1 Logic

1.1 Negations

The *negation* of a statement is a statement that is true exactly when the original statement is false.

Exercise 1.1. Negate the sentence "Every horse is red." Try to phrase your answer in a shortest possible minimally awkward English sentence that doesn't include the phrases 'it is not true that...' or 'it is not the case that', etc...

To formalize what's going on above, it's convenient to adopt some mathematical shorthand. Let's write $\forall$ for 'for all', $\exists$ for 'there exists', and denote the negation of a statement P by *not* P. Then

$$not\,(\forall X, P(X)) = \exists X \text{ such that } not\,P(X), \tag{1}$$

$$not\,(\exists X \text{ such that } P(X)) = \forall X, not\,P(X). \tag{2}$$

Here, we write $P(X)$ to indicate that the P is a statement that accepts an input X, and the truth of P depends on what X is. For example, P could be the statement *"X is a politician"*, so then P(Kamala Harris) is true, while P(my cat) is false. Note that the statement "Every horse is red." can be rewritten as "For all horses H, H is red." Using (1) above, check that your answer in Exercise 1.1 is correct.

Exercise 1.2. Negate the sentence "In every war there is a hero that does not die." *Hint: this is a 'for all' statement that has a "there exists" statement inside.*

Remark (Vacuous truth)**.** Is the following statement true?

> "All the people on Jupiter are friendly."

Well, the negation would be "There is a person on Jupiter that's not friendly," which is certainly false since there are no people on Jupiter. So, the statement above should be true! This is an example of what's called *vacuous truth*, where a statement about some kind of object is automatically true because there are no such objects, so you can't find a counterexample. In some sense, it's a mathematical convention to say that the negation of "$\forall X, P(X)$" should be "$\exists X$ such that $P(X)$" even when there are no X's, but it makes everything simpler to interpret things in this way.

Similarly, let's write $P \wedge Q$ for P *and* Q, and write $P \vee Q$ for P *or* Q. So, $P \wedge Q$ is true when *both* P and Q are true, while $P \vee Q$ is true if *at least one* of the two is true. This interpretation of 'or' is typical in math; the statement P *or* Q *but not both* is a different thing, usually referred to as *exclusive or*.

Fact 1.3. $not\,(P \wedge Q) = (not\,P) \vee (not\,Q)$

Proof. To prove such a statement, we check for each pair of truth values for P and Q whether the two sides of the equation have the same truth value. For instance, if both P and Q are true, then both sides of the equation are false. One way to organize this kind of case by case analysis is with a *truth table*. For example,

P	Q	$P \wedge Q$	$(not\,P) \vee (not\,Q)$
T	T	T	F
T	F	F	T
F	T	F	T
F	F	F	T

indicates the possible truth values (T or F) of the statements $P \wedge Q$ and $(not\,P) \vee (not\,Q)$, depending on whether P, Q are true or false. Since these values are negations of each other, the Fact follows.

In general, some words of explanation might be necessary in order to convince the reader of the correct value for a given entry. For instance, if P is false and Q is true, then $not\,P$ is true and $not\,Q$ is false, so at least one of the two is true, i.e. $(not\,P) \vee (not\,Q)$ is true. The justifications of the other entries above are similar in length. □

Exercise 1.4. Write a truth table showing that $not\,(P \vee Q) = not\,P \wedge not\,Q$.

Exercise 1.5. Negate the following statements, using the rules discussed above.

(a) For every horse that jumps higher than the holy rabbit, there exists a lion that runs faster than the ugly goat and wants to eat that horse.

(b) There exists a real number x such that for all real numbers y, we have $x \geq y$.

1.2 Implications

We write $P \implies Q$, read P *implies* Q, if whenever P is true, Q is true. Some other ways to describe this situation are Q, *if* P, and P *only if* Q. As a more concrete example, the following English sentences all have the same meaning:

(a) The fact that she is a lawyer implies that she passed the bar exam.

(b) If she is a lawyer, then she passed the bar exam.

(c) She passed the bar exam, if she's a lawyer.

(d) She is a lawyer only if she passed the bar exam.

The fact that the last sentence has the same meaning as the others can be confusing, partly because it's a grammatical structure that isn't used so often in English. However, one can reformulate it as "The only way that she can be a lawyer is if she passed the bar exam," which makes it more clearly the same as the others.

We write $P \iff Q$, read 'P *if and only if* Q' when $P \implies Q$ and $Q \implies P$, i.e. when P is true exactly when Q is true.

Remark (Vacuous truth)**.** We talked about vacuous truth of 'for all' statements in the previous section. A similar phenomenon arises when you have a statement P that's always false and you ask whether the statement '$P \implies Q$' is true. Really, you can interpret this as a 'for all' statement: '$P \implies Q$' is the same as saying 'for all possible conditions in which P is true, Q is also true.' So, if there are no conditions where P is true, the statement $P \implies Q$ is called *vacuously true*.

Exercise 1.6. Write a truth table that shows that $not\,(P \implies Q) = P \wedge (not\,Q)$.

Exercise 1.7. Negate the sentence "If it's raining outside, then either I will bring an umbrella, or if my raincoat is back from the cleaners, I will wear it."

Exercise 1.8. Suppose I am the coach of our dodgeball team and you all are the players. I tell you "If we win tonight, then I will buy you pizza tomorrow." When can you rightly claim to have been lied to?

The **converse** of an implication $A \implies B$ is $B \implies A$, while the **contrapositive** of $A \implies B$ is $not\,B \implies not\,A$.

Exercise 1.9. Provide an example of a true statement whose converse is false.

Exercise 1.10. Find the contrapositive of the following statements:

(a) If n is an even natural number, then $n + 1$ is an odd natural number.

(b) If it rains today, then I bring my umbrella.

Theorem 1.11. *Assume A and B are statements. Then $A \implies B$ if and only if $not\,B \implies not\,A$. That is, an implication is equivalent to its contrapositive.*

Proof. Here is the relevant truth table.

A	B	$A \implies B$	$not\, B \implies not\, A$
T	T	T	T
T	F	F	F
F	T	T	T
F	F	T	T

It should also sort of make sense that $A \implies B$ means that 'B is true whenever A is'. So, the only way that B can be false is if A is false. □

Exercise 1.12. Let A represent "6 is an even number" and B represent "6 is a multiple of 4." Express each of the following in ordinary English sentences and state whether the statement is true or false.

(a) $not\, A$

(b) $A \wedge B$

(c) $A \vee B$

(d) $(not\, A) \vee B$

(e) $A \wedge (not\, B)$

(f) $A \implies B$

(g) $B \implies A$

(h) The converse of $(not\, B) \implies A$

(i) The contrapositive of $A \implies B$

The upshot of Theorem 1.11 is that if you want to prove a conditional proposition, you can prove its contrapositive instead. For instance, an integer n is *even* if $n = 2k$ for some $k \in \mathbb{Z}$, and is *odd* otherwise. Prove the following using the contrapositive, being very literal and careful to use the definition above.

Exercise 1.13. Assume x and y are integers. If xy is odd, then both x and y are odd.

Exercise 1.14. Let $x, y \in \mathbb{R}$. Show that if $\forall \epsilon > 0$, we have $|x - y| < \epsilon$, then $x = y$.

Although we will use the implication symbols $\implies$, $\iff$, $\impliedby$, we mostly won't use the symbols $not, \wedge, \vee$ anymore. These are only used in logic texts, and we'll just write 'not' into English sentences, and write 'or' and 'and' instead of $\wedge, \vee$.

1.3 The importance of definitions

Exercise 1.15. Pair up with a friend or divide into groups. Try to write down a precise definition of one of the following geometric objects. Then have your friend or another group try to challenge your definition by coming up with either an example of some object that fits your definition literally, but isn't the object you're describing, or an example of the requested object that doesn't fit your definition.

(a) a circle,

(b) a line.

If you finish these, you could also try defining a 'polygon'!

2 Sets

Loosely, a *set* is a collection of objects, called its *elements*. If S is a set, we write $x \in S$ if x is an element of S. Here, $\in$ can be read as 'is in'. Sets are often presented in one of the following forms:

- A complete list of its elements: the set $S = \{1, 2, 3, 4, 5\}$ contains precisely the five smallest positive integers.
- A list of some of its elements, with ellipses to indicate unnamed elements: e.g., the set $S = \{3, 4, 5, \ldots, 100\}$ contains the positive integers from 3 to 100, including 6 through 99, even though these latter are not explicitly written. We also have the familiar examples of the *natural numbers*

  $$\mathbb{N} := \{1, 2, 3, \ldots\}$$

  and the *integers*

  $$\mathbb{Z} := \{\ldots, -2, -1, 0, 1, 2, \ldots\},$$

  as well as the set of all real numbers $\mathbb{R}$ and the set of rational numbers $\mathbb{Q}$ (i.e. the set of all integer fractions p/q), which we'll assume you have seen before.

  In this course, you'll often see ':=' written instead of '=' when we want to emphasize that something is a definition*, rather than an equality of previously defined objects.*
- By specifying a rule that picks out certain elements of a larger set, and puts them into a new set. The notation we use is:

  { things in the bigger set | conditions they have to satisfy to be in the subset }.

  As an example, we can write the set of even natural numbers as

  $$\{x \in \mathbb{N} \mid x = 2k \text{ for some } k \in \mathbb{N}\}.$$

  Similarly, the set of all real numbers whose squares are less than 3 is written

  $$\{x \in \mathbb{R} \mid x^2 < 3\}.$$
- An alternative way to describe sets is via a rule (a 'function') that produces certain elements in a set from others. Such a description is written as

  { rule producing something from certain elements | those elements }.

For example, the even numbers are exactly those number that are doubles of integers, so the set of even numbers can be written as

$$\{2k \mid k \in \mathbb{Z}\}.$$

We say two sets A and B are equal if they contain precisely the same elements, that is, if $x \in A$ if and only if $x \in B$.

Definition 2.1. The *empty set* is the set with no elements, and it is denoted $\emptyset$.

Exercise 2.2. Is it true that every element of the empty set is a whistling, flying purple cow?

2.1 Subsets of sets

We say that a set A is a *subset* of a set B, written $A \subset B$, if every element of A is also an element of B, that is, if $x \in A$, then $x \in B$. Note that $A = B$ if and only if $A \subset B$ and $B \subset A$.

Exercise 2.3. How many subsets does the empty set have?

Exercise 2.4. Let $A = \{1, \{2\}\}$. Is $1 \in A$? Is $2 \in A$? Is $\{1\} \subset A$? Is $\{2\} \subset A$? Is $1 \subset A$? Is $\{1\} \in A$? Is $\{2\} \in A$? Is $\{\{2\}\} \subset A$? Explain.

Definition 2.5. Let A be a set. The *power set* of A is the sets of all subsets of A and is denoted $\mathcal{P}(A)$. That is, $\mathcal{P}(A) = \{B \mid B \subset A\}$.

Exercise 2.6. Let $A = \{1, 2, 3\}$. Identify $\mathcal{P}(A)$ by explicitly listing its elements.

Exercise 2.7. Suppose that A is a set with n elements. How many elements does $\mathcal{P}(A)$ have? *An informal argument is fine. As a hint, to make a subset, you have to decide for each element of A whether to put it in the subset or not.*

For each non-negative integer n, let $\binom{n}{k}$ denote the number of subsets of $\{1, \ldots, n\}$ of size k. If $k < 0$ or $k > n$ then let $\binom{n}{k} = 0$. Here, $\binom{n}{k}$ is pronounced n *choose* k.

Exercise 2.8. Calculate $\binom{4}{2}$ by listing all the 2-element subsets of $\{1, 2, 3, 4\}$.

Exercise 2.9. Show that $\binom{n}{k} = \binom{n}{n-k}$ for all integers $0 \leq k \leq n$.

These numbers also come up naturally when you expand out powers of sums, via the following expression, which is called the *binomial formula*[1]

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

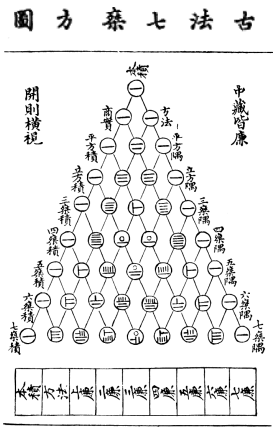
To see why this is true, note that when you multiply together n copies of $(x+y)$, and expand it out, each term of the result is created by selecting either x or y from each $(x+y)$, and then multiplying all your selections together. The number of $x^k y^{n-k}$ terms is the number of ways to select k of the x's, which is $\binom{n}{k}$. Accordingly, the numbers $\binom{n}{k}$ are often called *binomial coefficients*.

Exercise 2.10. If we plug in $x = y = 1$ above, we get $2^n = \sum_{k=0}^{n} \binom{n}{k}$. Explain why this latter equality is true using your solution to Exercise 8.14.

Exercise 2.11. Show that $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$ for $1 \leq k \leq n$. *Hint: take a subset of* $\{1, \ldots, n\}$ *and break into cases depending on whether it contains* n *or not.*

The preious exercise gives a nice way of calculating $\binom{n}{k}$, via *Pascal's triangle*, pictured below in arabic numerals and in Chinese rod numerals. The triangle was actually studied centuries before Pascal by various people around the world, but most western texts disregard this fact and still call it Pascal's triangle.

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1
1 7 21 35 35 21 7 1



Exercise 2.12. Figure out what the above triangle has to do with the numbers $\binom{n}{k}$, and prove your answer. Then write out the next row.

[1]Loosely, the expression $x + y$ is a certain type of combination of two things that is called a 'binomial' (literally: two names) which is why the above is called the binomial formula.

