Mathematics

Clémence Chanavat

November 19, 2023

Contents

These notes are a first draft, so they are most likely full of typos. If something is not clear, seems wrong, or you simply have a question, feel free to drop an email at clemence.chanavat@gmail.com. The material will also probably evolve throughout the course. Have fun, and (re)read Section 1.3 whenever you are stuck.

1 Logic

Mathematics is a language made of sentences. These sentences are written using the syntax and rules of logic. In this chapter, we learn the basics of logic's grammar, and how to make well formed sentences. Then, given a well formed sentence, we see a general procedure to understand wether a given sentence is true or not, that is, we will learn how to make proofs.

1.1 Connective

We start with a few list of symbols, that constitute the basic alphabet of mathematics. It is primordial to know them, they are called *connectives*. We have

1. The *and*, noted $\wedge$.
2. The *or*, noted $\vee$.
3. The *not*, noted $\neg$.
4. The *implies*, noted $\Rightarrow$.
5. The *equivalent*, noted $\iff$.

As in any language, these symbols have meanings, that somewhat correspond to the intuition. To understand it, let us abbreviate with the letter C, the sentence *the cat is orange*, and with D the sentence *the dog has three legs.* Then, we write

$$C \wedge D$$

to mean that both the cat is orange *and* the dog has three legs. Thus, whenever we see the cat, it is orange, and whenever we see the dog, it has three leg.

Next is the or. It is a little bit different than what we are used to in the English language. We write

$$C \vee D$$

to mean that the cat is orange, *or* the dog has three leg. This means that at least one of the three following statements is true:

1. The cat is orange,
2. The dog has three legs,
3. The cat is orange, and the dog has three legs.

In math, the or connective needs not to be exclusive. In $C \vee D$, both C and D can be true at the same time. It means that after seeing the cat and the dog, we are guarantee that at least the cat will be orange, or the dog will have three legs, at least one of the two statements will be true.

The next connective of interest is not. We write

$$\neg D$$

to mean that the dog has *not* three legs, that is whenever we will see the dog, it will have some number of legs and we are guarantee that this number is not three. It can be one, it can be four, it can be something else, but it will not be three.

Exercise 1: Argue that $\neg(C \wedge D)$ says the same thing as $\neg C \vee \neg D$.

We move on to equivalence. We write

$$C \iff D$$

to mean that knowing that the cat is orange is the same thing as knowing that the dog has three legs. It means that if we go first to see the cat and conclude it is orange, then we do not need to go to the dog to see it has three leg. *we already know it.* Conversely, if we go first see the dog, and it has three legs, then we are sure that the cat is orange. We also say C *if and only if* D.

Exercise 2: Argue that $C \iff D$ says the same thing as $D \iff C$.

Finally, we have the implication. It is the most used of them all. We write

$$C \Rightarrow D$$

to mean that *if* the cat is orange, *then* the dog has three legs. It says that if we go see the cat and conclude it is orange, then we are guarantee that the dog will have three legs. However, and this is very important, if we do not see that the cat is orange, then we *cannot* say anything at all about the dog, it might, or might not, have three legs. Unfortunately, the intuitive meaning of the implication does not really reflect well any construct in the English language. The implication is better understood as a function, $C \Rightarrow D$ is a procedure that takes a proof that the cat is orange, and transforms it into a proof that the dog has three legs. Therefore, if the cat is not orange, we will never be able to call feed this procedure with a proof of orangeness of the cat, and therefore this procedure will not let us prove that the dog has three leg. However, there might perfectly well exists another procedure proving that the dog has three legs by other means, not involving the color of the cat.

Exercise 3: Argue that $(C \Rightarrow D) \wedge (D \Rightarrow C)$ says the same thing as $C \iff D$. Is $C \Rightarrow D$ saying the same thing as $D \Rightarrow C$?

Notice that when we will do math later, we will freely use the symbol themselves, or their equivalent English terminology. In particular, we write "if X then Y" more often than "$X \Rightarrow Y$", but keep in mind that they are the same thing. We summarize these constructions with truth tables, you should refer to these when in doubt on what a sentence mean. Here is how to read it. The number 1 means True, the number 0 means False. In the table for $\vee$, on a given row, a column gives a particular truth value to C, to D, and to the resulting $C \vee D$. For instance, if C is true, D is false, we see that $C \vee D$ is true.

C	$\neg C$
0	1
0	0

,

C	D	$C \wedge D$
0	0	0
0	1	0
1	0	0
1	1	1

,

C	D	$C \vee D$
0	0	0
0	1	1
1	0	1
1	1	1

,

C	D	$C \iff D$
0	0	1
0	1	0
1	0	0
1	1	1

,

C	D	$C \Rightarrow D$
0	0	1
0	1	1
1	0	0
1	1	1

1.2 Quantifiers

So far, we cannot really say much. We need to introduce two new symbols:

$$\forall, \exists.$$

They will allow us to quantify, to say that all things in a big thing share the same property, or that there is some thing in a big thing that has a property. However, there is some subtleties that come with those symbols, we need to use free variables. Earlier, we said D means that the dog has three legs. There is no room in this formula, everything is fixed. Let us to do something, and replace "three" with the letter n, that we declare to be an unspecified natural number. Now, we write

$$D(n)$$

to mean "the dog has n legs", for some number n, that we deliberately *not* specify. This n is called a free variable, it can potentially be any natural number, and it is good to think of it as being *all* the natural number at the same time. Now, if we take a natural number, say 7, then we will write $D(7)$ to specify the unknown number n with 7, and $D(7)$ means that the dog has seven legs. Notice that the previous sentence D has now become $D(3)$.

Is the sentence $D(n)$ true or false? It does not make sense to ask this question. We cannot ask for the truth value of a sentence with free variables, we first need to specify a behavior for our free variable, and this is done with the quantifier. We write

$$\exists n, D(n)$$

to mean that there *exists* at least a value of n (like 4, 9, or seven billion) such that the dog has n legs. For instance, we know $\exists n, D(n)$ is true because when we will look at the dog, we will count its number of legs, and see (probably) that there is $n = 4$. We say that 4 is a *witness* of $\exists n, D(n)$.

Next, we write

$$\forall n, D(n)$$

to mean that *for all* choices of number n, the dog will have precisely this number of legs. Here, this is quite absurd, because the dog has one and only one number of leg. But consider the following:

$$\forall n, (5 \leq n \Rightarrow \neg D(n)).$$

It means that for all number n, if the number n is greater or equal to 5, then the dog has not n legs. This feels more true, as we know indeed that the dog has four legs.

Exercise 4: Is $\exists n, \neg D(n)$ true? Can you rewrite $\neg \exists n, \neg D(n)$ as something with less symbols?

1.3 Practical sentences, and how to do proofs

This was only the tip of the iceberg. Logic is a very powerful language that allows us to communicate with math. Ultimately, we want to do proofs. This section is a practical place that you are invited to read, and re-read every time you are confused with things.

1.3.1 What is a proof?

What is a proof? A proof is a way to testify that a mathematical sentence is true. We saw in the previous part how to construct mathematical sentences, but we didn't see how to prove them.

Definition 1: A *proof* is a succession of mathematically sound steps. Whenever you write a proof, you always have a *bag of hypothesis*. This bag of hypothesis is a collection of mathematical sentences that you locally assume as true. The bag will grow during the proof (we will see how), and the goal of a proof is to reach a certain mathematical sentence (the conclusion) by mean of logical steps that combine the hypothesis in the bag. The bag is implicit, and is never explicitly described in textbooks. It consists of:

1. The statements that are always true (i.e. the theorem we already proved),
2. The axioms that our objects of interest satisfy (for instance, if there is some $r \in \mathbb{Q}$ in the proof, then by definition, we can say $r = \frac{p}{q}$ with p, q integers).
3. The local assumptions, that is, the things we assumed to be true for the sake of our proof, but that are not always true. What does it mean? We will understand it better when we will talk about implication, but for instance consider the statement "if n is even, then $n+1$ is odd". Then to prove that, we will take some n, and assume for the sake of the proof, that n is even. However, we are *not* claiming that all n's are even, we are only assuming that locally, for our precise needs, n is even. It is like when we write a function in programming language. For instance if we write the function:

   ```
   float my_proof (int n) {
       ...
   }
   ```

   then we do not say that there is always (outside of the scope of the function) some `n` of type `int`, we merely say that during the construction of the function `my_proof`, we are allowed to use a variable of type `int` (this correspondence is very deep, functions and proofs are the same thing).

Do not forget your hypothesis, they are what you need to write your proofs. When you are stuck in a proof, first look if any of your local hypothesis are useful to move towards the goal. If you do not see anything relevant, look at the recent and related theorems we proved. Is there any that would significantly transform the *state* of the proof into something more interesting to work with? Can the application of a theorem give new hypothesis that will be useful to move forwards?

1.3.2 Constructing proofs

Now that we know what a proof is supposed to do, we summarize here how to construct them. Say we want to prove a mathematical statement, then it is a formula written in the language of logic. This formula have a certain shape, and depending on its shape, we will apply certain proof techniques. Thus, we will see that math is in fact a very mechanical procedure. To do a proof, it suffices to pattern match with the following list. It is good to reflect on why those really prove what we want to prove, given what we saw in the previous chapter.

- If you need to prove a statement of the form

$$A \wedge B,$$

then you will do two proofs, first you will prove A, then you will prove B. For instance,

$$4 \text{ is even } \wedge 7 \text{ is odd}.$$

Proof. We have that $4 = 2k$ with $k = 2$, thus is even. Next, $7 = 2k + 1$ with $k = 3$, so is odd. □

- If you need to prove a statement of the form

$$\forall x \in X, P(x),$$

that is a statement that starts with a forall, then *your proof will begin by "let $x \in X$"*. That is, when a statement starts with forall, you should pick an element of the thing we quantify over, and keep this element in your pocket. It is yours now. When you say "let $x \in X$", then you have an x, and this x belongs to X, so it enjoys all the properties of a being in X. Then, with this x in hand, we now prove $P(x)$ (where the x in $P(x)$ is that one x that we just picked in X). For instance,

$$\forall x \in \{t \in \mathbb{R} \mid t(t-2) = 0\}, x + 1 \text{ is odd }.$$

Proof. Let $x \in \{t \in \mathbb{R} \mid t(t-2) = 0\}$, then $x^2(x-2) = 0$, so $x = 0$, or $x = 2$, that is $x+1 = 1$ or $x + 1 = 3$, in both cases, $x + 1$ is odd. □

- If you need to prove a statement of the form

$$A \vee B,$$

then you can chose whichever you prefer, you can prove A, or you can prove B, you chose. Proving A, proves $A \vee B$, and proving B also proves $A \vee B$ (look at the truth table). However, in practice, it is not that simple, because A and B can both depend on the same parameter, where for some of its value A is true while B is false, and for the rest of the values, B is true while A is false. In that case (which is honestly the most frequent one), the trick is to do the following. We want to prove $A \vee B$. Suppose A is true, then great, we proved $A \vee B$, if not, then that means that A is false, and let us prove B. This works nicely, because now to prove B, we have a new hypothesis in our pocket, namely, that A is false. Let us see an example.

$$\forall n \in \mathbb{N}, n \text{ is even, or } n \text{ is odd}.$$

Proof. Let $n \in \mathbb{N}$ (let us not forget the previous point!!). Then we see that we will not be able to prove that n is even, or n is odd, because we do not have enough information on n. Therefore, we do the trick. Suppose n is not even, then (this is the definition of odd), n is odd. □

- If you need to prove a statement of the form

$$\exists x \in X, P(x),$$

then there is no general method. The only way is that you, mathematician, work and construct an x in X making $P(x)$ true. You will have to bring to existence a particular element of the set that satisfies the property P. In general, not all elements of X satisfy P, the statement $\exists x \in X, P(x)$ says that there exists at least one in X that does, and the role of the proof is to find it. For instance,

$$\exists n \in \mathbb{N}, n + 7 = 9.$$

Proof. By taking $n = 2$, we see that $2 + 7 = 9$. □

- There is a common upgrade to the exists quantifier. We write $\exists!$ to mean *there exists a unique*. To prove that there exists a unique, the standard way is to decompose the proof in two steps. First, we prove that there indeed exists something, and then, we assume that we have another thing, and prove that in fact it has to be the one we just exhibited from the existence. For instance,
$$\exists! n \in \mathbb{N}, n + 7 = 9.$$

  Proof. We already saw that $n = 2$ proves the existence. Suppose we have an $m \in \mathbb{N}$ such that $m + 7 = 9$, then $m = 9 - 7 = 2$, proving the uniqueness. □

- However, sometimes it is useful to prove that something is unique, *without proving it exists.* The way to do is to assume that we have two such things a, b, and then prove that $a = b$.

