Some semi-random background content

Contents

Chapter 1

Sets and functions

In this chapter we quickly recall some basic notions and notation about sets and functions. You should be familiar with most of it from first year.

1.1 Sets

For us, a set is simply be a collection of objects. These objects can be anything and are called the elements of the set. (This is a very intuitive notion of set, that will be sufficient for us. A precise study of the notion of set is beyond the scope of this course and is rather non-trivial. If you feel so inclined you may try to look up something called Russell's paradox)

There are two main notations used for defining sets. In the first one, a set is given by listing all its elements, and in the second one the set is given by describing its elements.

Set given by listing its elements

This is done for instance as follows:

$$\{1, -\pi, \star\},\ \{1, 2, 3, 4\}, \{\text{the colour blue}, -1, \text{the sun (the actual one)}\}.$$

The ellipsis $(\ldots)$ can be used to indicate that the list of elements goes on "in the obvious way", for instance:

$$\{1, 2, 3, 4, \ldots\}$$

denotes the set of all positive integers (also denoted by $\mathbb{N}$).

Set given by describing its elements

This is usually done by giving a property that characterizes the elements of the set. For instance

$$\{n \mid n \text{ is a positive integer, and 3 divides n}\}$$

denotes the set $\{3, 6, 9, 12, \ldots\}$ (and reads "the set of all n such that n is a positive integer and 3 divides n"). . More generaly, this is done by indicating a letter that will be used in the property, and then indicating the property: The notation

$$\{x \mid \text{some property involving } x\}$$

(sometimes written

$$\{x : \text{some property involving } x\}$$

)
means "the set of all elements (which we call x in the following property) such that the given property is true for x". In this case, if we want to know if some object (any kind of object) is in the set, we temporarilly call it x and check if the property is true for this value of x. If it is, the element belongs to the set, if not, the element does not belong to the set.

Both notations use the curly brackets { and }. They indicate that you are defining a set.

Example 1.1. *1.* $\mathbb{N} = \{1, 2, 3, \ldots\}$ *the set of all positive integers.*

2. $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, 3, \ldots\}$ *the set of all integers.*

3. $\mathbb{Q} = \{a/b \mid a, b \text{ are integers}, b \neq 0\}$*, the set of all rational numbers.*

4. $\mathbb{R}$ *denotes the set of all real numbers (it is more complicated to give a precise description of it).*

Some variations on the above notation are common and are usualy clear, for instance

$$\{x^2 \mid x \in \mathbb{R} \text{ and } x > 3\}$$

denotes the set of all squares of elements of $\mathbb{R}$ that are greater than 3 (it is equal to $(9, \infty)$).

Remark 1.2. *An important feature of the definition of set is that a set is completely determined by the elements that it contains. In other words, if two sets contain the same elements, they are equal:*

$$\{1, 2\} = \{1, 2, 2, 2\} = \{2, 1\}, \ \{1, 2, 3\} \neq \{1, 2, 3, 4\}.$$

As you can see, it means in particular that:

1. Repeating an element does not change the set;

2. The order in which the elements are listed does not matter.

Definition 1.3. *The empty set is the set that does not contain any element. Using the notation introduced above, it can be written* $\{\}$*, but the most common notation for it is* $\emptyset$ *(an "O" with a diagonal bar).*

The notation $\boldsymbol{a \in B}$ is short for "a is in B" (or "a belongs to B", or "a is an element of B"). It means that B is a set, and that a is an element of B. We write $a \notin B$ to denote that a is not an element of B. So, for instance:

$$3 \in \mathbb{N},\ -1 \notin \mathbb{N},\ 4 \notin \{-10, \star, \pi\}.$$

If S and T are both sets, the notation $\boldsymbol{S \subseteq T}$ (or $S \subset T$) means that every element of S is also in T, and we say that S **is contained in** T (or that S is **included in** T, or that S **is a subset of** T). Similarly, the notation $S \not\subseteq T$ means that S is not a subset of T, i.e., that there is at least one element of S that is not in T. For instance

$$\{1,2\} \subseteq \{1,2,3\},\ \{1,2\} \subseteq \{1,2\},\ \mathbb{N} \subseteq \mathbb{Q},$$

but

$$\{-1,2\} \not\subseteq \mathbb{N},\ \mathbb{Q} \not\subseteq \{1,2,3\}.$$

Remark 1.4. *Observe that if A is any set, then the empty set is a subset of A. Why? Because if it were not the case, then there would be an element in the empty set that is not in A (by definition of being a subset). But the empty set does not contain any element, so it is not possible.*

