Lecture Notes - Logic, Sets, Functions, Numbers

MA Math Camp 2023
Columbia University

Andrea Ciccarone*

This Version : August 4, 2023

Contents

*The present lecture notes were largely based on math camp materials from César Barilla, Palaash Bhargava, Paul Koh, and Xuan Li. All errors in this document are mine. If you find a typo or an error, please send me an email at ac4790@columbia.edu.

1 An Introduction to Mathematical Logic

1.1 Logical Statements

Definition 1.1. *An **statement** is a assertion to which we can attribute a truth value, i.e true (T) or false (F).*

A statement is the basic logical object of any mathematical reasoning. For instance, "$2 = 3$" is a statement (which happens to be false), but "$x = 3$" is technically *not* a statement because it is ambiguous – we do not know what x is in this context, so this assertion is neither true nor false. Observe that we don't need to know whether a statement is in fact true or false, just that it has to be one or the other.

Two statements are said to be **logically equivalent** if they have the same truth value.

Remark 1.1. *For instance, any two statements that are true are logically equivalent. This can lead to seemingly uninteresting statements, for instance* $1 = 1$ *and* $2 = 2$ *are both true, so they are logically equivalent (and this is not very interesting). This is, however, a useful concept when the the truth value of a statement is not know ex ante. For instance, for an arbitrary* $x \in \mathbb{R}$*, the statements* $2x + 1 \geq 2$ *is equivalent to the statement* $x \geq 1/2$ *: one is true if and only if the other is true. Rewriting things in another equivalent form to assess truth value is a tool that you have probably manipulated countless times, but it's useful to think about it formally – especially when dealing with more complex problems.*

Statements can be joined using **logical connectors**. Let p and q two statements, we define :

- "p and q" : the statement that is true if both p and q are true;
- "p or q" : the statement that is true if either p or q is true (or both);

Observe that the mathematical "or" is distinct from the common language "or" which can sometimes be exclusive (this or that but not both) while the mathematical "or" is always inclusive (this or that or both). This leads to the following truth table :

p	q	p and q	p or q
T	T	T	T
T	F	F	T
F	T	F	T
F	F	F	F

Sometimes, we notate p and q as $p \wedge q$ and p or q as $p \vee q$.

Another natural concept is the **negation**. Let p a statement, we define "Not p", also notated as $\neg p$ as the statement that is true if and only if p is false.

Let p and q two arbitrary statements. We define the statement "p implies q", notated as "$p \Rightarrow q$" as $\neg p$ or q. The intuition is that "$p \Rightarrow q$" is true if, when p is true, then q is also true; i.e either q is true or p is not true. Observe that this implies that if p is false, then the statement $p \Rightarrow q$ is always true.

p	q	$\neg p$	$p \Rightarrow q$
T	T	F	T
T	F	F	F
F	T	T	T
F	F	T	T

Often, we juste write p implies q to signify that the statement "p implies q" is true.

The equivalence $p \Leftrightarrow q$ is defined as the statement "$p \Rightarrow q$ and $q \Rightarrow p$". Observe that $p \Leftrightarrow q$ is true if and only if p and q are logically equivalent.

p	q	$p \Rightarrow q$	$q \Rightarrow p$	$p \Leftrightarrow q$
T	T	T	T	T
T	F	F	T	F
F	T	T	F	F
F	F	T	T	T

Proposition 1.1. *Let P, Q, R three statements. The following statements are logically equivalent :*

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

All the equivalences can be proven by writing out truth tables explicitly; this is left as a useful exercise. The equivalences $\neg(P \vee Q) \Leftrightarrow (\neg P \wedge \neg Q)$ and $\neg(P \wedge Q) \Leftrightarrow (\neg P \vee \neg Q)$ are sometimes referred to as De Morgan's law.

Exercise 1.1. *Prove Proposition 1.1. Write out each equivalence in words. Try to emphasize the logical intuition in each one.*

1.2 Quantifiers

We can define a notion of statements that are either true or false conditionally on a certain variable – called predicates. We are slightly anticipating on the next section by using the primitive notion of a set, which is defined as a collection of objects (called elements).

Definition 1.2. *A **predicate** is a statement that contains named variables (belonging to given sets), and which can be true or false depending on the value of the variables that occur in it.*

For example, "$x = 2$" for $x \in \mathbb{R}$ is a predicate. Specifying the set to which variables can belong is part of the predicate. For instance, "$x = 2$" for $x \in \mathbb{N}$ is different predicate from the previous one.

Predicates allow us to define the notion of quantifier, which serves to specify in some sense "how many" elements in the set under consideration verify the predicate. Let $P(x)$ a predicate depending on a variable x belonging to a set E. We define the following two statements :

- "$\forall x \in E, P(x)$" : this is true if and only if regardless of which element a of E we replace x by, $P(a)$ is true.
- "$\exists x \in E, P(x)$" : this is true if and only if we can find one element a of E such that $P(a)$ is true.

$\forall$ is called the *universal quantifier* and read as "for all" ($\forall x \in E$ is read as "for all x in E"); $\exists$ is called the *existential quantifier* and is read as "there exists". Observe that existence here is understood as existence of at least one element (there might be other). We sometimes notate existence of a unique element as $\exists!$. We also sometimes use $\nexists$ for the quantifier "there does not exist" : "$\nexists x \in E$, $P(x)$" is true if and only if there is no a in E such that $P(a)$ is true.

Exercise 1.2. *Prove that the following two statements are equivalent :*

- $\nexists x \in E, P(x)$
- $\forall x \in E, \neg P(x)$

It is also important to always remember that in a predicate the letter variable is a label or placeholder that can be relabeled – it is *not* a particular choice within this set, which is precisely why we introduced quantifiers. For instance, "$\forall x, P(x)$" and "$\forall y, P(y)$" are equivalent statements.

Proposition 1.2. *The statement $\neg(\forall x \in E, P(x))$ is equivalent to $(\exists x \in E, \neg P(x))$.*
The statement $\neg(\exists x \in E, P(x))$ is equivalent to $(\forall x \in E, \neg P(x))$.

This proposition (whose proof is left as an exercise), shows how to negate quantifiers. In words, the first equivalence says : it is false that $P(x)$ is true for all $x \in E$ if and only if there exists one x such that $P(x)$ is false. It is good to build a strong intuition of those logical relationships formally because it will keep things grounded when working with more complicated problems. The following very important remark explains how to deal with multiple quantifiers.

Remark 1.2. *When there is more than one quantifier,* **the order of quantifiers matters** *in general. Consider the following example statements :*

- $\forall x \in \mathbb{R}_+, \exists y \in \mathbb{R}_+, y = x^2$ *: this is true. Indeed, this states that any non-negative number x has a positive square.*
- $\exists y \in \mathbb{R}_+, \forall x \in \mathbb{R}_+, y = x^2$ *: this is false. This states that all non-negative numbers have the same square – there exists y such that for all x, $x^2 = y$.*

This example highlights that inverting quantifiers can drastically change a statement. In the first statement, when we write $\forall x, \exists y$, this implicitly allows the choice of y to depend on the choice of x : for any x that we choose, we can find a y such that $P(x, y)$. When instead we write $\exists y, \forall x$, we state that there exists a y such that $P(x, y)$ is true for any arbitrary selection of x. The two statements are clearly very different in general, so it is important to always pay attention to the order of logical chains – even more so when they become even longer.

However, when the same *quantifier is used successively, then we can invert the order without loss :*

$$(\forall x \in E, \forall y \in F, P(x, y)) \Leftrightarrow (\forall y \in F, \forall x \in E, P(x, y))$$
$$(\exists x \in E, \exists y \in F, P(x, y)) \Leftrightarrow (\exists y \in F, \exists x \in E, P(x, y))$$

In short, you can invert the order of two $\exists$ or two $\forall$ but ***not the order of a $\forall$ and a $\exists$.***

It is important to be able to have a clear intuition of the difference between statements with inverted quantifiers. Being able to easily read, understand and write formal logical statements will make it easier to structure proofs and build a reasoning on solid mathematical grounds. You should also be able to negate statements with multiple quantifiers.