- If you want to prove a statement of the form
$$A \Rightarrow B,$$
then you *your proof will always begin by "assume A". You put the hypothesis A in your bag of hypothesis*, and you use it to move towards B. Maybe you will not use A right away, but remember when you are stuck that there is this A lying around! For instance,
$$\forall n \in \mathbb{N}\left((n \text{ is even}) \Rightarrow (n \mod 4) \in \{0, 2\}\right).$$

  Proof. Let $n \in \mathbb{N}$, and assume n is even. We do the euclidean division of n by 4, we have $n = 4q + r$ with $0 \leq r < 4$. By definition $(n \mod 4)$ is the reminder r, so we want to prove that $r \in \{0, 2\}$, which is to say $r = 0$ or $r = 2$. *By hypothesis, n is even*, so $n = 2k$ for some integer k, therefore we have $2k = 4q + r$. Reordering this expression, we obtain $r = 2(k - 2q)$, thus r is even, and $0 \leq r < 4$, hence $r = 0$ or $r = 2$. □

  Notice also in this proof how the application of the theorem of euclidean division added another hypothesis in our bag, namely that $0 \leq r < 4$, that we later used to proved that $r = 0$ or $r = 2$.

- There is another equivalent way to prove the statement $A \Rightarrow B$, it is to prove $\neg B \Rightarrow \neg A$. Sometimes it is easier to work this way. We call it the *contrapositive*.

- Beware that we said $A \Rightarrow B$ is the same thing as $\neg B \Rightarrow \neg A$, and that we did NOT say something else, that people wrongly do instead.

- If you want to prove $A \iff B$, then you prove first $A \Rightarrow B$, and second $B \Rightarrow A$. Of course, using contrapositive, you can also first prove $A \Rightarrow B$, then $\neg A \Rightarrow \neg B$.

- If you need to prove a statement of the form
$$\neg A,$$
then the standard way is to assume that A is true, so putting it in our bag of hypothesis, and then deduce a contradiction. For instance,
$$\neg(3 \text{ is even})$$

  Proof. Suppose 3 is even, then $3 = 2k$ for some integer k, but solving for k, we get $k = \frac{3}{2}$, which is not an integer, contradiction. □

  We see here that the contradiction we reached was provided by "k is an integer", and "k is not an integer". A lot of contradictions arise this way, whose general form is $A \wedge \neg A$. So to prove that there is a contradiction, a common way is to prove that something and the negation of the same thing are both true at the same time. This is why for instance $0 = 1$ is a contradiction. Indeed, we can prove $\neg(0 = 1)$, so if we also prove $0 = 1$, then this is the general contradiction shape with $A \equiv$ "$0 = 1$".

Here is some more ways to prove things.

Definition 2: If A is true, and $A \Rightarrow B$ is true, then B is true. This principle is valid, and is known as the *modus ponens*.

This is in fact why we prove theorem. Indeed, this principle means that we can use the theorems we proved in the past. Suppose we proved a very cool theorem, whose statement is given by A, then (as we proved it), A is true. Suppose now we want to prove some seemingly related statement B, then according to modus ponens, to prove B, it suffices to prove $A \Rightarrow B$. By the above, to prove $A \Rightarrow B$, we assume first that A is true (we already know it, but why not), and we add it in our bag of hypothesis, *so we can use it in our proof*, then we proceed to prove B. In practice, it means that we can use the theorem A during the proof of B. This is the deep reason why we can use theorem, and this is why the theorems we already proved are allowed to be in our bag of hypothesis, because of the modus ponens. For a more formal construction of this, see the cut-elimination procedure.

Definition 3: The *induction principle* tells us

$$(P(0) \wedge (\forall n \in \mathbb{N}, P(n) \Rightarrow P(n+1))) \Rightarrow \forall n \in \mathbb{N}, P(n).$$

This is the general induction we are already used to. Suppose we want to prove something of the shape

$$\forall n \in \mathbb{N}, P(n),$$

Then, if we can prove

$$P(0) \wedge (\forall n \in \mathbb{N}, P(n) \Rightarrow P(n+1)),$$

by modus ponens, and the induction principle, it will prove $\forall n \in \mathbb{N}, P(n)$. Now, to prove something $A \wedge B$, we first do A, then B. In our case, we first prove $P(0)$, this is the *base case*. Then we prove $\forall n \in \mathbb{N}, P(n) \Rightarrow P(n+1)$, this is the *inductive step*. According to our rules, to prove something like that, we start the proof by "let $n \in \mathbb{N}$", and we prove $P(n) \Rightarrow P(n+1)$. According to the rules again, to prove that, we assume $P(n)$ and need to prove $P(n+1)$. In short, the proof of $\forall n \in \mathbb{N}, P(n) \Rightarrow P(n+1)$ starts with: "Let $n \in \mathbb{N}$, and assume $P(n)$ is true", then we do math to prove $P(n+1)$. This is how we already knew induction.

Exercise 5: Use induction to prove the (very useful in practice) strong induction principle,

$$(P(0) \wedge (\forall n \in \mathbb{N}, (\forall m \leq n, P(m)) \Rightarrow P(n+1))) \Rightarrow \forall n \in \mathbb{N}, P(n).$$

Exercise 6: Use the strong induction principle to prove that any natural number $n > 1$ decomposes as a product of prime numbers.

1.3.3 Destructing proofs

Now that we saw how to prove something, it is time to do the opposite, and see how to use something we already proved. Indeed, hypothesis in our bag have certain shapes. We now see how to use deconstruct those shapes to apply them in our proofs. Thus, we place ourselves in the situation where we are doing a proof, and during this proof (as always) we have a bag of hypothesis.

- Suppose we have an hypothesis $A \wedge B$ in the bag. Then at any time during the proof, we can use A and we can use B. It thus means that having $A \wedge B$ in our bag of hypothesis is the same thing as having A, and also having B.

- Suppose we have an hypothesis $A \vee B$ in the bag. Then if we want to use this hypothesis, we need to do two times the proof. One using assuming that A is true, and another assuming that B is true. Indeed, when we have $A \vee B$, we do not know what the universe will give us A or B? Hence, we take care, and make sure everything will be ok whatever she gives us.

- Suppose we have an hypothesis $A \Rightarrow B$ in our bag, then if we can find some A in the bag, we get a B for free!

- Suppose we have an hypothesis $\forall x \in X, P(x)$ in our bag. Then whenever we have an element $x \in X$, we know for free that $P(x)$. For instance, let us prove $\neg(\forall n \in \mathbb{N}, n \text{ is even})$.

  Proof. Suppose $\forall n \in \mathbb{N}, n$ is even, then "$\forall n \in \mathbb{N}, n$ is even" is in our bag of hypothesis. Thus using it with $n = 1$, we obtain 1 is even. Contradiction. $\square$

- Suppose we have an hypothesis $\exists x \in X, P(x)$ in our bag. Then, we can deconstruct it into *some* $x_0 \in X$, and a guarantee that $P(x_0)$. However, we do *not* choose which x_0 we get. The hypothesis gives us some, but does not tell us which one (of the potentially many possible) it is. Therefore, the only interest of x_0 is that it satisfies P.

- Suppose we have an hypothesis $\neg A$ in our bag. Then if we can prove that A is true, we have a contradiction.

- However in practice, we often distribute the negation. Indeed, we have (in classical logic), the following rules:

  1. $\neg(P \vee Q) \iff (\neg P \wedge \neg Q)$,
  2. $\neg(PQ) \iff (\neg P \vee \neg Q)$,
  3. $\neg(\forall x \in X, P(x)) \iff (\exists x \in X, \neg P(x))$,
  4. $\neg(\exists x \in X, P(x)) \iff (\forall x \in X, \neg P(x))$.

  Using that, we see that we can *push* the negation into the very core of the statement, until we ultimately reach an atomic symbol, often a $=$, or a $\in$. For instance, instead of using directly the hypothesis $\neg(\forall n \mathbb{N}, \exists k \in \mathbb{N}, n = 2k)$, we will rather use

  $$\begin{aligned}\neg(\forall n \mathbb{N}, \exists k \in \mathbb{N}, n = 2k) &\iff \exists n \in \mathbb{N}, \neg(\exists k \in \mathbb{N}, n = 2k)\\ &\iff \exists n \in \mathbb{N}, \forall k \in \mathbb{N}, \neg(n = 2k)\\ &\iff \exists n \in \mathbb{N}, \forall k \in \mathbb{N}, n \neq 2k,\end{aligned}$$

  meaning that we have a $n_0 \in \mathbb{N}$, such that for each $k \in \mathbb{N}$, n_0 is not equal to $2k$, that is n_0 is not 0, nor 2, nor 4, nor 6, nor ... i.e. this just claims that there exists an odd number.

1.3.4 On some of the notations

As in any language, maths has its own abbreviation, and tacit rules. In the following list, we try to hint some of the ways we better construct mathematical sentences, what are their implicit assumptions, and how to deal with them.

- In most textbook, you will barely see the connective $\Rightarrow$, and somehow it is the one that is the most widely used. This is because the implication are all implicit, and it is the reader's job to reconstruct them. For instance, when we write

  $$\text{If } P, \text{ then } Q,$$

  or when we write

  $$\text{Suppose } P, \text{ then } Q,$$

  or when we write

  $$\text{Assume } P, \text{ then } Q,$$

  then we really mean $P \Rightarrow Q$, this is just a way to write it in english. Therefore, when you will have to prove a statement "if P, then Q", you will have to use the techniques described for the $\Rightarrow$ connective.

- The forall quantifier is also very often implicit, or written in plain english. For instance, we will write "check that for all x in X, $P(x)$". Worse, it is also very common to say "prove that if $x \in X$, then $P(x)$. This sentence translate formally as "$\forall x \in X, P(x)$". Therefore, we use the "if ... then ..." construction to really mean a forall. This means no harm, as forall is secretly a generalization of the implication. Maybe you will feel that once you get use to math enough.

- The exists $\exists$ quantifier also read as "there is", "for some", etc.

- When dealing with quantifiers, we also often say "such that". For instance $\exists y \in X, \forall x \in X, x = y$ would read as *there is some y in X* such that *for all x in X, $x = y$*. Exercise: convince yourself that "$\exists y \in X, \forall x \in X, x = y$" is in in fact the definition of $\exists! x \in X$.

- When we say "it suffices that A for B", it means that $A \Rightarrow B$. For instance, it suffices that n is divisible by 4 to be even. When we say "it is necessary that A for B", it means that $B \Rightarrow A$. For instance, it is necessary that n is even for n to be divisible by 4. If we say "A is sufficient and necessary for B", then it means $A \iff B$.

- When we write ”i.e.”, we mean that what comes before is the same thing as what come after, i.e. we mean $\iff$. For instance in a proof, we will find ”$n + 1 = m$, i.e. $n = m - 1$”. We also use ”that is” in place of ”i.e.” (in fact, i.e. is ”id est” in Latin, which translates to ”that is”).

- In these notes, we sometimes use the notation $:=$ to define a new object, like an abbreviation for the new thing. Thus, when we write $a = b$, it means that the thing a we were already working with is equal to the thing b we were already working with, but when we write $a := b$, we are giving a the new name a to the thing b we were already working with. Of course, after writing $a := b$, we also have $a = b$. This notation might seem useless, but is often convenient to remember that we are not claiming any mathematical truth, just a way to abbreviate something longer.

- This is not purely notational, but it is very useful to typecheck! If you are a physicist, then you know it already (we do not add meters with kilograms). If you program in a typed language then you know it already (a variable of type `int` cannot be a variable of type `string`). Math is a typed language. If something does not typecheck, then it is either an abuse of notation (try to make the notation more explicit if you are not familiar with it enough), or an error.

Unfortunately, this list is far from being exhaustive. Worse, each subfield of math uses its own abbreviations, notations, and way of saying things. Mathematic is a language that contains many many subvarieties, and it is often difficult, when learning a new subject, to familiarize oneself with its sometimes peculiar rules. Notice however that no domain of math says something false (as far as mathematician are aware). That is, everything will *always* have a precise and well defined meaning. It might juts take some time and effort to see it. Once you are use to it, you understand the power of notation, and the great things you can do with it.

2 Sets and functions

We arrive to our first objects of interest, sets and functions. We cannot really give a precise definition of what a set is, it is a very far reaching question, and we will content ourselves (that will be enough for our applications), of a very intuitive definition. I am saying here that we will base the entire building of mathematics on something that we do not define precisely. This is crazy, but in fact, this is also what is happening in general. However, we try to reduce the part that we leave to intuition to a smaller and more specific chunk that we then build around, and we can study further. Here, we will be shaky, and treat sets as primitive objects with given sets of rules and syntax, that we will familiarize with.

2.1 Sets

Definition 4: A *set* is something that can contain other things. If a set is name X, the syntax

$$x \in X$$

means a thing, named x, is inside the set X. We call x an *element* of X. If there are a bunch of things $x_1, \dots, x_n$, and we want to make a set out of them, we use the syntax:

$$\{x_1, \cdots, x_n\}.$$

We will see later more advanced construction to make better sets.