We say that S is a **proper subset** of T if $S \subseteq T$ and $S \neq T$ (i.e. every element of S is in T, and there is at least one element of T that is not in S). This is denoted by $S \subsetneq T$ or $S \subsetneqq T$.

Finaly, if S is a finite set, we denote by $|S|$ the number of elements in S. If S is infinite, we write $|S| = \infty$. This number $|S|$ is called **the cardinality of S**.

1.2 Union and intersection

If A and B are two sets, we can build:

1. The **union** of S and T, denoted $\boldsymbol{S \cup T}$. It is the set that consists of all the elements of S together with all the elements of T:

$$S \cup T = \{x \mid x \in S \text{ or } x \in T\}.$$

For instance

$$\{-3,4\} \cup \mathbb{N} = \{-3,1,2,3,4,5,\ldots\},$$
$$\{1,2\} \cup \{\star,-6\} = \{1,2,\star,-6\}.$$

2. The **intersection** of S and T, denoted $\boldsymbol{S \cap T}$. It is the set that consists of all the elements that are in both S and T:

$$S \cap T = \{x \mid x \in S \text{ and } x \in T\}.$$

For instance

$$\{-3, 4\} \cap \mathbb{N} = \{4\},$$
$$\{1, 2\} \cap \{\star, -6\} = \emptyset.$$

1.3 Cartesian product

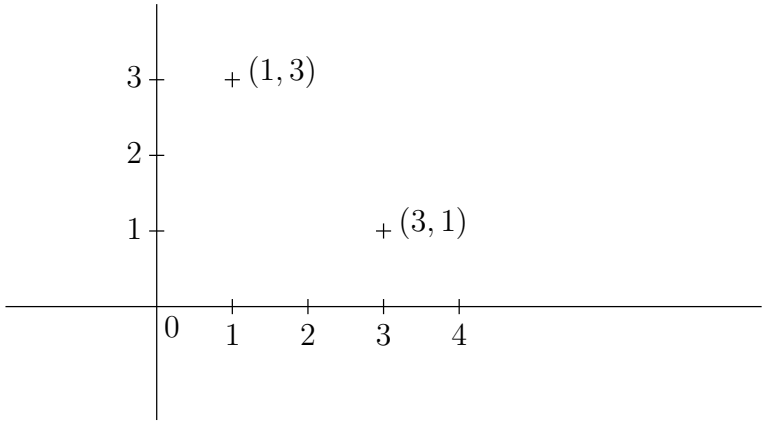
(Named after René Descartes. We will simply say "product".)

A **pair of elements** (more precisely, an ordered pair of elements) consists of two elements (which can be absolutely anything, as usual), where one of them is the first element and the other the second element. So the order of the elements is important.

If a and b are these elements, the pair where a is the first element and b the second is denoted by (a, b). Similarly, the pair where b is the first element and a the second is denoted by (b, a).

The order matters, so $(10, 3) \neq (3, 10)$.

You should already be familiar with this concept: Pairs of real numbers are used to identify points in the plane. As you can see, the order does matter:



The concept of pairs is very useful, and there is no reason to only use it with real numbers. It is also used to build the elements of what is called the product of two sets:

Definition 1.5. *If S and T are sets, the* **product** *of S by T (in this order), denoted $S \times T$, is the set of all pairs having as first element an element from S and as second element an element from T, so*

$$S \times T = \{(x, y) \mid x \in S \text{ and } y \in T\}.$$

Example 1.6.

$$\{-3, \star\} \times \{1, 2, 3\} = \{(-3, 1), (-3, 2), (-3, 3), (\star, 1), (\star, 2), (\star, 3)\}.$$

Observe that if S and T are finite, then $|S \times T| = |S| \cdot |T|$.

