



SAPIENZA
UNIVERSITÀ DI ROMA

Mathematical Analysis 1

Volume 2 — Numbers and sets

The second volume: the sets and the numbers on which analysis is built. Sets, number sets and intervals, ordered fields, supremum and the completeness axiom, roots, powers and logarithms, summations and progressions, factorials and binomial coefficients, the Fibonacci sequence, infinite sets, complex numbers. Definitions, theorems with proofs, worked examples; the exercises of the part, with solutions, are in the volume Exercises 2.

Fabio Furini

fabiofurini.github.io

Department of Computer, Control and Management Engineering — Sapienza University of Rome

Online version and downloadable material:
fabiofurini.github.io/mathematical-analysis-1/
github.com/fabiofurini/mathematical-analysis-1

Academic year 2026–2027

Contents

1	Sets	3
1.1	Informal introduction to set theory	3
1.2	Number sets	6
1.3	Relations between sets	11
1.4	Operations on sets	16
1.5	Further topics	33
2	Number sets and intervals	37
2.1	The five main number sets	37
2.2	Intervals	40
3	Ordered fields, supremum/infimum and the completeness axiom	42
3.1	Properties R_1 and R_2	42
3.2	Commensurable magnitudes	43
3.3	Geometric representation of $\mathbb{Q}$	44
3.4	Total order relations	45
3.5	Property R_3	46
3.6	Ordered fields	47
3.7	Bounded sets, maxima and minima	48
3.8	Upper/lower bounds and suprema/infima	53
3.9	Axiomatic definition of the real numbers	55
3.10	Geometric representation of $\mathbb{R}$	57
4	Roots, powers, logarithms and modular arithmetic	59
4.1	Roots, powers, logarithms	59
4.2	Modular arithmetic	66
5	Summations, geometric and arithmetic progressions	71
5.1	Summations	71
5.2	Geometric progressions	76
5.3	Arithmetic progressions	81
6	Factorials, binomial coefficients and triangle inequality	86
6.1	Factorials	86
6.2	Binomial coefficients	87
6.3	Absolute value	91
7	The Fibonacci sequence	93
7.1	The Fibonacci sequence	93

8	Cardinality of infinite sets	98
8.1	Countable cardinality	98
8.2	Cardinality of the continuum	101
9	Complex numbers	104
9.1	Definition of $\mathbb{C}$ and field structure	104
9.2	Algebraic form of complex numbers	106
9.3	Trigonometric form of complex numbers	112
9.4	Quadratic equations	120

Sets

1.1 Informal introduction to set theory

- Set theory is based on the following three key concepts:

1. Sets

The notion of set is generally taken as **primitive** (that is, not reducible to more elementary concepts). The words *collection*, *class*, *aggregate*, *family* are used as synonyms of set.

Example 1.1: sets

For example, we have the set of points of a plane, the set of students of a university, or the set of stars of a galaxy.

2. Elements

A set is determined by its *elements*, in the sense that a set is defined when we have a **criterion** to establish whether a given element is or is not an element of this set.

There are sets with a *finite number of elements* (for example the set of students of a university) and sets with an *infinite number of elements* (for example the set of points of the plane).

Sets are usually denoted by capital letters. The elements of a set are usually denoted by lowercase letters.

3. Membership

The concept of *membership* links elements to sets. When an object is an element of a set, we say that the element *belongs* to the set. To indicate that an element x belongs to a set A we write:

$$x \in A$$

To indicate that an element x does *not* belong to a set A we write:

$$x \notin A$$

1.1.1 Informal definition of sets

1. A first way to define sets is **definition by listing** (roster notation)

A set can be defined **by listing**, that is, by listing the elements that belong to it between curly brackets. This technique assumes that the set has a finite number of elements.

Example 1.2: definition by listing

- The notation:

$$A = \{a, b, c\}$$

means that the set A has as elements the three letters a , b and c . For example, a belongs to A , that is, $a \in A$, while $d \notin A$.

- The set of vowels of the Latin alphabet can be defined by listing:

$$V = \{a, e, i, o, u\}$$

- When listing the elements of a set, each element is written **only once**: a set does not contain the same object more than once.
- The elements of a set are **not ordered**: $\{1, 2, 3\}$ and $\{3, 1, 2\}$ describe the same set, because the order in which we list the elements is irrelevant.

2. A second way to define sets is **definition by a property**:

A set can be defined **by a property** as follows:

$$A = \{ x \in U : p(x) \text{ is true} \}$$

where $p(x)$ is the property that the element x of the set U must have in order to belong to the set A . This technique can be used to define sets with a finite or even an infinite number of elements.

Example 1.3: definition by a property

Consider for example the set of letters of the Latin alphabet:

$$U = \{a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z\}$$

Using for example the property $p(x)$ defined as “ x is a vowel” we can define the following set of vowels:

$$A = \underbrace{\{x \in U : x \text{ is a vowel}\}}_{\{a, e, i, o, u\}}$$

Note that to define a set A by a property we need a set U to which all the elements of the set A we want to define belong. The set U plays the role of the **universal set**.

Definition 1.1: of universal set

A **universal set** (or simply **universe**), denoted by U , is a set fixed in advance that contains all the objects under consideration in a given context: all the sets considered in that context have elements that belong to U .

It is important that the property $p(x)$ we use makes sense for every x of the set U (universal set), and therefore is true or false (without ambiguity of meaning) for every particular $x \in U$; the set A will then consist of exactly those x belonging to U for which the property $p(x)$ is true.

Fixing a single universal set for the whole discussion is convenient, but it is not necessary. We can start from **any set already known** B , which may change from one definition to the next:

Given a set B already known, the notation

$$A = \{ x \in B : p(x) \}$$

defines the set A made of exactly those elements of B for which the property $p(x)$ is true ("is true" is usually not written). The universal set is the special case in which $B = U$ is the same for all the definitions in a given context.

- This is the form used everywhere in mathematics. The property can be written in words or with a formula, and it can consist of several conditions, separated by commas or by "and".
- When the starting set is clear from the context, it is often left implicit and we simply write $\{ x : p(x) \}$. For example, when discussing real numbers, $\{ x : x > 0 \}$ denotes the set of positive real numbers. The starting set is still fixed: it is only omitted from the notation.

Example 1.4: properties with different starting sets

- Starting from $B = \{a, b, c, d, e\}$ and from the same property " x is a vowel" we obtain a set different from the one of the previous example:

$$\{x \in B : x \text{ is a vowel}\} = \{a, e\}$$

- Starting from $C = \{1, 2, 3, 4, 5, 6\}$ we can define:

$$\{x \in C : x \text{ is even}\} = \{2, 4, 6\}$$

- The conditions can be more than one:

$$\{x \in C : x \text{ is even, } x > 2\} = \{4, 6\}$$

In all cases the starting set is chosen each time and is not the universe of the whole discussion.

3. A third way to define sets is **definition by construction**: instead of choosing the elements among those of a set that already contains them, we **build** them from the elements of a set already known.

Given a set B already known, a set can be defined **by construction** as follows:

$$A = \{ \text{expression built from } x : x \in B \}$$

that is, A is made of all the objects obtained by applying the same rule to each element x of B . There is no need to know in advance a set containing the elements of A .

Example 1.5: definition by construction

- The set of the squares of the elements of $\{1, 2, 3\}$:

$$\{ x^2 : x \in \{1, 2, 3\} \} = \{1, 4, 9\}$$

- The set of the doubles of the elements of $\{1, 2, 3, 4\}$:

$$\{ 2x : x \in \{1, 2, 3, 4\} \} = \{2, 4, 6, 8\}$$

If the rule produces the same object from different elements, the object is counted only once: for example $\{ x^2 : x \in \{-1, 0, 1\} \} = \{0, 1\}$.

One must be careful when defining sets, since contradictions may arise. There is a formal definition of the concept of set, developed to avoid contradictions, but it is beyond the scope of this course.

- In all the ways seen so far we always start from a set already known. The notation $\{ x : p(x) \}$ understood as “all the objects for which $p(x)$ is true”, with no starting set at all (not even an implicit one), does **not** in general define a set.
- For example, with the property $p(x)$ defined as “ $x \notin x$ ” we would obtain the set of all sets that do not contain themselves, which however is not a set in the formal definition of sets. Admitting this set would generate the contradiction: “the set of all sets that do not belong to themselves belongs to itself if and only if it does not belong to itself” (**Russell’s paradox**, discussed in the further topics at the end of the chapter).

1.2 Number sets

Definition 1.2: numeral system

A **positional numeral system** is a way of encoding numbers using a **sequence of digits**, where each digit contributes differently to the number depending on its position. The number of distinct digits is the **base** of the system.

The **decimal expansion** of a number is a numerical representation in base 10; it expresses a number as a sum of powers of 10, and it can be finite or infinite depending on the type of number.

- Informal definition of the main number sets:

1. We denote by $\mathbb{N}$ the set of **natural numbers**, that is, the set of numbers that can be written as **decimal expansions without decimal point and without sign**. We will use the informal notation:

$$\mathbb{N} = \{0, 1, 2, 3, 4, \dots\}$$

We adopt the convention that **zero is a natural number**, that is $0 \in \mathbb{N}$ (other textbooks exclude zero from $\mathbb{N}$: it is only a convention, but it must be stated once and for all). We denote by $\mathbb{N}_{>0}$ the set of **positive natural numbers**:

$$\mathbb{N}_{>0} = \mathbb{N} \setminus \{0\} = \{1, 2, 3, 4, \dots\}$$

Definition 1.3: of even number

A natural number n is said to be **even** if there exists a natural number m such that:

$$n = 2m$$

Definition 1.4: of odd number

A natural number n is said to be **odd** if there exists a natural number m such that:

$$n = 2m + 1$$

Example 1.6: even and odd numbers

- 0 is even, because $0 = 2 \cdot 0$, and 6 is even, because $6 = 2 \cdot 3$.
- 1 is odd, because $1 = 2 \cdot 0 + 1$, and 7 is odd, because $7 = 2 \cdot 3 + 1$.

The same two definitions are used for the integers, by requiring m to be an integer: for example $-4 = 2 \cdot (-2)$ is even and $-3 = 2 \cdot (-2) + 1$ is odd.

The **well-ordering axiom** states that every non-empty subset of $\mathbb{N}$ has a minimum.

- It is a property that is **assumed as an axiom** and that characterizes $\mathbb{N}$: it does not hold in the other number sets, since $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ are themselves non-empty subsets without a minimum.
- It is the axiom that justifies the **principle of mathematical induction**, which we will deal with later on, and it guarantees the existence of a smallest element whenever a property is satisfied by at least one natural number.

Observation 1.5

Every natural number is either even or odd, and no natural number is at the same time even and odd.

Proof. We prove the two statements separately.

Every natural number is either even or odd. Consider the set

$$S = \{ n \in \mathbb{N} : n \text{ is neither even nor odd} \}$$

and suppose by contradiction that S is not empty. By the well-ordering axiom S has a minimum, which we denote by n_0 . We have that:

- $n_0 \neq 0$, because $0 = 2 \cdot 0$ is even;
- $n_0 \neq 1$, because $1 = 2 \cdot 0 + 1$ is odd.

Hence $n_0 \geq 2$ and therefore $n_0 - 2 \in \mathbb{N}$. Since $n_0 - 2 < n_0$ and n_0 is the minimum of S , the number $n_0 - 2$ does not belong to S , that is, it is even or odd:

- if $n_0 - 2 = 2m$ with $m \in \mathbb{N}$, then $n_0 = 2m + 2 = 2(m + 1)$ and hence n_0 is even;
- if $n_0 - 2 = 2m + 1$ with $m \in \mathbb{N}$, then $n_0 = 2m + 3 = 2(m + 1) + 1$ and hence n_0 is odd.

In both cases $n_0 \notin S$, contradicting the fact that n_0 is the minimum of S . Hence S is empty.

No natural number is at the same time even and odd. If a natural number n were even and odd, there would exist $m, k \in \mathbb{N}$ with

$$n = 2m \quad \text{and} \quad n = 2k + 1 \quad \implies \quad 2(m - k) = 1$$

But $m - k$ is an integer: if $m \leq k$ the left-hand side is less than or equal to 0, while if $m \geq k + 1$ the left-hand side is greater than or equal to 2. In no case can the left-hand side be equal to 1, and therefore such an n does not exist. □

2. We denote by $\mathbb{Z}$ the set of **integers**, that is, the set of numbers that can be written as **decimal expansions without decimal point and with a sign**. We will use the informal notation:

$$\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \pm 4, \dots\}$$

3. We denote by $\mathbb{Q}$ the set of **rational numbers**, that is, the set of numbers that can be written as **finite or infinite periodic decimal expansions**. In other words, it is the set of numbers that can be written as a fraction $\frac{p}{q}$ where p is an integer and q is a natural number different from zero.

Example 1.7: rational numbers

- For example, with $p = 2$ and $q = 5$ we have the fraction $\frac{2}{5}$, whose decimal expansion is 0.4.
- For example, with $p = 4$ and $q = 10$ we have the fraction $\frac{4}{10}$, whose decimal expansion is again 0.4. Note that a rational number can be written with more than one fraction.
- For example, with $p = 13$ and $q = 30$ we have the fraction $\frac{13}{30}$, whose decimal expansion is $0.4\bar{3} = 0.4333\dots$

However, we can represent every rational number different from 0 by a single fraction $\frac{p}{q}$ by choosing $p \in \mathbb{Z}$ and $q \in \mathbb{N}$ coprime (that is, relatively prime: p and q are not both divisible by the same integer greater than 1).

Observation 1.6

$$0.\bar{9} = 1$$

Proof. There are different proofs of this remark, based on different mathematical techniques.

- (a) A simple proof follows directly from the definition of 1 divided by 3; indeed, we have:

$$\begin{aligned}\frac{1}{3} &= 0.\bar{3} \\ \frac{1}{3} \cdot 3 &= 0.\bar{3} \cdot 3 \\ 1 &= 0.\bar{9}\end{aligned}$$

- (b) Using algebraic arguments we can write:

$x = 0.999\dots$	
$10x = 9.999\dots$	multiplying by 10
$10x = 9 + 0.999\dots$	separating the integer part
	from the fractional part
$10x = 9 + x$	by definition of x
$9x = 9$	subtracting x
$x = 1$	dividing by 9

(c) A proof by contradiction is the following:

$$\begin{aligned}
 0.\bar{9} &\neq 1 \\
 0.\bar{9} \cdot 9 &\neq 1 \cdot 9 \\
 0.\bar{9} \cdot 9 + 0.\bar{9} &\neq 1 \cdot 9 + 0.\bar{9} \\
 0.\bar{9} \cdot 9 + 0.\bar{9} &\neq 9.\bar{9} \\
 0.\bar{9} \cdot (9 + 1) &\neq 9.\bar{9} \\
 0.\bar{9} \cdot (10) &\neq 9.\bar{9} \\
 9.\bar{9} &\neq 9.\bar{9} \quad \text{contradiction!}
 \end{aligned}$$

□

4. We denote by $\mathbb{R}$ the set of **real numbers**, that is, the set of numbers identified with finite or infinite decimal expansions, periodic or non-periodic.

Example 1.8: real numbers

- Consider for example the number

$$0.10110111011110\dots$$

obtained by putting after the decimal point one digit equal to 1, then 0, then two digits equal to 1, then 0, then three digits equal to 1 ... and so on. The string of digits after the decimal point is neither finite nor periodic: therefore this number is real but not rational.