Remark 5: Notice a subtle difference. When writing:

$$\text{let } x \in X$$

we are doing an action, we are taking some x in X. When writing:

$$\text{so we have } x \in X$$

we are asserting, we are making a statement, that x is in X. For instance, in

$$\text{let } x \in \{0, 1\}, \text{ and call } y := x^2, \text{ then } y \in \{0, 1\}$$

when we wrote $x \in \{0, 1\}$, we made an action, and when we wrote $y \in \{0, 1\}$, we made a statement.

We can create sets of almost (and this "almost" might be the most important "almost" of math) anything. For instance, here are a bunch of classical sets, that can be defined more precisely from smaller sets, but again, we do not have time to enter into such details.

Example 6: Here are some examples of sets:

- $\emptyset$ is the *empty set*, the set with nothing in it.
- $\mathbb{N}$ is the set of *natural numbers*, its elements are the numbers $0, 1, 2, \cdots$.
- $\mathbb{Z}$ is the set of *integers*, its elements the numbers $\cdots, -2, -1, 0, 1, 2, \cdots$.
- $\{a, b\}$ is the set with two elements, called for the occasion a and b.

Exercise 7: Prove that $\forall x \in \emptyset, x = 0$, and that $\forall x \in \emptyset, x = 1$. What is happening here, did we just prove $0 = 1$?

Definition 7: Let X and Y be two sets. We say that that X is *included* in Y (or that X is a *subset* of Y), and we write $X \subseteq Y$, if

$$\forall x \in X, x \in Y.$$

We that that X and Y are *equal*, and we write $X = Y$, if $X \subseteq Y$ and $Y \subseteq X$.

Exercise 8: Prove that two sets X and Y are equal is to say

$$\forall x, x \in X \iff x \in Y.$$

Exercise 9: Prove that $\mathbb{N} \subseteq \mathbb{Z}$.

Exercise 10: Prove that the sets $\{x, x\}$ and $\{x\}$ are equal. Therefore, sets do not care about duplicates.

Let us see a very useful way to build sets from other. It goes formally by the name of *replacement* axiom, and is one of the foundational tool of mathematics. We present it a little bit informally here.

Definition 8: Let X be a set, and $\phi(x)$ a formula that depends on a parameter x allowed to vary in X. Then we define the set

$$\{x \in X \mid \phi(x)\}$$

to be the subset of X whose elements are precisely those of X that makes ϕ true.

Example 9: We can use the formula $\phi(n) := n \geq 0$ to define the natural numbers from the integers, indeed:

$$\mathbb{N} = \{n \in \mathbb{Z} \mid n \geq 0\}.$$

Exercise 11: Using replacement, define the set of even number from the set of natural numbers.

There are very important operations that we can do with sets. We have union, intersection, and complement. They are analogous to the logical operations we saw previously, respectively the $\vee$, the $\wedge$, and the $\neg$.

Definition 10: Let X, Y be two sets. We define the *intersection* of X and Y by

$$X \cap Y := \{x \mid x \in X \text{ and } x \in Y\}.$$

We define the *union* of X and Y by

$$X \cup Y := \{x \mid x \in X \text{ or } x \in Y\}.$$

Let $A \subseteq X$. We define the *complement* of A in X to be the set

$$X \backslash A := \{x \in X \mid x \notin A\}.$$

The syntax $x \notin A$ is just a shorthand for $\neg(x \in A)$, the same way $x \neq y$ is shorthand for $\neg(x = y)$.

Exercise 12: Prove that, for all sets X, Y, Z, we have

$$X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z).$$

Prove that for all $A, B \subseteq X$, we have

$$X \backslash (A \cup B) = (X \backslash A) \cap (X \backslash B).$$

Does it remind you of something?

Next, we have the product of sets. It is axiomatic in the theory, so we cannot define it from smaller primitive.

Definition 11: Let X, Y be sets. The *product* of X and Y is the set

$$X \times Y := \{(x, y) \mid x \in X, y \in Y\},$$

constituted of all the pairs (x, y) for $x \in X$ and $y \in Y$. If we have more sets, we can generalize this construction, for instance, with three sets X, Y, Z, we define

$$X \times Y \times Z := \{(x, y, z) \mid x \in X, y \in Y, z \in Z\}.$$

In that case, we rather call (x, y, z) a *tuple*, so a pair is a 2-tuple.

Remark 12: Notice that the pair (x, y) is different from the pair (y, x). However, the set $\{x, y\}$ is the set $\{y, x\}$. Pairs care about order, sets do not. Tuples care about multiplicity, sets do not, the tuple (x, x) is very different from the tuples (x) or (x, x, x) but the sets $\{x, x\}$, $\{x\}$, and $\{x, x, x\}$ are all the same.

Example 13: We have the following products:

- $\{0, 1\} \times \{a, b, c\} = \{(0, a), (0, b), (0, c), (1, a), (1, b), (1, c)\}$.
- For all set X, $X \times \emptyset = \emptyset$.

Exercise 13: If X has n elements, and Y has m elements, how many elements has $X \times Y$?

Another primitive of sets is the power set. The power set of a set is another set that contains all the subsets of the set with started with.

Definition 14: Let X be a set. We define its *power-set* 2^X (also written $\mathcal{P}(X)$) to be the set

$$2^X := \{A \mid A \subseteq X\}.$$

Example 15: We have the following power sets.

- $2^{\{0,1\}} = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\}$.
- $2^{\emptyset} = \{\emptyset\}$.

Notice that $\emptyset$ is very different from $\{\emptyset\}$. The former has zero elements, while the latter has one.

Exercise 14: If X has n elements, how many elements has 2^X?

2.2 Functions

Functions are the most fundamental objects of mathematics. A function can describe all sorts of thing, it is something that takes an input and produces an output. It is very convenient to declare inputs and outputs to be sets. Then a function will take any element from the input set, and return some element in the output set.

Definition 16: Let X, Y be two sets. A function f between X and Y is a thing that, for all $x \in X$, gives an element $f(x) \in Y$. We write

$$f : X \to Y.$$

X is called the *domain* of f, and Y is called the *codomain*. If we want to specify further the behavior of the function, we can use the following syntax

$$\begin{aligned} f : X &\to Y \\ x &\mapsto f(x) \end{aligned}$$

Note that a function is an asymmetric notion, the domain and the codomain are highly non-interchangeable.

Example 17: Here are a bunch of functions, and some various way of syntactically defining them (which are all equivalent, we often use the one that is more convenient).

- *A function that doubles its input.*

$$\begin{aligned} f : \mathbb{N} &\to \mathbb{N} \\ x &\mapsto 2x \end{aligned}$$

- *A function that says if a number is 57.* Let $f : \mathbb{N} \to \{\text{true}, \text{false}\}$ be the function such that $f(x) = \text{true}$ if $x = 57$, and $f(x) = \text{false}$ else.
- *The function that does nothing.* Let $f : X \to X$ be the function sending x to itself.

The last example is so fundamental that it deserves its own definition.

Definition 18: Let X be a set, we call id_X the function defined by

$$\begin{aligned} \mathrm{id}_X : X &\to X \\ x &\mapsto x. \end{aligned}$$

We call it the *identity* on X.

Definition 19: Let $f, g : X \to Y$ be two functions, we say that $f = g$ if for all $x \in X$, we have $f(x) = g(x)$. This principle is called *extensionnality*.

Remark 20: Beware that the domain and the codomain are part of the data of a function, that is, the function

$$\begin{aligned} f : \mathbb{N} &\to \mathbb{N} \\ x &\mapsto x + 1 \end{aligned}$$

and the function

$$\begin{aligned} g : \mathbb{N} &\to \mathbb{Z} \\ x &\mapsto x + 1 \end{aligned}$$

are not the same, even though they have the same behavior and produce the same outputs.

Exercise 15: Why we do not have $f = g$ in Remark 20, even though extensionality of Definition 19 seems satisfied?

Exercise 16: Where does the definition of equality for functions comes from? To answer this, let $f : X \to Y$ be a function. The *graph* of f is the subset Γ_f of $X \times Y$ defined by

$$\Gamma_f := \{(x, f(x)) \mid x \in X\}.$$

Let $f, g : X \to Y$ be two functions. Prove that $f = g$ if and only if $\Gamma_f = \Gamma_g$.

Remark 21: Notice, when we defined the graph $\Gamma_f := \{(x, f(x)) \mid x \in X\}$ in Exercise 16, we used a notation that was not allowed by Definition 8. Rather, the formal definition should have been

$$\Gamma_f := \{z \in X \times Y \mid \exists x \in X, z = (x, f(x))\},$$

but we overload notation a little bit, in a way that we will not make formally explicit, but that hopefully make sense. Here is another example:

$$\{f(x) \mid x \in X\} := \{y \in Y \mid \exists x \in X, y = f(x)\}.$$

We now define some important data associated to a function.

Definition 22: Let $f : X \to Y$ be a function. Let $A \subseteq X$, the *image* of A through f is the subset of Y defined by

$$f(A) := \{y \in Y \mid \exists x \in A, f(x) = y\}.$$

Let $B \subseteq Y$, the *preimage* of B through f is the subset of X defined by

$$f^{-1}(B) := \{x \in X \mid f(x) \in B\}.$$

Be careful that we overload the notation $f(-)$ with, in place of elements of its domain, subsets of its domain, therefore if $x \in X$ and $A \subseteq X$, then writing $f(x)$ and $f(A)$ are two very distinct things.

Exercise 17: Let $f : \mathbb{N} \to \mathbb{N}$ be the function defined by $f(x) = 2x$. What is the set $f(\mathbb{N})$? Let $\mathcal{O}$ be the subset of $\mathbb{N}$ constituted of odd numbers. What is the set $f^{-1}(\mathcal{O})$?

Exercise 18: Let $f : X \to Y$ be a function, let $A, A' \subseteq X$ and $B, B' \subseteq Y$. Prove some of the following identities (they are very useful to know, or at least remember they exist).

- $f(A \cup A') = f(A) \cup f(A)$.
- $f(A \cap A') \subseteq f(A) \cap f(A')$.
- $f^{-1}(B \cup B') = f^{-1}(B) \cup f^{-1}(B')$.
- $f^{-1}(B \cap B') \subseteq f^{-1}(B) \cap f^{-1}(B')$.
- $f^{-1}(f(X)) = X$.
- $f^{-1}(f(A)) \supseteq A$.
- $f(f^{-1}(Y)) = f(X)$.
- $f(f^{-1}(B)) \subseteq B$.

It is even a better exercise to try to come up with a example where the full equality fails, for instance provide a function where we do not have $f(A \cap B) = f(A) \cap f(B)$. For a more exhaustive list of these relations, see this Wikipedia page.

We can serialize functions, that is if we have a function $f : X \to Y$, and a function $g : Y \to Z$, we can consider the function that does f, then g.

Definition 23: Let $f : X \to Y$ and $g : Y \to Z$ be function such that the codomain of f is the domain of g. We define the function $g \circ f$ to be $g \circ f(x) := g(f(x))$. We call it the *composition* of f and g.

Composition is what we call a partial operation, not all functions can be composed: the output of the first one needs to match the input of the second one.

Example 24: Let $f : \mathbb{N} \to \mathbb{N}$ be the function defined by $f(x) = 2x$, and $g : \mathbb{N} \to \mathbb{N}$ be the function defined by $g(x) = x + 1$. Compute the functions $g \circ f$ and $f \circ g$.

Definition 25: Let $f : X \to Y$ be a function. An *inverse* of f is a function $g : Y \to X$ such that $f \circ g = \mathrm{id}_Y$ and $g \circ f = \mathrm{id}_X$.

Lemma 26: If $f : X \to Y$ has an inverse, then it is unique.

Proof. (**Lemma 26**) Let $f : X \to Y$ be a function, and suppose we have two inverses $g, g' : Y \to X$. Let $y \in Y$, we have by definition $y = f \circ g(y)$, thus applying g' both sides gives

$$g'(y) = g' \circ f \circ g(y),$$

and thus, as $g' \circ f = \mathrm{id}_X$, we have

$$g'(y) = g' \circ f \circ g(y) = g(y).$$

We conclude that for all $y \in Y$, $g(y) = g'(y)$, thus by extensionality, $g = g'$. □

Therefore, Lemma 26 allows us to use the notation f^{-1} for *the* unique inverse of f, when it exists. Be careful that f^{-1} might not always exists, and is in conflict with the notation $f^{-1}(B)$ (which is always well defined), and both do not mean the same thing.

Exercise 19: Provide a function that has an inverse, and a function that does not have an inverse.

Exercise 20 (Conflict of notation): Let $f : X \to Y$ be a function that admits an inverse, and let $B \subseteq Y$. Prove that

$$f^{-1}(B) = f^{-1}(B),$$

where the $f^{-1}(B)$ on the left is the preimage of B through f, and $f^{-1}(B)$ on the right the the image of B through the function f^{-1}.