Exercise 1.3. *Write the negation of each of the following statement. Interpret each statement and state (if possible) which is true or false.*

1. $\forall x \in \mathbb{R}, x^2 \geq 0$
2. $\forall x \in \mathbb{R}, x^2 > 0$
3. $\forall x \in \mathbb{N}, \forall y \in \mathbb{R}, x + y \in \mathbb{R}$
4. $\exists M \in \mathbb{R}, \forall x \in \mathbb{R}, x \leq M$
5. $\forall x \in \mathbb{R}, \exists M \in \mathbb{R}, x \leq M$
6. $\forall x \in \mathbb{R}, \exists y \in \mathbb{R}_+, \exists z \in \mathbb{R}_+, x = y - z$
7. $\exists y \in \mathbb{R}_+, \forall x \in \mathbb{R}, \exists z \in \mathbb{R}_+, x = y - z$
8. $\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, \exists \theta \in \mathbb{R}, |x - y|^2 \leq \theta |x - y|$
9. $\exists \theta \in \mathbb{R}, \forall x \in \mathbb{R}, \forall y \in \mathbb{R}, |x - y|^2 \leq \theta |x - y|$
10. $\forall (x, y) \in \mathbb{R}^2, x + y = 0 \Rightarrow (x = 0 \text{ and } y = 0)$
11. $\forall (x, y, z) \in \mathbb{R}^3, x^2 + y^2 + z^2 = 0 \Rightarrow (x = 0 \text{ and } y = 0 \text{ and } z = 0)$

We can also connect quantifiers and implications/equivalence structures.

Proposition 1.3. *Let $P(x), Q(x)$ two predicates depending on $x \in E$. The following statements are true :*

$$\Big((\forall x \in E, P(x)) \text{ or } (\forall x \in E, Q(x)) \Big) \Rightarrow \forall x \in E, (P(x) \text{ or } Q(x))$$

$$\Big((\exists x \in E, P(x)) \text{ or } (\exists x \in E, Q(x)) \Big) \Leftrightarrow \exists x \in E, (P(x) \text{ or } Q(x))$$

To see that the reverse implication in the first statement does not hold in general, consider the following example : $x \in \mathbb{R}$, $P(x) : x \geq 0$, $Q(x) : x \leq 0$. Clearly for all $x \in \mathbb{R}$, either $x \geq 0$ or $x \leq 0$. But we do not have either $P(x)$ or $Q(x)$ true for all $x \in \mathbb{R}$.

1.3 Proof Structures

Formal logic allows us to define methods to prove things. This makes rigorous many usual ideas : for example if we know that two statements are equivalent, proving that one is true will mean that the other is also true. This is useful because it might be to easier to prove some equivalent statement rather than directly show what we are interested in.

It is important to have multiple blueprints of proof structures in mind to be able to choose the most adapted (and the easiest) to what we are trying to achieve. Once again, although the presentation here is simple, it is extremely valuable to have a clear understanding of logical articulations and proof structures to tackle more complicated problems. Often, proofs will have to be chained and use complex intermediary statements to eventually reach the result of interest.

Below is an overview of the main proof strategies.

1. **($\mathbf{P}$ and $\mathbf{P \Rightarrow Q}$) implies $\mathbf{Q}$**
   To show that Q is true, it might be easier to show that something that implies Q is true. We need to show two results : P is true and P implies Q. This will yield that Q is true.

2. **($\mathbf{P \Rightarrow Q}$) and ($\mathbf{Q \Rightarrow R}$) implies that $\mathbf{P \Rightarrow R}$**
   To show an implication ($P \Rightarrow Q$), we can show a *chain of implications*. This is a fundamental property : implications are transitive. Hence we can write $P \Rightarrow Q \Rightarrow R$. To show an implication, it is often practical to show a chain of intermediary, easier implications. This proof structures extends to longer chains : $P \Rightarrow Q_1 \Rightarrow Q_2 \Rightarrow ... \Rightarrow Q_n \Rightarrow R$. Clearly since equivalences are stronger than implications, it will not change the structure if there is an implication in the chain, e.g. $P \Rightarrow Q_1 \Rightarrow Q_2 \Rightarrow ... \Rightarrow Q_n \Leftrightarrow R$.

3. **($\mathbf{P \Leftrightarrow Q}$) and ($\mathbf{Q \Leftrightarrow R}$) implies that $\mathbf{P \Leftrightarrow R}$**
   The same way that we can chain implications, we can chain equivalences. Hence to show that $P \Leftrightarrow R$, we can show two intermediary (hopefully easier) equivalences $P \Leftrightarrow Q$ and $Q \Leftrightarrow R$.

4. **($\mathbf{P \Rightarrow Q}$) and ($\mathbf{Q \Rightarrow P}$) implies that $\mathbf{P \Leftrightarrow Q}$**
   Alternatively, we can just prove an equivalence ($P \Leftrightarrow Q$) by breaking up the equivalence using the definition : we prove separately that both P and Q imply each other. This is sometimes called proving a double implication.

5. **$\mathbf{P \Rightarrow Q}$ is equivalent to $\mathbf{\neg Q \Rightarrow \neg P}$ (Contraposition)**
   To prove $P \Rightarrow Q$, it is sometimes easier to prove the contraposition : if Q is not true, then P is not true either. The equivalence can be proven directly from the definition of the implication : when $P \Rightarrow Q$, it cannot be that Q is not true if P is true, so if Q is false it must be that P is false ($\neg Q \Rightarrow \neg P$). Contraposition is an extremely useful tool, either we ultimately want to prove an implication or to prove an intermediate implication in the previous techniques (double implication, proving that Q is true by proving that P is true, implication chains).

6. **If $\mathbf{\neg P \Rightarrow Q}$ and $\mathbf{\neg P \Rightarrow \neg Q}$, then P is true (Proof by Contradiction)**
   The proof by contradiction consists in assuming that a statement ($\neg P$) is true, and showing that this entails something which is false by definition (Q and $\neg Q$). This relies on proving two implications : P implies both Q and its negation $\neg Q$. Since a statement cannot be both true and false by definition, Q and $\neg Q$ cannot be simultaneously true, so if P was false this would entail a contradiction. Therefore P must be true. Be careful not to confuse proofs by contradiction and proofs by contraposition. Both assume the contradiction of a statement – but contraposition shows the negation of another statement to yield an implication (by contraposition), while a proof by contradiction shows an absurd result to conclude that the initial statement cannot be true.

Exercise 1.4. *Show the contrapositive principle ($P \Rightarrow Q \Leftrightarrow \neg Q \Rightarrow \neg P$) using a truth table.*

Example 1.1. *Let a, b, c three positive real numbers. We want to show that :*

$$a \leq b + c \;\Rightarrow\; \frac{a}{1+a} \leq \frac{b}{1+b} + \frac{c}{1+c}$$

The first option is to proceed by showing a chain of implications :

$$\begin{aligned} a \leq b+c &\Rightarrow \frac{1}{a} \geq \frac{1}{b+c} \\ &\Rightarrow \frac{1}{a}+1 \geq \frac{1}{b+c}+1 \\ &\Rightarrow \frac{a}{1+a} \leq \frac{b+c}{b+c+1} \leq \frac{b}{1+b}+\frac{c}{1+c} \end{aligned}$$

Another option is to start from the right (implied) statement and show a chain of equivalences with something that is implied by the left statement :

$$\begin{aligned} \frac{a}{1+a} \leq \frac{b}{1+b}+\frac{c}{1+c} &\Leftrightarrow \frac{a}{1+a} \leq \frac{b+c+2bc}{(1+b)(1+c)} \\ &\Leftrightarrow a+ab+ac+abc \leq b+c+2bc+ab+ac+2abc \\ &\Leftrightarrow a \leq b+c+2bc+abc \end{aligned}$$

The last statement is implied by $a \leq b+c$ since $b+c \leq b+c+2bc+abc$.

Even when writing proofs out properly with words and full sentences, the underlying logical structure should always be clearly defined and clearly identifiable by the reader. The next example highlights a classical case of a proof by contradiction.