- Other examples of real but not rational numbers are $\sqrt{2}$ and $\sqrt{3}$, or π and Euler's number e (Napier's constant), which have infinite non-periodic decimal expansions and are therefore real but not rational numbers.

Example 1.9: number sets defined by listing and by a property

- The set of the first five prime numbers can be defined by listing:

$$P = \{2, 3, 5, 7, 11\}$$

For example $2 \in P$, while $4 \notin P$.

- The set of **even numbers** can be defined by a property, with universal set $\mathbb{Z}$:

$$E = \{x \in \mathbb{Z} : \text{there exists } k \in \mathbb{Z} \text{ such that } x = 2k\}$$

For example $-4 \in E$, because $-4 = 2 \cdot (-2)$, while $3 \notin E$, because $3 = 2k$ would give $k = \frac{3}{2} \notin \mathbb{Z}$.

- The set of **positive real numbers** can be defined by a property, with universal set $\mathbb{R}$:

$$\mathbb{R}_{>0} = \{x \in \mathbb{R} : x > 0\}$$

It is a set with an infinite number of elements, which cannot be defined by listing.

- There are formal definitions of the sets of natural, integer and rational numbers that are beyond the scope of this course. Later we will give a formal definition of the real numbers.

1.3 Relations between sets

Definition 1.7: of equal sets

Two sets A and B are **equal** when they have the same elements. We write

$$A = B$$

and this means that every element that belongs to A also belongs to B and every element that belongs to B also belongs to A . If A and B are not equal we write $A \neq B$.

Example 1.10: order and multiplicity of the elements

- The concept of *order* among the elements is foreign to sets:

$$\{1, 2, 5\} = \{1, 5, 2\}$$

because the two sets have the same elements, the numbers 1, 2 and 5: the order in which we list the elements is irrelevant.

- The concept of *multiplicity of the elements* is also foreign to sets. For example, the set of solutions of the equation

$$x - 1 = 0$$

(whose only element is the number 1) is equal to the set of solutions of the equation

$$(x - 1)^2 = 0$$

The fact that the second equation has the solution $x = 1$ with algebraic multiplicity 2 does not change the set of its solutions: both sets are equal to $\{1\}$.

It may happen that only one of the two requirements expressed by the equality relation holds. For example, if we only know that every element of A is also an element of B , we can say that A is contained in B .

Definition 1.8: of subset

Given two sets A and B , we say that A is a **subset** of B , and we write

$$A \subseteq B,$$

if every element of A is also an element of B , that is, if for every x the implication $x \in A \implies x \in B$ holds. We read “ A is contained in B ” or “ A is a subset of B ”.

If we state that $A \subseteq B$, we do not exclude that $A = B$. If instead we want to state precisely that A is contained in B but does not coincide with B , we say that A is *strictly contained* in B .

Definition 1.9: of proper subset

Given two sets A and B , we say that A is a **proper subset** of B , or that A is **strictly contained** in B , and we write

$$A \subsetneq B,$$

if $A \subseteq B$ and $A \neq B$. In this case we speak of **strict inclusion**.

- In some texts the symbol $\subset$ denotes strict inclusion, in others it is a synonym of $\subseteq$. To avoid ambiguity we will use $\subseteq$ for inclusion and $\subsetneq$ for strict inclusion; when it appears, the symbol $\subset$ has the same meaning as $\subseteq$.

Example 1.11: inclusion and strict inclusion

- $\{1, 4\} \subseteq \{1, 3, 4\}$, because 1 and 4 belong to $\{1, 3, 4\}$. The inclusion is strict, $\{1, 4\} \subsetneq \{1, 3, 4\}$, because $3 \in \{1, 3, 4\}$ but $3 \notin \{1, 4\}$, so the two sets are not equal.
- $\{1, 3, 4\} \subseteq \{1, 3, 4\}$, but the inclusion is not strict, because the two sets are equal.
- $\mathbb{N} \subsetneq \mathbb{Z}$: every natural number is an integer, and $-1 \in \mathbb{Z}$ but $-1 \notin \mathbb{N}$.

Be careful not to confuse “belongs to” and “is contained in” (in symbols, $\in$ and $\subseteq$). In a sense, they are two ways of indicating that “something is inside something else”, but they have a fundamental logical difference:

- a set is contained in another set;
- an element belongs to a set.

Example 1.12: “belongs to” vs “is contained in”

For example:

$$3 \in \{1, 3, 4\};$$

$$\{3\} \subseteq \{1, 3, 4\};$$

$$\{1, 4\} \subseteq \{1, 3, 4\}.$$

In particular, do not confuse 3 (which is a number) with $\{3\}$, which is the set that contains the number 3 as its only element.

Sometimes we consider sets whose elements are other sets. In this case too, the symbols $\in$ and $\subseteq$ are not interchangeable, but must be used correctly.

Example 1.13: “belongs to” vs “is contained in”

For example, if we define the set

$$A = \left\{ \{1\}, \{2\}, \{1, 2\} \right\},$$

then it is correct to state that $\{2\} \in A$, because $\{2\}$ is a set that plays the role of an element of A , whereas it would be incorrect to say that $\{2\} \subseteq A$, because this would mean that $2 \in A$, while A has $\{2\}$ as an element but not 2 . Instead, the subset of A containing only the element $\{2\}$ is denoted by $\{\{2\}\}$ and we have $\{\{2\}\} \subseteq A$.

1.3.1 Properties of inclusion**Proposition 1.10: reflexive property of inclusion**

For every set A we have

$$A \subseteq A.$$

Proof. By the definition of subset we must check that every element of A is also an element of A , that is, that for every x the implication $x \in A \implies x \in A$ holds. The implication is true because the conclusion coincides with the hypothesis: if $x \in A$, then $x \in A$. Hence $A \subseteq A$. $\square$

Proposition 1.11: antisymmetric property of inclusion

Given two sets A and B , we have

$$A = B \iff A \subseteq B \text{ and } B \subseteq A.$$

Proof. ($\implies$) Suppose $A = B$. By the definition of equal sets every element of A belongs to B , hence $A \subseteq B$ by the definition of subset. In the same way every element of B belongs to A , hence $B \subseteq A$.

($\impliedby$) Suppose $A \subseteq B$ and $B \subseteq A$. From $A \subseteq B$ it follows that every element of A belongs to B , and from $B \subseteq A$ it follows that every element of B belongs to A . Hence A and B have the same elements, that is, $A = B$. $\square$

- The antisymmetric property gives the most common method to prove that two sets A and B are equal, called **double inclusion**: we take an arbitrary element of A and show that it belongs to B (that is, $A \subseteq B$), then we take an arbitrary element of B and show that it belongs to A (that is, $B \subseteq A$).

Proposition 1.12: transitive property of inclusion

Given three sets A , B and C , we have

$$A \subseteq B \text{ and } B \subseteq C \implies A \subseteq C.$$

Proof. Suppose $A \subseteq B$ and $B \subseteq C$, and let x be an arbitrary element of A . Since $A \subseteq B$, we have $x \in B$. Since $B \subseteq C$, from $x \in B$ it follows that $x \in C$. Hence every element of A belongs to C , that is, $A \subseteq C$. $\square$

1.3.2 Empty set, cardinality and power set**Definition 1.13: of empty set**

The **empty set** is the set that contains no elements. It is denoted by $\emptyset$.

Observation 1.14

Given a set A , we have:

$$\emptyset \subseteq A$$

Proof. By the definition of subset we must prove that every element that belongs to $\emptyset$ also belongs to A , that is, that for every x the implication

$$x \in \emptyset \implies x \in A$$

holds. The hypothesis $x \in \emptyset$ is false for every x , because no element belongs to $\emptyset$. An implication with a false hypothesis is true whatever the conclusion is: we say that the implication is **vacuously true**. The same can be seen by contradiction: if $\emptyset \subseteq A$ did not hold, there would exist an element of $\emptyset$ that does not belong to A ; but $\emptyset$ has no elements, so such an element does not exist. Hence the thesis holds. $\square$

Definition 1.15: of cardinality of a set

The number of elements of a set A is the **cardinality** of the set. It is denoted by $|A|$. If the cardinality of A is a natural number, the set A is called **finite**; otherwise it is called **infinite**.

- The empty set is finite and has cardinality $|\emptyset| = 0$.
- A set with only one element is called a **singleton**: for example $\{3\}$ is a singleton and $|\{3\}| = 1$.
- A subset with k elements of a set A is called a **k -subset** of A .

Example 1.14: cardinality and subsets with k elements

Given the set $A = \{1, 2, 3\}$ we have $|A| = 3$. The 2-subsets of A are $\{1, 2\}$, $\{1, 3\}$ and $\{2, 3\}$; the 1-subsets of A are the singletons $\{1\}$, $\{2\}$ and $\{3\}$; the only 0-subset of A is $\emptyset$ and the only 3-subset of A is A itself.

Definition 1.16: of power set

Given a set A , the set whose elements are all the subsets of A is called the **power set** of A and is denoted by the symbol $\mathcal{P}(A)$. The notation 2^A is also used.

- Every set A has two trivial subsets, namely A itself (by the reflexive property of inclusion) and the empty set $\emptyset$ (by the remark on the empty set); the two subsets coincide if A is empty.

Example 1.15: power set

For example, given

$$A = \{1, 2, 3\},$$

then

$$\mathcal{P}(A) = \left\{ \emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 3\}, \{1, 2, 3\} \right\}$$

Observation 1.17

Given a set A with n elements, the power set $\mathcal{P}(A)$ has 2^n elements:

$$|\mathcal{P}(A)| = 2^n = 2^{|A|}$$

Proof. If $n = 0$, the set A is empty and its only subset is $\emptyset$, hence $|\mathcal{P}(A)| = 1 = 2^0$. If $n \geq 1$, let $a_1, a_2, \dots, a_n$ be the elements of A . To build a subset S of A we must decide, for each of the elements $a_1, a_2, \dots, a_n$, whether or not S contains that element: hence we must make a choice between two options n times, and each choice is independent of the others.

- Every sequence of n choices determines a subset of A and, conversely, every subset S of A is obtained from a sequence of choices: for each i we choose to include a_i if and only if $a_i \in S$.
- Different sequences of choices give different subsets: if two sequences differ in the choice about the element a_i , then a_i belongs to one of the two subsets but not to the other.

The total number of possible subsets is therefore equal to the number of sequences of n

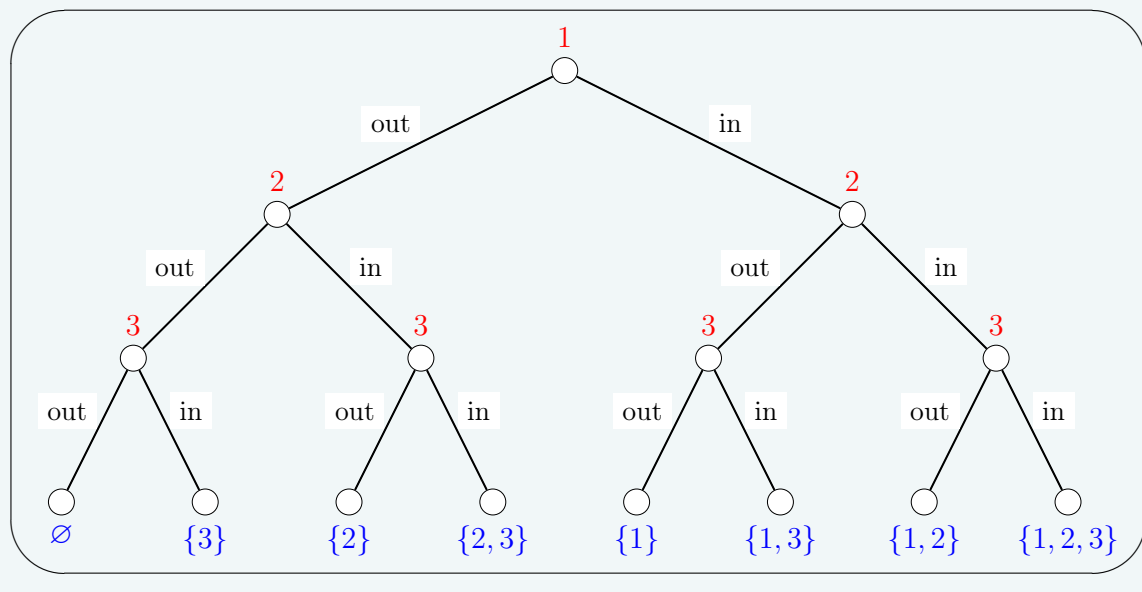
choices between two options, that is,

$$\underbrace{2 \cdot 2 \cdots 2}_{n \text{ times}} = 2^n.$$

□

Example 1.16: construction of the power set with a binary tree

Given the set $A = \{1, 2, 3\}$ with $n = 3$ elements, the cardinality of its power set is $2^3 = 8$. The construction of the power set $\mathcal{P}(A)$ can be visualized through the following **binary tree** (an undirected, connected and acyclic graph) in which at each level we decide whether or not to include the object in the subset:



1.4 Operations on sets

Definition 1.18: of intersection of sets

The *intersection* of two sets $A, B \subseteq U$ is the set defined by:

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

It is the set of elements that belong both to the first and to the second set.

Definition 1.19: of union of sets

The **union** of two sets $A, B \subseteq U$ is the set defined by:

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

It is the set of elements that belong to the first or to the second set, where “or” is meant in the non-exclusive sense (the set of elements that belong to A or to B or to both).

Definition 1.20: of difference of sets

The **difference** of two sets $A, B \subseteq U$ is the set defined by:

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

It is the set of elements that belong to the first but not to the second set. The symbol “ $\setminus$ ” can also be written “ $-$ ” by analogy with arithmetic subtraction.

Example 1.17: intersection, union and difference

Given the sets $A = \{1, 2, 3, 4\}$ and $B = \{3, 4, 5\}$, we have

$$A \cap B = \{3, 4\}, \quad A \cup B = \{1, 2, 3, 4, 5\},$$

$$A \setminus B = \{1, 2\}, \quad B \setminus A = \{5\}.$$

In particular $A \setminus B \neq B \setminus A$: in the difference the order of the two sets matters.

1.4.1 Complementary sets and disjoint sets

Example 1.18: universal sets

For example, in questions of arithmetic we could have $U = \mathbb{N}$; if we consider sets made up only of integers it is natural to choose $U = \mathbb{Z}$, while in questions of analysis we could have $U = \mathbb{R}$.

Definition 1.21: of set complementation and complementary sets

The **complementation** of a set $A \subseteq U$ is the set defined by:

$$\overline{A} = \{x \in U : x \notin A\}$$

This set is called the **complement** (complementary set) of A with respect to U .

- By the definition of difference, $U \setminus A = \{x \in U : x \notin A\}$, hence the complement of A is the difference between the universal set and A :

$$\overline{A} = U \setminus A.$$

Proposition 1.22: properties of the complement

For every set $A \subseteq U$, where U is the universal set, we have the following relations:

$$\begin{aligned}\overline{\overline{A}} &= A \\ A \cap \overline{A} &= \emptyset \\ A \cup \overline{A} &= U \\ \overline{\overline{U}} &= \emptyset \\ \overline{\emptyset} &= U\end{aligned}$$

Proof. All the sets involved are subsets of U : hence it is enough to consider the elements $x \in U$. We prove the five relations.

