning there exist $p, q \in \mathbb{Z}^+$ such that $x = 1/p - 2/q$. But now, since $q > 0$, it follows that $-2/q < 0$ and because $1/p \leq 1$, it means that $x < 1$, so 1 is an upper bound on A. Next, let $K < 1$, meaning that $1 - K > 0$ and so $2/(1 - K) > 0$. By the Archimedean Property of $\mathbb{R}$, there exists $q \in \mathbb{Z}^+$ such that $q > 2/(1 - K)$. This implies that $0 < 2/q < 1 - K$, which also means $1 - 2/q > K$. Hence, if we take $p = 1$, we have found an element of A that is larger than K, so K is not an upper bound on A.

We could prove that $\inf(A) = -2$, but the proof is very similar to the above so it omitted. □

Note: I didn't ask for a proof in Exercise 38 so don't worry if you didn't do one. Really, I just wanted you to test your intuition to learn if you can 'see' what the supremum/infimum should be just by looking at a set.

Exercise 39 Determine if there is an equivalent Axiom of Completeness for $\mathbb{Q}$, that is does every subset of rational numbers with an upper bound have a supremum in the rationals?

[**Hint:** The key part of the statement is that the supremum is in the **rationals**.]

Solution: There is **not** an equivalent to the Axiom of Completeness for the rations. Indeed, if we consider $A = \{r \in \mathbb{Q} : r^2 < 2\} \subseteq \mathbb{R}$, because $\mathbb{Q}$ is dense(ly-ordered), we will always be able to improve on the upper bound on A. There is no *least* upper bound in $\mathbb{Q}$ because $\sqrt{2} \notin \mathbb{Q}$. □

Exercise 40 A subset $A \subseteq \mathbb{R}$ is said to have the *predecessor property* if $a \in A$ implies that $a - 1 \in A$. Show that any subset with the predecessor property is unbounded below.

[**Hint:** Perform a proof by contradiction using Proposition 10.9.]

Solution: Assume to the contrary that A is bounded below. By Proposition 10.9, the number $L := \inf(A)$ exists. It is clear that $L + 1$ is **not** a lower bound on A (since $L + 1 > L$), meaning there exists $a \in A$ such that $a < L + 1$. However, this is equivalent to $a - 1 < L$ and, because A has the predecessor property, we know that $a - 1 \in A$, so L is **not** a lower bound, a contradiction to it being the infimum. □