Example 1.2. *We want to show that $\sqrt{2}$ is not a rational number*[1]*. Formally, we wish to show the proposition $P : \sqrt{2} \notin \mathbb{Q}$. We proceed by contradiction. Assume $\neg P$ i.e $\sqrt{2} \in \mathbb{Q}$. By definition, this means that $\sqrt{2}$ can be written as the ration of two integers m and n, i.e we can write $\sqrt{2} = m/n$ with $m, n \in \mathbb{N}$. Without loss of generality we take m/n to be irreducible (otherwise we just replace m and n by the corresponding irreducible integers). Then we must have $2n^2 = m^2$, which entails that m^2 is even so m must be even as well, i.e there exists m' such that $m = 2m'$. Hence $2n^2 = (2m')^2 = 4m'^2$, so $n^2 = 2m'^2$. This entails that n^2 is pair so there is n' such that $n = 2n'$. Therefore :*

$$\frac{m}{n} = \frac{2m'}{2n'} = \frac{m'}{n'}$$

Hence m/n is not irreducible, which is a contradiction (here the Q statement of the definition is the statement "m/n is irreducible").

All the proof structures previously introduced can naturally be used when the statements of interest involve quantifiers. We can still outline some specificities of statements with quantifiers when constructing proof strategies.

- As a general principle when proving that a statement involving quantifiers is true, it is important to remember that a variable corresponding to a $\forall$ has to remain arbitrary (any element of the set), while a variable corresponding to an $\exists$ can be chosen (possibly depending on the other variables).
- The same way that order matters in writing statements, order will matter in writing the corresponding proofs. For instance, to prove something of the form "$\forall x \in E, \exists y \in F, P(x, y)$", the proof should look something like "Consider an arbitrary $x \in E$... Choose a given $y \in F$ (because x is arbitrary but fixed, y can depend on x)... Then show that $P(x, y)$ is true". Keeping the order rigorous will ensure that no variable unduly depends on another when it should not.

[1] As a reminder, rational numbers are defined as the quotients of integers.

- To show an existence result $\exists x \in E, P(x)$, it is sufficient to find suitable element. If an example can be found directly, this is usually the easiest approach – this will give proofs of the form "Consider this particular $a \in E$, show that $P(a)$".
- To show that a result holds for all x, we need to find a proof strategy that is not contingent on which x we choose.
- Sometimes, it is convenient to proceed by disjunction of cases, i.e consider exhaustive subsets. If we want to show $\forall x \in E, P(x)$, we can take $A \subset E$ and $A^c := E \setminus A$ and then make two distinct proofs. If $x \in A$, then $P(x)$ is true; if $x \in A^c$, $P(x)$ is also true (possibly using a different proof or idiosyncratic arguments); we can conclude that $P(x)$ is true for all $x \in E$. We can naturally consider more than two cases but it is important to be sure that we exhaust all possible cases for x.
- To show the negation of a $\forall$, we need to show an existence result. To show the negation of an existence result, we need to show a $\forall$ result. Sometimes one might be simpler than the other and this can be used in e.g. contraposition.
- To show existence and uniqueness i.e $\exists! x \in E, P(x)$, we can do things in either order :
  - We first prove that there exists an $a \in E$ such that $P(a)$ is true (e.g. by finding or constructing it). We then verify that this a must be unique. Proofs by contradiction are quite typical for this step (although not the only approach) : assume there exists $a' \neq a$ such that both $P(a)$ and $P(a')$ are true and show that this entails a contradiction (often the contradiction comes from showing $a' = a$).
  - We can also proceed by analysis and synthesis. In the first step (analysis), assume there is an x such that $P(x)$ is true ; show that this must mean that $x = a$ for some a, i.e a is the only candidate. This in particular shows uniqueness. Then verify that indeed $P(a)$ is true (synthesis) : this shows existence and concludes the proof.

2 Sets

The concept of **set** is a fundamental notion. The definition of a set itself is intuitive and coincides with its common meaning : a **set** is a **collection of objects**.

Why is the notion of a set so important ? Essentially any mathematical statement is about saying something about some object(s) in a given set. The concept of set is what allows us to make rigorous statements about precisely defined objects – we have already began to see this when introducing quantifiers. It might seem simple, but it's important to never forget that no mathematical statement is complete or rigorous without explicitly defining to which set each object is allowed to belong.

For example, consider the statement : "Every chair in this room is blue". This is a statement about *every object* belonging to a set (the set of chairs in this room), specifying that they share a given property (being blue). This is not the same statement as "Every chair in the next room is blue". If we said "every chair is blue", not only is this yet another statement, but this might ambiguous – does this mean every chair in the world ? What set the statement applies to will clearly change whether it is true or false. Always keep this in mind : a statement's rigor is crucially dependent on defining the appropriate set.

A set is completely characterized by what objects it contains; we call them its **elements** or **members**. The most primitive way to define a set is by enumerating the objects it contains. We

use the curly bracket notation to denote this :

$$A = \{1, \pi, \text{orange}, \text{Economics}, \text{Columbia University}\}$$

We can also define sets by specifying the properties of the elements that they contain. We can again use the curly bracket notation

$$X = \{n \in \mathbb{Z} : n^2 = 4\}$$

We can and will also define sets directly using words – for instance "the set of integers greater than 4", or "the set of second order polynomials", etc. We allow for the possibility that a set contains no element. We call this set the **empty set**, and denote it $\emptyset$.

Objects in a set are called elements. To say that the object x **belongs to**, or **is an element of** the set X, we note $x \in X$.

Later, we will want to consider more complicated sets that might have some "structure" in an appropriately defined sense. For instance in some sets we might want to think about how far apart from each other elements in this set are – this will lead to define the notion of **distance** or **metric**. In doing so, we will try to talk about general properties of such spaces, so we will consider the abstract class of **metric spaces** (sets that can be endowed with a metric). Things get more complicated as we pile on more complex notions and consider richer objects with more structure, but it's important to never forget that any statement we make is always about a set and its objects.

For another example, consider the important topic of Optimization – which we will visit later. Optimization is about finding the maximal element of a set in some appropriate sense. For instance, in economics we often think about choice as maximization : among the set of possible choices, which object maximizes utility ? For this statement to be rigorous, it is crucial to have a good definition of the set of objects we are considering.

2.1 Inclusion

We say that A **is included in** B, or that A **is a subset of** B, and note $A \subseteq B$ (or $A \subset B$) if every member of A is a member of B: $x \in A \Rightarrow x \in B$. Any set is a subset of itself.

- We denote $A = B$ iff $A \subseteq B$ and also $B \subseteq A$.
- We denote $A \subsetneq B$ iff $A \subset B$ but $B \subset A$ is not true. In this case, we say that A is a **proper subset** of B. Alternatively, we can use $\subset$ to denote a proper subset and $\subseteq$ to denote a subset.
- The inclusion is **transitive** : if $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$.
- The set of all subsets of a set X is called the **power set** of X, and noted $\mathcal{P}(X)$ or 2^X.

Example 2.1. *The set of rational numbers is included in the set of real numbers and it is a proper subset :* $\mathbb{Q} \subsetneq \mathbb{R}$. *To prove the latter, we can for instance refer to Example 1.2 showing that* $\sqrt{2} \in \mathbb{R}$ *but* $\sqrt{2} \notin \mathbb{Q}$

Observe that the power set is a *set of sets*. Indeed, when considering what the objects of a set can be, we did not impose any restrictions – and sets of sets are frequently encountered. It is extremely important to always distinguish in a given context when we talk about a set as an element or the whole set. The following example illustrates this idea.

Example 2.2. *Let* $E := \{a, b\}$. *The power set of* E *is given by :*

$$\mathcal{P}(E) = \Big\{\emptyset, \{a\}, \{b\}, \{a, b\}\Big\}$$

The power set of $\emptyset$ is $\mathcal{P}(\emptyset) = \{\emptyset\}$, hence the power set of the power set of the empty set is $\mathcal{P}(\mathcal{P}(\emptyset)) = \{\emptyset, \{\emptyset\}\}$ Notice that the empty set ($\emptyset$) and the singleton set containing the empty set $\{\emptyset\}$ are not the same object – *by definition, the empty set has no element, while $\{\emptyset\}$ is a set containing one element (the empty set).*

Observe that all sets contain the empty set by convention i.e for any set E, $\emptyset \subseteq E$.