- Let $x \in U$. We have $x \in \overline{\overline{A}}$ if and only if $x \notin \overline{A}$, that is, if and only if it is not true that $x \notin A$, that is, if and only if $x \in A$. Hence $\overline{\overline{A}} = A$.
- If there existed $x \in A \cap \overline{A}$, we would have $x \in A$ and $x \in \overline{A}$, that is, $x \notin A$: this is impossible. Hence $A \cap \overline{A}$ has no elements, that is, $A \cap \overline{A} = \emptyset$.
- By the definition of union, $A \cup \overline{A}$ is made up of elements of U , hence $A \cup \overline{A} \subseteq U$. Conversely, let $x \in U$: if $x \in A$, then $x \in A \cup \overline{A}$; if $x \notin A$, then $x \in \overline{A}$ and hence $x \in A \cup \overline{A}$. Hence $U \subseteq A \cup \overline{A}$ and, by double inclusion, $A \cup \overline{A} = U$.
- $\overline{\overline{U}} = \{x \in U : x \notin U\}$ has no elements, because no element of U can fail to belong to U . Hence $\overline{\overline{U}} = \emptyset$.
- $\overline{\emptyset} = \{x \in U : x \notin \emptyset\}$, and every $x \in U$ satisfies $x \notin \emptyset$. Hence $\overline{\emptyset} = U$.

□

Proposition 1.23: identity laws

For every set $A \subseteq U$, where U is the universal set, we have:

$$\begin{aligned}A \cap \emptyset &= \emptyset \\ A \cap U &= A \\ A \cup \emptyset &= A \\ A \cup U &= U\end{aligned}$$

Proof. We prove the four relations.

- No x can belong both to A and to $\emptyset$, because $\emptyset$ has no elements. Hence $A \cap \emptyset = \emptyset$.
- If $x \in A \cap U$, then $x \in A$. Conversely, if $x \in A$, then $x \in U$ because $A \subseteq U$, and hence $x \in A \cap U$. By double inclusion $A \cap U = A$.
- We have $x \in A \cup \emptyset$ if and only if $x \in A$ or $x \in \emptyset$. The second possibility never occurs, hence $x \in A \cup \emptyset$ if and only if $x \in A$, that is, $A \cup \emptyset = A$.

- By the definition of union, $A \cup U$ is made up of elements of U , hence $A \cup U \subseteq U$. Conversely, if $x \in U$, then $x \in A \cup U$. By double inclusion $A \cup U = U$.

□

Definition 1.24: of disjoint sets

Two sets A and B are **disjoint** if they have no elements in common:

$$A \cap B = \emptyset$$

Example 1.19: disjoint sets

The sets $\{1, 2\}$ and $\{3, 4\}$ are disjoint, while $\{1, 2\}$ and $\{2, 3\}$ are not, because $\{1, 2\} \cap \{2, 3\} = \{2\}$. By the properties of the complement, every set $A \subseteq U$ and its complement $\bar{A}$ are disjoint.

Disjoint sets allow us to “split” a set into pieces that do not overlap. Given a family $\mathcal{F}$ of sets, the **union of the family** is the set of elements that belong to at least one of the sets of the family; it is denoted by

$$\bigcup_{B \in \mathcal{F}} B.$$

For a family made up of two sets B and C we recover the union $B \cup C$.

Definition 1.25: of partition

Given a set A , a family $\mathcal{F}$ of subsets of A is a **partition** of A if:

- every set of the family is nonempty: $B \neq \emptyset$ for every $B \in \mathcal{F}$;
- the sets of the family are pairwise disjoint: if $B, C \in \mathcal{F}$ and $B \neq C$, then $B \cap C = \emptyset$;
- their union is A :

$$A = \bigcup_{B \in \mathcal{F}} B,$$

that is, every element of A belongs to at least one set of the family.

The sets of the family $\mathcal{F}$ are called the **blocks** of the partition.

Observation 1.26

A family $\mathcal{F}$ of nonempty subsets of a set A is a partition of A if and only if every element of A belongs to **exactly one** set of the family.

Proof. ($\implies$) Let $\mathcal{F}$ be a partition of A and let $x \in A$. Since the union of the family is A , the element x belongs to at least one set of the family. If it belonged to two different sets $B, C \in \mathcal{F}$, we would have $x \in B \cap C$, against the fact that $B \cap C = \emptyset$. Hence x belongs to exactly one set of the family.

($\impliedby$) Suppose that every element of A belongs to exactly one set of the family. The sets of the family are nonempty by hypothesis. Every element of A belongs to at least one set of the family, hence $A \subseteq \bigcup_{B \in \mathcal{F}} B$; conversely, every element of the union belongs to a subset of A and hence to A : by double inclusion the union is A . Finally, if two different sets $B, C \in \mathcal{F}$ had a common element x , this element of A would belong to two sets of the family, against the hypothesis: hence $B \cap C = \emptyset$. $\square$

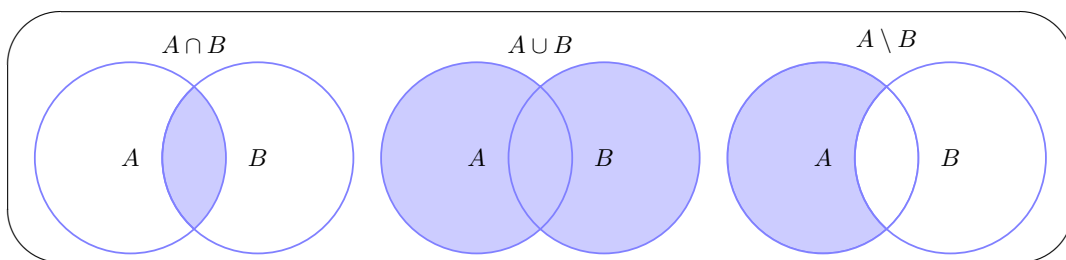
Example 1.20: partitions

- The family $\{\{1, 2\}, \{3\}, \{4, 5, 6\}\}$ is a partition of $A = \{1, 2, 3, 4, 5, 6\}$: the three sets are nonempty and every element of A belongs to exactly one of them.
- The family $\{\{1, 2\}, \{2, 3\}\}$ is not a partition of $\{1, 2, 3\}$: the element 2 belongs to two sets of the family.
- The family $\{\{1\}, \{2\}\}$ is not a partition of $\{1, 2, 3\}$: the element 3 does not belong to any set of the family.
- With universal set $\mathbb{Z}$, the set E of even numbers and its complement $\overline{E}$, that is, the set of odd numbers, form a partition of $\mathbb{Z}$: the two sets are nonempty ($0 \in E$ and $1 \in \overline{E}$) and, by the properties of the complement, $E \cap \overline{E} = \emptyset$ and $E \cup \overline{E} = \mathbb{Z}$.

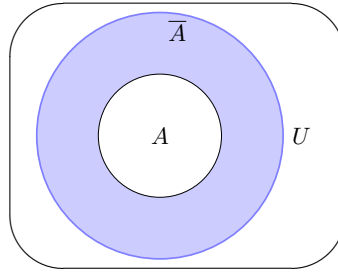
1.4.2 Venn diagrams

Venn diagrams are graphical representations in which sets are represented as regions of the plane

- Venn diagrams of intersection, union and difference:



- Venn diagram of the complement of a set:



1.4.3 Properties of operations on sets

Observation 1.27: properties of intersection

Given three sets A , B and C , intersection has the following properties:

- **Commutative:**

$$A \cap B = B \cap A$$

- **Associative:**

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

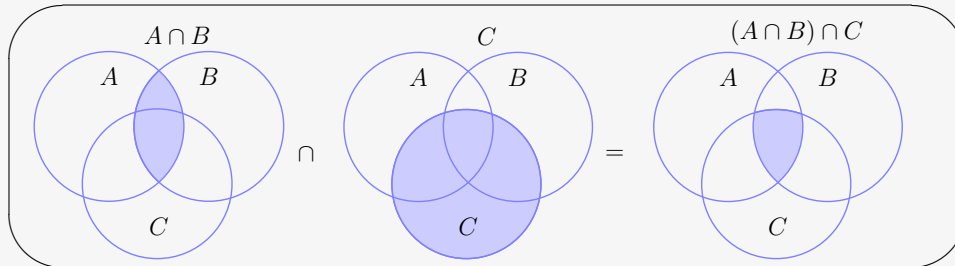
- **Idempotence:**

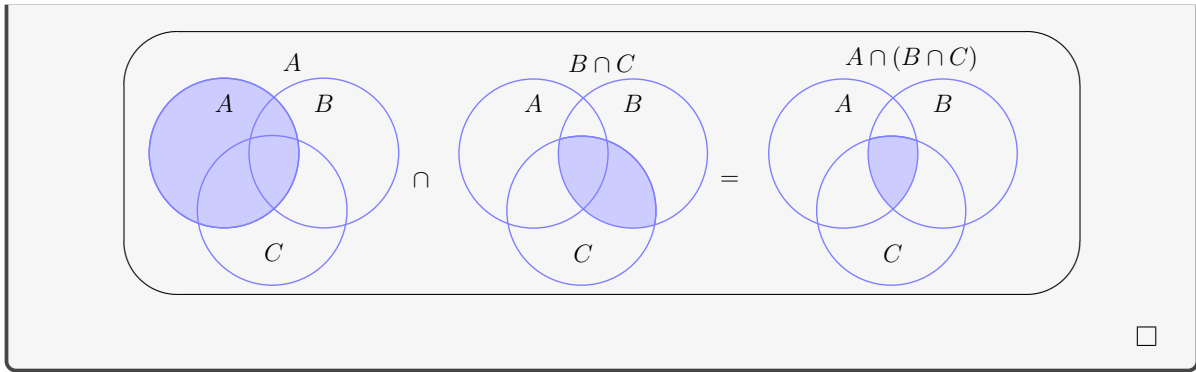
$$A \cap A = A$$

Proof. We prove the three properties by comparing the elements of the two sides.

- **Commutative:** $x \in A \cap B$ if and only if $x \in A$ and $x \in B$, that is, if and only if $x \in B$ and $x \in A$, that is, if and only if $x \in B \cap A$. Hence $A \cap B = B \cap A$.
- **Associative:** $x \in A \cap (B \cap C)$ if and only if $x \in A$ and $x \in B \cap C$, that is, if and only if x belongs to all three sets A , B and C , that is, if and only if $x \in A \cap B$ and $x \in C$, that is, if and only if $x \in (A \cap B) \cap C$.
- **Idempotence:** $x \in A \cap A$ if and only if $x \in A$ and $x \in A$, that is, if and only if $x \in A$. Hence $A \cap A = A$.

The graphical proof of the associative property of intersection is the following: in the first row we build $(A \cap B) \cap C$, in the second row $A \cap (B \cap C)$, and the two sets obtained on the right coincide.





Observation 1.28: properties of union

Given three sets A , B and C , union has the following properties:

- **Commutative:**

$$A \cup B = B \cup A$$

- **Associative:**

$$A \cup (B \cup C) = (A \cup B) \cup C$$

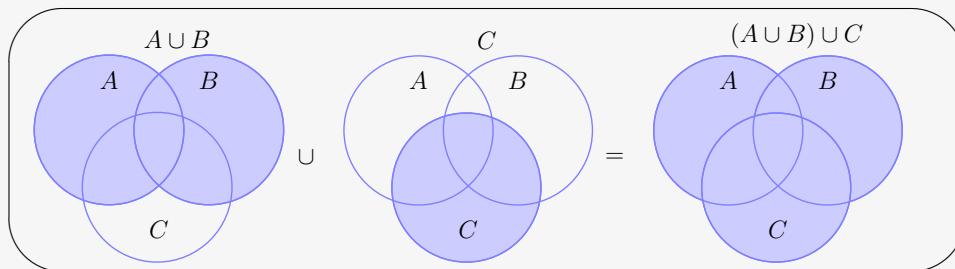
- **Idempotence:**

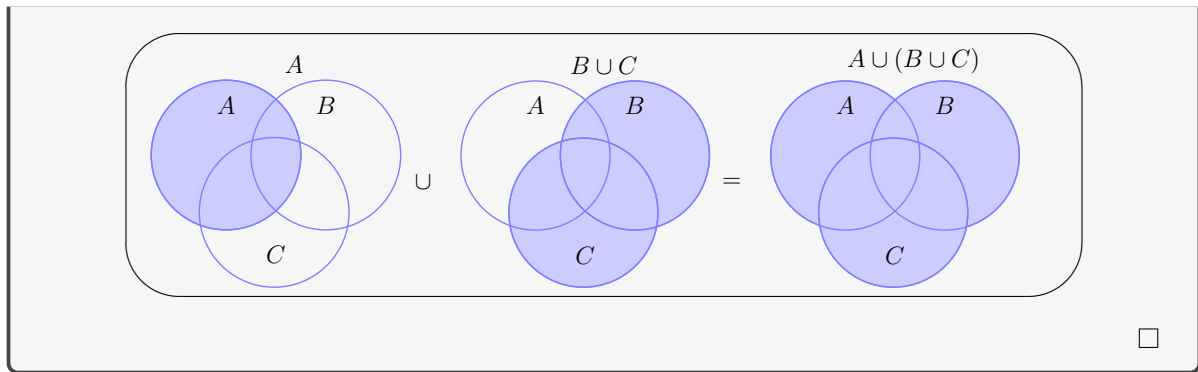
$$A \cup A = A$$

Proof. We prove the three properties by comparing the elements of the two sides.

- **Commutative:** $x \in A \cup B$ if and only if $x \in A$ or $x \in B$, that is, if and only if $x \in B$ or $x \in A$, that is, if and only if $x \in B \cup A$. Hence $A \cup B = B \cup A$.
- **Associative:** $x \in A \cup (B \cup C)$ if and only if $x \in A$ or $x \in B \cup C$, that is, if and only if x belongs to at least one of the three sets A , B and C , that is, if and only if $x \in A \cup B$ or $x \in C$, that is, if and only if $x \in (A \cup B) \cup C$.
- **Idempotence:** $x \in A \cup A$ if and only if $x \in A$ or $x \in A$, that is, if and only if $x \in A$. Hence $A \cup A = A$.

The graphical proof of the associative property of union is the following: in the first row we build $(A \cup B) \cup C$, in the second row $A \cup (B \cup C)$, and the two sets obtained on the right coincide.