If n is a natural number, n *factorial* is the number

$$n! = n \cdot (n-1) \cdots 1.$$

For example, $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$. Note that $n!$ is exactly the number of *permutations* of $1, 2, 3, \ldots, n$, i.e. the. number of ways to list these numbers in some order. For example, $3! = 6$ and there are six permutations of $1, 2, 3$:

$$123, 132, 213, 231, 312, 321.$$

The reason why $n!$ counts permutations is that you have n choices for the first element, $(n-1)$-choices for the second, etc...

Exercise 2.13. Show that $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. *Hint: put the denominator on the other side, and construct a permutation of* $1, \ldots, n$ *in multiple steps, starting by selecting a* k*-element subset.*

Exercise 2.14. Marcello hosts a party with 10 people, himself included. There is a toast and everyone clinks glasses. How many clinks are heard?

Exercise 2.15. Show that $\sum_{k=0}^{n} \binom{k}{i} = \binom{n+1}{i+1}$ for all integers $n \geq 0$ and $i \geq 0$.

2.2 Unions, Intersections and Products

Definition 2.16. Let A and B be two sets. The *union* of A and B is the set

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

Definition 2.17. Let A and B be two sets. The *intersection* of A and B is the set

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

Theorem 2.18 (Distribution of Union and Intersection)**.** *If A, B, and C are sets,*

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

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

Exercise 2.19. Prove part (a) above.

Definition 2.20. Two sets A and B are *disjoint* if $A \cap B = \emptyset$.

Exercise 2.21. Can a set be disjoint from itself?

Definition 2.22. Let A and B be two sets. The *difference* of A and B is the set

$$A \setminus B = \{x \in A \mid x \notin B\}.$$

When $B \subset A$, the set $A \setminus B$ is also called the *complement* of B in A.

Theorem 2.23. *(DeMorgan's Laws) Let X be a set, and let $A, B \subset X$. Then:*

(a) $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$

(b) $X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B)$

Exercise 2.24. Prove one of the two parts of DeMorgan's laws.

Exercise 2.25. If $X = \{\text{voters}\}$, $A = \{\text{libertarians}\}$ and $B = \{\text{republicans}\}$, what do DeMorgan's laws say?

Exercise 2.26. If S, T are sets, is it true that $\mathcal{P}(S) \cup \mathcal{P}(T) = \mathcal{P}(S \cup T)$?

Definition 2.27. If A, B are sets, their *(Cartesian) product* is the set

$$A \times B := \{(a, b) \mid a \in A, b \in B\}$$

of all 'ordered pairs' of elements from A and B, respectively. We can also take a product of any number of sets, e.g.

$$A_1 \times \cdots \times A_n := \{(a_1, \ldots, a_n) \mid a_i \in A_i \text{ for all i}\}.$$

Elements of $A_1 \times \cdots \times A_n$ are called *n-tuples*. Note that $(A \times B) \times C$ is basically the same as $A \times B \times C$, although technically one should write elements of the former set as $((a, b), c)$ instead of (a, b, c), but we'll ignore this from now on. Also, we define

$$A^n = \overset{n \text{ times}}{A \times \cdots \times A}.$$

Note that if $\mathbb{R}$ is the set of real numbers, $\mathbb{R}^2$ is then the plane. If

$$M = \{Subaru, Honda, Ford, Chevrolet, \ldots\}$$

is the set of all car makes, and

$$C = \{red, blue, \ldots\}$$

is the set of all (named) colors, then $M \times C$ is the set of all pairs of car makes with colors, which is useful set if you are buying a car.

Exercise 2.28. How many elements does $\{1, \ldots, m\} \times \{1, \ldots, n\}$ have?

Exercise 2.29. What is $\mathbb{N} \times \emptyset$?

Exercise 2.30. Show that for all sets A, B, C, D, we have

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D).$$

2.3 Math is broken

Let $\mathcal{S}$ be the set of all sets, and let $\mathcal{R} := \{A \in \mathcal{S} \mid A \notin A\}$. i.e. $\mathcal{R}$ is the set of all sets that don't contain themselves as elements.

Exercise 2.31. (Russel's Paradox) Find a contradiction in mathematics by studying whether $\mathcal{R}$ is an element of $\mathcal{R}$.

A colloquial restatement of this goes as follows. *In Seville, there is a barber who shaves all those men, and only those men, who do not shave themselves. So, who shaves the barber?*

Is math broken? What the exercise indicates is that it is problematic to assume that there is something like the 'set of all sets', and that we need a stricter definition of a set than 'some collection of elements'. Essentially, the way to resolve this is as follows. Starting out with some basic building blocks like the empty set, and say for simplicity $\mathbb{N}$, we only allow ourselves to look at sets constructed from these by natural set operations like unions, products, taking subsets, power sets, etc.... Writing down the rules precisely is pretty subtle, though, so we'll ignore all that and just naively assume that all reasonable expressions we write down do describe sets.

2.4 The halting problem

The following is sort of similar to Russell's paradox, although it doesn't have anything to do with sets.

Let's consider a computer program as a set of instructions that takes as input some text, and does something in response. A program *halts* if it eventually stops running. For example, consider a program **Admirer** that takes a text input X and then prints "I love X" on the screen. Compare this with a program **Stalker** that takes in an input X and then repeatedly writes "I love X" on the screen forever. The first program halts, while the second doesn't. Now each computer program can be itself encoded as text, say, so can be given as an input to another program. The *halting problem* asks if there's a single program that takes as an input a program P (encoded as text) together with another text input X, and prints "Yes" if P halts when fed the input X, and "No" if it doesn't. Here, you should imagine that X is always a string of characters, which P accepts as an input.

Exercise 2.32. Show that there is no such program. *Hint: Hoping for a contradiction, suppose there is a program that solves the halting problem, and call this program* Halt. *Write a program* Paradox *that takes as input a program P and prints "It doesn't*

halt on itself" if P doesn't halt when fed the input P, and prints "It halts!!!" repeatedly forever if P halts when fed the input P. What's the paradox?

3 Induction

Here's a well known formula for the sum of the first n natural numbers.

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}. \tag{3}$$

Let's say we want to prove this. For $n = 1$, the formula is $1 = 1(2)/2 = 1$, which is true. For $n = 2$, the formula is $1 + 2 = 3 = 2(2 + 1)/2$, which is true. How would you prove a statement like this in general? You could keep checking individual cases like this, but you'll never be able to verify case-by-case that the formula holds for all of the infinitely many natural numbers n.

The principle of *induction* says that once you know a particular case of this result, you can prove it for all higher cases by showing that whenever a particular case is true, so is the next one. Namely, note that

$$1 + 2 + \cdots + k + (k + 1) = (1 + 2 + \cdots + k) + (k + 1),$$

so if we already know that (3) holds for the case of $n = k$, this becomes

$$\begin{aligned} &= \frac{k(k+1)}{2} + (n+1) \\ &= \frac{k(k+1) + 2(k+1)}{2} \\ &= \frac{(k+1)(k+2)}{2}, \end{aligned}$$

which is exactly the right hand side of (3) in the case that $n = k + 1$. So, if we know that (3) is true when $n = k$, it also is true when $n = k + 1$. Together with the fact that the $n = 1$ case holds, it should be intuitive that this means that (3) is true for all n. Indeed, it holds for $n = 1$, hence it holds for $n = 2$, hence for $n = 3$, etc...

To formalize this logic, set

$$P(n) := \text{“}1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}\text{ ”}.$$

That is, $P(n)$ is *the statement that the above equality is true*. Note that $P(n)$ is not equal to $1+2+3+\cdots+n$, which is a number. It is a statement. So, saying $P(n) = 6$ makes no sense, but "$P(n)$ is true" and "$P(n)$ is false" do make sense. For instance, $P(3)$ is the statement that $1 + 2 + 3 = 3(4 + 1)/2$.

So, to prove (3), we first noted that $P(1)$ is true. We then showed that whenever $P(k)$ is true, the statement $P(k+1)$ is true, and appealed to our intuition to then say that $P(n)$ is true for all n. Here is a mathematical statement that says "We believe this approach should work!".

The Principle of Mathematical Induction (PMI). *Let $P(1), P(2), P(3), \ldots$ be a sequence of statements, one for each natural number. Assume the following:*

(a) $P(1)$ is true.

(b) For each $k \in \mathbb{N}$, if $P(k)$ is true, then $P(k+1)$ is true.

Then $P(n)$ is true for all $n \in \mathbb{N}$.

Use the PMI to do the following exercises. In doing so, first prove that what you want to show is true when $n = 1$ (this is called the 'base case'), and then show that whenever it is true for a given n, it is also true for $n + 1$ (this is the 'inductive case').

Exercise 3.1. Prove that for all positive integers n,

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(2n+1)(n+1)}{6}.$$

It's sometimes convenient to have the base case in the PMI not be $n = 1$, but something higher. Indeed, suppose that

(a) $P(m)$ is true, and

(b) for all $k \geq m$, if $P(k)$ is true, then $P(k+1)$ is true.

Then the same intuitive argument as above says that $P(n)$ should be true for all $n \geq m$. You can also see this directly from our earlier statement of the PMI, by applying the PMI to the new sequence of statements $Q(n) = P(n + m)$.

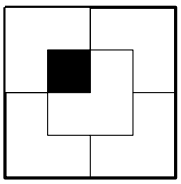
Exercise 3.2. (a) Show by induction that for all $n \geq 4$, we have $2n + 1 \leq 2^n$.

(b) Using induction again, and also part (a), show that for $n \geq 4$ we have $n^2 \leq 2^n$.

If m, n are natural numbers, we say that $m|n$, read m *divides* n, if there is some other natural number k such that $mk = n$.

Exercise 3.3. Show that if $n \in \mathbb{N}$, then $3 \mid 4^n - 1$. *Hint: for the inductive case, start with* $4^{k+1} - 1$ *and try to transform it into something involving* $4^k - 1$ *and things obviously divisible by* 3. *If you get stuck, remember that* $4 = 3 + 1$.

Exercise 3.4. A *triomino* is a 2×2 square with one square removed. Show that for any positive integer n, any $2^n \times 2^n$ checkerboard with one square removed can be tiled by triominos. *Note: in the inductive case, you* start *with a* $2^{n+1} \times 2^{n+1}$ *board with one square removed, and you don't get to pick which square it is. Do not start with a* $2^n \times 2^n$ *board and try to augment it. (Why?)*

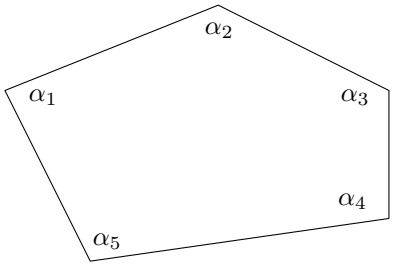


a tiling of a $2^2 \times 2^2$ checkerboard minus a square

Exercise 3.5. Let A be a set with n elements. Show using induction that the power set $\mathcal{P}(A)$ has 2^n elements, by setting $P(n)$ to be the statement *'for every set* A *of* n *elements,* $\mathcal{P}(A)$ *has* 2^n *elements'.*

Be careful with the beginning of the induction step—you need to *start with* a set of $n + 1$ elements.

Exercise 3.6. A standard fact from Euclidean geometry is that the interior angles of a triangle (measured in radians) sum to π. Use this to prove that for $n \geq 3$, the sum of all the interior angles in a convex n-gon is $\pi(n - 2)$.



$\sum_{i=1}^{5} \alpha_i = \pi(5 - 2) = 3\pi.$

Exercise 3.7 (Monochromatic cows)**.** What is wrong with the following proof?

Theorem. *All cows are the same color.*

Proof. We will show by induction that any group of n cows is monochromatic. By showing this is true for all n, we will conclude that all cows are the same color.

Base case. If there is only one cow in a group, then clearly all cows in that group have the same color.

Inductive case. Assume that any group of n cows is monochromatic. Consider a group consisting of $n+1$ cows. First, exclude the last cow and look only at the first n cows; all these are the same color since any group of n is monochromatic. Likewise, exclude the first cow and look only at the last n cows. These too, must also be of the same color. Therefore, the first cow in the group is of the same color as the cows in the middle, who in turn are of the same color as the last cow. Hence the first cow, middle cows, and last cow are all of the same color, and we have proven that our group of $n+1$ cows is monochromatic. By induction, any group of cows is monochromatic, so all cows are the same color. □

Exercise 3.8. There are n lions on an island, all lined up in front of a piece of meat. The meat is tranquilized, and a lion that eats the meat gulps it down in one bite, then falls asleep for a whole day, with the tranquilizer coursing through its blood, so essentially it becomes the tranquilized meat itself for that day. The lions are super-intelligent and all-knowing (for instance, they know the meat is tranquilized, and they know what will happen if they eat it), they will not cooperate or share the meat, they would rather starve than be eaten, and they are ultra-disciplined, so if they do eat then they do it in the order they lined up in, and they will not cut in line or fight. What is going to happen? *Hint: you might want to try the cases of $n = 1, 2, 3$ first in order to get a feel for things. Then formulate your conjecture about what is going to happen, and prove it using induction.*

3.1 The Well Ordering Principle

In some precise sense, the PMI is not something that you can prove using basic logic, without any extra assumptions. Rather, it is an *axiom*, an intuitive statement that we accept as being true. However, whether something is an axiom that you accept or a theorem that you can prove depends on what *other* axioms you accept as true.

Here is another reasonable sounding statement.

The Well Ordering Principle (WOP). *Every nonempty subset of $\mathbb{N}$ has a least element.*

Note that this is not true if we replace $\mathbb{N}$ with $\mathbb{Z}$. Which do you think is more intuitive, the PMI or the WOP?