We continue this section on functions by three very important notions.

Definition 27: Let $f : X \to Y$ be a function. We say that:

- f is *injective* if
$$\forall x, y \in X, f(x) = f(y) \Rightarrow x = y.$$
- f is *surjective* if
$$\forall y \in Y, \exists x \in X, f(x) = y.$$
- f is *bijective* if it is both injective and surjective.

Exercise 21: Prove that the function from $\mathbb{N}$ to $\mathbb{N}$ that adds 1 to a number is injective. Is it surjective? Prove that the function from $\mathbb{Z}$ to $\mathbb{Z}$ that add one to a number is bijective.

Exercise 22: Let $f : X \to Y$ be a function. Let $f' : X \to f(X)$ defined by letting $f'(x) := f(x)$. Prove that f' is surjective.

In fact (under the axiom of choice), being bijective is equivalent to having an inverse.

Proposition 28: Let $f : X \to Y$ be a function with $X \neq \emptyset$. We have

1. f is injective if and only if there exists $g : Y \to X$ such that $g \circ f = \mathrm{id}_X$.
2. f is surjective if and only if there exists $g : Y \to X$ such that $f \circ g = \mathrm{id}_Y$.
3. f is bijective if and only if it admits an inverse.

Proof. (**Proposition 28**) Suppose $f : X \to Y$ is injective. As $X \neq \emptyset$, select any $x_0 \in X$. To define $g : Y \to X$, take $y \in Y$, if there is some $x \in X$ such that $f(x) = y$, define $g(y) := x$, else define $g(y) := x_0$. By construction, for all $x \in X$, $g(f(x)) = x'$, where x' is such that $f(x') = f(x)$, by injectivity, this means $x = x'$, thus $g(f(x)) = x$. Conversely, suppose f has a left inverse g, and suppose $f(x) = f(y)$, then applying g both sides yields $g(f(x)) = g(f(y))$, that is $x = y$, so f is injective.

Next, suppose f is surjective. We construct $g : Y \to X$ as follow. For all $y \in Y$, we pick any element $x \in f^{-1}(\{y\})$, and we let $g(y) = x$. We can always pick such an element, as being surjective means precisely that for all $y \in Y$, the set
$$f^{-1}(\{y\}) = \{x \in X \mid f(x) = y\}$$
is non empty, so we can choose an element inside. (This last affirmation is quite subtle, to see that, Google "axiom of choice"). We then have, by construction, $f(g(y)) = y$, as $g(y) \in f^{-1}(\{y\})$. Conversely, if f admits a right inverse g, then for all $y \in Y$, $f(g(y)) = y$, so the element $g(y) \in X$ witnesses the existential quantifier for surjectivity.

Last, suppose f is bijective, then f is both surjective and injective, so by what we juts proved, there is a function $g : Y \to X$ such that $g \circ f = \mathrm{id}_X$, and a function $g' : Y \to X$ such that $f \circ g' = \mathrm{id}_Y$ (they need not to be the same so far). Let $y \in Y$, we have
$$g'(y) = g'(f \circ g(y)) = (g' \circ f)(g(y)) = g(y),$$
so $g = g'$, and thus f has an inverse. Conversely, suppose f has an inverse, then it is in particular a left inverse, so f is injective, and it is also a right inverse, so f is surjective, hence f is bijective. □

Lemma 29: Let $f : X \to Y$ and $g : Y \to Z$ be both bijections. Then $g \circ f$ is a bijection.

Proof. (**Lemma 29**) Call $h := f^{-1} \circ g^{-1} : Z \to X$, then one checks that $h \circ (g \circ f) = \mathrm{id}_X$ and $(g \circ f) \circ h = \mathrm{id}_Z$, therefore $g \circ f$ is a bijection. □

The following exercise emphasizes that the domain and codomain are *really* part of the data of a function.

Exercise 23: Determine if the following functions are injective, surjective, bijective, or none. We call $\mathbb{R}^+$ the set of real numbers greater or equal to 0.

- $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = x^2$.
- $f : \mathbb{R}^+ \to \mathbb{R}$ such that $f(x) = x^2$.
- $f : \mathbb{R} \to \mathbb{R}^+$ such that $f(x) = x^2$.

- $f : \mathbb{R}^+ \to \mathbb{R}^+$ such that $f(x) = x^2$.

The concept of bijection allows us to talk about finite sets, infinite sets, and cardinality. Let $n \in \mathbb{N}$ be a natural number. We write $[n]$ for the set $\{1, \cdots, n\}$, with convention that $[0] = \emptyset$.

The next lemma is very important. However, it does not really make sense to prove it formally, as we would need to make precise the foundations we are working with, but we did not. We are hence assuming it, and hope it makes sense intuitively.

Lemma 30 (Pigeonhole principle): Let $n, p \in \mathbb{N}$. Then

- There exists an injection $f : [n] \to [p]$ if and only if $n \leq p$.
- There exists a surjection $f : [n] \to [p]$ if and only if $p \leq n$.
- There exists a bijection $f : [n] \to [p]$ if anf only if $n = p$.

Definition 31: A set X is *finite* if there exists some $n \in \mathbb{N}$ such that X is in bijection with $[n]$. In that case, we say that X has cardinality n, we write $|X| = n$, and say that X has *cardinality* n. This is well defined, for if $f : X \to [n]$ is a bijection and $f' : X \to [p]$ is a bijection, then $f' \circ f^{-1} : [n] \to [p]$ is a bijection (by Lemma 29), so $n = p$ according to Lemma 30. Otherwise, we say that X is *infinite*.

Example 32: Let X, Y be sets.

- If X is finite and there is a surjection $X \to Y$, then Y is finite.
- If X is infinite and there is an injection $X \to Y$, then Y is infinite.
- If there is a bijection $X \to Y$, then X is finite if and only if Y is finite, and moreover in that case, the cardinality of X is the one of Y.

Remark 33: When X is a finite set, we often say $X = \{x_1, \ldots, x_n\}$ for some $n \in \mathbb{N}$. This notation is justified by the fact that X is finite, so we have some bijection $f : [n] \to X$, and x_i is just shorthand for $f(i)$. By injectivity, if $i \neq j$, then $x_i \neq x_j$, and by surjectivity, all elements of X are of the form x_i for some i.

2.3 Equivalence relations

Sometimes we have a set, and we would like to make things inside it *more* equal to each other. For instance, let us take the set $\mathbb{N}$. So far, if we take $n, m \in \mathbb{N}$, then $m = n$ if they are the same number. We would like to say that all even number are equal to each other, while all odd numbers are also equal to each other. The resulting set would then be a set with two elements, one for each parity. One element would represent all the even numbers, and the other one all the odd numbers. There is a very general way to do that called equivalence relation. We introduce it here because equivalence relations are pervasive in mathematics, and we will see them many times during this class. They are a little bit weird to talk about, and to define functions on them can be counter intuitive.

Definition 34: Let X be a set. A *binary relation* $\sim$ on X is a subset of $X \times X$. If $(x, y) \in \sim$, we simply write

$$x \sim y.$$

Definition 35: Let X be a set, and $\sim$ a binary relation on X. We say that

- $\sim$ is *reflexive* if
$$\forall x \in X, x \sim x.$$
- $\sim$ is *symmetric* if
$$\forall x, y \in X, x \sim y \Rightarrow y \sim x.$$
- $\sim$ is *reflexive* if for all $x \in X$, $x \sim x$.
$$\forall x, y, z \in X, (x \sim y \wedge y \sim z) \Rightarrow x \sim z.$$

A reflexive, symmetric, transitive relation is called an *equivalence relation*.

Example 36: The most famous equivalence relation of them all is simply the relation $=$. Indeed, $x = x$, if $x = y$, then $y = x$, and if $x = y$ and $y = z$, then $x = z$. It is good to think of equivalence relations as an extended equality.

Definition 37: Let $(X, \sim)$ be a set with an equivalence relation. We define the *equivalence class* of x, written $[x]$, or $\mathrm{cl}(x)$, to be the set

$$[x] := \{y \in X \mid x \sim y\}.$$

If $y \in [x]$, we say that y is a *representative* of $[x]$. Of course by reflexivity, x is a representative of $[x]$.

Lemma 38: Let $(X, \sim)$ be a set with an equivalence relation. We have that $x \sim y$, if and only if $[x] = [y]$.

Proof. (**Lemma 38**) Suppose $x \sim y$. Take any $z \in [x]$, then by definition $x \sim z$. By symmetry, also $y \sim x$, so by transitivity, $y \sim z$, hence $z \in [y]$. We proved $[x] \subseteq [y]$. Conversely, take $z \in [y]$, then $y \sim z$, and as $x \sim y$, by transitivity $x \sim z$ so $z \in [x]$, hence $[y] \subseteq [x]$, proving $[x] = [y]$. Conversely, if $[x] = [y]$, then as $y \in [y]$, also $y \in [x]$, so $x \sim y$. □

Definition 39: Let $(X, \sim)$ be a set with an equivalence relation. We define the *quotient* of X by $\sim$, written $X/\sim$, to be the set

$$X/\sim := \{[x] \mid x \in X\}.$$

We have a function $p : X \to X/\sim$, called the *canonical projection*, that sends x to $[x]$.

Remark 40: Suppose $(X, \sim)$ is a set with an equivalence relation, adn $f : X \to Y$ is a function such that for all $x, y \in X$, if $x \sim y$, then $f(x) = f(y)$. Then f defines a function $\bar{f} : (X/\sim) \to Y$ defined by $f([x]) = f(x)$. This does not depends on the choice of representative, for if $[x] = [y]$, by Lemma 38, $x \sim y$, so we have $f(x) = f(y)$, hence $\bar{f}([x]) = \bar{f}([y])$.

Exercise 24: Define $\sim$ on the natural number by letting $n \sim m$ if and only if m and n have same parity. Show that this is an equivalence relation, and that the set $\mathbb{N}/\sim$ has two elements. Show that the canonical projection $\mathbb{N} \to \{[0], [1]\}$ acts as the "mod 2" function, by seeing $[0]$ as 0, and $[1]$ as 1. This idea will be further generalized in the lesson on modular arithmetic.

Equivalence relations on a set X are precisely partitions of X. A partition of a set is a way of dividing it into disjoint pieces.

Definition 41: Let X be a set. A *partition* of X is a collection of nonempty subsets $\{U_i\}_{i\in I}$ of X, such that $\bigcup_{i\in I} U_i = X$, and if $i \neq j$, $U_i \cap U_j = \emptyset$.

Example 42:

- The set of odd number, and the set of even number form a partition of $\mathbb{N}$.
- The set $\{\{a, b\}, \{c\}\}$ forms a partition of $\{a, b, c\}$.
- The set $\{\{a, b\}, \{b, c\}\}$ is *not* a partition of $\{a, b, c\}$.
- The set $\{\{a, b\}, \{c\}\}$ is *not* a partition of $\{a, b, c, d\}$.

Proposition 43: Let X be a set. There is a bijection between the set of all equivalence relations on X, and the partition of X, given by the function that sends an equivalence relation to the set of all its equivalence classes.

Proof. (**Proposition 43**) Let $\sim$ be an equivalence relation on X. We let $\mathrm{cl}(X)$ be the set of equivalence classes of $\sim$, then $\bigcup_{C\in\mathrm{cl}(X)} C = X$. We show it forms a partition. Suppose $[x] \neq [x']$, and assume we have an $y \in [x] \cap [x']$, then by definition $x \sim y$ and $x' \sim y$, so by symmetry and transitivity, $x \sim x'$, so by Lemma 38, $[x] = [x']$, contradiction so $[x] \cap [x'] = \emptyset$, hence the elements of $\mathrm{cl}(X)$ are a partition of X. Conversely, suppose $\{U_i\}_{i\in I}$ is a partition of X. Define $\sim$ by

$$x \sim y \iff \exists i \in I, x \in U_i \wedge y \in U_i.$$

Let $x \in X$, as $\bigcup_{i \in I} U_i = X$, there exists some i such that $x \in U_i$, so $x \sim x$. The relation $\sim$ is also seen to be symmetric, and for transitivity, suppose $x \sim y$ and $y \sim z$, then there are some i, j such that $x, y \in U_i$ and $y, z \in U_j$. In particular $y \in U_i \cap U_j$, so $U_i \cap U_j \neq \emptyset$, hence by contrapositive, $i = j$, therefore $x, z \in U_i$ meaning $x \sim z$. □

Exercise 25: Let X be a set. Find the equivalence relation that corresponds the partition

$$\{\{x\} \mid x \in X\}.$$

Find the equivalence relation that corresponds the partition

$$\{X\}.$$

We give the following without proof (but it is an exercise, try first to do it for a two element partition).

Proposition 44: Let X be a finite set, and let $\{U_i\}_{i \in I}$ be a partition of X. Then

$$|X| = \sum_{i \in I} |U_i|.$$

We conclude by a canonical result that will appear here and there under similar forms during this course.