Observation 1.29: distributive properties (linking union and intersection)

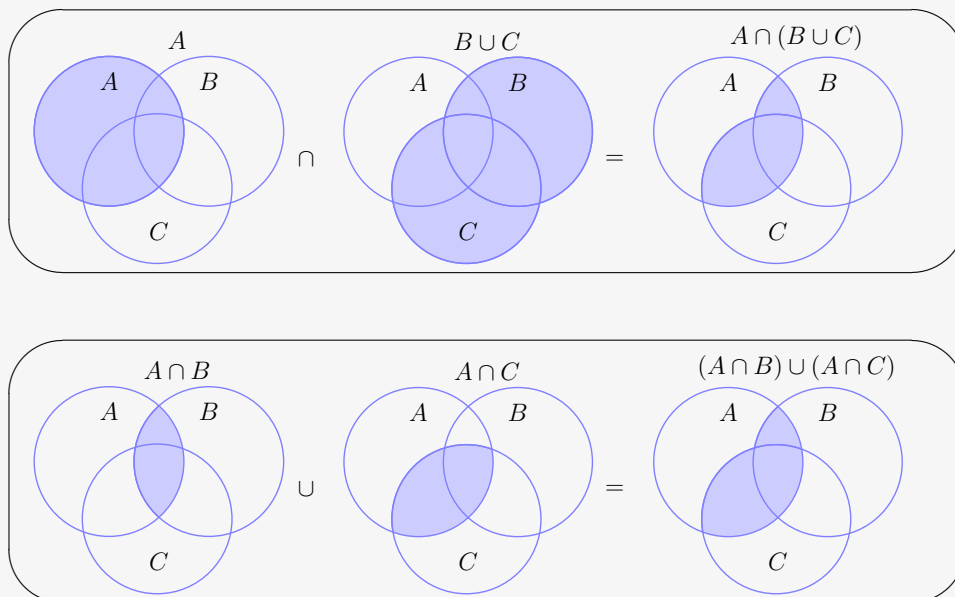
Given three sets A , B and C we have:

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

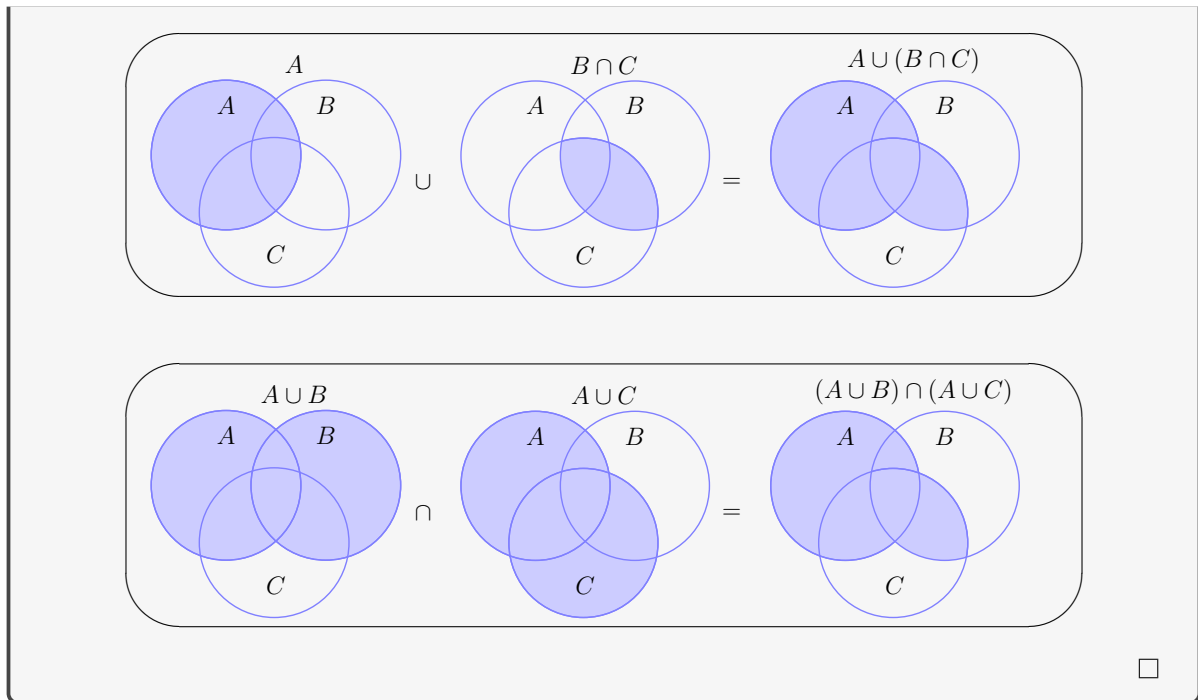
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

The distributive properties link union and intersection to each other.

Proof. The graphical proof of the first property is the following: in the first row we build the left-hand side $A \cap (B \cup C)$, in the second row the right-hand side $(A \cap B) \cup (A \cap C)$, and the two sets obtained on the right coincide.



The graphical proof of the second property is the following: in the first row we build the left-hand side $A \cup (B \cap C)$, in the second row the right-hand side $(A \cup B) \cap (A \cup C)$.

**Proposition 1.30: absorption laws**

Given two sets A and B we have:

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

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

Proof. We prove the two laws by double inclusion.

- If $x \in A \cap (A \cup B)$, then $x \in A$. Conversely, if $x \in A$, then $x \in A \cup B$, and hence $x \in A \cap (A \cup B)$. Hence $A \cap (A \cup B) = A$.
- If $x \in A$, then $x \in A \cup (A \cap B)$. Conversely, if $x \in A \cup (A \cap B)$, then $x \in A$ or $x \in A \cap B$; in the second case, again, $x \in A$. Hence $A \cup (A \cap B) = A$.

□

Proposition 1.31: inclusion, union and intersection

Given two sets A and B , the following statements are equivalent:

$$A \subseteq B, \quad A \cap B = A, \quad A \cup B = B.$$

Proof. We prove that $A \subseteq B$ is equivalent to each of the other two statements.

- $(A \subseteq B \implies A \cap B = A)$ Every element of $A \cap B$ belongs to A . Conversely, if $x \in A$, then $x \in B$ because $A \subseteq B$, and hence $x \in A \cap B$. By double inclusion

$$A \cap B = A.$$

- $(A \cap B = A \implies A \subseteq B)$ If $x \in A$, then $x \in A \cap B$, because $A \cap B = A$, and hence $x \in B$.
- $(A \subseteq B \implies A \cup B = B)$ Every element of B belongs to $A \cup B$. Conversely, if $x \in A \cup B$, then $x \in A$ or $x \in B$; in the first case $x \in B$ because $A \subseteq B$. In every case $x \in B$, and by double inclusion $A \cup B = B$.
- $(A \cup B = B \implies A \subseteq B)$ If $x \in A$, then $x \in A \cup B$, and hence $x \in B$ because $A \cup B = B$.

□

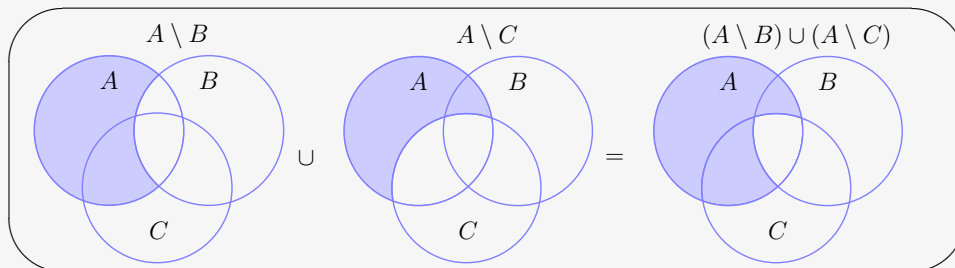
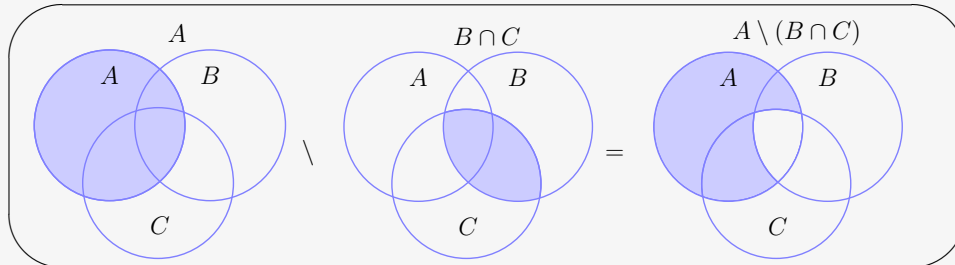
Proposition 1.32: De Morgan's laws (first version)

Given three sets A, B and C we have:

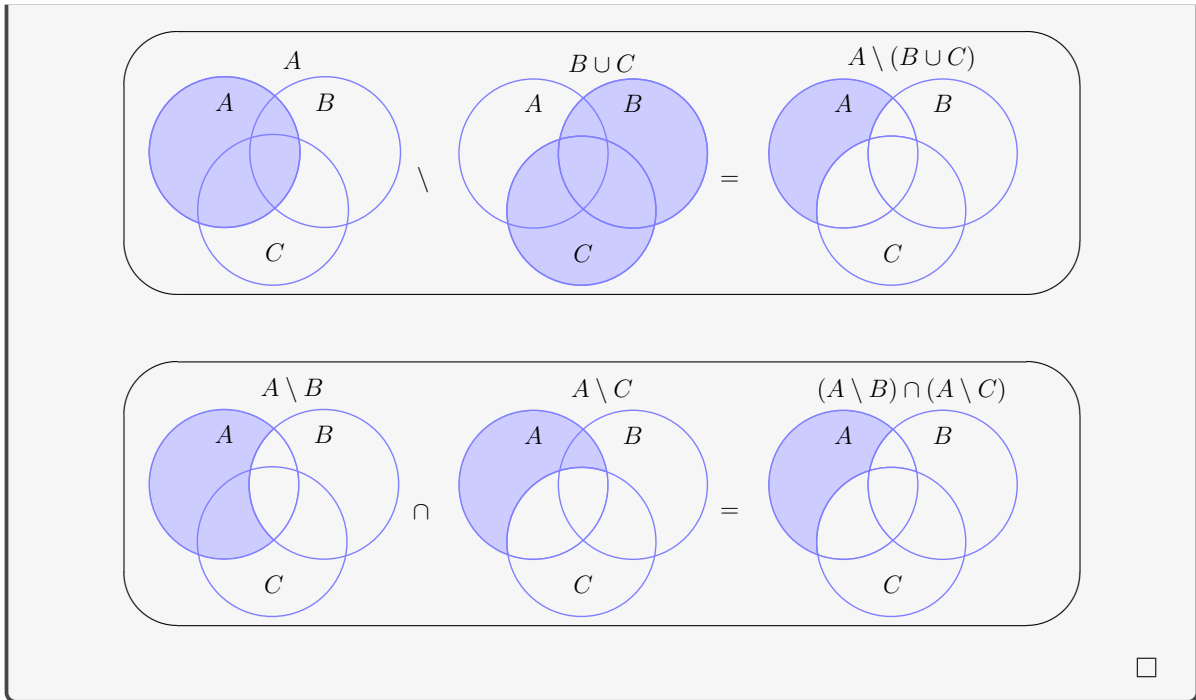
$$A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$$

$$A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$$

Proof. The graphical proof of the first property is the following: in the first row we build the left-hand side $A \setminus (B \cap C)$, in the second row the right-hand side $(A \setminus B) \cup (A \setminus C)$, and the two sets obtained on the right coincide.



The graphical proof of the second property is the following: in the first row we build the left-hand side $A \setminus (B \cup C)$, in the second row the right-hand side $(A \setminus B) \cap (A \setminus C)$.



Proposition 1.33: De Morgan's laws (second version)

Given the sets $B, C \subseteq U$, we have

$$\overline{B \cap C} = \overline{B} \cup \overline{C}$$

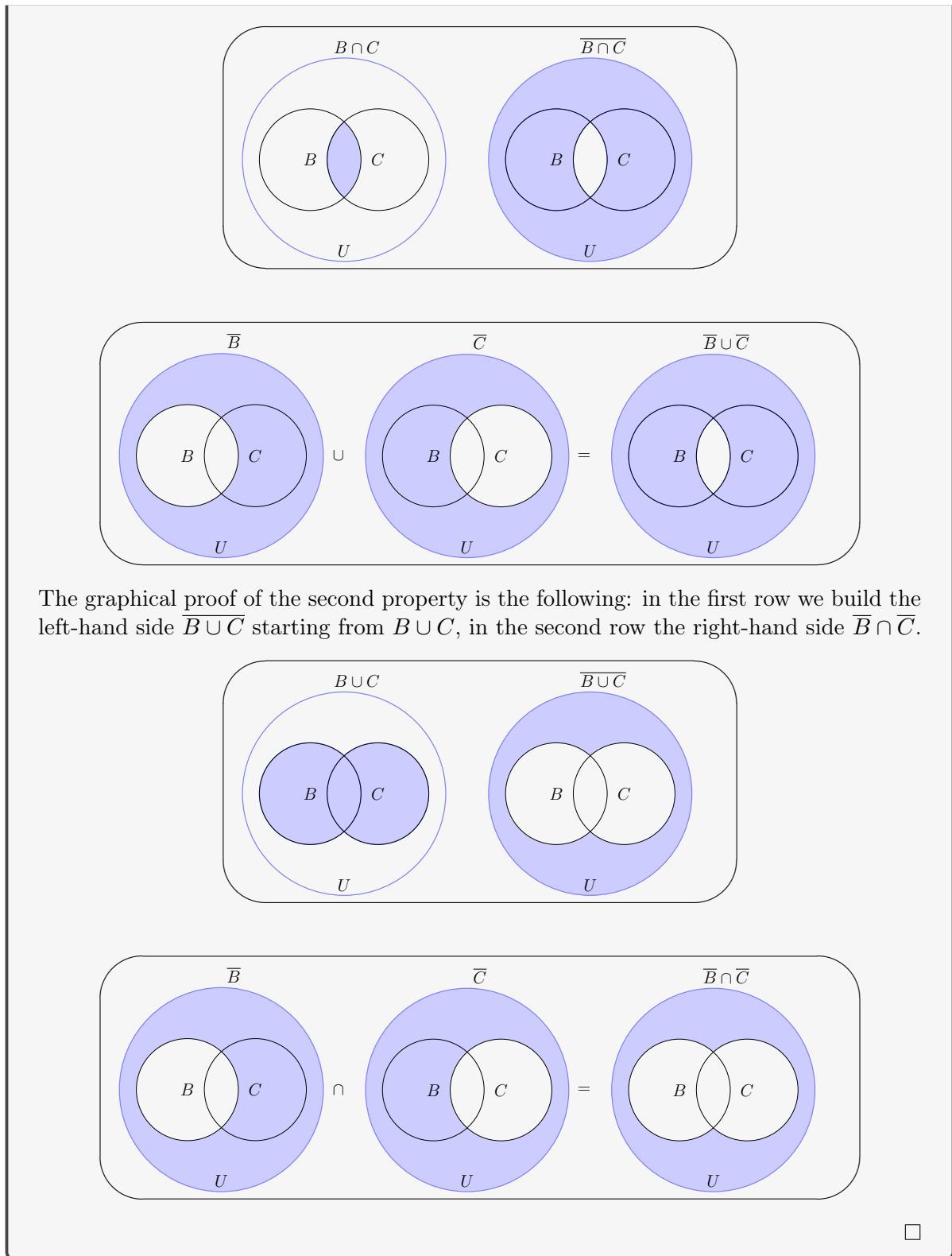
$$\overline{B \cup C} = \overline{B} \cap \overline{C}$$

Proof. These laws can be derived by setting A equal to U in the previous De Morgan's laws and recalling that $\overline{B} = U \setminus B$ and $\overline{C} = U \setminus C$, as follows:

$$\underbrace{U \setminus (B \cap C)}_{=\overline{B \cap C}} = \underbrace{(U \setminus B)}_{=\overline{B}} \cup \underbrace{(U \setminus C)}_{=\overline{C}}$$

$$\underbrace{U \setminus (B \cup C)}_{=\overline{B \cup C}} = \underbrace{(U \setminus B)}_{=\overline{B}} \cap \underbrace{(U \setminus C)}_{=\overline{C}}$$

Proof. The graphical proof of the first property is the following: in the first row we build the left-hand side $\overline{B \cap C}$ starting from $B \cap C$, in the second row the right-hand side $\overline{B} \cup \overline{C}$, and the two sets obtained on the right coincide.