Exercise 3.9. Show that the WOP implies the PMI. *Hint: start with some property $P(n)$ as given in Theorem 1, and assume the PMI fails for P. Your goal is to define some nonempty set $S \subset \mathbb{N}$, depending on P, and use the fact that S has a least element to get a contradiction.*

Exercise 3.10. Use the PMI to prove the WOP.

So, the two principles are equivalent: both are intuitive, neither one is something you can prove on its own, but you can use each to prove the other.

Sometimes, it's convenient to use the WOP instead of induction. Here are a couple of examples in which that is the case.

Definition 3.11. As above, if $m, n \in \mathbb{N}$, we say that m *divides* n if $n = km$ for some natural number $k \in \mathbb{N}$. Here, m is called a *divisor* of n. A natural number $p \geq 2$ is *prime* if its only divisors are 1 and itself. Any natural number $n \geq 2$ that is not prime can be written as $n = km$ for some $k, m < n$, and is called *composite*.

Exercise 3.12 (The Prime Factorization Theorem)**.** Show that any natural number n can be written as a product

$$n = p_1 \cdots p_k,$$

where $p_1, \ldots, p_k$ are all prime.

Here, if n is itself prime, the above product will only have one term. Also, if $n = 1$, let's say by convention that n is an 'empty product' of no primes, just so we don't have to say $n \geq 2$ in the exercise[2].

Exercise 3.13. A *round robin* is a tournament in which each player p plays each other player q exactly once. We write $p > q$ if player p wins, and $q > p$ otherwise. A *cycle* is a sequence of players $p_1, \ldots, p_k$ such that $p_1 < p_2 < \cdots < p_k < p_1$. The number k is called the *length* of the cycle. Show that in every round robin tournament, if there is a cycle, then there is a cycle of length 3. *Hint: Suppose we are given a tournament that has a cycle. The WOP says that there is a* shortest *cycle $p_1 < \cdots < p_k < p_1$.*

[2]There are good reasons to do this. For instance, you are probably familiar with exponential functions, and the fact that $e^0 = 1$. Well, e^n is the product of n copies of e, so this is saying that the product of no copies of e is 1. This is a good definition, since then the formula $e^n e^m = e^{n+m}$ works even when n or m is zero. Alternatively, when you are defining the exponential function you start out by defining e^n for natural n, as the n-fold product of e's. Since $(e^{1/n})^n$ should equal $e^1 = e$, the only correct definition for $e^{1/n}$ is the n^{th}-root of e. But as $n \to \infty$, we have $\sqrt[n]{e} \to 1$, so if you want the function $x \mapsto e^x$ to be continuous at 0, the only good definition for e^0 is 1.

3.2 Strong Induction

Here is a another variant of Theorem 1 that's similar to the WOP. Try to convince yourself that it's also a reasonable statement. We will refer to it as a Theorem here, because at the end of the section we'll prove it using the PMI.

Theorem 3.14 (Principle of strong induction)**.** *Suppose that for each $n \in \mathbb{N}$, we have a statement $P(n)$, and that the following two properties hold:*

(a) $P(1)$ is true.

(b) For every $k \in \mathbb{N}$, if $P(1), \ldots, P(k)$ are all true, then $P(k+1)$ is true as well.

Then $P(n)$ is true for all $n \in \mathbb{N}$.

Just like with normal induction, when applying strong induction you should feel free to start with your base case being 0 or any natural number m, if your goal is to prove that $P(n)$ is true for all $n \geq m$.

The *Fibonacci sequence* f_n is defined by letting $f_1 = f_2 = 1$ and setting $f_n = f_{n-1} + f_{n-2}$ for $n \geq 3$. The first few terms are $1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, \ldots$

Exercise 3.15. Show that for every $n \in \mathbb{N}$, f_n is even if $3|n$, and odd otherwise. *Hint: for the inductive case, you'll want to use the equation $f_{k+1} = f_k + f_{k-1}$. But this only works if $k \geq 2$. So, prove the claim for $n = 1, 2$ separately as dual base cases, and then do the inductive case starting with $k = 2$, so $k + 1 = 3$.*

A *polynomial* is a function of the form $f(x) = a_n x^n + \cdots + a_1 x + a_0$, where $a_n, \ldots, a_0 \in \mathbb{R}$. By discarding some of the terms, we can assume that the *leading coefficient* a_n of f is nonzero, and then the number n is called the *degree* of f. A real number x is called a *root* of f if $f(x) = 0$.

Theorem 3.16. *Any nonzero degree n polynomial has at most n roots.*

For example, the polynomials $f(x) = x^2 - 2$, $g(x) = x^2$, $h(x) = x^2 + 1$ all have degree 2. The roots of f are $x = \pm\sqrt{2}$, the polynomial g has only $x = 0$ as a root, and h has no roots. In all cases, the number of roots is at most 2, the degree. The only exception is that the zero polynomial $z(x) = 0$ has degree 0, which is at most 2, but *every* real number is a root of z, so z has infinitely many roots.

Exercise 3.17. Prove Theorem 3.16 by filling in the following outline.

Proof Outline. We'll do a strong induction proof. For the base case, consider a nonzero, degree zero polynomial $f(x) = a_0$. ($\cdots$ *insert solution here* $\cdots$)

For the inductive case, suppose that for $n = 1, \ldots, k$, any nonzero degree n polynomial has at most n roots. Now let $f(x) = a_{k+1}x^{k+1} + \cdots + a_1 x + a_0$ be a nonzero degree $k+1$ polynomial, and hoping for a contradiction, assume that $f(x)$ has $k+2$ distinct roots, which we'll call $x_1, \ldots, x_{k+2}$. Consider the polynomial

$$g(x) = f(x) - a_{k+1}(x - x_1)\cdots(x - x_{k+1}).$$

The degree of g is at most k, since ($\cdots$ *insert solution here* $\cdots$). Moreover, g is not the zero polynomial, since ($\cdots$ *insert solution here* $\cdots$).

So, by the inductive hypothesis, g has at most k roots. But ($\cdots$ *insert solution here* $\cdots$) are all roots of g, so this is a contradiction. Hence our initial assumption that f had $k+2$ roots had to be wrong, so f has at most $k+1$ roots as desired. This finishes the inductive case. □

Exercise 3.18. Use strong induction to prove that any $l \in \mathbb{N}$ can be written as

$$n = 2^{a_l} + 2^{a_{l-1}} + \cdots + 2^{a_1}$$

for distinct integers $a_l > a_{l-1} > \cdots > a_1 \geq 0$.

Note that if you then write a string of 1's and 0's where the 1's are in the $a_k, \ldots, a_1$ places, counting from the right, you get the binary expansion of n. For instance, $43 = 2^5 + 2^3 + 2^1 + 2^0$ is 101011 in binary. So, this exercise is showing that any positive integer can be written in binary.

The game of Nim involves two players and two piles of pennies. On each turn, a player removes some non-zero number of pennies from one of the piles. A player loses if on their turn, there are no pennies left in either pile.

Exercise 3.19. Show that if the two piles initially have the same number of pennies, there is a strategy in which the second player can always win!

You should prove using strong induction that for each $n \in \mathbb{N}$, the second player always has a winning strategy in a Nim game in which both players start with n pennies. To get some intuition, try doing the $n = 0, n = 1, n = 2$ cases first.

Finally, let's prove that strong induction works.

Exercise 3.20. Use the PMI to prove Theorem 3.14.

In practice, strong induction is essentially the same thing as the well ordering principle (WOP). To get some intuition for this, try thinking through Exercise 3.12 using strong induction or Exercises 3.18 and 3.15 using the well ordering principle instead of strong induction.

Remark 3.21. Conditions (a) and (b) in the principle of strong induction can actually be combined into the single statement

(c) If $k \in \mathbb{N}$ is a natural number, and $P(j)$ is true for all natural numbers $j < k$, then $P(k)$ is true.

Here, (c) may just look like a rephrasing of (b), and it sort of is, but it also vacuously implies (a). For above, if $k = 1$, then the statement *"$P(j)$ is true for all natural numbers $j < k$"* is always true, since there are no natural numbers less than 1, so the conclusion of (c) in the $k = 1$ case is that $P(1)$ is always true. This can be confusing, but it also explains why some strong induction proofs intuitively don't seem like they need a base case. To get some intuition for this, try to prove The Prime Factorization Theorem (Exercise 3.12) by strong induction. Essentially, the point will be that if you have a given natural number n, either it's prime and you're done, or it factors as the product of two smaller numbers, and those are products of primes by induction. If you think through this, what's really happening is that you're just factoring the factors, and factoring the factors of factors of n, until you end up with primes. It's not clear here where you're really using a base case, and the reason for that is that the base case is folded vacuously into the argument as it is in (c).

4 Number Theory

We will start in now with a bit of number theory. As always, $\mathbb{Z}$ will denote the set of integers. You are allowed to assume that all usual properties of addition, subtraction and multiplication of integers work in the usual ways, and interact as expected with the usual order $<$ on $\mathbb{Z}$. For instance, addition is commutative and multiplication distributes over addition, and if $a \leq b$ then $a + c \leq b + c$ for all c.

4.1 Division with remainder

In high school or earlier, you probably considered statements like *"if you divide 8 by 3, you get 2, but with a remainder of 2."* Here's a theorem that makes sense of this.

Theorem 4.1 (Division with remainder)**.** *If $a, b \in \mathbb{Z}$ and $b > 0$, then there exist unique integers q and r such that $a = bq + r$ and $0 \leq r < b$.*

Intuitively, in the setting above, dividing a by b gives q with a remainder of r.

Exercise 4.2. If $a = 7$ and $b = 314$, find q and r. What about if $a = -11$ and $b = 30$?

The following two exercises complete the proof of Theorem 1.

Exercise 4.3. Given a, b as in Theorem 1, let q the largest integer such that $bq \leq a$. Show that $a = bq + r$ where $0 \leq r < b$. *Hint: Set $r = a - bq$ then try to show it lies in the desired range.*

Exercise 4.4. Suppose q_1, r_1 and q_2, r_2 are integers with $0 \leq r_1, r_2 < b$ and

$$a = bq_1 + r_1 = bq_2 + r_2.$$

Show that $q_1 = q_2$ and $r_1 = r_2$. *Hint: show that $b(q_1 - q_2) = r_2 - r_1$ and then figure out what range of numbers the right side lies in.*

4.2 Greatest common divisors

Recall that if $a, b \in \mathbb{Z}$, then a *divides* b if there is some integer k with $b = ak$, and we call a a *divisor* of b, while b is a *multiple* of a. We write $a|b$ when this is the case.

Definition 4.5. Let $a, b \in \mathbb{Z}$, not both zero. A *common divisor* of a and b is defined to be any integer c such that $c|a$ and $c|b$. The largest common divisor is usually called the *greatest common divisor* of a and b, and is denoted $gcd(a, b)$.

Exercise 4.6. List all the common divisors of 18 and 24, and find $gcd(18, 24)$.

We say a, b are *relatively prime* or *co-prime* if $gcd(a, b) = 1$.

Exercise 4.7. Let f_n be the Fibonacci sequence. Show that f_n and f_{n+1} are relatively prime for all positive integers n.

In Exercise 4.6, you computed the gcd of 18 and 24 by just listing all the common divisors and taking the biggest one. This is not an effective method of computing gcd's of very large numbers though, since it's computationally intensive to find all divisors of a large number. In fact, this computational difficulty is what makes many common encryption schemes work, e.g. those used to encrypt web traffic.

Fortunately, there's a better way to compute gcds. The key is:

Exercise 4.8. If $a, b, q, r \in \mathbb{Z}$ and $a = qb + r$, then $gcd(a, b) = gcd(b, r)$.

How do we use this to compute gcds? Well, if you start with two (say, positive) numbers a, b, you can assume a is the bigger one, and then write $a = qb + r$ using division with remainder. The exercise says $gcd(a, b) = gcd(b, r)$, and this is easier to compute, since r is smaller than a. Moreover, if you then use division with remainder and the exercise *again*, you can perhaps reduce the computation of $gcd(b, r)$ to an even simpler computation, etc...

For example, say we want to compute $gcd(93, 36)$. We write:

$$\begin{aligned}
93 &= 2 \cdot 36 + 21 && \leftarrow \text{ reduces it to computing } gcd(36, 21)\\
36 &= 1 \cdot 21 + 15 && \leftarrow \text{ reduces it to computing } gcd(21, 15)\\
21 &= 1 \cdot 15 + 6 && \leftarrow \text{ reduces it to computing } gcd(15, 6)\\
15 &= 2 \cdot 6 + 3 && \leftarrow \text{ reduces it to computing } gcd(6, 3)\\
6 &= 2 \cdot 3 + 0 && \leftarrow \text{ says that 3 divides 6, so actually } gcd(6, 3) = 3!
\end{aligned}$$

Tracing through all these steps and applying the exercise each time then shows that $gcd(93, 36) = 3$. This procedure is called the *Euclidean algorithm*, after the Greek mathematician Euclid, who was alive around 300 B.C.E. Here's a formal statement.

Theorem 4.9 (The Euclidean Algorithm)**.** *Let $a, b \in \mathbb{Z}$ be positive integers, not both zero. Then we can apply division with remainder repeatedly to find q_i, r_i as follows:*

$$\begin{aligned}
a &= bq_1 + r_1 && 0 < r_1 < b\\
b &= r_1q_2 + r_2 && 0 < r_2 < r_1\\
r_1 &= r_2q_3 + r_3 && 0 < r_3 < r_2\\
&\vdots\\
r_{k-2} &= r_{k-1}q_k + r_k && 0 < r_k < r_{k-1}\\
r_{k-1} &= r_kq_{k+1}, &&
\end{aligned}$$

and where $r_k = gcd(a, b)$.