Exercise 2.1. *1. Write explicitly $\mathcal{P}(\mathcal{P}(\mathcal{P}(\emptyset)))$*

2. Write explicitly $\mathcal{P}(\mathcal{P}(\{a, b\}))$

Exercise 2.2. *Let E and F two sets. Show that $E \subseteq F \Leftrightarrow \mathcal{P}(E) \subseteq \mathcal{P}(F)$.*

2.2 Union and Intersection

The **intersection** of two sets A and B is the set of all elements that belong to both A and B[2]:

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

- If no element lies in both A and B, $A \cap B$ is still defined: defined to be the empty set $A \cap B = \emptyset$. We say that A and B are **disjoint**.
- The intersection is commutative: $A \cap B = B \cap A$.
- The intersection is associative: $(A \cap B) \cap C = A \cap (B \cap C)$.
- $A \cap B = A$ iff $A \subseteq B$.

The **union** of two sets A and B, noted $A \cup B$ is the set of all elements that belong to either A or B (possibly both):

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

- The union is commutative: $A \cup B = B \cup A$.
- The union is associative: $(A \cup B) \cup C = A \cup (B \cup C)$.
- $A \cup B = A$ iff $B \subseteq A$.

Exercise 2.3. *Let E, F two sets. Compare $\mathcal{P}(E \cup F)$ and $\mathcal{P}(E) \cup \mathcal{P}(F)$ (is one included in the other ?).*

More generally, we can talk about the union of any collection of sets, possibly infinite. Let Θ be a set (potentially infinite), and let $\{A_\theta\}_{\theta \in \Theta}$ be a collection[3] of sets indexed by $\theta \in \Theta$. That is, for each $\theta \in \Theta$, A_θ is a set in the collection. There are as many sets in the collection $\{A_\theta\}_{\theta \in \Theta}$ as there are elements in Θ.

Generalize the notation we defined previously to allow infinite union/intersection:

$$\bigcup_{\theta \in \Theta} A_\theta := \{x : \exists\ \theta \in \Theta \text{ s.t. } x \in A_\theta\}$$

$$\bigcap_{\theta \in \Theta} A_\theta := \{x : x \in A_\theta, \forall\ \theta \in \Theta\}$$

Notice that when the index set Θ has only two elements θ_1 and θ_2, the notations above reduces to pairwise union/intersection we defined at the beginning.

[2](The notation ":=" is for defining a new notation. The left-hand side is the new notation, and the right-hand side is its meaning.)

[3]Essentially, a collection of sets is a set of sets. We use the term "a collection of sets" to avoid confusion.

- The intersection is distributive with respect to the union:
  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
  $B \cap \left(\bigcup_{\theta \in \Theta} A_\theta \right) = \bigcup_{\theta \in \Theta} (B \cap A_\theta)$.
- The union is distributive wrt. the intersection:
  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
  $B \cup \left(\bigcap_{\theta \in \Theta} A_\theta \right) = \bigcap_{\theta \in \Theta} (B \cup A_\theta)$.

Remark 2.1. *As the notations might suggest, the union and the intersection are related to the logical "and" and "or" operators – although they should not be confused because they apply to different objects. In fact, the union and the intersection represent a particular kind of "or" and "and" in the context of sets. To make the connection explicit, observe that the following statements are equivalent :*

$$x \in A \cup B \Leftrightarrow (x \in A) \vee (x \in B)$$
$$x \in A \cap B \Leftrightarrow (x \in A) \wedge (x \in B)$$

Many properties of the union and intersection follow from the properties of the logical operators – for instance distributivity of union over intersection is a direct consequence of distributivity of or over and :

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

2.3 Difference and complement

Given two sets A and B, the **set difference** $A - B$ or $A \setminus B$ is the set of all x that belong to A but not to B:

$$A \setminus B := \big\{ x \in A : x \notin B \big\}$$

Notice that in the definition above B does not need to be a subset of A.

Let U be a set (usually the biggest relevant set in a given context), and A be a subset of U. Define the **complement** of A as

$$A^c := U \setminus A$$

Notice that the notation A^c makes sense only when the "universe" U has no ambiguity in the question. Whenever U currently used is ambiguous, we should be explicit about U by writing $U \setminus A$ instead of A^c. We denote $x \notin A$ for $x \in A^c$ ($x \in U$ implicitly). For any set A, we have :

$$(A^c)^c = A$$

Observe that this also follows from the logical propertie of the negation : $x \in (A^c)^c \Leftrightarrow \neg(\neg(x \in A)) \Leftrightarrow x \in A$.

Proposition 2.1 (De Morgan's law)**.** *Let $\{A_\theta\}_{\theta\in\Theta}$ be a collection of subsets of U. We have:*

(1) $\left(\bigcup_{\theta\in\Theta} A_\theta\right)^c = \bigcap_{\theta\in\Theta} A_\theta^c$

(2) $\left(\bigcap_{\theta\in\Theta} A_\theta\right)^c = \bigcup_{\theta\in\Theta} A_\theta^c$

The proof essentially follows from applying the formal version logic of De Morgan's law that we previously saw. Below is an explicit proof.

Proof. (1) $\subset$:

Take any $x \in \left(\bigcup_{\theta\in\Theta} A_\theta\right)^c = U \setminus \left(\bigcup_{\theta\in\Theta} A_\theta\right)$. By definition, we have $x \in U$, but $x \notin \bigcup_{\theta\in\Theta} A_\theta$.

Therefore $x \notin A_\theta$ for any $\theta \in \Theta$. Then $x \in A_\theta^c$ for any $\theta \in \Theta$, which implies $x \in \bigcap_{\theta\in\Theta} A_\theta^c$.

$\supset$:

Take any $x \in \bigcap_{\theta\in\Theta} A_\theta^c$. By definition, we have $x \in A_\theta^c$ for any $\theta \in \Theta$. Therefore $x \in U$ and $x \notin A_\theta$ for any $\theta \in \Theta$.

So we have $x \in U$ and $x \notin \bigcup_{\theta\in\Theta} A_\theta$, which implies $x \in \left(\bigcup_{\theta\in\Theta} A_\theta\right)^c$.

(2) By (1), we have

$$\left(\bigcap_{\theta\in\Theta} A_\theta\right)^c = \left(\bigcap_{\theta\in\Theta} (A_\theta^c)^c\right)^c = \left(\left(\bigcup_{\theta\in\Theta} A_\theta^c\right)^c\right)^c = \bigcup_{\theta\in\Theta} A_\theta^c$$

□

Exercise 2.4. *Let E a non-empty set and let A, B, C subsets of E. Show that :*

1. $A = B \Leftrightarrow A \cap B = A \cup B$
2. $A \cap B^c = A \cap C^c \Leftrightarrow A \cap B = A \cap C$
3. $\begin{cases} A \cap B \subseteq A \cap C \\ A \cup B \subseteq A \cup C \end{cases} \Rightarrow B \subseteq C$
4. *Define the symmetrical difference, denoted Δ, of A and B as :*

$$\begin{aligned} A\Delta B :&= \{(x \in A \text{ and } x \notin B) \text{ or } (x \in B \text{ and } x \notin A)\} \\ &= (A \cup B) \setminus (A \cap B) \\ &= (A \setminus B) \cup (B \setminus A) \end{aligned}$$

The three definitions above are equivalent. Show that :

$$A^c \Delta B^c = A \Delta B$$

2.4 Cartesian Product

Define the **Cartesian product** of set A and B as

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

The notation (a, b) is for an ordered pair. The pair is ordered in the sense that (a, b) is different from (b, a). In words, Cartesian product of A and B is the set of all ordered pairs whose first element is taken from A and whose second element is taken from B. Notice that $A \times B \neq B \times A$.

We can also talk about the Cartesian produce of multiple sets $A_1 \times A_2 \times \cdots \times A_n$, or simply $\prod_{i=1}^{n} A_i$, defined as

$$A_1 \times A_2 \times \cdots \times A_n := \big\{(a_1, a_2, \ldots, a_n) : a_i \in A_i, \forall\, i = 1, 2, \ldots, n\big\}$$

If all A_i's are the same, order does not matter and we can denote $A \times A \times \cdots \times A$ as A^n.