The same construction can be used for more than 2 sets:

Definition 1.7. *An ordered list of n elements is called an n-tuple (so a pair is a 2-tuple), and if $S_1, S_2, \ldots, S_n$ are sets, the product $S_1 \times S_2 \times \cdots \times S_n$ is the set of all n-tuples consisting of and element of S_1, an element of S_2, ..., an element of S_n. For instance, if S, T, U are sets:*

$$S \times T \times U = \{(x, y, z) \mid x \in S,\ y \in T,\ z \in U\}.$$

1.4 Functions

Given two sets S and T, a function (also called a map) from S to T is "something" that associates an element of T to each element of S. If the function is called f, we denote this by $f : S \to T$, and the element from T associated to an element a from S is denoted by $f(a)$ (and is called the image of a by f).

This "something" can be anything that associates an element of T to each element of S: A formula, a procedure explcitely described, or something completely arbitrary. A common compact notation for indicating how f is defined is:

$$f : S \to T,\ x \mapsto \text{"how to get } f(x) \text{ out of } x\text{"},$$

see examples below.

Examples 1.8.

1. *$f : \mathbb{N} \to \mathbb{N}$ given by $f(n) = n + 1$. A compact notation is for instance $f : \mathbb{N} \to \mathbb{N},\ n \mapsto n + 1$.*
2. *$g : \mathbb{R} \to \mathbb{R}$ given by $g(x) = \sin(x) + x^2 - 1$. A compact notation is for instance $g : \mathbb{R} \to \mathbb{R},\ x \mapsto \sin(x) + x^2 - 1$.*
3. *$f : \mathbb{Z} \to \mathbb{N}$ given by $f(-1) = 3$, $f(12) = 1$ and $f(x) = |x|$ if x is different from -1 and 12. A "compact" notation could be $f : \mathbb{Z} \to \mathbb{N},\ -1 \mapsto 3,\ 12 \mapsto 1,\ x \mapsto |x|$ if x is not -1 or 12, but would be rather cumbersome here.*
4. *$\varphi : \{-1, \star\} \to \mathbb{R}$ given by $\varphi(-1) = \pi$ and $\varphi(\star) = 0$.*
5. *$f : \mathbb{R} \to \mathbb{R},\ x \mapsto \sqrt{x}$ is not a map, because it is not defined for every element of $\mathbb{R}$ (only for the non-negative ones), but $f : [0, \infty) \to \mathbb{R},\ x \mapsto \sqrt{x}$ is a map.*
6. *If S is a set, the map $S \to S,\ x \mapsto x$ is called the identity map (on S) and is denoted by Id_S (or simply Id if the set S is clear from the context).*

Definition 1.9. *If $f : S \to T$ is a function, the set S is called the* **domain** *of f and the set T its* **codomain**. *The set of all the images of elements of S is called the* **image of** *f and is denoted $f(S)$:*

$$f(S) = \{f(x) \mid x \in S\}.$$

Definition 1.10. *Let f and g be two maps. We say that f and g are equal (and write $f = g$) if:*

1. *They have the same domain and codomain;*
2. *They agree on each element of the domain, i.e. if the domain is S: $f(x) = g(x)$ for every $x \in S$.*

Two properties of maps are particularly important:

Definition 1.11. *Let $f : S \to T$ be a map.*

1. *We say that f is* **injective** *if for all $x, y \in S$, if $x \neq y$ then $f(x) \neq f(y)$. It is equivalent to saying that for all $x, y \in S$, if $f(x) = f(y)$, then $x = y$.*
2. *We say that f is* **surjective** *if every element of T can be obtained as $f(x)$ for some $x \in S$. In other words, if $f(S) = T$.*
3. *A map that is both injective and surjective is called* **bijective**.

You may encounter the terminology "one to one" for injective and "onto" for surjective, but we will avoid using it.