Exercise 4.10. Use the Euclidean algorithm to find $gcd(18, 24)$, $gcd(75, -21)$ and $gcd(145, 690)$.

Exercise 4.11. Using Exercise 4.8, write a formal proof that the Euclidean Algorithm works. There's almost a proof above, but why does the algorithm always end in finitely many steps with a remainder of zero?

4.3 Bézout's Identity and $\mathbb{Z}$-linear combinations

The following is one of the most important properties of gcds. We'll use it in the next section to prove certain fundamental facts about prime numbers.

Theorem 4.12 (Bézout's Identity)**.** *If $a, b \in \mathbb{Z}$, not both zero, $\exists x, y \in \mathbb{Z}$ such that*

$$gcd(a, b) = xa + yb.$$

Given $a, b \in \mathbb{Z}$, an expression of the form $xa + yb$, where $x, y \in \mathbb{Z}$, is called a *$\mathbb{Z}$-linear combination* of a, b. The numbers x, y are called the *coefficients* of the linear combination. For instance, the following are $\mathbb{Z}$-linear combinations of 3 and 7:

$$-2 \cdot 3 + 5 \cdot 7 = 36, \qquad 0 \cdot 3 - 1 \cdot 7 = -7, \qquad 100 \cdot 3 + 2 \cdot 7 = 314.$$

Exercise 4.13. Find x, y such that $gcd(21, 27) = x21 + y27$.

Exercise 4.14. Show that any integer $n \geq 8$ can be written as $n = x \cdot 3 + y \cdot 5$ for some nonnegative $x, y \in \mathbb{Z}$. *Hint: use strong induction on n.*

Note that in Exercise 4.14, we're only using nonnegative coefficients. So, the exercise is showing, for instance, that any possible amount n of postage that is at least 8 cents can be made using some combination of 3 cent and 5 cent stamps. In general, though, $\mathbb{Z}$-linear combinations can have negative coefficients.

You can prove Theorem 4.12 using the following exercise, in combination with the Euclidean algorithm.

Exercise 4.15. Suppose that $a, b \in \mathbb{Z}$. If m, n are both $\mathbb{Z}$-linear combinations of a, b and $m = qn + r$, where $q, r \in \mathbb{Z}$, then r is also a $\mathbb{Z}$-linear combination of a, b.

Proof of Theorem 4.12. You can use induction and Exercise 4.15 to say that all the remainders that appear in the Euclidean algorithm are all $\mathbb{Z}$-linear combinations of the original a, b. So, this is the case for the final nonzero remainder, which is $gcd(a, b)$. □

One can use the proof above to *explicitly* find the coefficients x, y such that $xa+yb = gcd(a, b)$. It's almost more confusing to write an explicit procedure for this than to do it: the point is just to go through the Euclidean algorithm, and write each of the remainders that comes up as a $\mathbb{Z}$-linear combination of a, b, using the $\mathbb{Z}$-linear combinations for the previous remainders.

Here's a simple example. To compute $gcd(36, 26)$, we proceed as follows:

$$\begin{aligned} 36 &= 1 \cdot 26 + 10 \\ 26 &= 2 \cdot 10 + 6 \\ 10 &= 1 \cdot 6 + 4 \\ 6 &= 1 \cdot 4 + 2 \\ 4 &= 2 \cdot 2, \end{aligned}$$

so $gcd(36, 26) = 2$. To write 2 as a $\mathbb{Z}$-linear combination, you

$$\begin{aligned} 10 &= 36 - 26 \\ 6 &= 26 - 2 \cdot 10 = 26 - 2(36 - 26) = -2 \cdot 36 + 3 \cdot 26 \\ 4 &= 10 - 6 = (36 - 26) - (-2 \cdot 36 + 3 \cdot 26) = 3 \cdot 36 - 4 \cdot 26 \\ 2 &= 6 - 4 = (-2 \cdot 36 + 3 \cdot 26) - (3 \cdot 36 - 4 \cdot 26) = -5 \cdot 36 + 7 \cdot 26. \end{aligned}$$

Exercise 4.16. Use the Euclidean algorithm to find $gcd(105, 135)$ and x, y such that $x105 + y135 = gcd(105, 135)$.

Here's one corollary of Bézout's Identity. To state it, let $a, b \in \mathbb{Z}$ and let

$$S(a, b) := \{xa + yb \mid x, y \in \mathbb{Z}\}$$

be the set of all $\mathbb{Z}$-linear combinations of a and b.

Exercise 4.17. Show that $S(a, b) = \{n \cdot gcd(a, b) \mid n \in \mathbb{Z}\}$, i.e. $S(a, b)$ is exactly the set of all integer multiples of $gcd(a, b)$.

The following three exercises give an alternative proof of Bézout's Identity using the structure of $S(a, b)$. First, note that as long as a, b aren't both zero, the set $S(a, b)$ contains positive integers, e.g. it contains $\pm a$ and $\pm b$, and one of these four is positive. So, $S(a, b)$ contains a *least* positive integer $d(a, b)$, by the well ordering principle.

Exercise 4.18. Show that $d(a, b)$ is a common divisor of a, b.

Exercise 4.19. Suppose that c is a common divisor of a, b. Show that $c|d(a, b)$, and therefore $c \leq d(a, b)$.

Exercise 4.20. Use Exercises 4.18 and 4.19 to prove Bézout's Identity.

4.4 More primes

Recall that a natural number $p \geq 2$ is *prime* if its only positive divisors are 1 and itself. A natural number $n \geq 2$ is *composite* if n is not prime. Previously, we showed:

Theorem 4.21. *Any natural number n can be written as a product*

$$n = p_1 \cdots p_k,$$

where $p_1, \ldots, p_k$ are all prime.

Often, when we write n as such a product, we say that we have *factored* n. The terms in the product are the *factors*.

Exercise 4.22. If p is prime and p does not divide $a \in \mathbb{N}$, show that $gcd(p, a) = 1$.

Exercise 4.23. Let $a, b, n \in \mathbb{N}$ and assume $gcd(a, n) = 1$ and $n|ab$. Show that $n|b$. Then conclude that if p is prime and $p|ab$, then $p|a$ or $p|b$.

Exercise 4.24. Let p be prime and $a_1, \ldots, a_k$ be natural numbers. If $p|a_1 \cdots a_k$, show that $p|a_i$ for some i. *Hint: do induction on k.*

In particular, if the a_i in Exercise 4.24 are all prime, then we have that $p = a_i$ for some i. Using this, you should now prove:

Theorem 4.25 (The Fundamental Theorem of Arithmetic)**.** *Every integer $n \geq 1$ may be factored into a product of primes in a unique way up to the order of the factors. In other words, if $n = p_1 \cdots p_k = q_1 \cdots q_l$ are two prime factorizations of n, then $k = l$ and we can reorder the q_i's so that $p_1 = q_1, p_2 = q_2, \ldots, p_k = q_k$.*

Hint: it may be useful to do induction on n. The base case $n = 1$ is trivial.

FIX ME

Exercise 4.26. Suppose that $n = p_1^{e_1} \cdots p_k^{e_k}$, where $p_1, \ldots, p_k$ are distinct primes and the exponents are integers $e_i \geq 0$, and let $d \in \mathbb{N}$. Show that $d|n$ if and only if

$$d = p_1^{j_1} \cdots p_k^{j_k}$$

for some $j_i \in \mathbb{Z}$ with $0 \leq j_i \leq e_i$ for all $i = 1, \ldots, k$. *Hint: you should do this in two directions. First, show that if $d = p_1^{j_1} \cdots p_k^{j_k}$, then $d|n$. Then show that any divisor d of n has this form, using the uniqueness of prime factorizations (Theorem 4.25).*

Exercise 4.27. Suppose that $n = p_1^{e_1} \cdots p_k^{e_k}$ and $m = p_1^{f_1} \cdots p_k^{f_k}$, where $p_1, \ldots, p_k$ are distinct primes and $e_i, f_i \geq 0$. If

$$g := p_1^{\min\{e_1, f_1\}} \cdots p_k^{\min\{e_k, f_k\}},$$

show that $g = gcd(m, n)$. *Hint: use the previous problem.*

Here, $\min\{x, y\}$ is just the minimum of x, y, i.e. whichever one is smaller. Note that in the exercise, we can allow e_i or f_i to be zero. This allows us to apply the exercise to *any* pair of natural numbers n, m. For instance, we can write

$$60 = 2^2 \cdot 3^1 \cdot 5^1 \cdot 7^0, \quad 35 = 2^0 \cdot 3^0 \cdot 5^1 \cdot 7^1,$$

and then $gcd(60, 35) = 2^0 \cdot 3^0 \cdot 5^1 \cdot 7^0 = 5$.

A natural number n is a *perfect square* if there is some $a \in \mathbb{N}$ with $a^2 = n$. Here are some examples of perfect squares and their prime factorizations:

$$9 = 3^2, \quad 64 = 2^6 \quad 36 = 2^2 \cdot 3^2, \quad 400 = 2^4 \cdot 5^2.$$

Exercise 4.28. Show that n is a perfect square if and only if every prime factor occurs an even number of times in the (essentially unique) prime factorization of n.

A real number x is defined to be *rational* if there exist integers p and q such that $x = p/q$ and *irrational* otherwise.

Exercise 4.29. Show that if $n \in \mathbb{N}$ and $\sqrt{n}$ is rational, then n is a perfect square.

So in particular, $\sqrt{2}$ is irrational.

Exercise 4.30. Show that there are infinitely many primes. *Hint: hoping for a contradiction, suppose the only primes are $p_1, \ldots, p_k$ and consider $n = p_1 \cdots p_k + 1$.*

Exercise 4.31. Show that there are infinitely many primes of the form $4n + 3$. *Hint: you might find it useful to show that the product of two numbers of the form $4n + 1$ are also of that form. For example, $5, 9$ and $5 \cdot 9 = 45$ are all of the form $4n + 1$, for $n = 1, 2, 11$, respectively. Note that all primes except 2 are odd, and every odd prime is either of the form $4n + 1$ or $4n + 3$ for some n.*

Finally, try to digest the statement of the following theorem, for inspiration. I don't expect you to prove it, although if you do some sort of medal is in order.

Theorem 4.32 (The prime number theorem)**.** *Given a positive number x, let $\pi(x)$ be the number of primes that are less than or equal to x. Then*

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\ln x} = 1.$$

What does this mean? Well, if x is really large, this means that $\frac{\pi(x)}{x} \approx \frac{1}{\ln(x)}$. Here, $\frac{\pi(x)}{x}$ is the *percentage* of the numbers $1, \ldots, x$ that are prime. For instance, it turns out that there are 50,847,534 primes that are less than a billion. The percentage of numbers less than a billion that are prime is then

$$50,847,534/1,000,000,000 = .05847534,$$

while $1/\ln(1,000,000,000) = 0.04825494243$, which is pretty close. Note that as $x \to \infty$, the percentage of numbers less than or equal to x that are prime gets smaller and smaller. This is because primes are becoming sparser as you look at bigger numbers, since there are more numbers available to be factors.

5 Functions

A *function* is a rule that assigns to every element of a set A, an element of another set B. We write functions in the form $f : A \longrightarrow B$. Here, the element of B that the function assigns to an element $a \in A$ is written $f(a)$.

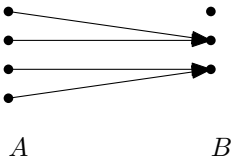
Definition 5.1 (Domain, codomain and range)**.** The *domain* of f is A, and its *codomain* is B. The *image*, or *range*, of f is the set

$$f(A) = \{f(a) \mid a \in A\}$$

Exercise 5.2. Find the domain, codomain and image of the function

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \quad f(x) = x^2 - 5.$$

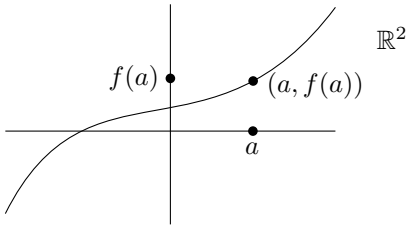
Here are two ways to picture a function. First, you can use dots to represent points in A and in B, and draw an arrow from each $a \in A$ to $f(a) \in B$.



Another way is to look at the function's graph. Namely, if $f : A \longrightarrow B$ is a function, then we define $graph(f)$ to be the subset of the product $A \times B$ given by

$$graph(f) := \{(a, f(a)) \mid a \in A\} \subset A \times B.$$

If $f : \mathbb{R} \longrightarrow \mathbb{R}$, then $graph(f)$ is the familiar subset of $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ you might remember from calculus.



However, you could also draw a picture of the graph of a function

$$f : A \longrightarrow B$$

of finite sets, say, by drawing the elements of the domain as a horizontal sequence of dots, say arranged along what would be the x-axis in a picture of $\mathbb{R}^2$, and drawing elements of the codomain vertically along what would be the y-axis, and then drawing dots at each value $(x, f(x))$, where $x \in A$.

Definition 5.3 (Surjective, injective and bijective)**.** A function $f : A \to B$ is *surjective*, or *onto*, if $f(A) = B$. It is *injective*, or *one-to-one* if for all $a_1, a_2 \in A$, we have $f(a_1) = f(a_2)$ only if $a_1 = a_2$. It is *bijective* if it is surjective and injective.