Theorem 45: Let $f : X \to Y$ be a function. There exists two (unique up to isomorphism) functions m, p such that p is surjective, m is injective, and $f = m \circ p$.

Proof. (**Theorem 45**) Let $\sim$ be the equivalence relation on X defined by $x \sim y$ iff $f(x) = f(y)$. We let $p : X \to X/\sim$ be the canonical projection, it is indeed surjective, and we let $m : (X/\sim) \to Y$ to be $\bar{f}$ as in Remark 40. If $m([x]) = m([y])$, then $f(x) = f(y)$, so $x \sim y$, hence by Lemma 38, $[x] = [y]$, so m is injective. Then we have $m \circ p(x) = m([x]) = f(x)$. □

3 Abstract algebra

Groups are mathematical structures that arise everywhere. Groups encode symmetries in structures, and symmetries are prevalent in math. As this class is computer-science oriented, we will first introduce monoids, which are objects slightly more general than groups, and that anyone in computer science already encountered. Typically, when we consider the regular expression $(ab)^*$, we are considering the free monoid on the alphabet $\{a, b\}$. A list is also an element in a certain monoid.

Before starting, we give a little bit of terminology that will be used throughout this chapter.

Definition 46: Let $f : X \to X$ be a function. We say that a subset $A \subseteq X$ is *closed* under f if whenever $a \in A$, $f(a) \in A$. Also, if we have $f : X \times X \to X$, and $A \subseteq X$, we say also that A is closed under f if for all $a, b \in A$, we have $f(a, b) \in A$.

3.1 Monoids

Definition 47: Let X be a set. A *binary operation* $\cdot$ on X is a function $\cdot : X \times X \to X$. Instead of writing $\cdot(x, y)$ for the application of $\cdot$ to (x, y), we typically write $x \cdot y$.

Example 48: This definition should not be new for you, it is just the abstract version of things we already know.

- The addition function $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is a binary operation.
- The multiplication function $\times : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is a binary operation.
- etc.

Definition 49: Let $(X, \cdot)$ be a set with a binary operation. We say that $(X, \cdot)$ is a monoid if

1. There exists a particular element $e \in X$, called the *neutral element*, such that:
$$\forall x \in X, e \cdot x = x = x \cdot e.$$
2. The binary operation is *associative*, that is:
$$\forall x, y, z \in X, x \cdot (y \cdot z) = (x \cdot y) \cdot z.$$
In that case, we are allowed to write $x \cdot y \cdot z$ to mean either $x \cdot (y \cdot z)$ or $(x \cdot y) \cdot z$, as they are equal.

Exercise 26: Prove that $(\mathbb{N}, +)$ is a monoid. Is $(\mathbb{R}, \times)$ a monoid? Prove that $(\mathbb{R}^*, \times)$ is a monoid, where by $\mathbb{R}^*$, we mean the set of real numbers with 0 removed. What is its neutral element?

Example 50: A very important monoid is the set with one element $\{e\}$. The binary law is (necessarily) defined by $e \cdot e = e$, and the neutral element is (necessarily) e. Check that this is indeed a monoid.

Lemma 51: Neutral elements are unique, that is, in a monoid $(X, \cdot, e)$, if there is an element $e' \in X$ such that
$$\forall x \in X, e' \cdot x = x = x \cdot e',$$
we have $e = e'$.

Proof. (Lemma 51) Suppose e' is another neutral element for all x we have
$$e' \cdot x = x,$$
so in particular letting $x = e$, we get $e' \cdot e = e$. Now, e is also neutral element, so for all x, we have $x \cdot e = x$, hence with $x = e'$, we get $e' = e' \cdot e$, so $e' = e$. □

Definition 52: Let $(X, \cdot, e)$ be a monoid with binary operation $\cdot$ and neutral element e. We say that X is *commutative* if
$$\forall x, y \in X, x \cdot y = y \cdot x.$$

Remark 53: Here are some common abuse of notation that we do in group theory. We say that X *is a monoid*, where we are supposed to say $(X, \cdot, e)$ is a monoid, the data of the binary law and the neutral element being part of the definition. As they are often implicit from the context, we tend to avoid it, and say simply that X is a monoid.
More often than not, when the monoid is commutative, we write its law $+$, and 0 its neutral element. Beware that these $+$ and 0 have a priori nothing to do with the $+$ and 0 of the natural numbers. It is just that this notation helps us remember that the monoid is commutative, as is the monoid $(\mathbb{N}, +, 0)$.
Also, when $(X, \cdot, e)$ is a monoid, we also like to write xy for $x \cdot y$, like we often write st for $s \times t$.

Definition 54: Let $(X, \cdot, e_X), (Y, \cdot, e_Y)$ be monoids, a function $f : X \to Y$ is a *morphism of monoids* if

$$\forall x, y, f(x \cdot y) = f(x) \cdot f(y),$$

and

$$f(e_X) = e_Y,$$

that is f preserves the monoid law, and it sends the neutral element to the neutral element.

Exercise 27: Prove that the exponential function $\exp : (\mathbb{R}, +) \to (\mathbb{R}^*, \times)$ is a morphism of monoids. Is the function

$$\begin{aligned} f : (\mathbb{N}, +, 0) &\to (\mathbb{N}, +, 0) \\ x &\mapsto x + 1 \end{aligned}$$

a morphism of monoids?

Definition 55: Let X be a monoid, and let $A \subseteq X$. We say that A is a *submonoid* if it contains the neutral element and is closed under the monoid law, that is $e \in A$, and for all $x, y \in A$, $xy \in A$.

Definition 56: Let $f : X \to Y$ be a morphism of monoids. We define the *kernel* of f to be the set

$$\ker(f) := f^{-1}(\{e_Y\}) = \{x \in X \mid f(x) = e_Y\} \subseteq X$$

and the *image* of f to be the set

$$\operatorname{im}(f) := \{f(x) \mid x \in X\} \subseteq Y.$$

(which is the same thing as the set-theoretical image of Definition 22).

The kernel and the image have in fact a structure of monoid, so when one has a morphism of monoid, one get two submonoids for free.

Lemma 57: Let $f : X \to Y$ be a morphism of monoid. Then $\ker(f)$ is a submonoid of X, and $\operatorname{im}(f)$ is a submonoid of Y.

Proof. (**Lemma 57**) By definition of a morphism of monoid, $f(e_X) = e_Y$, this means $e_X \in \ker(f)$. If $x, y \in \ker(f)$, then

$$f(xy) = f(x)f(y) = e_Y e_Y = e_Y,$$

so $xy \in \ker(f)$. This proves $\ker(f)$ is a submonoid of X.

Now for the image, again $f(e_X) = e_Y$, so $e_Y \in f(X) = \operatorname{im}(f)$. If $y, y' \in \operatorname{im}(f)$, then by definition there exist $x, x' \in X$ such that $f(x) = y$ and $f(x') = y'$, so $yy' = f(x)f(x') = f(xx')$, meaning that $yy' \in \operatorname{im}(f)$. This proves that $\operatorname{im}(f)$ is a submonoid of Y. □

Definition 58: Let X, Y be monoids. We define the product of X, Y to be the monoid whose underlying set is $X \times Y$, the neutral element is the couple (e_X, e_Y), and the law is defined pointwise, that is

$$(x, y) \cdot (x', y') := (x \cdot x', y \cdot y').$$

More generally, if $(X_i)_{i \in I}$ is a family of monoids, we define the product $\prod_{i \in I} X_i$ to be the monoid whose underlying set is $\prod_{i \in I} X_i$, whose neutral element is $(e_{X_i})_{i \in I}$, and whose law is defined pointwise by

$$(x_i)_{i \in I} \cdot (x'_i)_{i \in Y} := (x_i \cdot x'_i)_{i \in I}.$$

Exercise 28: Prove that if $(X_i)_{i\in I}$ is a family of monoids, indeed $\prod_{i\in I} X_i$ is a monoid.

Given a set X, how can we make it a monoid X^* such that elements of X are inside X^*? To see that, first suppose $X = \{x\}$, a set with an element. We construct a monoid X^*. As it is a monoid, it must have a neutral element, we call it e. We also want $x \in X^*$, so we put it there. So far our monoid X^* has elements e and x. But now, we can also consider $x \cdot x$, a priori, this element does not belong to X^*, but it should still exists, so we add it, and we call it xx for simplicity. Now our monoid has elements $\{e, x, xx\}$, and again we can consider $x \cdot (x \cdot x)$, or $(x \cdot x) \cdot x$. Those elements will have to be the same, so we add another element xxx to the monoid. And we continue forever. The end result will be that the elements of X^* are strings of x's, the monoid operation is concatenation, and the neutral element is the empty string. This indeed satisfies the axioms of monoid, as concatenating is associative, and concatenating the empty string to the left or the right of a word does not change it. Let us give a more general definition, when X is any set.

Definition 59: Let X be a set. We define X^* to be the monoid whose elements are finite strings $x_1 \dots x_n$ with $x_i \in X$, whose law is concatenation, and whose neutral element is concatenation. This indeed defines a monoid, as concatenation is associative, and concatenating with the empty string does not change a string.

Lemma 60: Let X be any set, and $(M, \cdot, e_M)$ be a monoid. Then any set-theoretical function $f : X \to M$ gives rise to a morphism of monoid $f^* : X^* \to M$ by letting

$$f^*(x_1x_2 \dots x_n) = f(x_1) \cdot f(x_2) \dots f(x_n),$$

where we allow $n = 0$, and we mean $f^*(e) = e_M$.

Proof. (**Lemma 60**) By definition, the map indeed maps the neutral element to the neutral element. If $x_1 \dots x_n, y_1 \dots y_m \in X^*$, then

$$f^*(x_1 \dots x_n y_1 \dots y_m) = f(x_1) \dots f(x_n) f(y_1) \dots f(x_m) = f^*(x_1 \dots x_n) f^*(y_1 \dots y_m).$$

□

3.2 Groups

Monoids are interesting objects, but if we ask moreover that every element has an inverse, a whole new world appears, it is the one of groups.

3.2.1 General theory of groups

Definition 61: A *group* $(G, \cdot, e)$ is a monoid together with a function $-^{-1} : G \to G$ that sends $x \in G$ to x^{-1}, called the *inverse* of x, and such that

$$\forall x \in G, x \cdot x^{-1} = e = x^{-1} \cdot x.$$

A group is *abelian* (or commutative) if its underlying monoid is commutative, see Definition 52.

Remark 62: The Remark 53 also applies for groups, for instance we will often write $+$ for the law of an abelian group. Furthermore, we extend this notation to $-x$ to mean x^{-1} in the case where the group is abelian.

Example 63:

- $(\mathbb{Z}, +, 0, -)$ is an abelian group.
- $(\mathbb{R}^*, \times, 1, x \mapsto 1/x)$ is an abelian group.
- Let X be a set, call $\mathrm{Bij}(X)$ the set of all bijective function $f : X \to X$. This set is a (non-abelian) group with composition. What is the inverse of a function? What is the neutral element?
- We let $\mathbb{Z}_2$ to be the set $\{0, 1\}$, with the binary law being addition modulo 2, so for instance $0 + 1 = 1$, and $1 + 1 = 0$. This is a group in which each element is its own inverse.

- More generally, we let $\mathbb{Z}_n$ to be the set $\{0, \ldots, n-1\}$ with law being addition mod n, so to compute $p+q$, we first do it as in $\mathbb{Z}$, and ten take the reminder modulo n.

- Let $\mathbb{U}_n$ be the set of nth roots of unity, that is

$$\mathbb{U}_n = \{\exp(\frac{2\pi k}{n}) \mid k \in \{0, \ldots, n-1\}\},$$

with law being multiplication. This is a group, which is in fact the same as $\mathbb{Z}_n$, more on this when we will do modular arithmetic.

- Dihedral groups are example of finite groups that are not abelian. The nth dihedral group represents the rotational and mirror symmetries of the regular $n$gone, so for instance the third dihedral groups is the symmetries of the equilateral triangle. It has 6 elements,

$$D_3 := \{r_0, r_1, r_2, s_0, s_1, s_2\},$$

where r_i means *rotate the triangle by* $i \times 120°$, (so r_0 is the neutral element) and s_i means *reflect the triangle along the ith median*. The law of groups is *doing the symmetry one after another, from the rightmost to the leftmost*, so for instance $s_2 s_1 = r_1$, as if we take a triangle, reflect it along the first median, then the second median, it amounts to rotating it by 120°. *TODO : add pictures, it would be better*. Check that this the group law is indeed not commutative.

- If you are interested in non-abelian finite groups, check out the classification of finite simple groups. It consist of finding all finite groups, that enjoy the property of being *simple* (it does not mean at all that the group is simple, but rather that its subgroups behave in some manageable way). It took humanity fifty years and tens of thousands of pages to prove that we found them all.