- The set of elements that belong to a set B or to a set C but not to both (meaning “or” in the exclusive sense) is called the **symmetric difference** of B and C and is denoted by $B \triangle C$:

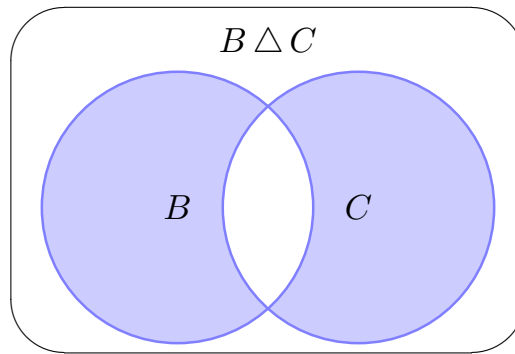
$$B \triangle C = (B \setminus C) \cup (C \setminus B).$$

It is also obtained by setting $U = B \cup C$ and taking the complement of the intersection

with respect to this universal set:

$$B \Delta C = (B \cup C) \setminus (B \cap C).$$

Proof. We prove that the two expressions of the symmetric difference coincide. We have $x \in (B \setminus C) \cup (C \setminus B)$ if and only if $(x \in B \text{ and } x \notin C) \text{ or } (x \in C \text{ and } x \notin B)$, that is, if and only if x belongs to only one of the two sets B and C . On the other hand, $x \in (B \cup C) \setminus (B \cap C)$ if and only if x belongs to at least one of the two sets but not to both, that is, again, if and only if x belongs to only one of the two sets. Hence the two sets have the same elements. $\square$



1.4.4 Properties of the cardinality of sets

Proposition 1.34: cardinality of the union of disjoint sets

If A and B are two disjoint finite sets, that is, $|A \cap B| = 0$, then

$$|A \cup B| = |A| + |B|.$$

More generally, if $A_1, A_2, \dots, A_m$ are pairwise disjoint finite sets, then

$$|A_1 \cup A_2 \cup \dots \cup A_m| = |A_1| + |A_2| + \dots + |A_m|.$$

Proof. Every element of $A \cup B$ belongs to A or to B , and no element belongs to both, because $A \cap B = \emptyset$. Counting first the $|A|$ elements of A and then the $|B|$ elements of B we therefore count every element of $A \cup B$ exactly once: $|A \cup B| = |A| + |B|$.

In the same way, every element of $A_1 \cup A_2 \cup \dots \cup A_m$ belongs to at least one of the sets $A_1, A_2, \dots, A_m$ and, since the sets are pairwise disjoint, to only one of them. Counting one after the other the elements of $A_1, A_2, \dots, A_m$ we therefore count every element of the union exactly once. $\square$

Proposition 1.35: cardinality of the union

For any two finite sets A and B , we have

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Proof. Every element of $A \cup B$ belongs to exactly one of the three sets

$$A \setminus B, \quad A \cap B, \quad B \setminus A :$$

if it belongs both to A and to B it lies in $A \cap B$, if it belongs only to A it lies in $A \setminus B$, if it belongs only to B it lies in $B \setminus A$. The three sets are therefore pairwise disjoint and their union is $A \cup B$, and by the cardinality of the union of disjoint sets

$$|A \cup B| = |A \setminus B| + |A \cap B| + |B \setminus A|.$$

In the same way, every element of A belongs to exactly one of the two disjoint sets $A \setminus B$ and $A \cap B$, and every element of B belongs to exactly one of the two disjoint sets $B \setminus A$ and $A \cap B$, hence

$$|A| = |A \setminus B| + |A \cap B|, \quad |B| = |B \setminus A| + |A \cap B|.$$

Adding the last two equalities we obtain

$$|A| + |B| = |A \setminus B| + |A \cap B| + |B \setminus A| + |A \cap B| = |A \cup B| + |A \cap B|,$$

that is, $|A \cup B| = |A| + |B| - |A \cap B|$. □

Proposition 1.36: inequality for the cardinality of the union

For any two finite sets A and B , we have

$$|A \cup B| \leq |A| + |B|.$$

Proof. By the cardinality of the union we have $|A \cup B| = |A| + |B| - |A \cap B|$. Since $|A \cap B| \geq 0$, we obtain

$$|A \cup B| = |A| + |B| - |A \cap B| \leq |A| + |B|.$$

□

Proposition 1.37: cardinality of a subset

Given two finite sets A and B with $A \subseteq B$, we have

$$|A| \leq |B|.$$

If moreover $A \subsetneq B$, then $|A| < |B|$.

Proof. Since $A \subseteq B$, every element of B belongs to exactly one of the two disjoint sets A and $B \setminus A$, and every element of these two sets belongs to B : hence $B = A \cup (B \setminus A)$ and, by the cardinality of the union of disjoint sets,

$$|B| = |A| + |B \setminus A| \geq |A|.$$

If moreover $A \subsetneq B$, then $A \neq B$ and hence, by the antisymmetric property of inclusion, $B \subseteq A$ cannot hold: there exists an element $b \in B$ with $b \notin A$, that is, $b \in B \setminus A$. Then $|B \setminus A| \geq 1$ and

$$|B| = |A| + |B \setminus A| \geq |A| + 1 > |A|.$$

□

1.4.5 Ordered pairs and Cartesian product

There is another operation on sets, which can be performed on any two sets (i.e., two sets not necessarily contained in the same universal set). To introduce it we need the concept of *ordered pair*: unlike the set $\{a, b\}$, in which the order of the elements is irrelevant, in an ordered pair it matters which element comes first.

Definition 1.38: of ordered pair

Given two elements a and b , the **ordered pair** (a, b) is the set

$$(a, b) = \{\{a\}, \{a, b\}\}.$$

The element a is called the **first component** and the element b the **second component** of the pair.

This definition, due to Kuratowski, expresses the ordered pair using only sets. What matters about the ordered pair is the following property.

Proposition 1.39: characteristic property of ordered pairs

Given the elements a, b, c and d , we have

$$(a, b) = (c, d) \iff a = c \text{ and } b = d.$$

Proof. ($\Leftarrow$) If $a = c$ and $b = d$, the sets $\{\{a\}, \{a, b\}\}$ and $\{\{c\}, \{c, d\}\}$ are written with the same elements, hence they are equal.

($\Rightarrow$) Suppose $\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$ and distinguish two cases.

- If $a = b$, then $\{a, b\} = \{a\}$ and the pair $(a, b) = \{\{a\}\}$ has only one element. Hence $\{\{c\}, \{c, d\}\}$ also has $\{a\}$ as its only element: from $\{c\} = \{a\}$ it follows that $c = a$, and from $\{c, d\} = \{a\}$ it follows that $d = a$. Hence $a = c$ and $b = a = d$.
- If $a \neq b$, the pair (a, b) has two distinct elements: $\{a\}$, which has one element, and $\{a, b\}$, which has two. Then also $c \neq d$, otherwise (c, d) would have only one

element. Since $\{a\} \in \{\{c\}, \{c, d\}\}$ and $\{a\}$ has only one element, while $\{c, d\}$ has two, we must have $\{a\} = \{c\}$, that is, $a = c$. Similarly $\{a, b\}$, which has two elements, cannot be equal to $\{c\}$, hence $\{a, b\} = \{c, d\} = \{a, d\}$. Since $b \in \{a, d\}$ and $b \neq a$, we obtain $b = d$.

□

- In particular, if $a \neq b$ then $(a, b) \neq (b, a)$: if $(a, b) = (b, a)$ held, the characteristic property would give $a = b$. Instead the sets $\{a, b\}$ and $\{b, a\}$ are always equal. If $a = b$, the two pairs (a, b) and (b, a) are the same pair.

Definition 1.40: of Cartesian product

Given two (not necessarily distinct) sets A and B , the set consisting of all ordered pairs (a, b) , with $a \in A$ and $b \in B$, is called the **Cartesian product** of A and B and is denoted by the symbol $A \times B$:

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

Proposition 1.41: cardinality of the Cartesian product

When A and B are two finite sets, the cardinality of their Cartesian product is:

$$|A \times B| = |A| \cdot |B|.$$

Proof. If $A = \emptyset$ there is no pair with first component in A , hence $A \times B = \emptyset$ and $|A \times B| = 0 = 0 \cdot |B|$. Otherwise let $a_1, a_2, \dots, a_m$ be the elements of A , with $m = |A|$, and for every $i \in \{1, 2, \dots, m\}$ consider the set of pairs with first component a_i :

$$R_i = \{(a_i, b) : b \in B\}.$$

- Every pair of $A \times B$ belongs to exactly one of the sets $R_1, R_2, \dots, R_m$, the one corresponding to its first component: by the characteristic property, pairs with different first components are different. Hence the sets $R_1, R_2, \dots, R_m$ are pairwise disjoint and their union is $A \times B$.
- Every set R_i has $|B|$ elements: by the characteristic property, $(a_i, b) = (a_i, b')$ if and only if $b = b'$, hence different elements of B give different pairs.

By the cardinality of the union of disjoint sets

$$|A \times B| = |R_1| + |R_2| + \dots + |R_m| = \underbrace{|B| + |B| + \dots + |B|}_{m \text{ times}} = |A| \cdot |B|.$$

□

Example 1.21: Cartesian product

$$\begin{aligned}\{a, b\} \times \{a, b, c\} &= \{(a, a), (a, b), (a, c), (b, a), (b, b), (b, c)\} \\ |\{a, b\}| &= 2, \quad |\{a, b, c\}| = 3, \quad |\{a, b\} \times \{a, b, c\}| = 2 \cdot 3 = 6\end{aligned}$$

The Cartesian product is not commutative: the pair (c, a) belongs to $\{a, b, c\} \times \{a, b\}$, but it does not belong to $\{a, b\} \times \{a, b, c\}$, because $c \notin \{a, b\}$. Hence $\{a, b, c\} \times \{a, b\} \neq \{a, b\} \times \{a, b, c\}$.

The Cartesian product extends to more than two sets.

Definition 1.42: of ordered n -tuple and of Cartesian product of n sets

Given $n \geq 3$ elements $a_1, a_2, \dots, a_n$, the **ordered n -tuple** $(a_1, a_2, \dots, a_n)$ is the ordered pair

$$(a_1, a_2, \dots, a_n) = ((a_1, a_2, \dots, a_{n-1}), a_n).$$

Given $n \geq 2$ sets $A_1, A_2, \dots, A_n$, their **Cartesian product** is the set of n -tuples

$$A_1 \times A_2 \times \dots \times A_n = \{(a_1, a_2, \dots, a_n) : a_i \in A_i \text{ for every } i \in \{1, 2, \dots, n\}\}.$$

- Applying repeatedly the characteristic property of ordered pairs we obtain

$$(a_1, a_2, \dots, a_n) = (b_1, b_2, \dots, b_n) \iff a_i = b_i \text{ for every } i \in \{1, 2, \dots, n\}.$$

Proposition 1.43: cardinality of the Cartesian product of n sets

If $A_1, A_2, \dots, A_n$ are finite sets, then

$$|A_1 \times A_2 \times \dots \times A_n| = |A_1| \cdot |A_2| \cdots |A_n|.$$

Proof. For $n = 2$ it is the cardinality of the Cartesian product of two sets. For $n \geq 3$, by the definition of n -tuple the n -tuples $(a_1, a_2, \dots, a_n)$ are exactly the pairs $((a_1, \dots, a_{n-1}), a_n)$ with $(a_1, \dots, a_{n-1}) \in A_1 \times \dots \times A_{n-1}$ and $a_n \in A_n$, that is,

$$A_1 \times A_2 \times \dots \times A_n = (A_1 \times A_2 \times \dots \times A_{n-1}) \times A_n.$$

By the cardinality of the Cartesian product of two sets

$$|A_1 \times A_2 \times \dots \times A_n| = |A_1 \times A_2 \times \dots \times A_{n-1}| \cdot |A_n|.$$

Repeating the same argument on $A_1 \times A_2 \times \dots \times A_{n-1}$, then on $A_1 \times A_2 \times \dots \times A_{n-2}$, and so on down to $A_1 \times A_2$, we obtain $|A_1| \cdot |A_2| \cdots |A_n|$. $\square$

- The Cartesian product of n sets equal to A is denoted by

$$A^n = \underbrace{A \times A \times \dots \times A}_{n \text{ times}}$$

and, if A is finite, by the previous proposition it has cardinality $|A^n| = |A|^n$. For example $\{0, 1\}^3$ has $2^3 = 8$ elements: the triples of 0 and 1 correspond to the “in” or “out” choices of the binary tree with which we built the power set of a set with 3 elements.

- A typical use of the Cartesian product is $\mathbb{R} \times \mathbb{R}$, which is abbreviated by the symbol $\mathbb{R}^2$ and denotes the set of ordered pairs of real numbers.
- Similarly, $\mathbb{R}^n$ (abbreviation of the Cartesian product of n sets equal to $\mathbb{R}$) is the set of *ordered n -tuples of real numbers*

$$\mathbb{R}^n = \{ (x_1, x_2, \dots, x_n) : x_i \in \mathbb{R}, i \in \{1, 2, \dots, n\} \}$$

Example 1.22: subsets of the plane

When we study subsets of the plane we choose $U = \mathbb{R}^2$ as universal set. For example, the set of points of the plane with both coordinates positive (the *first quadrant*) is

$$Q_1 = \{ (x_1, x_2) \in \mathbb{R}^2 : x_1 > 0 \text{ and } x_2 > 0 \}.$$

Since $(x_1, x_2) \in Q_1$ if and only if $x_1 \in \mathbb{R}_{>0}$ and $x_2 \in \mathbb{R}_{>0}$, we have $Q_1 = \mathbb{R}_{>0} \times \mathbb{R}_{>0}$.

1.5 Further topics

1.5.1 Russell’s paradox

“can a set be an element of itself or not?”

- For example, the set of all books in a library is not an element of itself (a set of books is not a book). On the other hand, reasoning naively, the collection of all sets with more than 20 elements would seem to be an element of itself, since it certainly has more than 20 elements. We will see at the end of this section that in the formal definition of sets this collection is not a set.
- Following this reasoning, two categories of sets can be defined:
 1. sets that are not elements of themselves
 2. sets that are elements of themselves

If we consider the set of all sets that are not elements of themselves, is it an element of itself or not?

Let us call this set S ; two hypotheses can be made:

- 1 If we suppose $S \in S$, then S contains itself as an element and therefore does not belong to S (since by definition a set belongs to S only if it does not contain itself as an element). Hence $S \notin S$, and we have a contradiction. We conclude that the hypothesis must be wrong.
- 2 If we suppose $S \notin S$, then S does not contain itself as an element and therefore belongs to S (since by definition a set belongs to S if it does not contain itself as an element). Hence $S \in S$ and we have another contradiction. We conclude that this hypothesis must be wrong too!

Russell's paradox: The set of all sets that do not belong to themselves belongs to itself if and only if it does not belong to itself.

- The formal definition of the concept of set is based on the **Zermelo-Fraenkel system of axioms**, abbreviated as **ZF**. This system of axioms includes the standard axioms of axiomatic set theory on which, together with the *axiom of choice*, all of ordinary mathematics is based.
- The **axiom of pairing** states that “Given two objects, there exists a set whose elements are the two objects”.
- The **axiom of regularity** states that “Every non-empty set A contains an element disjoint from A ”.