For example, the function drawn with arrows and dots above is neither injective nor surjective. The function $f : \mathbb{R} \longrightarrow [-5, \infty)$, $f(x) = x^2 - 5$ is surjective but not injective.

Exercise 5.4. Determine whether the following are injective, surjective or bijective.

(a) $f : \mathbb{Z} \to \mathbb{Z} \times \mathbb{Z},\ f(n) = (n, n)$

(b) $f : \mathbb{Z} \to \mathbb{Z},\ f(n) = \begin{cases} n & n \geq 0 \\ n + 5 & n < 0. \end{cases}$

(c) $g : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z},\ g(m, n) = m + n$

(d) $h : \mathbb{R} \times (\mathbb{R} \setminus \{0\}) \to \mathbb{R},\ h(x, y) = \frac{x}{y}$

Definition 5.5. The *composition* of $f : A \longrightarrow B$ and $g : B \longrightarrow C$ is the function

$$g \circ f : A \longrightarrow C,\ g \circ f(a) = g(f(a)).$$

Exercise 5.6. Write nice-looking formulas for the compositions $f \circ g$ and $g \circ f$, where

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R},\ f(x, y) = \frac{x}{y^2 + 1},\quad g : \mathbb{R} \longrightarrow \mathbb{R}^2,\ g(x) = \left(x^2, x\right).$$

Exercise 5.7. (a) Show that the composition of two injective functions is injective.

(b) Show that the composition of two surjective functions is surjective.

(c) Show that if $g \circ f$ is surjective, so is g.

(d) If $g \circ f$ is surjective, does f have to be surjective? (Either prove it or give a counterexample.)

Definition 5.8 (Image and preimage of subsets)**.** Let $f : A \to B$ be a function. Let $X \subseteq A$. Then the *image of X under f* is

$$f(X) = \{f(x) \mid x \in X\}$$

Let $Y \subseteq B$. Then the *preimage of Y under f* is

$$f^{-1}(Y) = \{a \in A \mid f(a) \in Y\}$$

Exercise 5.9. Identify the following images and preimages.

(a) $f([-1, 6])$, where $f : \mathbb{R} \longrightarrow \mathbb{R},\ f(x) = x^2$,

(b) $f^{-1}([4, 9])$, where $f : \mathbb{R} \longrightarrow \mathbb{R},\ f(x) = x^2$,

(c) $g^{-1}([1, 2] \times [3, 5])$, where $g : \mathbb{R} \longrightarrow \mathbb{R}^2,\ g(x) = (x, x^2)$.

Exercise 5.10. Let $f : A \to B$ be a function. In each of the following, decide if the statement is true. If so, prove it. If not, give an explicit counterexample and then try to prove it under an additional assumption that f is either surjective or injective.

(a) $f^{-1}(f(X)) = X$ for all $X \subseteq A$.

(b) $f(f^{-1}(Y)) = Y$ for all $Y \subseteq B$.

(c) $f(X_1 \cap X_2) = f(X_1) \cap f(X_2)$ for all $X_1, X_2 \subseteq A$.

(d) $f^{-1}(Y_1 \cap Y_2) = f^{-1}(Y_1) \cap f^{-1}(Y_2)$ for all $Y_1, Y_2 \subseteq B$.

Definition 5.11. Let $f : A \longrightarrow B$ be a bijection. The *inverse* of f is the function

$$f^{-1} : B \longrightarrow A,$$

where $f^{-1}(b)$ is the unique element $a \in A$ such that $f(a) = b$.

This is a slight abuse of notation, since we previously used the symbol f^{-1} to denote the preimage. However, it should always be clear from context whether we are referring to the preimage or to the inverse function.

Exercise 5.12. Show the following are bijections, and find the inverse.

(a) $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = (5x - 2)/12$.

(b) $g : \mathbb{N} \cup \{0\} \longrightarrow \mathbb{Z},\ g(n) = \begin{cases} 0 & n = 0 \\ \frac{n}{2} & n \text{ is even} \\ -\frac{n+1}{2} & n \text{ is odd.} \end{cases}$

(c) $h : \mathbb{R} \longrightarrow \{(x, y) \in \mathbb{R}^2 \mid y = 2x\},\ h(x) = (x, 2x)$.

Exercise 5.13. What is a function, really? That is, say that you're comfortable with all the stuff from our set theory sheet. Can you give a definition of a function using only the language of set theory, without saying vague things like 'rule that assigns'?

6 Equivalence Relations

Let A be a set. For every pair $a, b \in A$, let's write either $a \sim b$ or $a \not\sim b$, read as *'a is related to b'* and *'a is not related to b'*, respectively. An arbitrary way to do this for each pair is called a *relation* on A.

Example 6.1. If $a, b \in \mathbb{Z}$, declare $a \sim b$ if $a \leq b$, and $a \not\sim b$ otherwise.

A relation $\sim$ is an *equivalence relation* if the following properties are satisfied.

(1) for all $a \in A$, we have $a \sim a$ (reflexivity),

(2) for all $a, b \in A$, if $a \sim b$ then $b \sim a$ (symmetry),

(3) for all $a, b, c \in A$, if $a \sim b$ and $b \sim c$ then $a \sim c$ (transitivity).

If $\sim$ is an equivalence relation, we often read $a \sim b$ as 'a is equivalent to b', rather than using the word 'related', but either works. Note that the relation $a \sim b$ if $a \leq b$ is not an equivalence relation on $\mathbb{Z}$: we have $2 \leq 3$ but $3 \not\leq 2$, so the relation isn't symmetric.

Example 6.2 (Equality)**.** If A is any set, define $a \sim b$ if $a = b$. This is an equivalence relation on A. Namely, $a = a$ for all $a \in A$, so it's reflexive. If $a = b$, then $b = a$, so it's symmetric. If $a = b$ and $b = c$, then $a = c$, so it's transitive.

Exercise 6.3 (The trivial equivalence relation)**.** If A is any set, define $a \sim b$ for all $a, b \in A$. Show that $\sim$ is an equivalence relation on A.

Really, the whole point of equivalence relations is to have some notion of 'similarity' between elements of A that behaves kind of like equality.

Exercise 6.4. Are the following equivalence relations?

(a) Let L be the set of lines on the plane. For $a, b \in L$ let $a \sim b$ if a and b are parallel. *Here, lines are parallel if they are disjoint or identical. It's ok to write an informal argument for this one, just say why you think it is or isn't an equivalence relation.*

(b) Let $\mathcal{P}$ be the set of polygons in the plane. Set $P \sim Q$ if P is congruent to Q.

(c) For $a, b \in \mathbb{Z}$, let $a \sim b$ if $a - b$ is odd.

(d) Fix $n \in \mathbb{Z}$, and for $a, b \in \mathbb{Z}$ let $a \sim_n b$ if $a - b$ is a multiple of n.

(e) Let X be a set, and consider the relation $\sim$ on the power set $\mathcal{P}(X)$, where $A \sim B$ if $A \cap B \neq \emptyset$.

(f) Let $\mathcal{F}$ be the set of all functions $f : \mathbb{R} \longrightarrow \mathbb{R}$, and define $f \sim g$ if there is some finite subset $S \subset \mathbb{R}$ such that $f(x) = g(x)$ for all $x \notin S$.

Here is an especially important example. Suppose that $f : A \longrightarrow X$ is a function. Let's define a relation $\sim_f$ on A by declaring $a \sim_f b$ when $f(a) = f(b)$.

Exercise 6.5. Show that $\sim_f$ is an equivalence relation.

For example, if we take $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $f(x, y) = \sqrt{x^2 + y^2}$ then $a, b \in \mathbb{R}^2$ are related exactly when they have the same length. In general, this exercise says that whenever we define a relation by saying '*$a \sim b$ if a and b have the same blarhgh*', it will automatically be an equivalence relation, since we can take

$$blarhgh : A \longrightarrow \{ \text{ possible values of blarhgh } \}.$$

as our function in Exercise 6.5. Do any of the examples in Exercise 6.4 arise like this?

Exercise 6.6. Is the following theorem and proof correct?

Theorem 6.7. *Suppose $\sim$ is a relation on a set A that's symmetric and transitive. Then $\sim$ is reflexive.*

Proof. Let $a \in A$. Pick some $b \in A$ with $a \sim b$. Then $b \sim a$, by symmetry. Since $a \sim b$ and $b \sim a$, by transitivity we have $a \sim a$. □

Exercise 6.8. Suppose that $\sim_1, \sim_2$ are two equivalence relations on A. Define a new relation $\sim$ by setting $a \sim b$ when $a \sim_1 b$ and $a \sim_2 b$. Show that $\sim$ is an equivalence relation. What happens if you use 'or' instead of 'and' in the construction?

Exercise 6.9. Can you give a rigorous interpretation of an equivalence relation in terms of set theory? $\sim$ should be defined to be a certain subset of something.

6.1 Equivalence classes

Let A be a set and let $\sim$ be an equivalence relation on A. Then for $a \in A$, the *$\sim$-equivalence class of a* is defined as

$$[a]_\sim = \{x \in A \mid a \sim x\}$$

When the equivalence relation is understood, we will sometimes just write $[a]$ instead of $[a]_\sim$. For example, if ℓ is a line in $\mathbb{R}^2$ and $\sim$ is as in Exercise 6.4 (a), then $[\ell]$ is the set of lines parallel to a.

Exercise 6.10. Determine the following equivalence classes, by writing out a set theoretic description of all the elements in them.

(a) The $\sim_2$-equivalence class of 5 in Exercise 6.4 (c).

(b) The $\sim_f$-equivalence class of the point $(1,0) \in \mathbb{R}^2$, where f is the function defined just after Exercise 6.5.

Exercise 6.11. Suppose that $\sim$ is an equivalence relation on A. Show that

(a) if $a \sim b$ then $[a] = [b]$,

(b) if $a \not\sim b$ then $[a] \cap [b] = \emptyset$.

Definition 6.12. If $\sim$ is an equivalence relation on A, the set

$$A/\sim := \{[a] \mid a \in X\},$$

of all $\sim$-equivalence classes is called *the quotient of A by $\sim$*, or just the *quotient set*.

For example, suppose that $\sim$ is the equivalence relation on $A = \{1, \ldots, 6\}$ where each element is related to itself, $2 \sim 3$, and $4 \sim 5 \sim 6$. Then

$$A/\sim = \{\{1\}, \{2,3\}, \{4,5,6\}\}.$$

Here, it's a 'quotient' in that you're 'grouping/dividing A into pieces' as dictated by the equivalence relation.

Exercise 6.13. Define a map $\pi : A \longrightarrow A/\sim$, where $\pi(a) = [a]$. Show that the equivalence relation $\sim_\pi$ of Exercise 6.5 is just the original $\sim$. This shows that in fact, every equivalence relation comes from the construction in Exercise 6.5.

Example 6.14. Suppose that $\sim_n$ is the equivalence relation on $\mathbb{Z}$ where $a \sim_n b$ if $a - b$ is a multiple of n. We define

$$\mathbb{Z}/n\mathbb{Z} := \mathbb{Z}/\sim_n = \{[x] \mid x \in \mathbb{Z}\}.$$

Here, the notation '$\mathbb{Z}/n\mathbb{Z}$' comes from abstract algebra. So, what are the elements of $\mathbb{Z}/n\mathbb{Z}$? By division with remainder, every $m \in \mathbb{Z}$ has the form

$$m = qn + r, \quad q \in \mathbb{Z}, r \in \{0, \ldots, n-1\},$$

in which case $m \sim r$, so $[m] = [r]$. Moreover, if $r, s \in \{0, \ldots, n-1\}$ and $r \neq s$, then $s - r$ cannot be divisible by n, so $r \not\sim s$, implying $[r] \neq [s]$. In other words, the equivalence relation $\sim_n$ has exactly n equivalence classes, and

$$\mathbb{Z}/n\mathbb{Z} = \{[0], [1], [2], \ldots, [n-1]\}.$$

6.2 Defining the rational numbers

Many natural mathematical objects are constructed as quotient sets of equivalence relations. For instance, let's assume that we are children that know about integers and set theory. How do we define the set of rational numbers $\mathbb{Q}$? We say that a rational number is just an integer fraction, but there are coincidences like

$$\frac{a}{b} = \frac{-a}{-b} = \frac{2a}{2b}$$

and if you only know about integer arithmetic it's not clear what that horizontal line between the a and the b is supposed to mean. However, you can at least express what the coincidences are that you want to deal with just using integer arithmetic, since

$$\frac{a}{b} = \frac{c}{d} \iff ad = bc.$$

Exercise 6.15. Let $S = \{(a, b) \in \mathbb{Z} \times \mathbb{Z} \mid b \neq 0\}$ and define $(a, b) \sim (c, d)$ if $ad = bc$. Show that $\sim$ is an equivalence relation.

