Fund. Math. Topic 1: Set Theory

Oxbridge Academic Bootcamp 2025

Seppe J. Staelens

August 15, 2025

Introduction

We will start the course by introducing the basics of set theory, a branch of mathematics that lies at the foundation of all other mathematical concepts. We will discuss basic set operations, the power set and functions between sets. This will form the basis for what we will discuss in the other topics.

Lecture 1: Set Theory

A **set** is one of the most basic, yet abstract concepts in mathematics. A set is a *collection of objects* - summarized in the most non-mathematical way. A set consists of different elements, and this can really be anything, from integers to real numbers, or even matrices, functions and more abstract objects. Other mathematical structures, like groups, rings and vector spaces, are obtained from equipping sets with more properties and relations between them.

1.1 Basic definitions

For now, we aim to denote sets with capital letters and their elements with small letters. We say that

- x is an **element** of A if x is contained in the set A. This is denoted as $x \in A$.
- x is not an element of A otherwise. This is denoted as $x \notin A$.
- the set A is a **subset** of the set B if every element of A is in B as well. This is denoted as $A \subset B$.
- the set A and B are equal if they contain the same elements. Equality of sets is denoted as $A = B$. Given the above definition, this is equivalent to $A \subset B$ and $B \subset A$ holding at the same time.

Sets can be denoted in different ways. The most common way is to list the elements of the set between curly brackets, separated by commas. For example, the set of the first three natural numbers[1] can be written as

$$A = \{0, 1, 2\}. \tag{1}$$

[1]Note that we treat 0 as belonging to the natural numbers $\mathbb{N}$, which is not always the case. The set of natural numbers *without* 0 is denoted as $\mathbb{N}_0$ in our convention.

Another way to denote sets is by specifying a **set condition**. For example, the set of all even natural numbers can be written as

$$B = \{x \in \mathbb{N} \mid x \text{ is even}\}. \tag{2}$$

The vertical line precedes the condition(s) that the elements of the set must satisfy. We can use similar notation to specify a set by means of a formula. For example, the set of all perfect squares can be written as

$$C = \{n^2 \mid n \in \mathbb{N}\}. \tag{3}$$

The size of a set, also known as **the cardinality** of the set, is denoted as $\#A$. For finite sets, i.e. sets with a finite number of elements, this is simply the number of elements. A set with cardinality 1, i.e. a set with only one element, is called a **singleton (set)**. The (unique) set with no elements, i.e. cardinality 0, is called the **empty set**, and is denoted as $\emptyset$. For sets with an infinite amount of elements, like the integers, this is a bit more subtle. We will come back to this in Lecture 3.

1.2 Operations on sets

Given two sets A, B, we can define the following operations:

- The **union** of A and B, denoted as $A \cup B$, is the set of all elements that are in A or in B. In mathematical notation, this is
$$A \cup B = \{x \mid x \in A \text{ or } x \in B\}. \tag{4}$$
- The **intersection** of A and B, denoted as $A \cap B$, is the set of all elements that are in both A and B. In mathematical notation, this is
$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}. \tag{5}$$
If two sets have an empty intersection, i.e. $A \cap B = \emptyset$, we say that the sets are **disjoint**.
- The **set difference** of A and B, denoted as $A \setminus B$, is the set of all elements that are in A but not in B. In mathematical notation, this is
$$A \setminus B = \{x \mid x \in A \text{ and } x \notin B\}. \tag{6}$$
Note that this is not symmetric, i.e. $A \setminus B \neq B \setminus A$ (much like $a - b$ is in general not equal to $b - a$, for real numbers a, b).
- The **Cartesian product** of A and B, denoted as $A \times B$, is the set of all pairs (a, b) with $a \in A$ and $b \in B$. In mathematical notation, this is
$$A \times B = \{(a, b) \mid a \in A \text{ and } b \in B\}. \tag{7}$$
- If we consider some *Universe* U, a 'large' set containing all elements that could potentially be relevant for our problem, and a subset $A \subset U$, we can define the **complement** of A in U as
$$A^c = U \setminus A. \tag{8}$$
- The **power set** of A, denoted as $\mathcal{P}(A)$, is the set of all subsets of A. In mathematical notation, this is
$$\mathcal{P}(A) = \{B \mid B \subset A\}. \tag{9}$$
Note that this is *set of sets*, meaning that every element of $\mathcal{P}(A)$ is a set itself.

The first five of these operations are rather straightforward, and we discuss some examples in class. The power set, however, deserves a bit more attention.