Example 1.12.
1. *$f : \mathbb{R} \to \mathbb{R},\ x \mapsto x + 1$ is injective, since $f(x) = f(y)$ implies $x = y$. It is also surjective, since every element of $\mathbb{R}$ is in the image of f.*
2. *$f : \mathbb{R} \to [0, \infty),\ x \mapsto x^2$ is not injective (since different elements of $\mathbb{R}$ can have the same image) and is surjective. However, $g : \mathbb{R} \to \mathbb{R},\ x \mapsto x^2$ is not surjective. Observe that the maps f and g are different (they do not have the same codomain).*

Lemma 1.13. *Let $f : S \to T$ be an injective map. Then $|S| = |f(S)|$.*

Proof. We first consider the case where S is finite, so $S = \{x_1, \ldots, x_n\}$ where the x_i are all different. By definition $f(S) = \{f(x_1), \ldots, f(x_n)\}$, so $|f(S)| \leq n$. We now check that $f(x_1), \ldots, f(x_n)$ are all different: If it where not the case, we would have $f(x_i) = f(x_j)$ for some $x_i \neq x_j$, but it is impossible since f is injective. Therefore $f(S)$ has n elements.

If S is infinite: Let T be a subset of S with n elements (where $n \in \mathbb{N}$). Then by the argument above $f(T)$ has n elements. But $f(T) \subseteq f(S)$, so $f(S)$ has at least n elements, for every $n \in N$, i.e. $f(S) = \infty$. □

Corollary 1.14. *Let $f : S \to T$ be a map where S and T are finite with the same number of elements. Then f injective is equivalent to f surjective.*

Proof. If f is injective: By the previous lemma we know that $|f(S)| = |S|$. Since $|S| = |T|$ we get $|f(S)| = |T|$. So $f(S) \subseteq T$ and both have the same finite number of elements, which shows that $f(S) = T$, i.e. that f is surjective.

If f is surjective. Write $S = \{x_1, \ldots, x_n\}$, with the x_i all different. Then $f(S) = \{f(x_1), \ldots, f(x_n)\}$. Since $f(S) = T$ and $|T| = |S| = n$, we have $|f(S)| = n$, so the $f(x_i)$ must be all different, i.e. f is injective. □

Definition 1.15. *Let $f : S \to T$ and $g : T \to W$ be maps. The* **composition** *of f with g, denoted $g \circ f$ is the map*

$$g \circ f : S \to W,\ x \mapsto g(f(x)).$$

Observe that it is computed from right to left: You first apply f to x, then feed the result (i.e., $f(x)$) into g.

Definition 1.16. *Let $f : S \to T$. A map $g : T \to S$ is called the* **inverse** *of f if $f \circ g = \mathrm{Id}_T$ and $g \circ f = \mathrm{Id}_S$.*
(In other words: $f(g(x)) = x$ for every $x \in T$ and $g(f(y)) = y$ for every $y \in S$.) The map g is denoted by f^{-1}.

We say that f is **invertible** *if it has an inverse.*

Remark 1.17. *In the previous definition, we called g* the *inverse of f. We can do this because it is not very hard to check that if f is invertible, then it has only one inverse: If both g and g' are inverses of f, then we must have $g = g'$. This is left as an exercise (use that $g(f(y)) = y$ and that f is surjective).*

Proposition 1.18. 1. *Let f be a map. Then f has an inverse if and only if f is bijective.*

2. *Let $f : S \to T$ and $g : T \to U$ be bijective maps. Then $g \circ f$ is bijective and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$*

Chapter 2

Random remarks on mathematical notation, terminology, and proofs

2.1 "Or"

The meaning of the word "or" in English can be exclusive or inclusive. Consider the following statements:

- On a chessboard the rook can move horizontally or vertically.

  In this case the "or" is inclusive: both are possible.

- We will meet at 5pm or at 6pm.

  In this case the word is (usually) exclusive: it is one or the other, but not both.

In mathematics, the convention is to always use the inclusive or. In other words, if A and B are statements, then the statement

$$A \text{ or } B$$

is true in the following cases:

$$A \text{ is true, } B \text{ is true, both } A \text{ and } B \text{ are true.}$$

In particular, it is false exactly when both A and B are false.