Remark 2.2. *We commonly use lists separated by commas to signify that several elements belong to the same set : $a, b, c \in E$. If we view (a, b, c) as an triple instead, it belongs to the cartesian product of E with itself (twice) :*

$$(a, b, c) \in E^3$$

Although it is often equivalent to use either notation, it is important not to mix them up because they are conceptually different – lists introduce several separate elements belonging to the same set, while tuples (denoted by parenthesis) introduce one object *in a product set.*

3 Relations

Mathematically, a **(binary) relation** R **from** A **to** B is a subset of $A \times B$. Each ordered pair $(a, b) \in A \times B$ either has this relation R (if $(a, b) \in R$) or does not have this relation R (if $(a, b) \notin R$). We can also use the notation aRb to stand for $(a, b) \in R$.

For example, think of the set P of professors, the set S of students, and the "advisorship" relation between professors and students. For each professor-student pair $(p, s) \in P \times S$, professor p either advises student s or not. If p advises s, we put the pair (p, s) into the "advisorship" relation R, and finally we will obtain R as a set of professor-student pairs, i.e. a subset of $P \times S$. Also notice that in the "advisorship" relation, a professor is allowed advice multiple or no student, and a student is allowed to have multiple or no advisor.[4]

For a relation R from A to B, its inverse is a relation from B to A, defined as

$$R^{-1} := \big\{(b, a) \in B \times A : (a, b) \in R\big\}$$

Notice that we can always invert a relation to get another relation (in contrast to functions as we will see later).

A relation R from A to A itself is also called a **relation on** A. For example, the inclusion relation $\subset$ of sets can be viewed as a relation on the set of all sets.

Definition 3.1. *Let R be a relation on the set X.*

[4] It is also possible to talk about multilateral relations $R \subset A_1 \times A_2 \times \cdots \times A_n$. This generalized concept will be useful if we want to model, for example, which professor teaches which course on which day, in which case (p, c, d) is in R iff professor p teaches course c on day d. Throughout this math camp, however, we only use binary relations.

*(1) Relation R is **reflexive** iff*[5] *xRx for any $x \in X$.*

*(2) Relation R is **transitive** iff for any $x, y, z \in X$ s.t. xRy and yRz, we have xRz.*

*(3) Relation R is **anti-symmetric** iff for any $x, y \in X$ s.t. xRy and yRx, we have $x = y$.*[6]

*(4) Relation R is **complete** iff for any $x, y \in X$, either xRy or yRx.*

*(5) Relation R is **symmetric** iff for any $x, y \in X$ s.t. xRy, we have yRx.*

3.1 Orders

Definition 3.2. *A relation $\leq$ on X is a **pre-order** iff $\leq$ is reflexive and transitive. If $\leq$ is a pre-order, we often note $a < b$ when $a \leq b$ and $a \neq b$.*

Definition 3.3. *A relation $\sim$ on a set X, is an **equivalence relation** on X iff $\sim$ is reflexive, symmetric and transitive.*

Note that an equivalence relation is a pre-order that also satisfies the symmetry axiom. But the name *pre-order* comes from the fact that if we add the property of antisymmetry, we have a (partial) order.

Definition 3.4. *A relation $\leq$ on X is a **partial order** iff $\leq$ is reflexive, transitive, and anti-symmetric. In this case, we call $(X, \leq)$ a **partially ordered set**, or a **poset**.*

Definition 3.5. *A relation $\leq$ on X is a **total order** (or **linear order**) iff $\leq$ is complete, transitive, and anti-symmetric. In this case, we call $(X, \leq)$ a **totally ordered set**.*

Notice that both total order and partial order requires anti-symmetry, which rules out ties. This is in contrast to the economic concept of rational preference relation, which only requires completeness and transitivity, and therefore allows indifference between two alternatives.

It is easy to verify that the set inclusion relation $\subseteq$ is reflexive, transitive, anti-symmetric, but not complete.

Because completeness implies reflexiveness, a totally ordered set is a special case of partially ordered set. I'm going to state the following concepts in terms of poset to maintain the generality. In most applications, we will be working with total orders. For example, on the set $\mathbb{R}$ of real numbers, we can verify that the naturally defined relation

$$\leq := \Big\{(x, y) \in \mathbb{R}^2 : y - x \text{ is nonnegative}\Big\}$$

is a total order.

Based on a partial order $\leq$, we can define a few more relations for convenience:

- $<$: $x < y$ iff $x \leq y$ and $y \not\leq x$
- $\geq$: inverse of $\leq$
- $>$: inverse of $<$

[5] To state a definition, the convention is in fact to use "if" instead of "iff" when it actually means "iff". I simply use "iff" to avoid this ambiguity.

[6] By $x = y$, we mean x and y are the same element in X.

3.2 Upper/Lower Bound, Maximum/Minimum, and Supremum/Infimum

Definition 3.6. *Let $(X, \leq)$ be a poset, and let $A \subset X$.*

1. *$u \in X$ is an **upper bound** of A iff $u \geq x$, $\forall\ x \in A$. If such u exists, we say that the set A is bounded from above.*
2. *$l \in X$ is a **lower bound** of A iff $l \leq x$, $\forall\ x \in A$. If such l exists, we say that the set A is bounded from below.*
3. *$x^* \in A$ is a **maximum** of A iff x^* is an upper bound of A.*
4. *$x_* \in A$ is a **minimum** of A iff x_* is a lower bound of A.*

A maximum/minimum of A must be an element in A, and at the same time an upper/lower bound of A. Clearly, maximum/minimum of a set A may not exist, but when it exists, it must be unique due to anti-symmetry (exercise). So it makes sense to talk about "the" maximum/minimum if it exists, and we denote them as $\max A$ and $\min A$.

According to the definition, if A is the empty set, then any $u \in X$ is an upper bound, since the requirement ($u \geq x$ for any $x \in A$) is void.

A set may have many upper/lower bounds in general. However, there is a particular upper/lower bound of special interest.

Definition 3.7. *Let $(X, \leq)$ be a poset, and let $A \subset X$.*

1. *$u \in X$ is the **least upper bound**, or **supremum**, of A iff*
   (a) *u is an upper bound of A, and*
   (b) *$u \leq v$ for any upper bound v of A.*
2. *$l \in X$ is the **greatest lower bound**, or **infimum**, of A iff*
   (a) *l is a lower bound of A, and*
   (b) *$l \geq m$ for any lower bound m of A.*

By definition, the supremum is the minimum of the set of upper bounds, and the infimum is the maximum of the set of lower bounds. Therefore they are unique if exist, and so it makes sense to talk about "the" supremum/infimum. We denote them as $\sup A$ and $\inf A$.

Proposition 3.1. *(1) If a maximum exists, then it is the least upper-bound.*

(2) If a minimum exists, then it is the greatest lower-bound.

Notice that supremum/infimum may not exist even when upper/lower bounds exist. For example, let $X = \mathbb{R} \setminus \{0\}$ and $A = \{x \in \mathbb{R} : x < 0\}$. In the partially ordered set $(X, \leq)$, the set A has an upper bound (1 for example), but there is no least upper bound. We say that a poset $(X, \leq)$ has the **least upper bound (l.u.b.) property** iff any nonempty subset of X bounded from above has a least upper bound. As we have shown, the set $X = \mathbb{R} \setminus \{0\}$ endowed with the natural order $\leq$ does not have the l.u.b. property. However, the whole set of real numbers $\mathbb{R}$ endowed with $\leq$ has the l.u.b. property, due to the construction of real numbers.

4 Functions

There are several ways to define functions. The most general way is to view functions are special cases of relations.

Definition 4.1. *A relation f from X to Y is a **function** iff*

(1) $\forall\ x \in X$, $\exists\ y \in Y$ *s.t.* $(x, y) \in f$, *and*

(2) $\forall\ x \in X$ *and* $y_1, y_2 \in Y$ *s.t.* $(x, y_1) \in f$ *and* $(x, y_2) \in f$, *we have* $y_1 = y_2$.