Remark 64: What we said previously about monoid, can often be extended to groups. In particular, the neutral element is unique.

Lemma 65: Let G be a group, and let $x \in G$. Suppose we have $x' \in G$ such that $x' \cdot x = e$ or $x \cdot x' = e$, then $x' = x^{-1}$.

Proof. (Lemma 65) Suppose for instance $x' \cdot x = e$, then multiplying by x^{-1} both sides yields

$$x' \cdot x \cdot x^{-1} = e \cdot x^{-1},$$

which simplifies to $x' = x^{-1}$. □

Remark 66: Beware that it is very important that we know G to be a group for Lemma 65. Indeed, try to find an example of monoid where $x \cdot y = e$, but $y \cdot x \neq e$.

Exercise 29: Let $(G, \cdot, e)$ be a group, and let $x, y \in G$, prove that we always have the following identities (and remember that they exists):

- $e^{-1} = e$,
- $(xy)^{-1} = y^{-1}x^{-1}$.
- $(x^{-1})^{-1} = x$.

Definition 67: Let G be a group. A subgroup of G is a subset $H \subseteq G$ that contains the neutral element, is closed under the group operation, and under taking inverses.

In fact, the definition of subgroups contains redundant parts, to check that a subset is a subgroup, there is less work that we need to do.

Lemma 68: Let G be a group, and $H \subseteq G$. Then H is a subgroup if and only if it is non-empty, and whenever $x, y \in H$, we have $x \cdot y^{-1} \in H$.

Proof. (Lemma 68) Suppose H is a subgroup of $(G, \cdot, e)$, then it contains the neutral element, so is not empty. If $x, y \in H$, then by closure under taking the inverse, $y^{-1} \in H$, and as H is closed under the group law, we get $x \cdot y^{-1} \in H$. Conversely, as H is non-empty, we can chose $x \in H$, and by hypothesis, $x \cdot x^{-1} \in H$, that is $e \in H$. Now take any $x \in H$, we have $e \cdot x^{-1} \in H$, so $x^{-1} = e \cdot x^{-1} \in H$, so H is closed under taking inverses. Finally, let $x, y \in H$, we have by what we just shown that $y^{-1} \in H$, now applying the hypothesis with x and y^{-1}, we get $x \cdot (y^{-1})^{-1} \in H$, that is, $x \cdot y \in H$. □

Example 69: Here are some examples of subgroups.

- Let $n \in \mathbb{Z}$, call $n\mathbb{Z}$ the set
$$n\mathbb{Z} := \{nk \mid k \in \mathbb{Z}\},$$
for instance $2\mathbb{Z}$ is the set of even integers. Then $n\mathbb{Z}$ is a subgroup of $\mathbb{Z}$.

- If n divides p, then $p\mathbb{Z}$ is a subgroup of $n\mathbb{Z}$ (thus generalizing the previous example, as $\mathbb{Z} = 1\mathbb{Z}$)

- The dihedral groups D_3 of Example 63 has $\mathbb{Z}_3$ as a subgroup, for instance $\{r_0, r_1, r_2\} \subseteq D_3$ is a subgroup that is isomorphic to $\mathbb{Z}_3$.

Exercise 30: Let G be a group, and let $(G_i)_{i \in I}$ be a family of subgroup of G, check that the intersection $\bigcap_{i \in I} G_i$ is again a subgroup of G.

We now dive into the world of morphism of groups. We used the word *isomorphic* above, it means that two groups are the same, in a very precise and powerful way.

Definition 70: Let $f : G \to H$ be a function between two groups, we say that f is a *morphism of groups* if $f(e_G) = e_H$, for all $x, y \in G$, $f(xy) = f(x)f(y)$, and for all $x \in G$, $f(x^{-1}) = f(x)^{-1}$. That is, f is a function that respect the structure of a group, namely the neutral element, the multiplication, and taking the inverse.

Exercise 31: Prove that if $f : G \to H$ is a function between two groups such that for all $x, y \in G$, $f(xy) = f(x)f(y)$, then it is the case that $f(e_G) = e_H$, and that $f(x^{-1}) = f(x)^{-1}$. Therefore, Definition 70 is redundant, and it suffices to check that $f(xy) = f(x)f(y)$ for f to be a morphism of groups.

Example 71:

- Let G be a group, let 0 be the group with one element. There is a unique morphism $0 \to G$, and a unique morphism $G \to 0$.

- Let G be a group and $H \subseteq G$ be a subgroup, then the function $\iota : H \to G$ that sends $x \in H$ to itself is a morphism of group.

- Let $f : \mathbb{Z} \to \mathbb{Z}_2$ be the function sending even elements to 0 and odd elements to 1. It is a morphism of group.

Remark 72: As morphisms of groups are in particular functions, it makes sense to say that a morphism of group is injective, surjective, or bijective. In the last case, we say that f is an *isomorphism*. So an isomorphism between two groups is a morphism of group that is bijective, i.e. that is both injective and surjective. If $f : G \to H$ is an isomorphism of groups, we write $G \simeq H$, and it means that the two groups are the same, up to renaming the elements. Any group-theoretic thing that one can say about a group will be true for any isomorphic group to it.

Like in the case for monoids, we define the kernel and the image. Notice that the definition is the same as Definition 56.

Definition 73: Let $f : G \to H$ be a morphism of group. The *kernel* of f is the set

$$\ker(f) := f^{-1}(\{e_H\}) = \{x \in X \mid f(x) = e_H\} \subseteq G$$

and the *image* of f is the set
$$\mathrm{im}(f) := \{f(x) \mid x \in G\} \subseteq H.$$
(which is the same thing as the set-theoretical image of Definition 22).

Lemma 74: Let $f : G \to H$ be a morphism of groups. Then $\ker(f)$ is a subgroup of G, and $\mathrm{im}(f)$ is a subgroup of H.

Proof. (**Lemma 74**) We already know that $\ker(f)$ and $\mathrm{im}(f)$ are submonoids, according to Lemma 57. To check that it is a subgroup, it remains to see that they are closed under taking inverses. For that, let $x \in \ker(f)$, then $f(x^{-1}) = f(x)^{-1}$, because f is a morphism of groups, and by hypothesis, $f(x) = e_H$, therefore, $f(x^{-1}) = {e_H}^{-1} = e_H$. This proves $\ker(f)$ is a subgroup of G. Finally, take $y \in \mathrm{im}(f)$, that is $y = f(x)$ for some $x \in G$. We have $y^{-1} = f(x)^{-1} = f(x^{-1})$, meaning that $y^{-1} \in \mathrm{im}(f)$, so $\mathrm{im}(f)$ is a subgroup of H. □

Images and kernels are very convenient objects. Knowing them tells us when a morphism is injective, surjective, or an isomorphism.

Proposition 75: Let $f : G \to H$ be a morphism of groups. Then

- f is injective if and only if $\ker(f) = \{e_G\}$;
- f is surjective if and only if $\mathrm{im}(f) = H$;
- f is an isomorphism if and only if $\ker(f) = \{e_G\}$ and $\mathrm{im}(f) = H$.

Proof. (**Proposition 75**) Suppose f is injective, and let $x \in \ker(f)$, then $f(x) = e_H = f(e_G)$, so by injectivity, $x = e_G$. Conversely, suppose $\ker(f) = \{e_G\}$, and take $x, y \in G$ such that $f(x) = f(y)$. Multiplying by $f(y)^{-1}$ both sides, we get

$$e_H = f(y)f(y)^{-1} = f(x)f(y)^{-1} = f(x)f(y^{-1}) = f(xy^{-1}),$$

so $xy^{-1} \in \ker(f)$, and $\ker(f) = \{e_G\}$ by hypothesis, this means $xy^{-1} = e_G$, so multiplying by y, we find $xy^{-1}y = y$, i.e. $x = y$, and f is injective. Next, f surjective is by definition to say that $\mathrm{im}(f) = Y$, and finally, an isomorphism is a bijective group morphism, that is an injective and surjective group morphism. □

Example 76: Recall the examples of Example 71.

- Let G be a group. The map $0 \to G$ is injective, and the map $G \to 0$ is surjective. If G is not the zero group, then the composition $G \to 0 \to G$ is neither injective, nor surjective.
- If $H \subseteq G$ is a subgroup, then the map $\iota : H \to G$ is injective.
- The map $f : \mathbb{Z} \to \mathbb{Z}_2$ of Example 71 is surjective.

Definition 77: Let G be a group, let $H \subseteq G$ be a subgroup, and $x \in G$. We define the *left coset* of H with x, to be the set

$$xH := \{x \cdot h \mid h \in H\}.$$

Similarly, the *right coset* of H with x is

$$Hx := \{h \cdot x \mid h \in H\}.$$

Definition 78: Let $x \in G$ be an element of a group. Let $n \in \mathbb{Z}$, we define by induction x^n. If $n = 0$, we let $x^0 = e$. Suppose x^n has been defined, we define x^{n+1} to be $x^n \cdot x$. Suppose $n < 0$, then we define $x^n = (x^{-n})^{-1}$.

Exercise 32: Check that this definition satisfies the usual laws of powers, that is

- $x^n x^m = x^{n+m}$
- $(x^n)^m = x^{nm}$
- $x^{-n} = (x^n)^{-1}$

Argue that defining x^n for some x is in fact the same thing as defining a morphism of group $\mathbb{Z} \to G$ that sends 1 to x.

Remark 79: If G is abelian, then according to Remark 62, we will often write $+$ for the group law, and in that case, we will write nx for x^n, with Exercise 32 giving for instance $nx + mx = (n+m)x$, as expected.

Definition 80: Let G be a group, and let $S \subseteq G$ be a subset. We let $\langle S \rangle$ be the smallest subgroup of G containing S.

Proposition 81: Let $\mathcal{S} := \{H \subseteq G \mid H \text{ subgroup and } S \subseteq H\}$, then

$$\langle S \rangle = \bigcap_{H \in \mathcal{S}} H.$$

Proof. (**Proposition 81**) We have $\langle S \rangle \subseteq \bigcap_{H \in \mathcal{S}} H$, as $\bigcap_{H \in \mathcal{S}} H$ is a subgroup of G (see Exercise 30) containing S, and $\langle S \rangle$ is the smallest of them all. Conversely, $\langle S \rangle \in \mathcal{S}$, so $\bigcap_{H \in \mathcal{S}} H \subseteq \langle S \rangle$. □

Definition 82: Let $x \in G$, we denote by $\langle x \rangle$ the subgroup $\langle \{x\} \rangle$, and we call it the *subgroup generated by x*.

3.2.2 Finite group theory

Armed with those definition, we are ready to specialize ourself to groups that are finite.

Definition 83: A *finite group* is a group that is finite set. We call its cardinality the *order* of the group.

Exercise 33: For all $n \in \mathbb{N}$, the group $\mathbb{Z}_n$ is a finite groups of order n. The group with one element is the unique finite group of order 1. There are no group of order 0, as such a group would be empty, thus would not have a neutral element.

Definition 84: Let G be a group, we call the order of an element $x \in G$, if it exists, the smallest positive natural number such that $x^n = e$.

Lemma 85: In a finite group, the order of an element always exists.

Proof. (**Lemma 85**) Let G be a finite group, and $x \in G$. Let $\alpha : \mathbb{N}^* \to G$ be the function sending n to $\alpha(n) := x^n$. As G is finite, α is not injective (otherwise the set $\alpha(\mathbb{N})$ would be an infinite subset of G, contradicting that G is finite, see Example 32). Therefore there exists some $n \neq m$ such that $x^n = x^m$. Without loss of generality, suppose $n < m$. Then, multiplying by $(x^n)^{-1}$ both sides, we get $e = x^m (x^n)^{-1} = x^{m-n}$ with $0 < m - n$. Therefore, there exist some $n \in \mathbb{N}$ with $x^n = e$, and thus, we can consider the minimum of them. □

Remark 86: Therefore, as the order of an element is always defined for finite groups, we define the function

$$\begin{aligned} \operatorname{ord} : G &\to \mathbb{N} \\ x &\mapsto \operatorname{ord}(x) \end{aligned}$$

that sends an element to its order, and we extend this notation to groups themselves by letting $\operatorname{ord}(G)$ be the order of the group G.

The order of an element of a group is related to the order of the group itself, we have that $\operatorname{ord}(x) \mid \operatorname{ord}(G)$ ($\mid$ means divides). In order to prove that, we will prove a more general theorem saying that the order of any subgroup of G divides the order of G. This is known as Lagrange's theorem.

Definition 87: Let $H \subseteq G$ be a finite group and a subgroup. We define $[G : H]$, the *index* of H in G to be the cardinality of the (finite) set

$$\{xH \mid x \in G\}.$$

That is, the index of a subgroup is its number of distinct left cosets.

Theorem 88 (Lagrange): Let G be a finite group and and $H \subseteq G$ be a subgroup. We have

$$\operatorname{ord}(G) = [G : H] \operatorname{ord}(H).$$

In particular the order of H divides the order of G.