Example 2.1.

- *"$2 > 1$ or $2 > 0$" is true (this is the one that can seem odd if you are more used to the exclusive or);*
- *"$2 > 3$ or $2 > 0$" is true;*
- *"$2 > 1$ or $2 > 5$" is true;*
- *"$2 > 4$ or $2 > 5$" is false.*

2.2 Implication

We say that a statement A implies a statement B (or that B follows from A, and I am sure that there are other ways to put it) if the following holds:

If A is true, then B is true.

The mathematical abbreviation for this is:

$$A \Rightarrow B.$$

In words: "A implies B", or "B follows from A", or "if A is true then B is true" (again, there are probably other ways to say it).

Example 2.2.

- *If x is a prime number, then 10 does not divide x. It is the implication $A \Rightarrow B$, where A is "x is a prime number" and B is "10 does not divide x".*
- *If x is a positive real number, then x is a square. It is the implication $A \Rightarrow B$, where A is "x is a positive real number" and B is "x is a square".*

Note that it is of course possible to write false implications. For instance

If x is a human, then x has blue eyes.

It is obviously false since there are humans with (for instance) brown eyes.

So: In general, how can we see if an implication $A \Rightarrow B$ is false? We must find a case where A is true, but where B is false. (Indeed, the implication must then be incorrect, because if it were correct: since A is true, we would have that B is true.)

To go back to the previous example: Take a human x with brown eyes. Then A is true (x is a human), but B is false (x does not have blue eyes). Therefore $A \Rightarrow B$ is a false statement.

Finally, how do you prove that a statement in the form $A \Rightarrow B$ is true?

You start with the data that might also be given in the question, and you assume that A is true. And with all of this you find an argument that justifies that B is true.

2.3 Equivalence

We say that two statements A and B are equivalent when

(A implies B) and (B implies A).

The mathematical abbreviation for this is

$$A \Leftrightarrow B.$$

It means that if A is true, then B is true, and if B is true then A is true. It is exactly the same as saying "$A \Rightarrow B$ and $B \Rightarrow A$".

Some common ways to express it in words in mathematics texts: "A and B are equivalent", "A is equivalent to B", "A if and only if B".

How do you prove that a statement of the form $A \Leftrightarrow B$ is true? The most basic way (which works most of the time, but may not always be the most efficient) is to prove both $A \Rightarrow B$ and $B \Rightarrow A$.

Sometimes it is possible to find an argument of the form:

$$\begin{aligned} A &\Leftrightarrow C_1 \\ &\Leftrightarrow C_2 \\ &\vdots \\ &\Leftrightarrow C_k \\ &\Leftrightarrow B \end{aligned}$$

which mean that at each step, each statement is clearly equivalent to the previous one. But it requires taking great care that the steps are small enough that each successive equivalence is clear.

Finally, how can you see if a statement of the form $A \Leftrightarrow B$ is false? It would mean that ($A \Rightarrow B$ and $B \Rightarrow A$) is false, so that at least one (but maybe both) of the statements $A \Rightarrow B$, $B \Rightarrow A$ is false (see the previous section for this).

2.4 Quantifiers

The quantifiers are the symbols $\forall$ and $\exists$. They are both abbreviations. The first one means "for all" (or "for every") and the second one means "there exists" (or "there is").

The only small difficulty is to know the range of $\forall$ or $\exists$. By this I mean: If we say (for instance) $\forall z$, does is mean "for every integer z", or for "every function z", or... In general it is clear from the context, or it is written carefully in the statement.

Example 2.3. *For instance, if we know that we are working with real numbers, then the following statement is true*

$$\forall g \quad g \geq 0 \text{ or } g < 0.$$

But this statement is not true anymore if we are working with (for instance) functions from $\mathbb{R}$ to $\mathbb{R}$. Which means that we may need to be a bit more precise if the context does not give enough information.