1.2.1 Power set

The power set of a set A is the set of all subsets of A. Intuitively, the power set must be larger than the original set. Indeed, for every element in the set A, call it x, the singleton $\{x\}$ is a subset of A, and thus an element of the power set. This means that the power set contains at least as many elements as there are in A. But, do we know whether the power set is always larger than A?

Lemma 1. *(Binomial Theorem) For any natural number $n \in \mathbb{N}$ and any real numbers $a, b \in \mathbb{R}$, we have*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k. \tag{10}$$

This is a very useful formula, which can be proven in many ways (as we will see in the next lecture). The coefficient $\binom{n}{k}$ is called a **binomial coefficient**, and is defined as

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}\,, \tag{11}$$

where $n! = n \cdot (n-1) \cdot \dots \cdot 2 \cdot 1$ is the **factorial** of n. The binomial coefficient counts the number of ways to choose k elements from a set of n elements, and is a very important concept in *combinatorics*.

The summation symbol $\sum$ indicates that we take a sum over terms. For example, we have that

$$\sum_{k=0}^{n} k = 1 + 2 + 3 + \dots + (n-1) + n\,,$$

i.e. we take the sum over the first natural numbers. The summation symbol usually has a lower limit and upper limit, though the latter can be infinite. The k is called a *dummy index*, and is the variable of the summation, which changes in every term of the sum, between the lower and upper bounds. Note that the dummy index itself does not hold any meaning, and we could as well have used i. More specifically,

$$\sum_{k=0}^{n} k = \sum_{i=0}^{n} i\,.$$

We now use Lemma 1 to prove the following theorem.

Theorem 1. *Given a finite set A with $\#A = N$, the power set of A, $\mathcal{P}(A)$, has 2^N elements.*

Proof. We will count the subsets of A by considering subsets of size $k = 0, 1, \dots, N$.

For $k = 0$, there is only one subset of A with no elements, namely the empty set $\emptyset$. For $k = 1$, there are N subsets of A with one element, namely the singletons $\{x\}$ for $x \in A$. For $k = 2$, we need to find the number of subsets with 2 elements. Clearly, any subset with two elements is obtained by *choosing* two elements from A. We have seen right before that the number of ways this can be done is $\binom{N}{2}$, and therefore there are $\binom{N}{2}$ subsets of A with two elements.

Similarly, for $k = 3$, there are $\binom{N}{3}$ subsets of A with three elements, and so on. Finally, for $k = N$, there is only one subset of A with N elements, namely A itself.

Therefore, adding up all these numbers, we find that the total number of subsets of A is

$$\begin{aligned}
\#\mathcal{P}(A) &= 1 + N + \binom{N}{2} + \binom{N}{3} + \cdots + \binom{N}{N-1} + 1 \\
&= \binom{N}{0} + \binom{N}{1} + \binom{N}{2} + \binom{N}{3} + \cdots + \binom{N}{N-1} + \binom{N}{N} \\
&= \sum_{k=0}^{N} \binom{N}{k} \\
&= \sum_{k=0}^{N} \binom{N}{k} 1^{N-k} 1^{k} \\
&= (1+1)^N \qquad \text{by the binomial theorem} \\
&= 2^N.
\end{aligned}$$

□

Lecture 2: Functions

You have already discussed functions in high school and last week, in the context of functions of real numbers. For example, you probably have seen something like a linear function $f(x) = ax+b$, where x is a real number, and studied its graph - which is a straight line. Functions, however, are a much more general concept, and can be defined between any two sets.

A **function (map)** from a set X to a set Y is a map that assigns to each element $x \in X$ exactly one element $y \in Y$. This is denoted as $f : X \to Y$. This clearly encompasses the concept of a function that you are more familiar with, which is a function from $\mathbb{R}$ to $\mathbb{R}$. We discuss more general examples in class.

Usually we describe a function by specifying its **image**, i.e. given $x \in X$ we specify the element $f(x) \in Y$. A more extensive notation for the function would then be

$$f : X \to Y : x \mapsto f(x). \tag{12}$$

Note the different arrows in the expression above: they are not the same. Usually the set X is called the **domain** of the function, and the set Y is called the **codomain** of the function.

Given two functions $f : X \to Y$ and $g : Y \to Z$, we can define the **composition** of these functions as

$$g \circ f : X \to Z : x \mapsto g(f(x)). \tag{13}$$

This is a new function, which maps elements of X to elements of Z by first applying f and then g.

Given subsets $A \subset X$ and $B \subset Y$ and a function $f : X \to Y$, we can define the **image** of A under f as

$$f(A) = \{f(x) \mid x \in A\}. \tag{14}$$

Similarly, we can define the **preimage** of B under f as

$$f^{-1}(B) = \{x \in X \mid f(x) \in B\}. \tag{15}$$

Note that this is not the same as the inverse of the function f, which is a different concept.