Proof. (**Theorem 88**) We define the binary relation $\sim$ on G by letting

$$x \sim y \iff y^{-1}x \in H.$$

We prove that it is an equivalence relation. It is reflexive, as $e = xx^{-1} \in H$. If $x \sim y$, then $xy^{-1} \in H$, so (H is a subgroup), $(y^{-1}x)^{-1} \in H$, that is, $yx^{-1} \in H$, meaning that $y \sim x$, hence the relation is symmetric. Finally, if $x \sim y$ and $y \sim z$, then $y^{-1}x \in H$ and $z^{-1}y \in H$, so $z^{-1}yy^{-1}x \in H$, that is $z^{-1}x \in H$, i.e. $x \sim z$. This proves $\sim$ is an equivalence relation. We denote by $[x]$ the equivalence class of x under $\sim$. We prove that in fact $[x] = xH$. For that, let $y \in [x]$, then $x \sim y$, so $y \sim x$ that is $x^{-1}y \in H$, which is to say that there exists some $h \in H$ such that $x^{-1}y = h$, that we can rewrite to $y = xh$, so $y \in xH$. Conversely, if $y \in xH$, then $y = xh$ for some $h \in H$, so $x^{-1}y \in H$, i.e $y \in [x]$. The equivalence classes of $\sim$ are thus precisely the left cosets of H.

We now prove that all equivalence classes have the same cardinality, equal to the order of H. For that, we let $x \in G$, and by Example 32, it suffices to establish a bijection $H \to xH$. We construct such function f by sending $h \in H$ to $f(h) = xh$. We see that it is a bijection by considering $g : xH \to H$ sending any element $y \in xH$ to $x^{-1}y$. Then we compute $g \circ f(h) = x^{-1}xh = h$, so $g \circ f = \mathrm{id}_H$, and $f \circ g(y) = xx^{-1}y = y$, so $f \circ g = \mathrm{id}_{xH}$. We thus established $|[x]| = |xH| = \operatorname{ord}(H)$.

Now we let $\{x_1H, \dots, x_kH\}$ the all the left cosets of H, by definition there are $[G : H]$ many of them, that is $k = [G : H]$. Recall that we proved that these cosets are exactly the equivalence classes of $\sim$, so we apply by the formula of Proposition 44:

$$\operatorname{ord}(G) = \sum_{i=1}^{k} |[x_i]| = \sum_{i=1}^{k} |H| = (\sum_{i=1}^{k} 1)|H| = k|H| = [G : H] \operatorname{ord}(H).$$

$\square$

Lemma 89: Let $x \in G$ be an element of a group of order n, then $\langle x \rangle$, the subgroup generated by x, is

$$\{e, x, x^2, \cdots, x^{n-1}\}.$$

In particular, $\operatorname{ord}(x) = \operatorname{ord}(\langle x \rangle)$.

Proof. (**Lemma 89**) First, we have $\{e, x, x^2, \cdots, x^{n-1}\} \subseteq \langle x \rangle$, as $e \in \langle x \rangle$ as it is a subgroup, and $x \in \langle x \rangle$ by definition of the subgroup generated, thus, by closure under the group operation, also $x^n \in \langle x \rangle$. For the converse, it suffices to check that $\{e, x, x^2, \cdots, x^{n-1}\}$ is indeed a subgroup. Indeed, if it is the case, it would be a group containing x, so would contain $\langle x \rangle$.

Call $X = \{e, x, x^2, \cdots, x^{n-1}\}$, let $x^a, x^b \in X$, and do the euclidean division $a - b = qn + r$, with $0 \leq r < n$. We then have, using 32,

$$x^a(x^b)^{-1} = x^{a-b} = x^{qn+r} = x^{qn}x^r = (x^n)^q x^r = e^q x^r = x^r.$$

As $0 \leq r \leq n$, we indeed have $x^a(x^b)^{-1} = x^r \in X$. This proves X is a subgroup.

To conclude that $\operatorname{ord}(x) = \operatorname{ord}(\langle x \rangle)$, we need to show that $|\{e, x, x^2, \cdots, x^{n-1}\}| = n$, which is not immediate. To see the subtle problem, recall that with our notation, the set $\{a, a\}$ has only one element! Thus, we need to see that when $0 \leq i, j < n$ such that $i \neq j$, $x^i \neq x^j$. We can assume that $i < j$, and $x^i = x^j$ is to say that $x^{j-i} = e$. As the order of x is the smallest number k such that $x^k = 0$, we have $n \leq j - i$, meaning that $n \leq n + i \leq j$, but we assumed $j < n$. This is a contradiction, so $x^i \neq x^j$. $\square$

Corollary 90: Let G be a finite group, and let $x \in G$, then $\operatorname{ord}(x) \mid \operatorname{ord}(G)$.

Proof. (**Corollary 90**) Let $\langle x \rangle$ be the subgroup generated by $\{x\}$, and let $n = \operatorname{ord}(x)$. We apply Theorem 88 and get $\operatorname{ord}(\langle x \rangle)$ divides $\operatorname{ord}(G)$, so by Lemma 89, $\operatorname{ord}(\langle x \rangle) = \operatorname{ord}(x) = n$ divides $\operatorname{ord}(G)$. $\square$

Exercise 34: Let G be a group. We say that a subgroup is a *trivial subgroup* if it is $\{e_G\}$ or G itself. Prove that a group has only trivial subgroups if and only if $G \simeq \mathbb{Z}_p$ for some prime number p. Deduce that if G and H are two groups of order p a prime number, then $G \simeq H$.

Exercise 35: Let G be a group such that for all $x \in G$, $\operatorname{ord}(x) = 2$ or $x = e$. Prove that G is abelian.

Exercise 36: Let G be a finite group, and let $x \in G$. Prove that $x^k = e$ if and only if $\text{ord}(x)$ divides k.

Exercise 37: Let $f : G \to H$ be a morphism of group. Prove that for all $x \in G$, $\text{ord}(f(x)) \mid \text{ord}(x)$. Prove that if f is moreover injective, then $\text{ord}(f(x)) = \text{ord}(x)$.

3.2.3 Cyclic groups

Previously, we used several times the idea that the set $\{0, \dots, n-1\}$, with the addition modulo n, was a group. For instance, this fact is hidden in the proof of Lemma 89. We also introduced them without real proof in Example 63. We now construct these groups very formally, using the important idea of quotient groups, that we only develop in the case of abelian groups, although it is completely possible to make it more general by introducing *normal* subgroups. Then, we use group theoretical tools to generalize the fact that if p is prime, and $1 \leq a < p$, then a^p is equal to 1 modulo n. Recall that the RSA algorithm is based on this fact.

Remark 91: Let A be an abelian group, and $B \subseteq A$ be a subgroup. We write $a + B$ for the left coset, and we notice that, as A is abelian (this is false in general), the right coset $B + a$ is equal to the left coset $a + B$. Thus for abelian groups, we will speak of cosets, without specifying left or right.

Definition 92: Let A be an abelian group, and $B \subseteq A$ be a subgroup. We define A/B to be the set of cosets $\{a + B \mid a \in A\}$. We also define the function $[-] : A \to A/B$ sending a to $[a] = a + B$.

Lemma 93: The set A/B is a group, called the *quotient group*, whose laws are induced by the one of A, and the map $[-] : A \to A/B$ is a group morphism, called the canonical projection. More precisely, the laws of the quotient group are such that

- The neutral element is $[0] = B$,
- The inverse is defined with $-[a] = [-a]$,
- The addition is defined by $[a] + [a'] = [a + a']$.

Proof. (Lemma 93) We need to prove that the definition of the laws do not depend on the representative we chose. Suppose $[a] = [a']$, then we show that $[-a] = [-a']$, but this follows from the definition of a coset. Indeed, if $c \in [-a]$, then $c = -a + b = -(a - b) = -(a + (-b))$, so $-c \in [a]$, thus $-c \in [a']$, hence by a similar reasoning, $c \in [-a']$. We thus have $[-a] \subseteq [-a']$, and as the role of a and a' is symmetrical, it implies $[-a] = [-a']$. Next, suppose $[a] = [a']$, we aim to show that $[a + c] = [a' + c]$. We have (crucially, we use here commutativity) $a + c + B = c + a + B = c + (a + B) = c + (a' + B) = a' + c + B = [a' + c]$. Finally, one checks that $[0]$ is indeed the neutral element. □

Lemma 94: The kernel of the canonical projection $A \to A/B$ is B.

Proof. (Lemma 94) Suppose $[a] = [0]$, then $a \in [0] = 0 + B = B$. Conversely, for all $b \in B$, $b + B = B$, so $[b] = [0]$. □

The next proposition allows us to define maps out of a quotient. It is very reminiscent of Remark 40, because we are in fact doing the same thing, just in another mathematical realm. Remark 40 happened for quotients on sets, and the next result is for quotients on abelian groups. The underlying construction is however the same, and if you are interested by how those universal constructions happen, you should learn about *category theory*.

Proposition 95: Let A, C be abelian groups, let $B \subseteq A$ be a subgroup. Then for all group morphism $f : A \to C$ such that for all B in C, $f(b) = 0$, there exists a unique group morphism $\bar{f} : A/B \to C$ such that $f(a) = \bar{f}([a])$.

Proof. (Proposition 95) We let $\bar{f} : A/B \to C$ defined by $\bar{f}([a]) = f(a)$. Suppose $[a] = [a']$, then $a + b = a' + b'$ for some $b, b' \in B$, thus $a - a' = b' - b \in B$. Then

$$\bar{f}([a]) - \bar{f}([a']) = f(a) - f(a') = f(a - a') = 0,$$

as $a - a' \in B$, and f sends elements of B to 0 by hypothesis. This proves that f is unique and well defined. To show that this is a morphism of group, we see that $\bar{f}([0]) = f(0) = 0$, and

$$\bar{f}([a + a']) = f(a + a') = f(a) + f(a') = \bar{f}([a]) + \bar{f}([a']).$$

$\square$

Remark 96: Let $f : A \to C$ be a morphism of abelian groups. Then by Proposition 95, if $B \subseteq A$ is a subgroup such that $B \subseteq \ker(f)$, we can consider the map $\bar{f} : A/B \to C$, and we often call this map f again. We say that the map $f : A \to C$ *passes to the quotient*.

The next theorem is the analogue for groups of Theorem 45.

Theorem 97 (First isomorphism theorem): Let $f : A \to B$ be a morphism of abelian groups, then

$$A/\ker(f) \simeq \operatorname{im}(f),$$

where the isomorphism is given by $\bar{f}$ of Proposition 95.

Proof. (**Theorem 97**) Suppose $\bar{f}([a]) = 0$, then $f(a) = 0$, so $a \in \ker(f)$, thus $[a] = [0]$. This proves $\bar{f}$ is injective. Let $y \in \operatorname{im}(f)$, then $y = f(a)$ for some $a \in A$, hence $\bar{f}([a]) = f(a) = y$. This proves $\bar{f}$ is surjective. Thus $\bar{f} : A/\ker(f) \to \operatorname{im}(f)$ is an isomorphism. $\square$

After those abstract consideration, we define more concrete objects, called the cyclic group. We already saw them, and we also see them every day. The twenty-four hours of the days are isomorphic the $\mathbb{Z}_2 4$, the twenty-fourth cyclic group.

Definition 98: Let $n \in \mathbb{N}$, the *nth cyclic group* $\mathbb{Z}_n$ is the quotient $\mathbb{Z}/n\mathbb{Z}$.

Remark 99: If $n = 0$, then $n\mathbb{Z} = \{0\}$, so $\mathbb{Z}_0 = \mathbb{Z}/\{0\} \simeq \mathbb{Z}$, and if $n = 1$, then $1\mathbb{Z} = \mathbb{Z}$, so $\mathbb{Z}_1 \simeq \{0\}$. Therefore, we are mostly interested with cyclic groups for $n \geq 2$.

As it can be difficult to deal with quotient groups, we take a little, and important, detour via modular arithmetic, and prove that the objects we define are in fact the cyclic groups. We recall the theorem of Euclidean division.

Theorem 100 (Euclidean division): Let $n, m \in \mathbb{Z}$ be integers. There exists a unique couple (q, r) with $q \in \mathbb{Z}$ and $0 \leq r < n$ such that

$$n = bm + r.$$

Uniqueness of such a couple allows us to define

Definition 101: Let $n \geq 1$. We define the function $\mod n : \mathbb{Z} \to \{0, \ldots, n-1\}$ sending any $m \in \mathbb{Z}$ to the (unique) number $m \mod n$ such that

$$n = mq + (m \mod n).$$

This is well defined according to Theorem 100.

Definition 102: Let $a, b \in \mathbb{Z}$, when $(a \mod n) = (b \mod n)$, write

$$a \equiv b\ [n].$$

Exercise 38: Check that $- \equiv -\ [n]$ of Definition 102 defines an equivalence relation on $\mathbb{Z}$.