So here are two other ways to write the same statement, with the very same meaning, but written a bit differently to make sure that is is clear that g is a real number (there are of course other ways to write it. The only important thing is to be perfectly clear):

$$\forall g \quad g \text{ is a real number} \Rightarrow (g \geq 0 \text{ or } g < 0),$$

$$\forall g \text{ real number} \quad g \geq 0 \text{ or } g < 0.$$

Observe that changing the "name" of the quantified element does not change the statement. For instance, the statement in the previous exercise is exactly the same as this one:

$$\forall t \quad t \geq 0 \text{ or } t < 0.$$

Example 2.4. *Another example that you should know from analysis, if f is a function from $\mathbb{R}$ to $\mathbb{R}$ and $a \in \mathbb{R}$, then*

$$\forall \epsilon > 0 \ \exists \eta > 0 \quad (\textit{such that}) \ \forall x \quad |x - a| < \eta \Rightarrow |f(a) - f(x)| < \epsilon,$$

expresses that f is continuous at a. The "such that" that should be added in English after a "there exists" to get a proper sentence, is sometimes ommitted in mathematical statements.

Important: When there are several $\forall$ and $\exists$ following each other (in any order), the convention is that each $\exists$ depends on all the previous data, including all the previous $\forall$ and $\exists$. It is actually the standard convention in English. Let us look at two examples:

1. Consider the sentence:

   $$\forall \text{ foot } \ \exists \text{ a shoe that will fit the foot.}$$

   It is of course understood that the shoe depends on the foot. We make the same interpretation in maths.

2. Going back to the example with the continuity: It says that if you take any $\epsilon > 0$ you can find $\eta > 0$ such that some property holds.

   The η is allowed to depend on f, a and ϵ: if you take a different f, a or ϵ, you will in general need to come up with a different η.

2.5 "Let X be such that..."

A common way to state a question (in an exercise sheet or an exam) is to use a sentence similar to:

> Let X be such that –some property–. Prove –something about X–.

In mathematics, this formulation always means that X is completely random and that the only thing you know about it is that it has the property indicated. In particular, it is not an invitation for you to choose a specific X.

Let us illustrate this with the following exercise:

Exercise 2.5. *Let p be a integer such that p is prime and greater than 2. Show that $2p$ is not prime.*

Example of a correct answer / proof: $2p$ is not prime because it is divisible by 1, 2 and p (which are all different).

Example of an incorrect answer / proof: We take $p = 3$. Then $2p = 6$ and we know that 6 is not prime (it is divisible by more than just 1 and 6).

The problem in the second "proof" is that you took a specific value for p, while the question should have been proved for a general p.

2.6 Proof by contradiction

This is a particular way to prove a statement. Let us call this statement S. In a proof by contradiction of S, you first assume that S is false. Then, using this extra information, you try to reach a statement that is false (the contradiction). It means that S must be true.

What happened?

You want to show that the statement S is true. Since S is either true or false, you have only 2 cases:

1. S is true: Nothing to do, it is what you want.
2. S is false: Now you look closely at what this would imply. So you use this extra knowledge to try to deduce more things. If you manage to deduce something that is false, it means that this case cannot occur (and sso that S is true).

Example 2.6. *One of the standard examples of a proof by contradiction is the proof that there are infinitely many prime numbers:*

We assume that there are only finitely many prime numbers (so: we are placing ourselves in the second case, our objective is then to find something false out of it). Let $p_1, \ldots, p_n$ be the full list of all prime numbers. Then $(p_1 p_2 \ldots p_n) + 1$ is not divisible by any of $p_1, \ldots, p_n$, which is impossible because an integer should be divisible by at least one prime number. This is a contradiction, proving that there are infinitely many prime numbers. (I.e., we showed that the second case is impossible since it leads to a contradiction.)

Why is it sometimes convenient to try to use a proof by contradiction? Because we can then use one more hypothesis (namely, that S is false). It gives us a little bit more material to work with.

One final remark: The fact that this particular type of proof has a name should not suggest that there are only a few precise types of proofs that you need to stick to. There are infinitely many possible ways to write proofs. But this way is common enough that it was convenient to give it a name.