Requirement (1) is an existence statement, and (2) is a uniqueness statement. Together they require that for each x, there exists one and only one y s.t. (x, y) has the relation f. That is, a function f is a special relation in which one $x \in X$ corresponds to one and only one $y \in Y$. As a result, it is unambiguous to use the notation $f(x)$ to denote the unique element $y \in Y$ s.t. $(x, y) \in f$. We call $f(x)$ the value of f evaluated at x. However, we are silent about how many x's a y may correspond to.

We use the notation $f : X \to Y$ to denote a function from X to Y, and we call X the **domain** of f, and Y the **codomain** of f. A function is also called a **mapping**.

We will sometimes denote Y^X for the set of functions from X to Y. Consider two functions $f, g \in Y^X$. We write $f = g$ to signify that the functions are equal, meaning that they take the same value at all points : $\forall x \in X, f(x) = g(x)$. It is important not to confuse the notation/equalities $f = g$ and $f(x) = g(x)$: the former is an equality on the *set of functions* Y^X, while the latter is an equality in Y (at a point). The latter is also a weaker statement since equality at a particular point is implied by equality everywhere.

Exercise 4.1. *Let I an interval of $\mathbb{R}$ and $f : I \to \mathbb{R}$ a function defined over I taking values in $\mathbb{R}$. Write mathematical statements (using quantifiers) to express the following statements :*

a. f takes the value zero

b. f is the zero function (takes the value zero everywhere)

c. f is not a constant function

d. f never takes the same value twice

4.1 Image and Inverse Image

Define the **image** of a set $S \subset X$ under f as

$$f(S) := \{f(x) : x \in S\}$$

i.e. the set of elements y in Y s.t. there exists $x \in S$ with $f(x) = y$.

Define the **inverse image** of a set $T \subset Y$ under f as

$$f^{-1}(T) := \{x \in X : f(x) \in T\}$$

i.e. the set of elements x in X s.t. $f(x)$ is in T.

The image of the whole domain X under f, $f(X)$, is called the **range** of f. Notice that $f(X)$ may be a proper subset of the codomain Y, since there may exist y's that correspond to no x.

We have the following results.

Claim 4.1. *Let $f : X \to Y$, $S_1, S_2 \subset X$ and $T_1, T_2 \subset Y$.*

(1a) $S_1 \subset S_2 \Rightarrow f(S_1) \subset f(S_2)$

(1b) $T_1 \subset T_2 \Rightarrow f^{-1}(T_1) \subset f^{-1}(T_2)$

(2a) $f(S_1 \cup S_2) = f(S_1) \cup f(S_2)$

(2b) $f^{-1}(T_1 \cup T_2) = f^{-1}(T_1) \cup f^{-1}(T_2)$

(3a) $f(S_1 \cap S_2) \subset f(S_1) \cap f(S_2)$

(3b) $f^{-1}(T_1 \cap T_2) = f^{-1}(T_1) \cap f^{-1}(T_2)$

(4a) No result for $f(S^c)$ *and* $\big(f(S)\big)^c$

(4b) $f^{-1}(T^c) = \big(f^{-1}(T)\big)^c$[7]

Proof. (1a), (1b), (2a), (3b), and (4b) are left as exercises.

(2b)

$\subset$:

Take any $x \in f^{-1}(T_1 \cup T_2)$. By definition of inverse image, we have $f(x) \in T_1 \cup T_2$. Then either $f(x) \in T_1$ or $f(x) \in T_2$. By definition of inverse image again, either $x \in f^{-1}(T_1)$ or $x \in f^{-1}(T_2)$. Therefore, $x \in f^{-1}(T_1) \cup f^{-1}(T_2)$.

$\supset$:

Take any $x \in f^{-1}(T_1) \cup f^{-1}(T_2)$. Then either $x \in f^{-1}(T_1)$ or $x \in f^{-1}(T_2)$. That is, either $f(x) \in T_1$ or $f(x) \in T_2$. Therefore, $f(x) \in T_1 \cup T_2$, which implies $x \in f^{-1}(T_1 \cup T_2)$.

(3a)

Take any $y \in f(S_1 \cap S_2)$. By definition of image, there exists $x \in S_1 \cap S_2$ s.t. $f(x) = y$. Therefore we have $x \in S_1$ and $x \in S_2$, and thus $y = f(x) \in f(S_1)$ and $y = f(x) \in f(S_2)$. So $y \in f(S_1) \cap f(S_2)$ □

4.2 Composition of Functions

Definition 4.2. *Let $f : X \to Y$ and $g : Y \to Z$. Then the* ***composite function*** *of f and g, denoted as $g \circ f$, is a function from X to Z s.t. for each $x \in X$, the value of the function $g \circ f$ is defined as $(g \circ f)(x) := g\big(f(x)\big)$.*

The composite function $g \circ f$ first applies f and maps x to $f(x)$ in Y, and then applies g and maps $f(x)$ to $g\big(f(x)\big)$ in Z. Notice that the composite operator is associative, i.e. $(h \circ g) \circ f = h \circ (g \circ f)$, since for each $x \in X$, the two functions both map x to $h\Big(g\big(f(x)\big)\Big)$.[8] As a result, we can use notations like $h \circ g \circ f$ with no ambiguity. However, the composite operator is not commutative, and so we cannot switch the order of h, g, and f.

[7] Notice that in the LHS the complement is taken in Y, while in the RHS the complement is taken in X. The equality really means $f^{-1}\big(Y \backslash T\big) = X \backslash f^{-1}(T)$.

[8] Two functions are the same iff they have the same domain and codomain, and for each element in the domain, they maps it to the same element in the codomain.

4.3 Injections, Surjections, and Bijections

Definition 4.3. *Consider a function $f : X \to Y$.*

*(1) f is an **injective function**, or **injection**, iff*

$$\forall\ x_1, x_2 \in X \ s.t.\ f(x_1) = f(x_2)\,,\ we\ have\ x_1 = x_2$$

*(2) f is a **surjective function**, or **surjection**, iff $f(X) = Y$*

*(3) f is a **bijective function**, or **bijection**, iff f is both injective and surjective.*

Injectivity requires that for all $y \in Y$, there exists *at most* one $x \in X$ such that $f(x) = y$ (uniqueness). Surjectivity requires that for all $y \in Y$, there exists *at least* one $x \in X$ such that $f(x) = y$ (existence). Bijectivity requires that for all $y \in Y$ there exists a unique $x \in X$ such that $f(x) = y$.

Exercise 4.2. *Let $f \in Y^X$. Show that f is injective if and only if for any $A, B \in \mathcal{P}(E)$:*

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

Some books use the term "*one-to-one* function" for injections, "function from X *onto* Y" for surjections, and "*one-to-one correspondence*" for bijections. We are going to stick to our terminologies.

If we view a function f as a relation from X to Y, we can think of its inverse relation f^{-1} from Y to X. Clearly, f^{-1} may fail to be a function. It is not difficult to see that f^{-1} is a function iff f is a bijection. That is the reason why we also call bijections **invertible functions**, since their inverse is still a function.

Definition 4.4. *Let $f : E \to F$ a bijection. By definition for all $\forall y \in F, \exists! x \in E, f(x) = y$. Define $f^{-1} : F \to E$ such that $f^{-1}(y) = x$ where $f(x) = y$. We call f^{-1} the reciprocal bijection of f.*

Remark 4.1. *It is important not to confuse the* inverse image of a set *$f^{-1}(A)$ – which is itself a set, and always exist for an arbitrary function f – and the reciprocal bijection f^{-1} understood as a function, which only exists when f is bijective. We commonly use the same notation because they are naturally related : when f is bijective the inverse image of A is the image of A by f^{-1}.*

Theorem 4.1. *Let f a bijection from E to F, then f^{-1} is a bijection from F to E. Furthermore :*

$$f^{-1} \circ f = Id_E \quad f \circ f^{-1} = Id_F$$

Where Id_E denotes the identity function defined by :

$$\begin{aligned} Id_E : E &\to E \\ x &\mapsto x \end{aligned}$$

A converse of the previous holds : if we find a function such that the composition with f yields the identity, then f is bijective (this also demonstrates that the reciprocal bijection is uniquely defined).