Observation 1.44

No set is an element of itself.

Proof. Let A be a set. By the axiom of pairing, applied to the two objects A and A , there exists the set $\{A, A\}$, which coincides with $\{A\}$ since sets do not contain repeated objects: hence $\{A\}$ is a set (it is a special case of a pair), and it is non-empty.

We now apply the axiom of regularity to the set $\{A\}$: there must exist an element of $\{A\}$ disjoint from $\{A\}$. Since the only element of $\{A\}$ is A , A is disjoint from $\{A\}$, that is, $A \cap \{A\} = \emptyset$.

If $A \in A$ held, since also $A \in \{A\}$, we would have $A \in A \cap \{A\}$, against the fact that $A \cap \{A\} = \emptyset$ (definition of disjoint sets). Hence $A \notin A$. □

- In particular, the collection of all sets with more than 20 elements is not a set in ZF: if it were, it would have more than 20 elements and would therefore be an element of itself, against the remark just proved.
- The other ZF axioms are not part of the basic course in Mathematical Analysis.

1.5.2 A further proof via the Archimedean property

- A further proof of

$$0.\overline{9} = 1$$

starts from the assumption that two numbers are equal if and only if their difference is equal to zero, and it is based on computing the value of $1 - 0.\overline{9}$.

- This proof is based on the fact that 0 is the only non-negative number less than all the reciprocals of the positive integers, or equivalently that there is no number greater than every natural number. This is the **Archimedean property**, which holds for the rational and the real numbers.

The **Archimedean property** states that for every real number x there exists a natural number n such that:

$$n > x$$

- Equivalently: for every real number $\varepsilon > 0$ there exists a natural number $n > 0$ such that $\frac{1}{n} < \varepsilon$ (it is enough to take n greater than $\frac{1}{\varepsilon}$). As a consequence, the only non-negative number that is less than or equal to $\frac{1}{n}$ for every n is zero: this is the form in which the property is used in the proof below.
- It is not an additional axiom: in the chapter *Ordered fields, supremum/infimum and the completeness axiom* we will prove it starting from the supremum property of the real numbers.

Proof. We write the number $0.999\dots$ with n digits after the decimal point as $0.(9)_n$, hence $0.(9)_1 = 0.9$, $0.(9)_2 = 0.99$, $0.(9)_3 = 0.999$, and so on.

Given $\frac{1}{10^n} = 0.0\dots 01$, with n digits after the decimal point, the addition rules for decimal numbers imply

$$0.(9)_n + \frac{1}{10^n} = 1 \quad \text{moreover} \quad 0.(9)_n < 1, \forall n \in \mathbb{N}.$$

We must prove that 1 is the smallest number that is not less than all the $0.(9)_n$. For this it is enough to prove that, if a number x is not greater than 1 and not less than all the $0.(9)_n$, then $x = 1$.

So let x be such that

$$0.(9)_n \leq x \leq 1$$

for every positive integer n . Hence

$$1 - 1 \leq 1 - x \leq 1 - 0.(9)_n$$

which, using basic arithmetic and the first equality established above, simplifies to

$$0 \leq 1 - x \leq \frac{1}{10^n}$$

This implies that the difference between 1 and x is less than the reciprocal of any positive integer. Hence this difference must be zero, and therefore $x = 1$; which in turn implies

$$0.999 \dots = 1$$

□

Number sets and intervals

2.1 The five main number sets

2.1.1 The natural numbers

$\mathbb{N}$ is the set of **natural numbers**, that is, the numbers used for counting:

$$\mathbb{N} = \{0, 1, 2, 3, 4, \dots\}$$

- We recall the convention adopted in these notes: **zero is a natural number**, that is $0 \in \mathbb{N}$. Other textbooks exclude zero from $\mathbb{N}$: it is only a convention, but it must be stated once and for all.
- We denote by $\mathbb{N}_{>0}$ the set of **positive natural numbers**:

$$\mathbb{N}_{>0} = \mathbb{N} \setminus \{0\} = \{1, 2, 3, 4, \dots\}$$

In the chapter [Sets](#) we have already introduced, for the natural numbers:

- the definitions of **even** number ($n = 2m$) and **odd** number ($n = 2m + 1$), with $m \in \mathbb{N}$, and the fact that every natural number is either even or odd, but never both;
- the **well-ordering axiom**, stating that every non-empty subset of $\mathbb{N}$ has a minimum.

2.1.2 The integers

$\mathbb{Z}$ is the set of **integers** (also called signed numbers), that is:

$$0, 1, -1, 2, -2, \dots$$

2.1.3 The rational numbers

$\mathbb{Q}$ is the set of **rational numbers**, that is, of the fractions:

$$\frac{n}{m}, \quad \text{where } n, m \in \mathbb{Z}, \text{ and } m \neq 0.$$

Rational numbers can also be written in decimal form. For example:

$$\frac{3}{4} = 0.75, \quad \text{while} \quad \frac{1}{3} = 0.333\dots = 0.\bar{3}.$$

A rational number, written in decimal form, has a **finite or infinite periodic decimal expansion**: after the decimal point there is a finite number of digits, or an infinite number of digits which, from a certain point on, repeat *periodically*.

The same rational number can have two different decimal expansions. The classical example is the equality

$$0.\bar{9} = 1$$

for which we gave four different proofs in the chapter [Sets](#), and which we therefore do not repeat here.

2.1.4 The real numbers

$\mathbb{R}$ is the set of **real numbers**, that is, those which are identified with **finite or infinite decimal expansions, periodic or non-periodic**: after the decimal point there may be any string of digits, possibly *infinite* and *non-periodic*.

That numbers of this last kind exist (that is, real but not rational), can be understood by reflecting on examples such as:

$$0.10110111011110\dots$$

After the decimal point, the previous number has: one digit equal to 1, then 0, then two digits equal to 1, then 0, then three digits equal to 1 ... and so on. It is clear that this rule defines a decimal number precisely. On the other hand, the string of digits after the decimal point is neither finite nor periodic: therefore this number is irrational.

2.1.5 The complex numbers

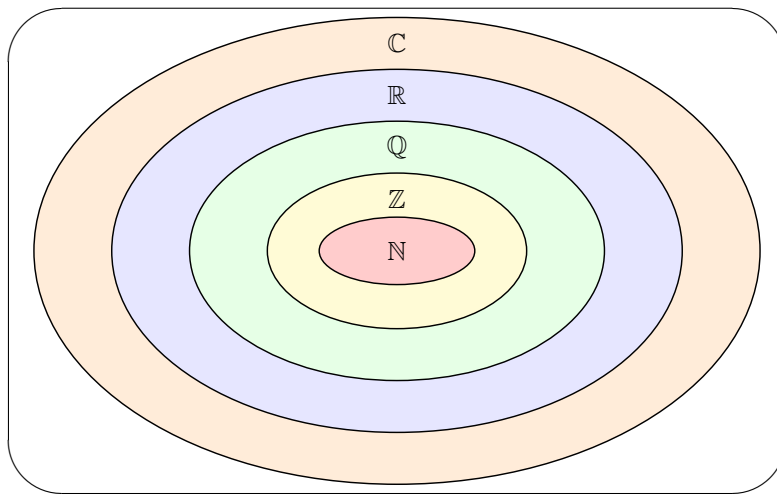
$\mathbb{C}$ is the set of **complex numbers**, that is, numbers of the form $a + ib$, where a, b are real numbers, and i is the *imaginary unit*, that is, a number whose square is -1 .

Among the number sets introduced, the following inclusions hold:

$$\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R} \subsetneq \mathbb{C}.$$

As indicated by the symbols, all the inclusions are strict:

- there exist integers that are not natural numbers (the negative numbers),
- there exist rational numbers that are not integers (the proper fractions),
- there exist real numbers that are not rational (the irrational numbers),
- there exist complex numbers that are not real (the imaginary numbers).



As for $\mathbb{N}_{>0}$, we denote by

$$\mathbb{Z}_{>0}, \quad \mathbb{Q}_{>0}, \quad \mathbb{R}_{>0}$$

the sets of **positive** integer, rational and real numbers, and by

$$\mathbb{Z}_{\geq 0}, \quad \mathbb{Q}_{\geq 0}, \quad \mathbb{R}_{\geq 0}$$

the sets of **non-negative** integer, rational and real numbers (that is, positive or zero). For example:

$$\mathbb{R}_{>0} = \{x \in \mathbb{R} : x > 0\} \quad \text{and} \quad \mathbb{Z}_{\geq 0} = \{x \in \mathbb{Z} : x \geq 0\} = \mathbb{N}$$

2.2 Intervals

Definition 2.1: of bounded interval

Given two real numbers a, b , a **bounded interval** with endpoints a and b is one of the following sets:

$$\begin{aligned} [a, b] &= \{x \in \mathbb{R} : a \leq x \leq b\}, & [a, b) &= \{x \in \mathbb{R} : a \leq x < b\} \\ (a, b] &= \{x \in \mathbb{R} : a < x \leq b\}, & (a, b) &= \{x \in \mathbb{R} : a < x < b\} \end{aligned}$$

- As can be seen, the square (round) bracket at one of the two endpoints indicates that this endpoint is included in (excluded from) the interval.
- The intervals $[a, b]$ are called **closed**; the intervals (a, b) are called **open**.

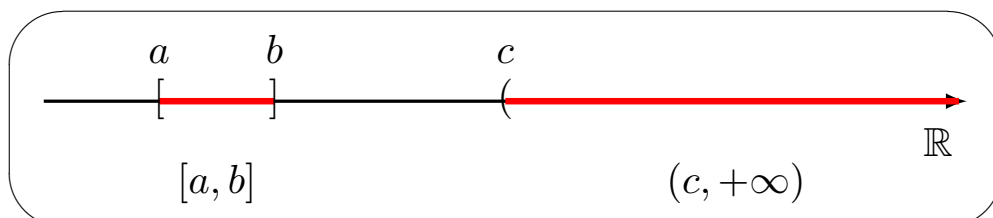
Definition 2.2: of unbounded interval

Given a real number a , an **unbounded interval** (or **half-line**) with endpoint a is one of the following sets:

$$\begin{aligned} [a, +\infty) &= \{x \in \mathbb{R} : x \geq a\}, & (a, +\infty) &= \{x \in \mathbb{R} : x > a\} \\ (-\infty, a] &= \{x \in \mathbb{R} : x \leq a\}, & (-\infty, a) &= \{x \in \mathbb{R} : x < a\} \end{aligned}$$

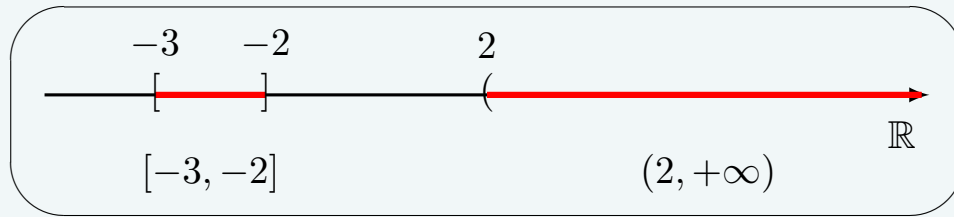
- The symbols $-\infty$ and $+\infty$ **are not real numbers**: they are not endpoints that can belong to the set, and therefore next to them one always writes a round bracket.
- The whole line is an unbounded interval as well:

$$\mathbb{R} = (-\infty, +\infty)$$



Example 2.1: of intervals

The graphical representation of the closed bounded interval $[-3, -2]$ and of the open unbounded interval $(2, +\infty)$ is the following:



The first one contains both of its endpoints, the second one does not contain its endpoint 2 and is not bounded above.

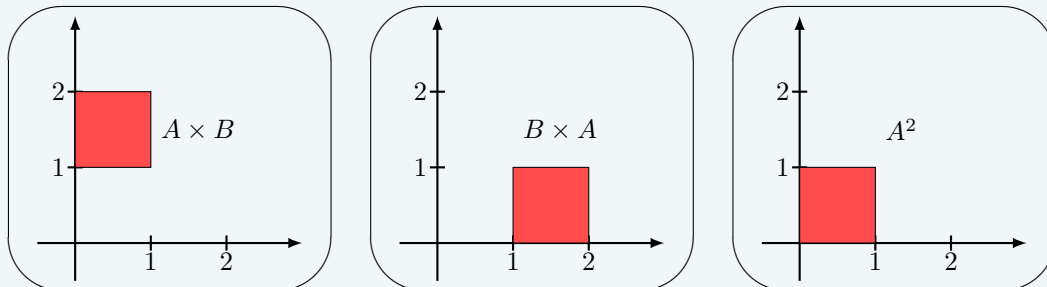
It can be proved that the intervals, bounded or unbounded, are exactly the subsets I of $\mathbb{R}$ that satisfy the following property (called **connectedness**):

$$\forall x_1, x_2, x_3 \in \mathbb{R} \text{ such that } x_1 < x_2 < x_3, \text{ if } x_1, x_3 \in I, \text{ then } x_2 \in I$$

- In what follows we will sometimes consider the Cartesian product of two (or more) intervals, which can be given the geometric meaning of a rectangle (in two dimensions) or of a parallelepiped (in three dimensions).

Example 2.2: Cartesian product of intervals

Let $A = [0, 1]$, $B = [1, 2]$; the figure



illustrates the sets $A \times B$, $B \times A$, $A \times A$, also denoted by A^2 . In general, $A \times B$ is different from $B \times A$.

Ordered fields, supremum/infimum and the completeness axiom

3.1 Properties R_1 and R_2

- We now begin to study more closely the **structure of the number sets** that we introduced earlier, and in particular: the set $\mathbb{Q}$ of **rational numbers** and the set $\mathbb{R}$ of **real numbers**.
- We denote by a, b, c three generic **rational** or **real** numbers.

$R_1 \rightarrow$ In $\mathbb{Q}$ and $\mathbb{R}$ an operation is defined (called **addition** or **sum**) which has the following four properties:

- 1 $\forall a, b, \quad a + b = b + a \quad (\text{commutative property})$
- 2 $\forall a, b, c, \quad (a + b) + c = a + (b + c) \quad (\text{associative property})$
- 3 there exists a **neutral element for the sum**, denoted by 0, such that:

$$\forall a, \quad a + 0 = a$$

- 4 for every a there exists an **inverse** element of a **with respect to the sum**, called the **opposite** of a and denoted by $-a$, such that:

$$a + (-a) = 0$$

$R_2 \rightarrow$ In $\mathbb{Q}$ and $\mathbb{R}$ an operation is defined (called **multiplication** or **product**) which has the following four properties:

- 1 $\forall a, b, \quad a \cdot b = b \cdot a \quad (\text{commutative property})$
- 2 $\forall a, b, c, \quad (a \cdot b) \cdot c = a \cdot (b \cdot c) \quad (\text{associative property})$

3 there exists a **neutral element for the product**, denoted by 1, such that:

$$\forall a, \quad a \cdot 1 = a$$

4 for every $a \neq 0$ there exists an **inverse element of a with respect to the product**, called the **reciprocal** of a and denoted by a^{-1} , such that:

$$a \cdot a^{-1} = 1$$

- The operations of sum and product are linked by the following property:

$$\forall a, b, c, \quad (a + b) \cdot c = a \cdot c + b \cdot c \quad (\text{distributive property})$$

- From properties R_1 and R_2 follows the **possibility of performing the four fundamental operations without restrictions**:

1. **addition** and **multiplication** (defined above)

2. **subtraction**, by setting:

$$a - b = a + (-b)$$

3. **division**, by setting:

$$a/b = a \cdot b^{-1} \quad \text{provided that } b \neq 0$$

3.2 Commensurable magnitudes

In mathematics, the word **magnitude** denotes a **property** of a **geometric figure** that can be **measured**. Examples of magnitudes are: the length of a segment, the measure of an angle, the area of a plane figure or the volume of a solid.