We now just define $\mathbb{Q} := S/\sim$. This is very nice, but currently it is just a set that we for some reason are calling by a suggestive name. In order for the set to really behave like rational numbers, we need to say what it means to add and multiply elements of the set. For instance, suppose we have $x, y \in \mathbb{Q}$ and want to define $x + y$. Well, by definition we have $x = [(a, b)]$ and $y = [(c, d)]$ for some $a, b, c, d \in \mathbb{Z}$, where $b, d \neq 0$. We'd like to make our definition so that it satisfies the usual law

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd},$$

so the way to do it using our current notation is to write:

$$\text{if } x = [(a, b)] \text{ and } y = [(c, d)], \text{ then } x + y := [(ad + bc, bd)]. \tag{4}$$

There's a subtlety here. Suppose for a moment that we try to define a similar operation called $\oplus$ on $\mathbb{Q}$, where $x \oplus y := [(a + c, b + d)]$. (This is the naive way that children without our sophistication might try to add fractions.) Now of course, you probably will object and say 'that's not a good way to do it', but can you say concretely why it's not a good definition? Let's get a feeling for how this new sort of addition works. Here are two examples:

$$\text{If } x = [(0, 1)] \text{ and } y = [(1, 2)], \text{ then } x \oplus y := [(1, 3)].$$
$$\text{If } x = [(0, 2)] \text{ and } y = [(3, 6)], \text{ then } x \oplus y := [(3, 8)].$$

This all seems very nice until you notice that the x's in the two examples are actually the same, as are the two y's. Namely, $(0,1) \sim (0,2)$ and $(1,2) \sim (3,6)$, so the associated equivalence classes are the same. However, $(1,3) \not\sim (3,8)$, so you get different definitions of $x \oplus y$ depending on which elements you are using to represent the equivalence classes! That's ridiculous. In the situation, we say that the problem is that $\oplus$ is not *well-defined*.

Exercise 6.16. Is the function $f : \mathbb{Q} \longrightarrow \mathbb{Z}$, $f([(a,b)]) = b$ well defined?

Exercise 6.17. Consider the equivalence relation $\sim$ on the set $\mathcal{F}$ of all functions $f : \mathbb{R} \longrightarrow \mathbb{R}$ from Exercise 6.4 (f). Is the following function well defined?

$$Z : \mathcal{F}/\sim \longrightarrow \mathbb{R}, \quad Z([f]) = f(0)$$

Exercise 6.18. Show that the addition operation $+$ on $\mathbb{Q}$ is *well-defined* by (4), meaning that the equivalence class $x + y$ doesn't depend on the particular choices of representative elements $(a,b) \in x$ and $(c,d) \in y$.

Exercise 6.19. Define multiplication on $\mathbb{Q}$ and show it is well-defined.

Exercise 6.20. Prove the distributive property: $x(y+z) = xy+xz$ for all $x, y, z \in \mathbb{Q}$.

7 Modular Arithmetic

Fix $n \in \mathbb{N}$, and let $\sim_n$ be the equivalence relation on $\mathbb{Z}$ given by $a \sim_n b$ if $n|a-b$. As described in Example 6.14, we define

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

Define operations $+, \cdot$ on $\mathbb{Z}/n\mathbb{Z}$ as follows.

$$[a] + [b] = [a + b].$$
$$[a] \cdot [b] = [a \cdot b].$$

Theorem 7.1. *The operations $+$ and $\cdot$ above are well-defined.*

One can show that $+, \cdot$ on $\mathbb{Z}/n\mathbb{Z}$ obey lots of the same properties that you know from integer arithmetic, like commutativity, associativity and distributivity. The elements [0] and [1] also function as 'additive and multiplicative identities', since

$$[0] + [a] = [0 + a] = [a], \quad [1] \cdot [a] = [1 \cdot a] = [a].$$

A set where you can add and multiply, satisfying the assumptions above, is called a *commutative ring*. You may see these later in an algebra class!

Exercise 7.2. Prove the distributive property in $\mathbb{Z}/n\mathbb{Z}$, namely that if $a, b, c \in \mathbb{Z}$ then $[a]([b] + [c]) = [a][b] + [a][c]$. *(This should be rather easy, and just uses the distributive property for integer arithmetic.)*

Exercise 7.3. Suppose that $m|n$. Writing the $\sim_n$-equivalence class of $a \in \mathbb{Z}$ as $[a]_n$, and the $\sim_m$-equivalence class as $[a]_m$, define a function

$$\pi : \mathbb{Z}/n\mathbb{Z} \longrightarrow \mathbb{Z}/m\mathbb{Z}, \quad \pi([a]_n) = [a]_m.$$

Here, we write the $\sim_n$ and $\sim_m$ equivalence classes of a as $[a]_n$ and $[a]_m$, to avoid confusing the two. Show that this function is well-defined, and explain how you could introduce someone who uses 24 hr clocks to 12 hr clocks.

Exercise 7.4. Is the function π in Exercise 7.3 well-defined if $n = 3, m = 2$, say?

In practice, when we work at a higher level with $\mathbb{Z}/n\mathbb{Z}$, we usually drop the brackets from our notation and represents its elements as $0, \ldots, n-1$, bearing in mind that really these numbers represent their associated equivalence classes. Using this notation here's a multiplication table for $\mathbb{Z}/3\mathbb{Z}$:

$\cdot$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

For instance, the bottom right entry reflects that $2 \cdot 2 = 1 \in \mathbb{Z}/3\mathbb{Z}$, which is code for the fact that $[2] \cdot [2] = [4] = [1]$, since $4 \sim_3 1$.

Exercise 7.5. Write out completely a 'multiplication table' for $\mathbb{Z}/4\mathbb{Z}$.

Exercise 7.6. In $\mathbb{Z}/7\mathbb{Z}$, what is $3 \cdot 4 + 3 \cdot 6 \cdot 5 \cdot (2 + 6) - 1$? You should be able to do this in your head, without a calculator and without ever even thinking about a number bigger than 30.

Note that if $[a] \in \mathbb{Z}/n\mathbb{Z}$, then $[a] + [-a] = [a - a] = [0]$. We say here that $[-a]$ is an *additive inverse* for $[a]$, since it's an element we can add to $[a]$ to get back to the additive identity. Similarly, a *multiplicative inverse* for $[a]$ is an element $[b] \in \mathbb{Z}/n\mathbb{Z}$ that we can multiply $[a]$ by to get back to the multiplicative identity:

$$[a] \cdot [b] = [1] \in \mathbb{Z}/n\mathbb{Z}.$$

Note that $[a] \cdot [b] = [1]$ means $[ab] = [1]$ which means $n|ab - 1$.

While every element of $\mathbb{Z}/n\mathbb{Z}$ has an additive inverse, not every element has a multiplicative inverse!

Exercise 7.7. Using the table from Exercise 7.5, find out which elements of $\mathbb{Z}/4\mathbb{Z}$ have multiplicative inverses.

Theorem 7.8. *An element $a \in \mathbb{Z}/n\mathbb{Z}$ has a multiplicative inverse in $\mathbb{Z}/n\mathbb{Z}$ if and only if the greatest common divisor $gcd(a, n) = 1$.*

As a hint for the proof, remember our work on $\mathbb{Z}$-linear combinations.

7.1 Sunzi's Theorem

Instead of writing that $a \sim_n b$ as above, people will write often

$$a \equiv b \pmod{n},$$

which should be read as *a is congruent to b mod n.*

Exercise 7.9. Suppose that $d|n$ and $a \equiv b \pmod{n}$. Then $a \equiv b \pmod{d}$. *(This is just a reinterpretation of Exercise 7.3.)*

So for instance, suppose that $a \equiv 3 \pmod{4}$. Then it follows immediately that $a \equiv 3 \equiv 1 \pmod{2}$. Hence, any number that is 3 mod 4 is odd. This leads us to:

Exercise 7.10. Show that there's no integer x such that we have both

$$x \equiv 5 \pmod{6} \quad \text{and} \quad x \equiv 6 \pmod{8}.$$

So, when can you arbitrarily prescribe two congruences?

Theorem 7.11 (Sunzi's Theorem)**.** *Suppose that $a, b \in \mathbb{Z}$, $n, m \in \mathbb{N}$ and that the $gcd(m, n) = 1$. Then there is some $c \in \mathbb{Z}$ such that both*

$$c \equiv a \ (mod\ m) \quad \text{and} \quad c \equiv b \ (mod\ n). \tag{5}$$

Furthermore, if c, c' both satisfy (5)*, then we have*

$$c \equiv c' \ (mod\ mn).$$

This theorem first appeared (without a formal proof) in a math textbook by Sunzi Suanjing, written sometime between 200 and 400 CE. In the west, it's usually called *The Chinese Remainder Theorem*, since it was first introduced to Europeans in some expository work about Chinese mathematics. However, we'll refer to it as above.

Proof. Suppose that $a, b \in \mathbb{Z}$, $m, n \in \mathbb{N}$ and $gcd(m, n) = 1$. Then $1 = xm + yn$ for some $x, y \in \mathbb{Z}$. Set $c = bxm + ayn$. Then we have

$$b - c = b - bxm - ayn = b(1 - xm) - ayn = byn - ayn = (b - a)yn,$$

so $b - c$ is divisible by n, implying $c \equiv b \pmod{n}$. Similarly, $c \equiv a \pmod{m}$.

Now suppose both c, c' satisfy (5). Then $c \equiv c' \bmod m$ and also mod n. In other words, $c - c'$ is divisible by both m and n. From your homework, it follows that $c - c'$ is divisible by $lcm(m, n) = nm/gcd(m, n) = mn$. So, $c \equiv c' \pmod{mn}$. □

Exercise 7.12. Find a number congruent to 7 (mod 10) and congruent to 8 (mod 11).

Exercise 7.13. Is there a number congruent to 3 (mod 4) and also to 8 (mod 12)?

Finally, you should prove the following reinterpretation of Sunzi's Theorem.

Corollary 7.14. *Suppose that $n, m \in \mathbb{Z}$ and $gcd(n, m) = 1$. Then the map*

$$f : \mathbb{Z}/mn\mathbb{Z} \longrightarrow \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}, \quad f([a]_{mn}) = ([a]_m, [a]_n)$$

is a bijection. (See Exercise 7.3 for the notation.)

8 Cardinality

Definition 8.1. Given two sets A, B we say that A, B have the *same size*, or alternatively *same cardinality*, written $A \approx B$, if there is a bijection $f : A \longrightarrow B$.

This $\approx$ satisfies the following three properties.

(a) For any set A, we have $A \approx A$, via the identity bijection $i : A \longrightarrow A$, $i(a) = a$.

(b) If $A \approx B$ then $B \approx A$, since any bijection $f : A \longrightarrow B$ has an inverse $f^{-1} : B \longrightarrow A$, which is also a bijection.

(c) If $A \approx B$ and $B \approx C$, then $A \approx C$. Indeed, bijections $A \longrightarrow B$ and $B \longrightarrow C$ compose to give a bijection $A \longrightarrow C$.

So if the 'set of all sets' were a set, which it is not, $\approx$ would define an equivalence relation on it. When A, B are finite, we have $A \approx B$ exactly when A, B have the same number of elements, since then those elements can be matched up to give the desired bijection.

You may be used to thinking that all infinite sets have cardinality ∞, so should all be equivalent under $\approx$, but we'll see in this worksheet that there are many different 'sizes' of infinity.

Exercise 8.2. Show that $\mathbb{N} \approx \mathbb{N} \cup \{0\}$.

Exercise 8.3. Show that $\mathbb{Z} \approx \mathbb{N}$.

8.1 Countability

Definition 8.4. If $A \approx \mathbb{N}$, we say that A is *countably infinite*. We say A is *countable* if it is finite or countably infinite.

For example, $\mathbb{Z}$ is countably infinite, and $\{1, 8, \heartsuit\}$ and $\mathbb{Z}$ are both countable. Note that a set A is countably infinite exactly when its elements can be arranged into an infinite list, i.e. when A can be written in the form

$$A = \{a_1, a_2, \ldots ...\}.$$

Indeed, if A is countable, then there's a bijection $f : \mathbb{N} \longrightarrow A$, and if we set $a_i := f(i)$, then $A = \{a_1, a_2, \ldots ...\}$ as above. Conversely, if $A = \{a_1, a_2, \ldots ...\}$, then we can define a bijection $f : \mathbb{N} \longrightarrow A$ by setting $f(i) = a_i$.

Here's a first example of an uncountable set.

Theorem 8.5. *There's no surjection $f : \mathbb{N} \longrightarrow (0,1)$. In particular, there's no bijection, so $(0,1)$ is uncountable.*

The proof is Cantor's famous diagonal argument.

Proof. Let $f : \mathbb{N} \longrightarrow (0,1)$ be a function. We'll show there's some $x \in (0,1)$ such that $x \neq f(i)$ for all $i \in \mathbb{N}$. This will show f isn't surjective.

Let's write out decimal expansions of all the numbers $f(i) \in (0,1)$ as follows.

$$\begin{aligned} f(1) &= .a_{11}\, a_{12}\, a_{13}\, a_{14} \ldots \\ f(2) &= .a_{21}\, a_{22}\, a_{23}\, a_{24} \ldots \\ f(3) &= .a_{31}\, a_{32}\, a_{33}\, a_{34} \ldots \\ f(4) &= .a_{41}\, a_{42}\, a_{43}\, a_{44} \ldots \\ &\vdots \end{aligned}$$

We want to construct some $x \in (0,1)$ that's not equal to any of these. So, set

$$x = .x_1\, x_2\, x_3 \ldots, \quad x_i = \begin{cases} 3 & a_{ii} = 4 \\ 4 & a_{ii} \neq 4. \end{cases}$$

For example, suppose we have

$$\begin{aligned} f(1) &= .3869 \ldots \\ f(2) &= .0482 \ldots \\ f(3) &= .4490 \ldots \\ f(4) &= .2224 \ldots \\ &\vdots \end{aligned}$$

Then $x = .4343 \ldots$. By construction $x_i \neq a_{ii}$, so x and $f(i)$ differ in the i^{th} decimal place, and hence aren't equal. (Note that since x doesn't have 0's and 9's in its decimal expansion, it has a *unique* decimal example, so to check it's not equal to any of the $f(i)$, it suffices to check that the decimal expansions above are different. Contrast this with the two decimal expansions for $.10000 \ldots = .099999 \ldots$.) $\square$

Exercise 8.6. Draw the graph of a bijection $f : (0,1) \longrightarrow \mathbb{R}$, and conclude that $\mathbb{R}$ is also uncountable.