2.1 Injection, surjection, bijection

To properly define some concepts in this section, we first briefly discuss quantifiers:

- The **existential quantifier** $\exists$ is used to denote that there exists at least one element in a set that satisfies a certain property. For example, with $X \subset \mathbb{R}$, the expression $\exists x \in X : x > 2$ means that there is at least one element x in X that is larger than 2.
- The **universal quantifier** $\forall$ is used to denote that all elements in a set satisfy a certain property. For example, with $X \subset \mathbb{R}$, the expression $\forall x \in X : x > 2$ means that all elements x in X are larger than 2.
- The **unique quantifier** $\exists!$ is used to denote that there exists exactly one element in a set that satisfies a certain property. For example, with $X \subset \mathbb{R}$, the expression $\exists! x \in X : x > 2$ means that there is exactly one element x in X that is larger than 2.

We use these quantifiers in the following definitions. Given a function $f : X \to Y$, we can give the following definitions:

- The function f is called an **injection** if for every $y \in Y$, there is at most one $x \in X$ such that $f(x) = y$. In mathematical notation, this can be expressed as[2]

$$\forall x_1, x_2 \in X : x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2). \tag{16}$$

[2] Note the use of the "$\Rightarrow$" symbol, to which we come back in Lecture 2.

- The function f is called a **surjection** if for every $y \in Y$, there is at least one $x \in X$ such that $f(x) = y$. In mathematical notation, this can be expressed as

$$\forall y \in Y, \exists x \in X : f(x) = y. \tag{17}$$

- The function f is called a **bijection** if it is both an injection and a surjection. In mathematical notation, this can be expressed as

$$\forall y \in Y, \exists! x \in X : f(x) = y. \tag{18}$$

2.2 The inverse function

The concept of a bijection becomes important when we want to define the **inverse function** of a function $f : X \to Y$.

Definition 1. *The inverse function of a* ***bijection*** $f : X \to Y$ *is a function* $f^{-1} : Y \to X$ *that assigns to each element* $y \in Y$ *the unique element* $x \in X$ *such that* $f(x) = y$*. In particular, we must have that*

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

Indeed, remember that the key requirement for a map to be a function is that it assigns to each element $x \in X$ exactly one element $y \in Y$. Suppose now that f is not injective, i.e. there are two different elements $x_1, x_2 \in X$ such that $f(x_1) = f(x_2) = y$ for some $y \in Y$. Then, we cannot define an inverse function $f^{-1} : Y \to X$ that satisfies Def. 1. Indeed, suppose that $f^{-1}(y) = x_1$: we would have that $f^{-1}(f(x_2)) \neq x_2$.

Similarly, if f is not surjective, i.e. there is some $y \in Y$ such that there is no $x \in X$ with $f(x) = y$. This again means we cannot define an inverse function $f^{-1} : Y \to X$ that satisfies Def. 1. Thus, we conclude that f must be a bijection in order for an inverse function to exist.

Example 1. *Consider the function* $f : \mathbb{R} \to \mathbb{R} : x \mapsto x^2$*. This function is not injective, as both* $f(1) = 1$ *and* $f(-1) = 1$*. Additionally,* f *is not surjective, as there is no* $x \in \mathbb{R}$ *such that* $f(x) = -1$*. This means we cannot define an inverse function for* f*. However, if we restrict the domain and codomain of* f *to* $\mathbb{R}^+$*, i.e. the positive real numbers, we get*

$$f^* : \mathbb{R}^+ \to \mathbb{R}^+ : x \mapsto x^2. \tag{19}$$

This function now is a bijection, meaning that at this point we can define an inverse function $f^{-1} : \mathbb{R}^+ \to \mathbb{R}^+$*. Knowing that* $\forall x \in \mathbb{R}^+ : \sqrt{x^2} = x$*, we know that this must be*

$$f^{-1} : \mathbb{R}^+ \to \mathbb{R}^+ : y \mapsto \sqrt{y}. \tag{20}$$

References

These notes are based on my own knowledge of these basic mathematical concepts, and the writing has been accelerated by the use of *GitHub copilot* and its implementation in VSCode. Inspiration (and most of Lecture 3) has been taken from the course notes for *"Bewijzen en Redeneren"* (Proving and Reasoning), used in the first year of the Bachelor of Mathematics at the KU Leuven, at the time taught by Prof. Arno Kuijlaars - also the author of the lecture notes.