Definition 3.1: of homogeneous magnitudes

Two magnitudes are **homogeneous** if they have the same dimension, i.e., if they can be expressed using the same unit of measurement.

- The length of a segment and the height of a solid are two homogeneous magnitudes; indeed, both can be expressed with the same unit of length.
- The area of a square and the volume of a cone are not homogeneous magnitudes; indeed, area is expressed in square meters or in other units of surface, which however cannot be used to express volume.

Definition 3.2: of commensurable and incommensurable magnitudes

Two homogeneous magnitudes a and b are called **commensurable** if their ratio is a rational number, i.e., if $a/b \in \mathbb{Q}$. They are instead called **incommensurable** if their ratio is not a rational number, i.e., if $a/b \notin \mathbb{Q}$.

Example 3.1: of commensurable magnitudes

The volume of a sphere and the volume of the cylinder circumscribed about the sphere are commensurable. In formulas, the volume V_s of the sphere as a function of its radius r is given by the formula:

$$V_s = \frac{4}{3} \pi r^3$$

The cylinder circumscribed about it has base radius equal to the radius r of the sphere and height h equal to twice the radius. The volume V_c of this cylinder is therefore given by the formula:

$$V_c = \pi r^2 h = \pi r^2 2r = 2\pi r^3$$

Taking the ratio we obtain:

$$\frac{V_s}{V_c} = \frac{\frac{4}{3} \pi r^3}{2 \pi r^3} = \frac{2}{3} \in \mathbb{Q}$$

Example 3.2: of incommensurable magnitudes

The length of a circle and the length of its diameter are incommensurable. In formulas, denoting by d the diameter of the circle and by c the length of the circle, we have:

$$c = \pi d$$

Now, taking the ratio between the length of the circle and its diameter, we obtain:

$$\frac{c}{d} = \frac{\pi d}{d} = \pi \notin \mathbb{Q}$$

A proof that $\pi \notin \mathbb{Q}$ will be given later.

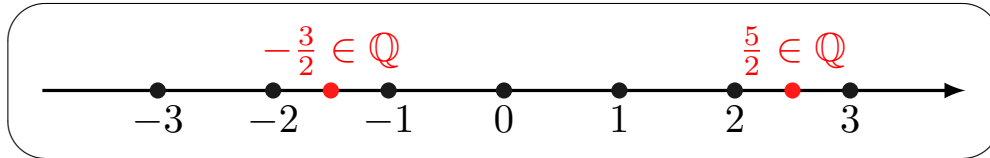
3.3 Geometric representation of $\mathbb{Q}$

A geometric representation of $\mathbb{Q}$ can be obtained by associating with each rational number a point of the **Euclidean line** (which can be imagined as a line in the plane of infinite length)

- With an arbitrarily chosen point of the Euclidean line we associate 0, and with another point, distinct from the first, we associate 1, thus identifying the oriented segment 01, which constitutes the **unit of measurement**

- At this point we have a **one-to-one correspondence** between the rational numbers and those points P of the line that are endpoints of oriented segments $0P$ **commensurable** with 01 . Indeed, if the rational number p (its **abscissa**) is associated with the point P , the segment $0P$ measures p times the unit segment 01 , and therefore the ratio of the two lengths is:

$$\frac{0P}{01} = \frac{p}{1} = p \in \mathbb{Q}$$



3.4 Total order relations

Definition 3.3: of total order relation

A **binary relation** on a set A is a subset R of $A \times A$; we write $a R b$ if $(a, b) \in R$. The relation R is a **partial order relation** if it is reflexive ($a R a$ for every a), antisymmetric ($a R b$ and $b R a$ imply $a = b$) and transitive ($a R b$ and $b R c$ imply $a R c$). It is a **total order relation** if, in addition, $a R b$ or $b R a$ for every $a, b \in A$; the pair (A, R) is then called a **totally ordered set**.

Observation 3.4

The sets of rational numbers $\mathbb{Q}$ and of real numbers $\mathbb{R}$ with the relation “less than or equal to” (“ $\leq$ ”) are totally ordered sets.

- The relations “less than or equal to” (“ $\leq$ ”) for the set of rational or real numbers:

$$R_{\leq} = \left\{ (a, b) : a, b \in \mathbb{Q} \quad \text{and} \quad a \leq b \right\}, \quad R_{\leq} = \left\{ (a, b) : a, b \in \mathbb{R} \quad \text{and} \quad a \leq b \right\}$$

are **partial order relations**. That is, they satisfy the following properties:

$$\forall a, \quad a \leq a \quad (\text{reflexive property})$$

$$\forall a, b, \quad a \leq b \quad \text{and} \quad b \leq a \Rightarrow a = b \quad (\text{antisymmetric property})$$

$$\forall a, b, c, \quad a \leq b \quad \text{and} \quad b \leq c \Rightarrow a \leq c \quad (\text{transitive property})$$

Moreover, the relations “less than or equal to” (“ $\leq$ ”) for $\mathbb{Q}$ and $\mathbb{R}$ are **total relations** since:

$$\forall a, b, \quad a \leq b \quad \text{or} \quad b \leq a$$

- Consequently, the pair consisting of the set $\mathbb{Q}$ or $\mathbb{R}$ and the corresponding relation $R_{\leq}$ is a totally ordered set.

Observation 3.5: trichotomy property

For every pair of numbers a, b (rational or real) **exactly one** of the three relations holds:

$$a < b, \quad a = b, \quad a > b$$

- This is not a new property: it is a consequence of the fact that “ $\leq$ ” is a **total** order relation, as the proof below shows. For this reason we do not count it among the axioms R_1, R_2, R_3 and R_4 .

Proof. Recall that $a < b$ means $a \leq b$ and $a \neq b$.

At least one of the three relations holds. Since the relation is total, we have $a \leq b$ or $b \leq a$:

- if $a \leq b$ and $a \neq b$, then $a < b$;
- if $b \leq a$ and $a \neq b$, then $a > b$;
- if $a = b$ the second relation holds.

At most one of the three relations holds. The relations $a < b$ and $a = b$ cannot hold together, because the first one requires $a \neq b$; for the same reason $a > b$ and $a = b$ cannot hold together. Finally, if $a < b$ and $a > b$ held together, we would have $a \leq b$ and $b \leq a$, and hence $a = b$ by the antisymmetric property, contradicting $a \neq b$. $\square$

3.5 Property R_3

- We now highlight the properties of the ordering of rational or real numbers:

$R_3 \rightarrow$ In $\mathbb{Q}$ and $\mathbb{R}$ the total order relation “less than or equal to” (“ $\leq$ ”) is defined, compatible with the algebraic structure ^a, that is:

$$\begin{aligned} 1. \quad & \forall a, b, c, \quad a \leq b \Rightarrow a + c \leq b + c \\ 2. \quad & \forall a, b \text{ and } \forall c > 0, \quad a \leq b \Rightarrow a \cdot c \leq b \cdot c \end{aligned}$$

^aan algebraic structure is a set, called the underlying set (of the structure), equipped with one or more operations

- We observe that all the rules of algebraic computation follow from the properties: R_1 , R_2 , R_3 .

Example 3.3

For example, all the usual procedures used to solve inequalities are a consequence of the axioms and of the algebraic properties of the sum and the product, expressed by R_1 and R_2 , **and of the compatibility of the order relation with these operations**, expressed by R_3 . It is indeed R_3 that allows one to add the same number to both sides of an inequality and to multiply them by the same positive number without changing the direction of the inequality.

3.6 Ordered fields

Definition 3.6: of (ordered) field

A **field** is a set on which two operations (sum and product) are defined, satisfying the properties R_1 and R_2 and the **distributive property**.

An **ordered field** is a field on which a total order relation satisfying the property R_3 is also defined.

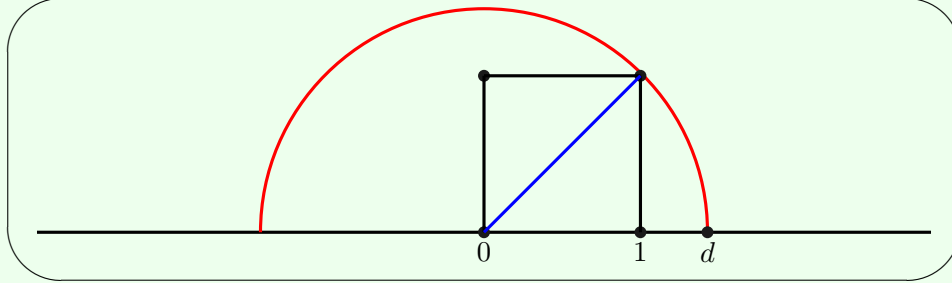
- The distributive property is part of the definition of a field: it is the only property that links the sum and the product to each other, and without it the two operations would remain independent.
- Everything we have said so far about the operations of sum and product and about the total order relation “less than or equal to” (“ $\leq$ ”) holds both for the set of rational numbers and for the set of real numbers.

Observation 3.7

The sets of rational numbers $\mathbb{Q}$ and of real numbers $\mathbb{R}$ are ordered fields.

- From this point of view, therefore, $\mathbb{Q}$ and $\mathbb{R}$ appear to have similar properties. In what follows we will highlight the **property that substantially distinguishes $\mathbb{R}$ from $\mathbb{Q}$** and that makes $\mathbb{R}$ the **right setting for developing mathematical analysis**.
- We begin by observing that the **set of rational numbers is inadequate for expressing the lengths of segments** (but also areas, volumes, times, speeds, etc.)

As we have seen, there exist magnitudes that are not **commensurable** with each other. The classic example is given by the diagonal and the side of a square: if the side measures 1, the abscissa d that measures the diagonal is not a rational number.



Indeed, by the Pythagorean Theorem we have:

$$d^2 = 1^2 + 1^2 = 2$$

but, as we have seen, there is no rational number whose square is 2. Hence the point d on the line does not represent any rational number. This means that, after “occupying” the points of the line with the rational numbers, there are still some **empty places** left on it.

3.7 Bounded sets, maxima and minima

Definition 3.8: of bounded set

Let E be a set contained in $\mathbb{Q}$ or in $\mathbb{R}$. The set E is called **bounded** if there exist two numbers m and M such that:

$$\forall x \in E, \quad m \leq x \leq M$$

- E is **bounded below** if there exists a number m such that:

$$\forall x \in E, \quad m \leq x$$

- E is **bounded above** if there exists a number M such that:

$$\forall x \in E, \quad x \leq M$$

- It makes no difference whether we require m and/or M to belong to $\mathbb{Q}$ or to $\mathbb{R}$, since the condition only concerns the existence of such numbers. This is not an immediate consequence of the definition, though, because $\mathbb{Q}$ is a proper subset of $\mathbb{R}$: it follows from the **Archimedean property** (recalled in the chapter [Sets](#) and proved later in this chapter). Indeed, if $M \in \mathbb{R}$ satisfies $x \leq M$ for every $x \in E$, there exists a natural number $n > M$, and hence $x \leq M < n$ for every $x \in E$, with $n \in \mathbb{Q}$; in the same way, if $m \in \mathbb{R}$

satisfies $m \leq x$ for every $x \in E$, there exists a natural number $n > -m$, that is $-n < m$, and hence $-n < x$ for every $x \in E$, with $-n \in \mathbb{Q}$. Conversely, every rational number is also a real number, and therefore the two requirements single out the same bounded sets.

Definition 3.9: of maximum and minimum of a set

An element $x_M \in E$ is called the **maximum** of E if:

$$\forall x \in E, \quad x \leq x_M$$

An element $x_m \in E$ is called the **minimum** of E if:

$$\forall x \in E, \quad x_m \leq x$$

- Note that the maximum and the minimum, by definition, **belong** to the set E .

Observation 3.10

If a set E has a maximum, this is **unique**; in the same way, if it has a minimum, this is unique. We then write $\max E$ and $\min E$.

Proof. Let x_M and x'_M be two maxima of E . Since x'_M belongs to E and x_M is a maximum of E , we have $x'_M \leq x_M$. Exchanging the roles, since x_M belongs to E and x'_M is a maximum of E , we have $x_M \leq x'_M$. From the antisymmetric property of the order relation it follows that:

$$x'_M \leq x_M \quad \text{and} \quad x_M \leq x'_M \quad \implies \quad x_M = x'_M$$

The proof for the minimum is identical: if x_m and x'_m are two minima of E , then $x_m \leq x'_m$ and $x'_m \leq x_m$, and hence $x_m = x'_m$. $\square$

The existence of the maximum and of the minimum of a set implies that the set is bounded, that is:

$$\text{existence of maximum and minimum} \implies \text{bounded set} \quad (3.1)$$

but the converse is not true (we will give a counterexample later):

$$\text{bounded set} \not\Rightarrow \text{existence of maximum and minimum} \quad (3.2)$$

Hence the fact that a set is bounded is a necessary but not sufficient condition for the set to have a maximum and a minimum. Moreover, the existence of a maximum and a minimum is a sufficient but not necessary condition for a set to be bounded.

From the **contrapositive** of (3.1) we have:

$$\text{unbounded set} \implies \text{the maximum does not exist or the minimum does not exist}$$

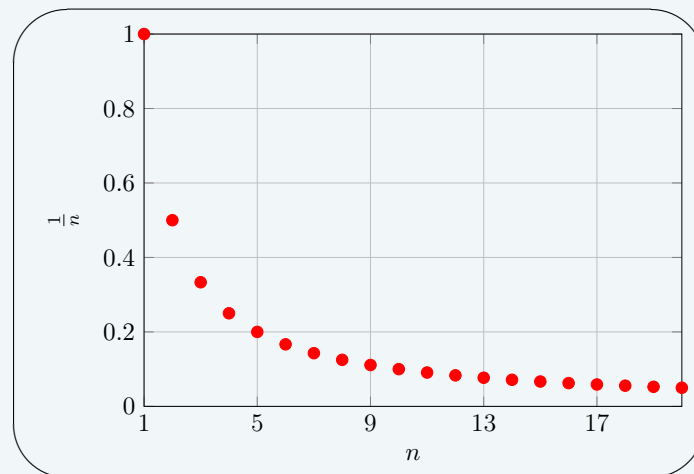
that is, if a set is not bounded, it has no maximum or it has no minimum (it may well have one of the two: $\mathbb{N}$ is unbounded and nevertheless has minimum 0).