Exercise 8.7. (a) If $B \subset \mathbb{N}$, show that B is countable. *Hint: suppose that B is infinite. Given $b \in B$, let $g(b)$ be the number of elements of B that are less than or equal to b. Show that $g : B \longrightarrow \mathbb{N}$ is a bijection.*

(b) Using part (a), show very quickly that if A is countable and $f : B \longrightarrow A$ is injective, then B is countable. In particular, a subset of a countable set is countable.

Exercise 8.8. (a) If $f : \mathbb{N} \longrightarrow B$ is surjective, show that B is countable. *Hint: if $b \in B$, let $g(b) = \min\{x \in A \mid f(x) = b\}$.*

(b) Using part (a), show that if A is countable and $f : A \longrightarrow B$ is surjective, then B is countable.

Exercise 8.9. Show that if A, B are both countably infinite, so is $A \cup B$.

Exercise 8.10. (a) Show that $\mathbb{N} \times \mathbb{N}$ is countably infinite. *Hint: there are two approaches I can think of. The easiest one to write down is to let p_i be the i^{th} prime and consider p_i^j. For a geometric approach, think about snakes! Make an $\mathbb{N} \times \mathbb{N}$ grid, and starting at $(1, 1)$, try to wind through it.*

(b) Using (a), show that if A, B are both countably infinite, so is $A \times B$.

A version of the following was suggested by F. Dong on his Intro to Abstract Math Final Exam, in Fall 2024! It gives another possible solution to Exercise 8.10 (a).

Exercise 8.11. Let $f : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}$, $f(x, y) = (x+y)^2 + x$. Show that f is injective.

Exercise 8.12. Using the previous exercises and our definition of $\mathbb{Q}$ as a quotient set, show that $\mathbb{Q}$ is countably infinite.

Exercise 8.13. Show that the set $\mathbb{R} \setminus \mathbb{Q}$ of all irrational numbers is uncountable.

8.2 Ordering sets by cardinality

If there is an injection $f : A \longrightarrow B$, we write $A \preceq B$. If $A \preceq B$ but $A \not\approx B$, i.e. there's an injection $A \longrightarrow B$ but there is no such bijection, we write $A \prec B$. For example, $\mathbb{N} \prec \mathbb{R}$ because the inclusion $i : \mathbb{N} \longrightarrow \mathbb{R}$, $i(n) = n$ is an injection, but we proved in Exercise 8.6 that there's no bijection $\mathbb{N} \longrightarrow \mathbb{R}$.

Note that to prove that $A \prec B$, it does *not* suffice to produce an injection that is not a bijection. For example, $f : \mathbb{N} \longrightarrow \mathbb{N}$, $f(x) = x + 1$ is an injection that is not a bijection, but $\mathbb{N} \approx \mathbb{N}$. The point is to construct an injection, and then show separately there is no possible (unrelated) bijection.

Exercise 8.14. In this exercise, we show that if A is a set, then $A \prec \mathcal{P}(A)$.

(a) Construct an injective function $i : A \longrightarrow \mathcal{P}(A)$, showing that $A \preceq \mathcal{P}(A)$.

(b) Suppose that $f : A \longrightarrow \mathcal{P}(A)$ is a function. Show that the subset

$$X = \{a \in A \mid a \notin f(a)\} \subset A$$

has the property that $X \neq f(b)$ for all $b \in A$. Conclude that $A \not\approx \mathcal{P}(A)$, which together with (a) implies that $A \prec \mathcal{P}(A)$.

(c) Suppose now that $A = \mathbb{N}$. To each subset $B \subset \mathbb{N}$, associate an infinite string $b_1 b_2 b_3 \ldots$ of 0's and 1's, where $b_i = 1$ if $i \in B$, and $b_i = 0$ otherwise. Explain how from this perspective, the proof in (b) above is exactly the same as Cantor's diagonal argument, from Theorem 8.5.

The following theorem may look obvious at first, but it's really not!

Theorem 8.15. *(Schroeder–Bernstein) If $A \preceq B$ and $B \preceq A$, then $A \approx B$.*

Proof. We can assume A, B are disjoint, take injections

$$f : A \longrightarrow B,\ g : B \longrightarrow A,$$

and for convenience we combine them to give an injection

$$h : A \cup B \longrightarrow A \cup B, \quad h(x) = \begin{cases} f(x) & x \in A \\ g(x) & x \in B. \end{cases}$$

If $x \in A \cup B$ lies in the image of h, we define $h^{-1}(x) \in A \cup B$ to be the unique element of $A \cup B$ that h takes to x. Then given $x \in A \cup B$, we can repeatedly apply h^{-1}, giving elements $h^{-1}(x), h^{-2}(x), \ldots$. Let's call these elements the *ancestors* of x. If all these ancestors are always in the image of h, we can keep going forever. However, if there's some ancestor $h^{-n}(x)$ that doesn't lie in the image of h, it has no further ancestors, and we call $h^{-n}(x)$ the *original ancestor*. We now define a bijection

$$H : A \longrightarrow B, \quad H(a) = \begin{cases} g^{-1}(a) & \text{if } a \text{ has an original ancestor, which lies in } B \\ f(a) & \text{otherwise.} \end{cases}$$

We first claim that H is injective. So, assume $a \neq a'$ and $H(a) = H(a')$. Since f, g are injective, we can assume a, a' are in the two separate cases in the definition of

H. So, a has an original ancestor that lies in B, and a' doesn't, but $g^{-1}(a) = f(a')$. However, here $a' = f^{-1}(g^{-1}(a)) = h^{-2}(a)$, so if a has an original ancestor in B, then a' will too, and we're done.

Next, we want to show that H is surjective. Pick some $b \in B$. Suppose first that b has an original ancestor, which lies in B. Let $a = g(b)$. Then a has the same original ancestor, which lies in B, so $H(a) = g^{-1}(a) = b$ and we're done. On the other hand, say that b either doesn't have an original ancestor, or it has one that lies in A. Then in particular b has ancestors, i.e. it lies in the image of h, so we can set $a := h^{-1}(b) = f^{-1}(b)$. This a also either doesn't have an original ancestor or has one that lies in A, so $H(a) = f(a) = b$. □

Alternative proof. The following is the same proof as above, but phrased more visually. As before, suppose we have injective functions $f : A \longrightarrow B$ and $g : B \longrightarrow A$. Draw elements of A and B as dots: $\bullet$ for A and $\odot$ for B. Draw an arrow from each $a \in A$ to $f(a)$, and from each $b \in B$ to $f(b)$. If you pick a point x in $A \cup B$, you can look at all the points you get from that point by tracking forwards and backwards along the arrows. Let's call such a set of points a *lineage*. Since every dot has exactly one arrow pointing out of it, and at most one arrow pointing into it, lineages look like

(a) a biinfinite line $\cdots \to \bullet \to \odot \to \bullet \to \odot \to \cdots$,

(b) a circle $\bullet \to \odot \to \cdots \to \bullet \to \odot \to \bullet$, where the first and last $\bullet$ are the same,

(c) a ray $\bullet \to \odot \to \bullet \to \odot \to \cdots$, or

(d) a ray $\odot \to \bullet \to \odot \to \bullet \to \cdots$

Think of $A \cup B$ as broken up into all the possible lineages above. Then we can make a bijection $H : A \longrightarrow B$ by defining it separately on each lineage, and matching up the $\bullet$'s to $\odot$'s in that lineage. In cases (a)-(c), just define $H(\bullet)$ to be the $\odot$ to the immediate right. In case (d), just define $H(\bullet)$ to be the $\odot$ to the immediate left. □

Exercise 8.16. Show that $\mathcal{P}(\mathbb{N}) \approx \mathbb{R}$. *Hint: for convenience, produce injections in both directions. You will probably find using binary or decimal expansions useful.*

Exercise 8.17. If $A \preceq B$ and $B \prec C$, show that $A \prec C$. *Hint: Amazingly, this isn't obvious. Use the Schroeder-Bernstein theorem.*

Exercise 8.18. Let $\mathcal{A}$ be a set whose elements are sets. Show that there is some set B such that $A \prec B$ for all $A \in \mathcal{A}$.

This exercise suggests an unimaginable number of different infinite cardinalities. Namely, given a set A, let $\mathcal{P}^n(A)$ be the n^{th} *iterated power set* of a set A, defined by

$$\mathcal{P}^n(A) := \mathcal{P}(\cdots \mathcal{P}(\mathcal{P}(A)) \cdots).$$

Starting with $\mathbb{N}$, we can construct a sequence of sets as follows:

$$\mathbb{N} \prec \mathcal{P}(\mathbb{N}) \prec \mathcal{P}^2(\mathbb{N}) \prec \ \cdots \prec B \prec \mathcal{P}(B) \prec \mathcal{P}^2(B) \prec \cdots \prec C \prec \mathcal{P}(C) \prec \cdots,$$

where here, B is some set that is bigger in size than all $\mathcal{P}^n(\mathbb{N})$, and then C is defined similarly with B instead of $\mathbb{N}$. Of course, this goes on forever, even after we run out of letters in the alphabet, and the exercise above shows that even after you repeat this procedure forever, there's STILL a bigger set than everything so far constructed. Then you can repeat the process with that bigger set, and continue...

Exercise 8.19. Write out all the elements of $\mathcal{P}(\mathcal{P}(\emptyset))$. Then try to come up with a formula for the number of elements in the set $\mathcal{P}^n(\emptyset)$.

9 Complex numbers

The real numbers $\mathbb{R}$ are great, but they could be better in some algebraic ways! For instance, some polynomial equations can't be solved, e.g. $x^2 = -1$ has no real solutions. To fix this, we introduce a new number called i, with the property that $i^2 = -1$, and we see what we get when we throw this in with all the other real numbers.

We'd like be able to add and multiply, so we're forced to also consider expressions of the form $x+iy$, where $x, y \in \mathbb{R}$, which we call *complex numbers*. Assuming addition and multiplication of complex numbers satisfy the usual rules, we should define

$$(x + iy) + (x' + iy') := (x + x') + i(y + y'), \tag{6}$$

$$(x + iy)(x' + iy') := (xx' - yy') + i(yx' + xy'). \tag{7}$$

Here, to justify the second equality, expand out the left hand side and use the fact that $i^2 = -1$. We let $\mathbb{C}$ denote the set of all complex numbers.

Fact 9.1. *Defined as in* (6) *and* (7)*, addition and multiplication on* $\mathbb{C}$ *satisfy:*

(a) for $z, w \in \mathbb{C}$*, we have* $z \cdot w = w \cdot z$ *and* $z + w = w + z$*, (commutativity)*

(b) for $z, u, w \in \mathbb{C}$*, we have* $z \cdot (u \cdot w) = (z \cdot u) \cdot w$*, (associativity)*

(c) for $z, u, w \in \mathbb{C}$*, we have* $z \cdot (u + w) = z \cdot u + z \cdot w$*. (distributivity)*

You can try to prove these properties if you like, just using the analogous properties of addition and multiplication of real numbers.

Definition 9.2. Suppose that $z = x + iy \in \mathbb{C}$. Then

- the *real part* of z is $Re(z) := x$,
- the *imaginary part* of z is $Im(z) := y$,
- the *(complex) conjugate* of z is $\bar{z} := x - iy$,
- the *absolute value* or *modulus* of z is $|z| := \sqrt{x^2 + y^2}$.

Geometrically, we imagine $\mathbb{C}$ as the plane. The real numbers form one axis and the *imaginary numbers* iy, where $y \in \mathbb{R}$, form another axis. If $z \in \mathbb{C}$, then $Re(z)$ and $Im(z)$ are the coordinates of z. The absolute value $|z|$ is the distance from the origin to z. And the complex conjugate $\bar{z}$ is obtained by reflecting z through the real axis.

Exercise 9.3. Suppose that $z, w \in \mathbb{C}$. Show that

(a) $|z|^2 = z\bar{z}$,

Solution. Suppose $z = x + iy$. Then

$$z\bar{z} = (x+iy)(x-iy) = x^2 + iyx - iyx - i^2y = x^2 + y^2 = |z|^2. \qquad \square$$

(b) $\overline{z+w} = \bar{z} + \bar{w}$, and $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$,

Solution. Let's do multiplication. If $z = x + iy$ and $w = x' + iy'$ then

$$\overline{z \cdot w} = \overline{(x+iy)(x'+iy')} = \overline{xx' - yy' + i(yx' + xy')} = xx' - yy' - i(yx' + xy')$$
$$\bar{z} \cdot \bar{w} = (x-iy)(x'-iy') = xx' - iyx' - iy'x + i^2yy' = xx' - yy' - i(yx' + xy'),$$

and these are the same. $\square$

(c) $|zw| = |z||w|$.

Solution. We have $|zw|^2 = zw\overline{zw} = zw\bar{z}\bar{w} = z\bar{z} \cdot w\bar{w} = |z|^2|w|^2$, so taking square roots finishes it. $\square$

(d) Show that if $z \in \mathbb{C}$ then $|Re(z)| \leq |z|$ and $|Im(z)| \leq |z|$.

Solution. Suppose $z = x+iy$. Then $|z| = \sqrt{x^2+y^2} \geq \sqrt{x^2+0} = |x| = |Re(z)|$. The other one is similar. $\square$

(e) Show that if $z \in \mathbb{C}$ then $z + \bar{z} = 2Re(z)$, while $z - \bar{z} = 2Im(z)$.