Theorem 4.2. *Let $f : E \to F$. If there exists $g : F \to E$ such that $g \circ f = Id_E$ and $f \circ g = Id_F$, then f is bijective and $g = f^{-1}$.*

The proof is left as a useful exercise.

Exercise 4.3. *Show that ther does not exist any surjective function from E into $\mathcal{P}(E)$ (this is a famous result due to Cantor). Hint : consider $\phi : E \to \mathcal{P}(E)$ and assume by contradiction that it is surjective, then consider the set $A := \{x \in E, x \notin \phi(x)\}$.*

We can also relate compositions of bijections to bijections of compositions with the following theorem.

Theorem 4.3. *Let $f : E \to F$ and $g : F \to G$ two bijective functions. Then $g \circ f$ is bijective and :*

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}$$

Exercise 4.4. *Prove Theorem 4.3.*

4.4 Monotonic Functions

When both the domain and the codomain are ordered sets, we can talk about monotonicity of a function. A monotonic function is simply a function that preserve, or inverse, the order.

Definition 4.5. *Let $(X, \leq_X)$ and $(Y, \leq_Y)$ be posets, and consider a function from X to Y.*

*(1) f is **weakly increasing** iff $x \leq_X x'$ implies $f(x) \leq_Y f(x')$.*

*(2) f is **weakly decreasing** iff $x \leq_X x'$ implies $f(x) \geq_Y f(x')$.*

*(3) f is **strictly increasing** iff $x <_X x'$ implies $f(x) <_Y f(x')$.*

*(4) f is **strictly decreasing** iff $x <_X x'$ implies $f(x) >_Y f(x')$.*

5 Numbers

In this section, we introduce and give some elements of construction for the usual number sets :

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$$

5.1 Natural Numbers

The set $\mathbb{N}$ of **natural numbers** is a fundamental set that is defined axiomatically – through what is known as Peano's axioms. Informally, $\mathbb{N}$ is constructed recursively by "counting" : take a starting element 0, then the next element or successor of 0 (denoted as 1), the successor of the successor of 0 (denoted as 2), and so on. Shortly put, $\mathbb{N} := \{0, 1, 2, \ldots\}$.

Formally, we claim that there exists a set $\mathbb{N}$ that contains an element denoted 0 and is equipped with a function $s : \mathbb{N} \to \mathbb{N}$, where $s(n)$ is called the successor of n and denoted $n + 1$, such that :

(i) s is a bijection from $\mathbb{N}$ to $\mathbb{N}^* := \mathbb{N} \setminus 0$

(ii) **Induction principle :** Let A a subset of $\mathbb{N}$ verifying the two following properties :

a. $0 \in A$

b. $\forall n \in \mathbb{N}, (n \in A \Rightarrow n + 1 \in A)$

then $A = \mathbb{N}$

Observe that the recursion or induction principle which is sometimes seen as a property is actually the very definition of natural numbers : if you can move from one to the next and you contain the initial point, the you contain all points of the set. An immediate consequence is the principle of induction for proving statements indexed by natural numbers.

Proof by Induction Principle : Let $P(n)$ a predicate depending on $n \in \mathbb{N}$. If :

1. $P(0)$ is true
2. And $\forall n \in \mathbb{N}$, $P(n) \Rightarrow P(n+1)$

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

This is a very powerful proof technique, which directly follows from the construction of natural numbers when considering the set $A := \{n \in \mathbb{N}, P(n) \text{ is true }\}$. Instead of proving directly that $P(n)$ is true for all n, we can prove that it holds for 0 (initialization) and then prove that if $P(n)$ is true, then $P(n+1)$ is also true (induction).

There are natural variations on the induction principle :

- Induction starting at a non-zero index : for $n_0 \in \mathbb{N}$, if $P(n_0)$ is true and for all $n \geq n_0$, $P(n) \Rightarrow P(n+1)$, then $P(n)$ is true for all $n \geq n_0$
- Double induction : if $P(0)$ and $P(1)$ are true and for all $n \geq 0$, $P(n)$ and $P(n+1) \Rightarrow P(n+2)$, then $P(n)$ is true for all $n \in \mathbb{N}$.
- Strong induction : if $P(0)$ is true and for all $n \geq 0$, $(\forall k \leq n, P(k)) \Rightarrow P(n+1)$, then $P(n)$ is true for all $n \in \mathbb{N}$.
- Finite induction : let $P(n)$ depend on $n \in \{0, ..., n_0\}$; if $P(0)$ is true and for all $n \in \{0, ..., n_0\}$, $(\forall k \leq n, P(k)) \Rightarrow P(n+1)$, then $P(n)$ is true for all $n \in \{0, ..., n_0\}$

Exercise 5.1. *1. Use a proof by induction to show that :*

$$\sum_{k=0}^{n} k = \frac{n(n+1)}{2}$$

2. Similarly show that

$$\sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$$

Where we recall that the indexed sum notation $\sum_{k=0}^{n} u_k$ is a shorthand for $u_0 + u_1 + ... + u_N$ where $(u_k)_{k\in\mathbb{N}}$ is a sequence indexed by k.

Given the construction of $\mathbb{N}$ we can define **sequences** which we can view either as collections indexed by integers $(u_k)_{k\in\mathbb{N}}$ or equivalently as mappings from $\mathbb{N}$ into some set E. Given the induction principle, we can define sequences uniquely by induction : let E a set, $a \in E$ an element and $f : E \to E$ a function; there exists a unique sequence $(v_n)_{n\in\mathbb{N}}$ of elements of E such that :

$$\begin{cases} u_0 = a \\ \forall n \in \mathbb{N}, u_{n+1} = f(u_n) \end{cases}$$

We can also show results on maxima and minima of subsets of $\mathbb{N}$.

Theorem 5.1. *Any non-empty subset of $\mathbb{N}$ has a unique minimum. Any non-empty subset of $\mathbb{N}$ that is bounded above has a maximum.*

5.2 Signed Integers and Rationals

The set $\mathbb{Z}$ of (signed) **integers** consists of natural numbers and their negative counterparts, i.e. $\mathbb{Z} := \{\ldots, -2, -1, 0, 1, 2, \ldots\}$. In all rigor, this is constructed by defining properly the usual addition $+$ on $\mathbb{N}$ and defining for all $n \in \mathbb{N}$ $-n$ to be the unique element such that $n + (-n) = 0$.

The set $\mathbb{Q}$ of **rational numbers** consists of ordered pairs (m, n) of integers, and treats (m, n) and (m', n') as the same element iff $m \cdot n' = m' \cdot n$. $\mathbb{Q} := \{m/n : m.n \in \mathbb{Z}, m, n \text{ are coprime}\}$.[9]

5.3 Real numbers

We have seen before that $\mathbb{Q}$ do not contain "all numbers" – numbers like π or $\sqrt{2}$ do not belong to $\mathbb{Q}$ and are as such called "irrational". The real line is constructed by "completing" the set of rational numbers in some appropriate sense, i.e filling in the gaps in $\mathbb{Q}$

We will not define/construct the real line explicitly, as this is outside the scope of this class. Instead, we state the key property of the real line:

Axiom 5.1. $(\mathbb{R}, \leq)$ *has the least upper bound property.*

$(\mathbb{Q}, \leq)$ does not have the least upper bound property. See Chapter 1 of Rudin's book for a detailed discussion[10]

5.4 Complex numbers

The set of complex numbers $\mathbb{C}$ is at its core just the set $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ the set of all ordered pairs of $\mathbb{R}$, equipped with the usual addition and its particular notion of inner product :

$$(a, b) + (a', b') = (a + a', b + b')$$
$$(a, b) \cdot (a', b') = (aa' - bb', ab' + a'b)$$

The set really becomes $\mathbb{C}$ once we define those operations on it: an addition and a multiplication, like on $\mathbb{R}$. Instead of noting $z = (a, b) \in \mathbb{R}^2$, we note $z = a + ib$, where i is the **imaginary unit**. We call a the **real part**, noted $Re(z)$, and b the **imaginary part** of $a + ib$, noted $Im(z)$. We see the real line as the subset of $\mathbb{C}$ whose numbers have imaginary part equal to zero. The operations are defined so that they generalize the operations on $\mathbb{R}$, which we can see by writing them in complex form :

- The addition: $(a + ib) + (c + id) = (a + b) + i(b + d)$.
- The multiplication : we define $i^2 = -1$ so that $(a + ib)(c + id) = (ac - bd) + i(bc + ad)$.