Exercise 39: Let $a, b, c, d \in \mathbb{Z}$, with $a \equiv b\ [n]$ and $c \equiv d\ [n]$. Check that

- $n \equiv 0\ [n]$.
- $a + c \equiv b + d\ [n]$
- $a \times c \equiv b \times d\ [n]$

Definition 103: Calling $k \mapsto \overline{k}$ the maps $(- \mod n)$ of Definition 101, we define:

$$\overline{k} + \overline{l} := \overline{k+l}$$
$$\overline{k} \times \overline{l} := \overline{kl}.$$

Exercise 39 proves that it does not depend on the representative.

Proposition 104: Let $n \geq 2$. The set $\{\overline{0}, \dots, \overline{n-1}\}$ together with the law $+$ of 103 is a group with neutral element $\overline{0}$, and inverse $-\overline{k} = \overline{-k}$. Moreover, this group is isomorphic to $\mathbb{Z}_n$.

Proof. (**Proposition 104**) The fact that it is a group follows from Exercise 39. Let $f : \mathbb{Z} \to \{\overline{0}, \dots, \overline{n-1}\}$ be the morphism of abelian groups sending k to $\overline{k}$. If $f(k) = \overline{0}$, then $\overline{k} = \overline{0}$, thus $k \equiv 0\ [n]$, meaning that $\ker(f) = n\mathbb{Z}$. Moreover, f is surjective, so $\operatorname{im}(f) = \{\overline{0}, \dots, \overline{n-1}\}$. By the first isomorphism theorem,

$$\mathbb{Z}/\ker(f) = \mathbb{Z}/n\mathbb{Z} \simeq \operatorname{im}(f) = \{\overline{0}, \dots, \overline{n-1}\}.$$

□

Remark 105: Generally, there is no harm in removing completely the notation $\overline{k}$, and simply write k. Therefore, in $\mathbb{Z}_2$, we will write $1 + 1 = 0$ and $-1 = 1$.

3.3 Rings

A ring is a group and a monoid behaving well together. When two interesting structures interact in interesting ways, is where mathematics happens. Ring theory revolutionized the realm of mathematics last century, and they belong to the toolbox of any person interested in math.

Definition 106: A *ring* $(R, +, \cdot, 0, 1)$ is the data of a set R together with

1. A commutative group structure $(R, +, 0)$,
2. A monoid structure $(R, \cdot, 1)$ (i.e. $\cdot$ is associative and has a neutral element called 1),

such that multiplication distributes over addition, that is, for all $a, b, c \in R$,

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c),$$

and

$$(b + c) \cdot a = (b \cdot a) + (c \cdot a).$$

Remark: we also make the convention that $0 \neq 1$ in these lecture notes.

Example 107: We already know examples of rings:

- $(\mathbb{Z}, +, \cdot, 0, 1)$ is a ring.
- $(\mathbb{R}, +, \cdot, 0, 1)$ is a ring.

Remark 108: As indeed rings are generalizations of the structure of $\mathbb{Z}$ and $\mathbb{R}$, we will use the same convention of notations than for those two sets. It means that we often omit the symbol $\cdot$, and we give it higher priority than $+$, so instead of writing

$$(x \cdot y) + z,$$

we will more concisely write

$$xy + z.$$

As usual, we will say "let R be a ring", when we should rather say "let $(R, +, -, \cdot, 0, 1)$" be a ring.

Lemma 109: Let R be a ring, then for all $a \in R$, $0 \cdot a = 0 = a \cdot 0$.

Proof. (**Lemma 109**) By distributivity, and by 0 being neutral for $+$,

$$a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0.$$

Subtracting $a \cdot 0$ both sides leads $0 = a \cdot 0$. The other distributive law proves $0 \cdot a = 0$. □

Definition 110: A ring is *commutative* if its underlying monoid structure is commutative. Notice that the group structure is already abelian in a ring.

Definition 111: Let R, S be rings. A *ring morphism* $f : R \to S$ is a function of the underlying sets that is bot a monoid morphism and a group morphism. It means that for all $a, b \in R$:

$$\begin{aligned} f(a+b) &= f(a) + f(b), \\ f(a \cdot b) &= f(a) \cdot f(b), \\ f(1) &= 1, \\ f(0) &= 0. \end{aligned}$$

Notice that the symbols $(+, \cdot, 1, 0)$ on the left-hand side are things happening in R, and in S in the right-hand side.

Exercise 40: Prove that we can drop the condition $f(0) = 0$ for a ring morphism, i.e. prove that it already follows from the first three equations.

Definition 112: Let $(R, \cdot, +)$ be a commutative ring. An *ideal* I of R is a subgroup I of the underlying abelian group $(R, +)$ such that

$$\forall r \in I, \forall x \in R, r \cdot x \in I.$$

Example 113: The sets $n\mathbb{Z}$ are ideals of $\mathbb{Z}$.

The same way we can quotient an abelian group by another group, we can quotient a commutative ring by an ideal. We will not delve into the details of the precise construction, but quotienting $\mathbb{Z}$ by the ideal $n\mathbb{Z}$ gives the ring $\mathbb{Z}/n\mathbb{Z}$. Its underlying group structure is given by the cyclic group structure we studied in the previous chapter.

Definition 114: Let R be a ring. We say that $r \in R$ is *invertible*, or an *unit*, if there exists some $r' \in R$ such that $r \cdot r' = 1 = r' \cdot r$. We denote by $R^\times$ the sets of all units.

Remark 115: Let $(R, \cdot, +, 0, 1)$ be a ring. Let $r \in R$, then r is always invertible for the law $+$, as we have $r + (-r) = 0 = (-r) + r$, but it is not always invertible for the $\cdot$ law. See the next example.

Example 116:

- The units of the ring $\mathbb{Z}$ are 1 and -1. Indeed, in $\mathbb{Z}$, $a \cdot b = 1$ means that either $a = -1 = b$ or $a = 1 = b$.
- We claim that $(\mathbb{Z}/n\mathbb{Z})^\times = \{k \in \mathbb{N} \mid 0 \leq k \leq n, \gcd(k, n) = 1\}$. Why? Hint: a proof is hidden in the one of Exercise sheets. Or simply use Bézout identity.

Proposition 117: Let R be a ring. Then $(R^\times, \cdot, 1)$ is a group.

Proof. (**Proposition 117**) Let $r, q \in R$ be units. Let r', q' be their inverse. Then $(rq)(q'r') = 1 = (q'r')(rq)$ so $rq \in R^\times$. Let $r \in R^\times$. Suppose r', r'' are two inverses for r. Then $r'r = 1 = r''r$, multiplying on the right by r', we get $r'rr' = r''rr'$, but $rr' = 1$, so $r' = r''$, thus inverses are well defined. Associativity and neutrality of 1 follows directly from the ring axioms. □

Corollary 118 (Euler's theorem): Call ϕ Euler's totient function. Let $n \geq 2$, let $a \in \mathbb{N}$ with $\gcd(n, a) = 1$, then

$$a^{\phi(n)} \equiv 1\,[n].$$

Proof. (**Corollary 118**) Let G be a finite group of size n, let $a \in G$, then by Lagrange, $\mathrm{ord}(a) \mid n$, so $a^n = a^{p\,\mathrm{ord}(a)} = e$. Now apply this with $G := \mathbb{Z}/n\mathbb{Z}$. Let $a \in \mathbb{N}$, with $\gcd(n, a) = 1$. We know by Example 116 that $a \in (\mathbb{Z}/n\mathbb{Z})^\times$, and that $(\mathbb{Z}/n\mathbb{Z})^\times$ has $\phi(n)$ elements, so $a^{\phi(n)}$ has to be the neutral of $(\mathbb{Z}/n\mathbb{Z})^\times$, which is 1. □

Remark 119: Let R be a ring, then there is always a unique ring morphism $\iota : \mathbb{Z} \to R$, that sends $1 \in \mathbb{Z}$ to $1_R \in R$. As such, whenever we are working in a ring R, we can always invoke the elements $n \in \mathbb{Z}$. For instance, the element $3 \in \mathbb{Z}$ is send to $1_R + 1_R + 1_R$.

Definition 120: Let $0 \leq k \leq n$ be natural numbers. We define the *binomial coefficient*

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}.$$

Lemma 121 (Pascal's rule): Let $0 \leq k \leq n$ be natural numbers, then

$$\binom{n-1}{k} + \binom{n-1}{k-1} = \binom{n}{k}.$$

Proof. (**Lemma 121**) We compute

$$\begin{aligned}
\binom{n-1}{k} + \binom{n-1}{k-1} &= \frac{(n-1)!}{k!(n-1-k)!} + \frac{(n-1)!}{(k-1)!(n-k)!} \\
&= (n-1)!\left(\frac{n-k}{k!(n-k)!} + \frac{k}{k!(n-k)!}\right) \\
&= (n-1)!\left(\frac{n}{k!(n-k)!}\right) \\
&= \frac{n!}{k!(n-k)!} \\
&= \binom{n}{k}.
\end{aligned}$$

□

Proposition 122 (Binomial theorem): Let $(R, +, \cdot, 0, 1)$ be a commutative ring. Let $x, y \in R$, let $n \in \mathbb{N}$, then

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

Proof. (**Proposition 122**) By induction. If $n = 0$, then $(x+y)^0 = 1 = \binom{0}{0}x^0y^0$. If $n = 1$, then $(x+y)^1 = x + y = \binom{1}{0}y + \binom{1}{1}x$. Let $n \in \mathbb{N}$, and suppose that

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

We compute

$$\begin{aligned}
(x+y)^{n+1} &= (x+y)^n(x+y) \\
&= \left(\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}\right)(x+y) \\
&= \sum_{k=0}^{n} \binom{n}{k} x^{k+1} y^{n-k} + \sum_{k=0}^{n} \binom{n}{k} x^k y^{n+1-k} \\
&= x^{n+1} + y^{n+1} + \sum_{k=0}^{n-1} \binom{n}{k} x^{k+1} y^{n-k} + \sum_{k=1}^{n} \binom{n}{k} x^k y^{n+1-k} \\
&= x^{n+1} + y^{n+1} + \sum_{k=1}^{n} \binom{n}{k-1} x^k y^{n-(k-1)} + \sum_{k=1}^{n} \binom{n}{k} x^k y^{n+1-k} \\
&= x^{n+1} + y^{n+1} + \sum_{k=1}^{n} \left(\binom{n}{k-1} + \binom{n}{k}\right) x^k y^{n+1-k} \\
&= \binom{n+1}{n+1} x^{n+1} y^0 + \binom{n+1}{0} x^0 y^{n+1} + \sum_{k=1}^{n} \binom{n+1}{k} x^k y^{n+1-k}
\end{aligned}$$

$$= \sum_{k=0}^{n+1} \binom{n+1}{k} x^k y^{n+1-k}.$$

□

Remark 123: If we are working in a non-commutative ring, then the binomial identity of Proposition 122 holds as soon as $xy = yx$.

Exercise 41: Let p be a prime number. Let $x, y \in \mathbb{Z}/p\mathbb{Z}$.

1. Let $0 < k < p$. Show that p divides $\binom{p}{k}$.
2. Deduce that $(x+y)^p \equiv x^p + y^p \ [p]$.

Definition 124: A *field* is a ring $(F, \cdot, +, 1, 0)$ such that $F^\times = F\backslash\{0\}$. That is, a ring is a field if every element besides 0 is invertible.

Example 125:

- The rings $\mathbb{R}$ and $\mathbb{Q}$ are fields.
- The ring $\mathbb{Z}$ is not a field.
- The ring $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime.

Exercise 42: Let R be a commutative ring. We say that $a \in R$ is a *divisor of 0* if there exist a non-zero $b \in R$ such that $a \cdot b = 0$.

1. Suppose R is a field. Show that the only divisor of 0 is 0 itself.
2. Find a ring R whose only divisor of 0 is 0, but is not a field.
3. Let $a \in R$. Show that a is a divisor of 0 if and only if the map $f : R \to R$ that sends x to $f(x) := a \cdot x$ is not injective.
4. For those familiar with linear algebra, call $\mathcal{M}_n(\mathbb{R})$ the ring of square matrices of size n. Let $M \in \mathcal{M}_n(\mathbb{R})$. Show that M is a divisor of 0 if and only if it is not invertible.

Exercise 43: We define the

$$\mathbb{Z}[\sqrt{2}] := \{a + b\sqrt{2} \mid a, b \in \mathbb{Z}\}.$$

- Show that $\mathbb{Z}[\sqrt{2}]$ is a ring with the usual $\times$ and $+$ operation. In fact there is nothing special about 2: show that $\mathbb{Z}[\sqrt{n}]$ is a ring for all $n \in \mathbb{N}$.
- Show that $3 + 2\sqrt{2}$ is a unit in $\mathbb{Z}[\sqrt{2}]$.