Solution. If $z = x + iy$ then $z + \bar{z} = x + iy + x - iy = 2x = 2Re(z)$. The other part is similar. $\square$

Theorem 9.4 (The Triangle Inequality)**.** *If $z, w \in \mathbb{C}$ then $|z+w| \leq |z| + |w|$.*

Try to visualize this geometrically. If you form a triangle in $\mathbb{C}$ with vertices $0, z, z+w$, then the side lengths are $|z|, |w|$, and $|z+w|$, respectively. Any one side length is at most the sum of the other two, which is the statement of the triangle inequality.

Exercise 9.5. Let's prove the triangle inequality algebraically!

(a) Given $v, w \in \mathbb{C}$, show that $|z+w|^2 = |z|^2 + |w|^2 + 2Re(z\bar{w})$. *Hint: it'll be useful to use Exercise 9.3.*

(b) Prove the triangle inequality.

We'd also like to be able to divide complex numbers, so suppose that $z, w \in \mathbb{C}$ and $w \neq 0$. What should z/w be? Well, if division works as expected, we should get

$$\frac{z}{w} = \frac{z}{w} \cdot \frac{\bar{w}}{\bar{w}} = \frac{z \cdot \bar{w}}{w\bar{w}} = \frac{z \cdot \bar{w}}{w\bar{w}} = z \cdot \frac{\bar{w}}{|w|^2}, \tag{8}$$

and here $\bar{w}/|w|^2$ is a complex number divided by a real number, which makes sense as you can just divide the real and imaginary parts of $\bar{w}$ by $|w|^2$ separately. For example,

$$\frac{2+3i}{3+5i} = \frac{2+3i}{3+5i} \cdot \frac{3-5i}{3-5i} = \frac{(2+3i)(3-5i)}{34} = \frac{6+9i-10i-15i^2}{34} = \frac{21}{34} - \frac{1}{34}i.$$

You can then check that if you define division by

$$\frac{z}{w} := z \cdot \frac{\bar{w}}{|w|^2},$$

it satisfies all the usual properties you'd like, for instance

$$\frac{z}{w} + \frac{u}{v} = \frac{zv + uw}{wv}, \quad \frac{z}{w} \cdot \frac{u}{v} = \frac{zu}{wv} \tag{9}$$

Exercise 9.6. Compute $(1+2i)/(4-2i)$.

Exercise 9.7. Prove one of the two properties in (9).

Finally, we mention the following important theorem, which illustrates why complex numbers are sometimes better than real numbers.

Theorem 9.8 (The Fundamental Theorem of Algebra (FTA))**.** *Every nonconstant polynomial $p(z) = a_n z^n + a_{n-1}x^{n-1} + \cdots + a_0$, where $a_i \in \mathbb{C}$, has a root $z \in \mathbb{C}$.*

The nonconstant assumption rules out polynomials like $p(z) = 5$, which certainly has no roots. The proof is a bit too difficult for this class, but if you're interested, then take a complex analysis class! Note that the FTA isn't true with $\mathbb{R}$ instead of $\mathbb{C}$, since $p(x) = x^2 + 1$ has no roots. But the FTA also justifies our development of $\mathbb{C}$: we started the section by saying that $x^2 = -1$ has no solutions, so we should throw in another number i that is a solution to this equation. But we could just as well have started with $x^4 - 6x^3 + 15x^2 - 18x + 10 = 0$, which it turns out also doesn't have any real solutions, and tried to add in a solution to this. The FTA says that actually, as soon as we add in i, we have solutions to *all* nonconstant polynomial equations.

9.1 The exponential function and polar coordinates

Definition 9.9. If $z \in \mathbb{C}$, we define $e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots \in \mathbb{C}$.

Why does this series converge? It's essentially the same as proving convergence when z is a real number, which you probably did in a Calculus class at some point. One just has to develop convergence of sequences and series for complex numbers instead of real numbers, and then one can use the ratio test, for instance, which for complex series says that $\sum_{n=0}^{\infty} z_n$ converges if $\lim_{n\to\infty} |z_{n+1}|/|z_n| < 1$.

Fact 9.10. *If $z, w \in \mathbb{C}$, we have $e^z e^w = e^{z+w}$.*

Proof Sketch. The point is to expand out the product of two convergent infinite series $\sum_{n=0}^{\infty} z_n$ and $\sum_{n=0}^{\infty} w_n$. You'd intuitively expect their product to be a series in which you sum up every product $z_i w_j$, where $i, j \geq 0$, exactly once. The series

$$\sum_{n=0}^{\infty} c_n, \quad c_n := \left(\sum_{i=0}^{n} z_i w_{n-i} \right)$$

is called the *Cauchy product* of the two original series, and does exactly this: the product $z_i w_j$ is the i^{th} term in the sum defining c_{i+j}, and appears nowhere else. It turns out that as long as one of the two series involved converges absolutely (e.g., if $\sum_{n=0}^{\infty} |z_n|$ converges) then the Cauchy product converges and we have

$$\left(\sum_{n=0}^{\infty} z_n \right) \cdot \left(\sum_{n=0}^{\infty} w_n \right) = \sum_{n=0}^{\infty} c_n.$$

Let's apply this to the problem at hand. We have:

$$\begin{aligned} e^z e^w &= \left(\sum_{n=0}^{\infty} \frac{z^n}{n!} \right) \left(\sum_{n=0}^{\infty} \frac{w^n}{n!} \right) \\ &= \sum_{n=0}^{\infty} \sum_{i=0}^{n} \frac{z^i}{i!} \frac{w^{n-i}}{(n-i)!} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{i=0}^{n} \binom{n}{i} z^i w^{n-i} \\ &= \sum_{n=0}^{\infty} \frac{(z+w)^n}{n!} \\ &= e^{z+w}, \end{aligned}$$

where the second to last step is the Binomial Theorem. □

Theorem 9.11 (Euler's Theorem)**.** *If $\theta \in \mathbb{R}$, we have $e^{i\theta} = \cos(\theta) + i\sin(\theta)$.*

Proof. We just manipulate the series definition of $e^{i\theta}$ into a form involving the Taylor series expansions of $\cos(\theta)$ and $\sin(\theta)$, as follows:

$$\begin{aligned} e^{i\theta} &= 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \frac{(i\theta)^5}{5!} + \cdots \\ &= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots\right) + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \cdots\right) + \\ &= \cos(\theta) + i\sin(\theta). \quad \square \end{aligned}$$

Corollary 9.12. $e^{i\pi} = -1$

Proof. Just plug in $\theta = \pi$ to Euler's Theorem. $\square$

More generally, if $r > 0$ and $\theta \in \mathbb{R}$, consider the complex number

$$z = re^{i\theta} = r\cos(\theta) + ir\sin(\theta).$$

Here, $|z| = r$, and θ measures the angle from the positive real axis to the the line segment from 0 to z. Since we can pick r and θ arbitrarily, we have:

Fact 9.13. *Every $z \in \mathbb{C}$ can be written as $z = r \cdot e^{i\theta}$ for some $r \geq 0$ and $\theta \in \mathbb{R}$.*

Here, $z = r \cdot e^{i\theta}$ is called writing z in *polar coordinates*.

Exercise 9.14. (a) If $re^{i\theta} = 1$, show that $r = 1$ and $\theta = 2\pi n$ for some $n \in \mathbb{Z}$.

(b) Using (a), show that if $re^{i\theta} = r'e^{i\theta'}$, then $r = r'$, and if $r \neq 0$ then we have $\theta' - \theta = 2\pi n$ for some $n \in \mathbb{Z}$.

Solution. For (a), note that $r = |re^{i\theta}| = |1| = 1$, and if $1 = e^{i\theta} = \cos(\theta) + i\sin(\theta)$ then $\cos(\theta) = 1$ and $\sin(\theta) = 0$, which only happens when θ is a multiple of 2π.

For (b), $\frac{r}{r'}e^{i(\theta-\theta')} = 1$, so $\frac{r}{r'} = 1$ and $\theta - \theta' = 2\pi n$ for some n by (a). $\square$

One advantage of polar coordinates is that the formula for complex multiplication is simple than that given in (7): from Fact 9.10 we get

$$(re^{i\theta}) \cdot (r'e^{i\theta'}) = (rr')e^{i(\theta+\theta')}.$$

In particular, using the multiplication formula and induction one can prove that

$$z = re^{i\theta}, n \in \mathbb{N} \quad \Longrightarrow \quad z^n = r^n e^{in\theta},$$

which is sometimes called DeMoivre's Theorem.

Exercise 9.15. If $z = re^{i\theta}$, write $1/z$ in polar coordinates.

Solution. $1/z = \frac{1}{r}e^{i(-\theta)}$ □

Exercise 9.16. Since $e^{i\theta}e^{i\theta'} = e^{i(\theta+\theta')}$, Euler's Theorem says that

$$(\cos(\theta) + i\sin(\theta))(\cos(\theta') + i\sin(\theta')) = \cos(\theta+\theta') + i\sin(\theta+\theta') \tag{10}$$

Show how to derive from this the angle sum formulas for cosine and sine.

9.2 Roots of unity

An n^{th} *root of unity* is a complex number $z \in \mathbb{C}$ such that $z^n = 1$. We let

$$U_n := \{z \in \mathbb{C} \mid z^n = 1\}$$

be the set of all n^{th} roots of unity. Note that if $z \in U_n$ then $|z|^n = |z^n| = |1| = 1$, so $|z| = 1$, and hence $U_n \subset S^1$, the unit circle in $\mathbb{C}$.

Exercise 9.17. If $z, w \in U_n$, show that $zw \in U_n$ and $1/z \in U_n$.

In the language of abstract algebra, this shows U_n is a *group*.

Solution. If $z^n = w^n = 1$ then $(zw)^n = z^n w^n = 1$ and $(1/z)^n = 1/z^n = 1/1 = 1$. □

Exercise 9.18. In polar coordinates, $U_n = \{e^{2\pi i\frac{k}{n}} \mid k = 1, \ldots, n\}$.

Solution. First, let's note that for each $k = 1, \ldots, n$, we have

$$(e^{2\pi ik/n})^n = e^{2\pi ik} = \cos(2\pi ik) + i\sin(2\pi ik) = 1.$$

Conversely, suppose that $z^n = 1$. Write $z = re^{i\theta}$, with $\theta \in [0, 2\pi)$, say. Then $r^n e^{i\theta n} = 1 = e^{i\cdot 0}$, so by Exercise 9.14 (a) we have $r = 1$ and $\theta \cdot n = 2\pi k$ for some $k \in \mathbb{Z}$, implying $\theta = 2\pi k/n$, but since $\theta \in [0, 2\pi)$ we have $k \in \{1, \ldots, n\}$. □

The *order* of $z \in U_n$ is the smallest $d \in \mathbb{N}$ such that $z^d = 1$. For example, we have $-1, i \in U_4$, since $(-1)^4 = i^4 = 1$, but $-1 \in \mathbb{C}$ has order 2, while i has order 4.

Exercise 9.19. If $z \in U_n$ has order d, show that $d|n$. *Hint: division with remainder.*

Solution. Write $n = qd + r$, where $0 \leq r < d$. Then $1 = z^n = z^{qd+r} = (z^d)^q z^r = z^r$, which contradicts that d is the smallest such that $z^d = 1$. □

Exercise 9.20. Show that $e^{2\pi i\frac{k}{n}} \in U_n$ has order $n/\gcd(k,n)$. *Hint: show that the order is the minimal d such that $n|k \cdot d$, and then argue that $k \cdot d = lcm(k,n)$.*

Solution. First, note that

$$\left(e^{2\pi i\frac{k}{n}}\right)^d = 1 \iff e^{2\pi i\frac{kd}{n}} = 1 \iff \frac{kd}{n} \in \mathbb{Z} \iff n|kd,$$

so the order is the minimal such d. But for such a d, the product kd is the minimal multiple of k that's also a multiple of n, so $kd = lcm(k,n) = kn/gcd(k,n)$, and dividing by k, we're done. □

An element $z \in U_n$ is called *primitive* if it has order n, rather than something smaller. Note that for each n, the element $e^{2\pi i\frac{1}{n}} \in U_n$ is primitive.

Exercise 9.21. Suppose that $z \in U_n$ is primitive. Show that $U_n = \{z, z^2, \ldots, z^n\}$.

That is, if $z \in U_n$ is primitive then *all* elements of U_n are powers of z. As a hint, try to show that all the above powers of z are distinct.

Solution. It suffices to show that all the above powers are distinct, since they're all in U_n and U_n has n elements. But if $z^i = z^j$, with $j > i$, say, then $z^{j-i} = 1$ and $0 < j - i < n$, contradicting that z^n is the first power of z that equals 1. □

Exercise 9.22. Set $z = \frac{1}{2} - i\frac{\sqrt{3}}{2}$. Show that z is a root of unity, and find its order. Then write z^{100} in the form $x + iy$.

Exercise 9.23. If $z \in \mathbb{C}$, $m, n \in \mathbb{N}$, $z^m = 1$ and $z^n = 1$, show that $z^{gcd(m,n)} = 1$.

Exercise 9.24. For each $n \geq 2$, show that the sum of all elements of U_n is zero. *Hint: there are multiple ways to do this. Set $c = e^{2\pi i/n}$. One method is to note that multiplying by c just rearranges the elements of U_n. Another method is to observe that you're trying to sum $1 + c + c^2 + \cdots + c^{n-1}$. It's a good idea here to multiply that whole expression by $1 - c$. You probably did something like this in Calculus II..*

Exercise 9.25. Show that the product of all elements of U_n is 1 if n is odd, and -1 if n is even. *Hint: one way to do it is to note that if $z \in U_n$, so is $1/z$. Another method is to just multiply out the product of the elements using polar coordinates.*