We define the **conjugate** of a complex number $z = a + ib$ to be $\bar{z} = a - ib$.
We define the **modulus** of a complex number $z = a + ib$ to be $|z| = \sqrt{a^2 + b^2}$. The modulus extends the notion of absolute value $|x| = \max(x, -x)$ on $\mathbb{R}$.
In order to define division note, when $z = (c, d) \neq (0, 0) \in \mathbb{C}$, $(c, d)^{-1} = \left(\frac{c}{c^2 + d^2}, \frac{-d}{c^2 + d^2}\right)$ and hence $(a, b)/(c, d) = (a, b)(c, d)^{-1}$
One can verify that $z\bar{z} = |z|^2$.

[9] m, n are coprime if the only positive integer that divides both m and n is 1.

[10] There are several equivalent ways to construct real numbers. Some people use the l.u.b. property directly as a part of the definition of real numbers, while some people construct real numbers without explicitly using the l.u.b. property, but later derive it as a property of real numbers. All we need to know for now is that $(\mathbb{R}, \leq)$ has the l.u.b. property.

6 Countability and Cardinality

6.1 Countability of Sets

It is natural to use the number of elements in a set to get an idea about the size of it. However, this approach does not work if the set has infinitely many elements, and we need a more sophisticated approach.

Definition 6.1. *A set X is **countably infinite** iff there exists a bijection between $\mathbb{N}$ and X.*

The next claim states that countably infinite sets are the "smallest" infinite sets.

Claim 6.1. *If X is countably infinite, then any infinite subset $Y \subset X$ is also countably infinite.*

Proof. * Because X is countably infinite, by definition there exists bijection $f : \mathbb{N} \to X$. Consider the inverse image of Y under f, $f^{-1}(Y)$, which is an infinite set of natural numbers. Let n_1 be the smallest number in $f^{-1}(Y)$, n_2 be the second smallest number in $f^{-1}(Y)$... Define the function $g : \mathbb{N} \to Y$ as $g(i) := f(n_i)$ for each $i \in \mathbb{N}$. By construction, g is a bijection from Y to $\mathbb{N}$, and so Y is countably infinite. □

Definition 6.2. *A set X is **countable** iff it is either finite or countably infinite. A set X is **uncountable** iff it is not countable.*

Intuitively, uncountable sets are infinite sets which we can not find a bijection from $\mathbb{N}$ to. Therefore, they are considerably "larger" than countably infinite sets.

Theorem 6.1. *The countable union of countable sets is countable.*

Proof. * We need to extend the previous proof to an infinitely countable union. Consider an infinitely countable collection of set $\{S_1, S_2, ...\}$. For each set S_j, since S_j is countable, we can index its argument as $(x_1^j, x_2^j, x_3^j, ...)$. Place each sequence $(x_1^j, x_2^j, x_3^j, ...)$ as the j^{th} column of a matrix with infinitely many rows. The problem is that this time, the matrix has also infinitely many columns, so we cannot follow our previous proof. However, Georg Cantor, the late XIXth-century mathematician who founded set theory, got a smart idea: snake along the diagonals to define the sequence $(x_1^1, x_2^1, x_1^2, x_1^3, x_2^2, x_3^1, ...)$. This way, we count all the arguments of $\bigcup_n S_n$. (Again, some elements may repeat themselves; just drop them). □

Using this proposition, it is easy to show that $\mathbb{Q}$ is countable.

Corollary 6.1. *$\mathbb{Q}$ is countable.*

Proof. * Partition $\mathbb{Q}$ by the value of the denominators of fractions: $\mathbb{Q} = \bigcup_{q \in \mathbb{N}, q \neq 0} Q_p$ with $Q_p = \{\frac{p}{q}, p \in \mathbb{N}\}$. Each Q_p is countable, so $\mathbb{Q}$ is a countable union of countable sets. □

Proposition 6.1. *The cartesian product of finitely many countable sets is countable.*

Proof. * It is enough to prove it for the cartesian product of two sets; the result will then follow by induction. The proof is a corollary of the result on the countable union of countable sets. Let S and T be two countable sets. Note $(x_1, x_2, ...)$ the elements of S. Write the cartesian product as $S \times T = \bigcup_n T_n$, where $T_n = \{(x_n, y), y \in T\}$. Each T_n is countable, so $S \times T$ is a countable union of countable sets. □

Claim 6.2. *$\mathbb{R}$ is uncountable.*

6.2 Distinguishing between infinities*

Definition 6.3. ** Let S and T two sets.*

- *S and T have **equal cardinality** if there exists a bijection between them.*
- *S has **higher cardinality** than T if there exists an injection from T onto S.*
- *S has **strictly higher cardinality** than T if it has a higher but not equal cardinality than T.*

We can prove Claim 6.2 by showing it has strictly higher cardinality than $\mathbb{N}$.

Proof. * First let's show that $S := \{0,1\}^{\mathbb{N}}$ (an element of S is a sequence of 0 and 1) is uncountable. By contradiction, assume it is countable. Then there exists a bijection from $\mathbb{N}$ to S; call it $(s^1, s^2, ..., s^n, ...)$, where each s^j is an element of S, i.e. a sequence of 0 and 1. Place the sequence s^j as the j^{th} column of an infinite matrix. Now, consider the diagonal of this matrix: it is a sequence of 0 and 1, that is an element of S; call it d. Define the sequence $s^* \in S$ such that for all i, s^*_i is the complement of d_i ($s^*_i = 0$ if $d_i = 1$, and conversely). But s^* is different from all $s^j, j \in \mathbb{N}$ since $s^j_j \neq s^*_j$, which means $s^* \notin S$, a contradiction.

Then we prove $\mathbb{R}$ is in bijection with $\{0,1\}^{\mathbb{N}}$ is in two steps. First we show that $\mathbb{R}$ is in bijection with the interval $(0,1)$, then that $(0,1)$ is in bijection with $\{0,1\}^{\mathbb{N}}$.

- For the first step, note that $x \mapsto \tan(\pi x - \frac{\pi}{2})$ works. Or for that matter, any bijection map from $\mathbb{R} \to (0,1)$ would work.
- For the second step, the idea is to use the binary expression of a real number: any real $x \in (0,1)$ can be written as $x = \sum_{n=1}^{\infty} a_n 2^{-n}$, so that the sequence $(a_n)_n \in \{0,1\}^{\mathbb{N}}$ characterizes x. There is a small difficulty because some numbers have actually two such representations: for instance $2^{-1} = \sum_{n=2}^{\infty} 2^{-n}$ are the same number. But we just need to agree not to use the second representation for such numbers.

Therefore, $\mathbb{R}$ is uncountable.

□

Proposition 6.2. ** $\mathbb{R}^n$ (and so $\mathbb{C}$) has the same cardinality as $\mathbb{R}$.*

Proof. * By induction, it is enough to prove that $\mathbb{R}^2$ has the same cardinality as $\mathbb{R}$. Since $\mathbb{R}$ is in bijection with $S = \{0,1\}^{\mathbb{N}}$, $\mathbb{R}^2$ is in bijection with S^2. Hence it is enough to find a bijection between S^2 and S. Consider the function $f : S^2 \to S$ that associates to the couple $((a_n)_n, (b_n)_n)$ the interlaced sequence $(a_1, b_1, a_2, b_2, ...)$. It is a bijection. □

Remark 6.1. *It is important not to confuse the notion of cardinality with the notion of* dimension*, which we will introduce later when reviewing linear algebra. For instance $\mathbb{R}$ and $\mathbb{R}^2$ have the same cardinality, but they do not have the same dimension. Dimension is another notion of "size" of a set which is adapted to vector spaces. Later, we will see yet another notion of size when talking about measure theory and the idea of "measure" of sets. It is important not to confuse these notions as they are tailored for different purposes.*