Example 3.4: of maxima and minima

Example	set E	minimum	maximum
I)	$\mathbb{N}$	0	does not exist
II)	even integers	does not exist	does not exist

Example 3.5: of maxima and minima

Example	set E	minimum	maximum
III)	$\left\{ \frac{1}{n} : n \in \mathbb{N}_{>0} \right\}$	does not exist	1

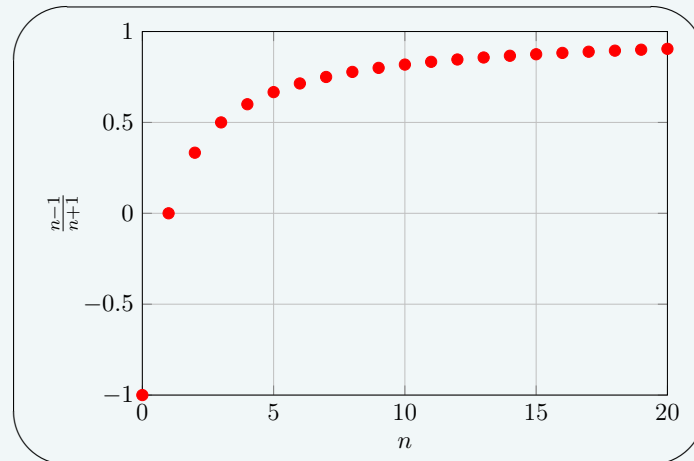


Example 3.6: of maxima and minima

Example	set E	minimum	maximum
IV)	$\left\{ \frac{n-1}{n+1} : n \in \mathbb{N} \right\}$	-1	does not exist

We have:

$$\frac{n-1}{n+1} = \frac{n+1-1-1}{n+1} = 1 - \frac{2}{n+1}$$



Note that sometimes, even though the set is bounded, it may have no maximum or no minimum.

Example 3.7: of maxima and minima

Example	set E	minimum	maximum
V)	$\{x \in \mathbb{R} : 27 \leq x^3\}$	3	does not exist

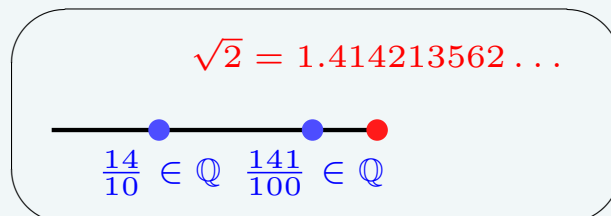
We have:

$$27 \leq x^3 \iff \sqrt[3]{27} = 3 \leq x \quad \text{and} \quad 3 \in E$$

Example 3.8: of maxima and minima

Example	set E	minimum	maximum
VI)	$\{x \in \mathbb{Q} : x \geq 0, x < \sqrt{2}\}$	0	does not exist

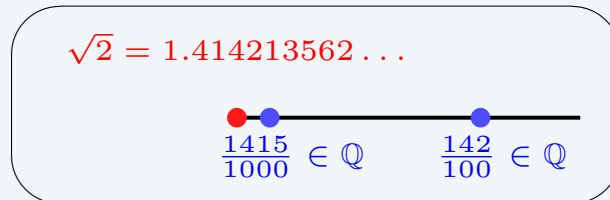
The value $\sqrt{2}$ is irrational, hence by truncating we obtain rational numbers that get closer and closer to $\sqrt{2}$, none of which has the property of being the maximum.



It is an example of a bounded set that has no maximum.

Example	set E	minimum	maximum
VII)	$\{x \in \mathbb{Q} : x > \sqrt{2}, x \leq 4\}$	does not exist	4

The value $\sqrt{2}$ is irrational, hence by truncating and rounding up we obtain rational numbers that get closer and closer to $\sqrt{2}$, none of which has the property of being the minimum.



It is an example of a bounded set that has no minimum.

Example 3.9: of maxima and minima of intervals

Example	set E	minimum	maximum
VIII)	$[-2, 1]$	-2	1
IX)	$(-\infty, 3)$	does not exist	does not exist

- In example VIII) the interval contains both of its endpoints: $-2 \in [-2, 1]$ and $-2 \leq x$ for every $x \in [-2, 1]$, hence -2 is the minimum; $1 \in [-2, 1]$ and $x \leq 1$ for every $x \in [-2, 1]$, hence 1 is the maximum.
- In example IX) the interval is not bounded below and therefore it has no minimum. It has no maximum either: given any $x \in (-\infty, 3)$, the number

$$y = \frac{x+3}{2} \quad \text{satisfies} \quad x < y < 3$$

indeed $y - x = \frac{3-x}{2} > 0$ and $y < 3 \Leftrightarrow x < 3$. Hence $y \in (-\infty, 3)$ and it is strictly greater than x : no element of the set can be the maximum.

Example 3.10: of the maximum and the minimum of a set defined by an inequality

Example	set E	minimum	maximum
X)	$\{x \in \mathbb{R} : -5 \leq 3x < 4\}$	$-\frac{5}{3}$	does not exist

It is convenient to **solve the inequality** first: dividing by $3 > 0$, and using the property

R_3 which preserves the direction of the inequality, we obtain

$$-5 \leq 3x < 4 \quad \Longleftrightarrow \quad -\frac{5}{3} \leq x < \frac{4}{3} \quad \text{that is} \quad E = \left[-\frac{5}{3}, \frac{4}{3} \right)$$

- $-\frac{5}{3} \in E$ and it is less than or equal to every element of E : it is therefore the minimum.
- The maximum does not exist: as in example IX), given $x \in E$, the number $y = \frac{1}{2}(x + \frac{4}{3})$ satisfies $x < y < \frac{4}{3}$ and is still an element of E strictly greater than x .

3.8 Upper/lower bounds and suprema/infima

Definition 3.11: of upper bound of a set

Let E be a set contained in $\mathbb{Q}$ or in $\mathbb{R}$; a number k belonging to $\mathbb{Q}$ or to $\mathbb{R}$ respectively is called an **upper bound** of E if:

$$\forall x \in E, \quad x \leq k$$

Definition 3.12: of lower bound of a set

Let E be a set contained in $\mathbb{Q}$ or in $\mathbb{R}$; a number k belonging to $\mathbb{Q}$ or to $\mathbb{R}$ respectively is called a **lower bound** of E if:

$$\forall x \in E, \quad k \leq x$$

We observe that a set bounded above (below) has many upper (lower) bounds. Moreover, the upper (lower) bounds do not necessarily belong to the set itself.

Observation 3.13

If k is an upper bound of E , then every number $h \geq k$ is still an upper bound of E ; if k is a lower bound of E , then every number $h \leq k$ is still a lower bound of E . Hence the set of the upper (lower) bounds of E , if it is not empty, is a **half-line**.

Proof. Let k be an upper bound of E and let $h \geq k$. For every $x \in E$ we have $x \leq k$ by definition of upper bound, and hence, by the transitive property,

$$x \leq k \leq h \quad \Longrightarrow \quad x \leq h$$

that is, h is an upper bound of E . The set of the upper bounds therefore contains, together with each of its elements k , all the numbers greater than k : it is thus a half-line unbounded above.

The proof for the lower bounds is analogous: if k is a lower bound of E and $h \leq k$, then for every $x \in E$ we have $h \leq k \leq x$ and hence $h \leq x$. $\square$

Observation 3.14

An upper bound of E that **belongs** to E is the maximum of E ; a lower bound of E that belongs to E is the minimum of E .

Proof. Let k be an upper bound of E with $k \in E$. By definition of upper bound we have $x \leq k$ for every $x \in E$: these are exactly the two conditions that define the maximum of E (belonging to E and being greater than or equal to all its elements), and hence $k = \max E$.

In the same way, if k is a lower bound of E and $k \in E$, then $k \leq x$ for every $x \in E$ and $k \in E$, that is $k = \min E$. $\square$

Definition 3.15: of supremum

The **supremum** of E (denoted by $\sup E$) is the minimum of the upper bounds of E .

Definition 3.16: of infimum

The **infimum** of E (denoted by $\inf E$) is the maximum of the lower bounds of E .

We observe that the supremum (infimum) may not exist (unbounded sets). We also observe that if the set has a maximum (minimum), it coincides with the supremum (infimum).

For unbounded sets the following **convention** is adopted: if E is not bounded above one writes

$$\sup E = +\infty$$

and if E is not bounded below one writes

$$\inf E = -\infty$$

The symbols $+\infty$ and $-\infty$ **are not numbers**: this notation is only a compact way of saying that E has no upper (lower) bounds.

Example 3.11: of suprema and infima of intervals

Example	set E	$\inf E$	$\sup E$	minimum	maximum
VIII)	$[-2, 1]$	-2	1	-2	1
IX)	$(-\infty, 3)$	$-\infty$	3	does not exist	does not exist
X)	$\{x \in \mathbb{R} : -5 \leq 3x < 4\}$	$-\frac{5}{3}$	$\frac{4}{3}$	$-\frac{5}{3}$	does not exist

- In examples VIII) and X) the minimum exists and therefore coincides with the infimum; in example VIII) the maximum exists as well and coincides with the supremum.
- In example IX) the set is not bounded below and therefore, by the convention just introduced, $\inf E = -\infty$. The number 3 is an upper bound, because $x < 3$ for every

$x \in E$; no number $k < 3$ is an upper bound, because the number $y = \frac{k+3}{2}$ satisfies $k < y < 3$ and therefore belongs to E . Hence 3 is the minimum of the upper bounds, that is $\sup E = 3$ (and it is not a maximum, because $3 \notin E$).

- In example X), with the same argument applied to $E = [-\frac{5}{3}, \frac{4}{3})$, we obtain $\sup E = \frac{4}{3} \notin E$.
- Finally, going back to example III), for $E = \{\frac{1}{n} : n \in \mathbb{N}_{>0}\}$ we have $\inf E = 0$ and $\sup E = \max E = 1$. Indeed 0 is a lower bound, because $\frac{1}{n} > 0$ for every $n \in \mathbb{N}_{>0}$; and no number $k > 0$ is a lower bound, because by the **Archimedean property** there exists $n \in \mathbb{N}_{>0}$ with $n > \frac{1}{k}$, and hence $\frac{1}{n} < k$. Hence 0 is the maximum of the lower bounds, but it is not the minimum because $0 \notin E$.

Example 3.12: of suprema and infima

Example	set E	$\inf E$	$\sup E$	minimum	maximum
I)	$\{x \in \mathbb{Q} : x \geq 0, x < \sqrt{2}\}$	0	does not exist in $\mathbb{Q}$	0	does not exist

Here the upper bounds are looked for **inside** $\mathbb{Q}$. We have:

$$\sqrt{2} \notin E \quad \text{and} \quad \sqrt{2} \notin \mathbb{Q}$$

hence the maximum does not exist; and the sup does not exist **in** $\mathbb{Q}$ either, because the rational upper bounds of E are the rational numbers greater than $\sqrt{2}$ and among them there is no smallest one (by truncating $\sqrt{2}$ from above one obtains smaller and smaller rational upper bounds). As the next example shows, in $\mathbb{R}$ the supremum does exist and is equal to $\sqrt{2}$.

Example	set E	$\inf E$	$\sup E$	minimum	maximum
II)	$\{x \in \mathbb{R} : x \geq 0, x < \sqrt{2}\}$	0	$\sqrt{2}$	0	does not exist

We have:

$$\sqrt{2} \notin E \quad \text{and} \quad \sqrt{2} \in \mathbb{R}$$

that is, the maximum does not exist, but $\sqrt{2}$ is the minimum of the upper bounds and therefore it is the sup.

3.9 Axiomatic definition of the real numbers

A totally ordered number set X is said to have the **supremum property** (least upper bound property) if:

$R_4 \rightarrow$ Every non-empty set $E \subset X$ that is bounded above has a supremum in X

- It is easy to prove that, if this property holds, then it is also true that every non-empty subset of X that is bounded below has an infimum.
- It is not required that X itself has a supremum (for example, in the case $X = \mathbb{Q}$ or $\mathbb{R}$ this is certainly false!), but that every non-empty subset of X that is bounded above has one.

Example VI shows that $\mathbb{Q}$ certainly does not have the supremum property. On the other hand, $\mathbb{R}$ has this property. ^a

^aHowever, we will not present a formal proof.

- In the **axiomatic definition** of $\mathbb{R}$, this property is part of the very definition of $\mathbb{R}$

Definition 3.17: (axiomatic) of the real numbers

We call $\mathbb{R}$ a set that satisfies the properties R_1, R_2 (together with the distributive property), R_3 and R_4 , that is, an ordered field that has the supremum property

Proposition 3.18: Archimedean property

For every real number x there exists a natural number n such that:

$$n > x$$

- This is the property that we recalled in the chapter [Sets](#) and used above. **It is not an additional axiom**: as the proof shows, it follows from the property R_4 .

Proof. Suppose by contradiction that there exists a real number x such that $n \leq x$ for every $n \in \mathbb{N}$. Then the set $\mathbb{N}$ is a non-empty subset of $\mathbb{R}$ bounded above (by x), and therefore, by the property R_4 , it has a supremum in $\mathbb{R}$:

$$s = \sup \mathbb{N}$$

Since s is the **minimum** of the upper bounds of $\mathbb{N}$, the number $s - 1$, which is smaller than s , is not an upper bound of $\mathbb{N}$: there exists therefore $n \in \mathbb{N}$ such that

$$n > s - 1 \quad \implies \quad n + 1 > s$$

But $n + 1$ is still a natural number, and this contradicts the fact that s is an upper bound of $\mathbb{N}$. The initial assumption is therefore absurd, and for every $x \in \mathbb{R}$ there exists $n \in \mathbb{N}$ with $n > x$. $\square$

- Equivalently: for every real number $\varepsilon > 0$ there exists $n \in \mathbb{N}_{>0}$ such that $\frac{1}{n} < \varepsilon$. Indeed it is enough to choose $n > \frac{1}{\varepsilon}$, and then $\frac{1}{n} < \varepsilon$. As a consequence, the only non-negative number that is less than or equal to $\frac{1}{n}$ for every $n \in \mathbb{N}_{>0}$ is zero.
- The supremum property is also called the **Dedekind axiom**, or **axiom of continuity**, or **completeness axiom**, and it can also be stated in the following equivalent form.

Let $\{A, B\}$ be a partition of $\mathbb{R}$ (i.e., A and B are non-empty disjoint sets whose union is $\mathbb{R}$); the partition is called a **cut** (section) if:

$$\forall a \in A \text{ and } \forall b \in B \text{ we have } a < b.$$

Then one proves that:

$R'_4 \rightarrow$ For every cut $\{A, B\}$ of $\mathbb{R}$ there exists a unique real number s (called the **separating element**) such that:

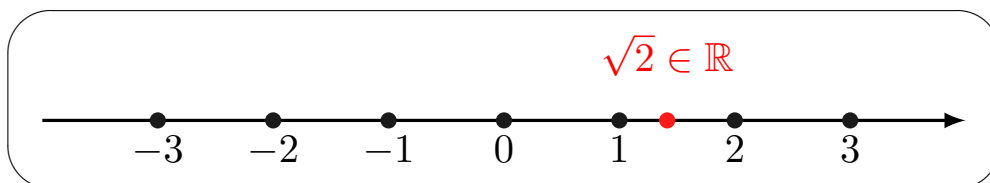
$$\forall a \in A \text{ and } \forall b \in B, \quad a \leq s \leq b$$

This separating element is $\sup A$ and $\inf B$, hence $\sup A = \inf B$

3.10 Geometric representation of $\mathbb{R}$

Can the set of real numbers $\mathbb{R}$ be put in one-to-one correspondence with the points of the Euclidean line?

- Let us now return to the problem of the incommensurability of the side and the diagonal of a square. When, with straightedge and compass, having fixed the unit side of the square, we construct its diagonal and transfer it onto the starting line, the arc drawn by the compass "splits the line in two", generating what we have called a cut; the separating element, which exists by axiom R_4 , geometrically represents the point of intersection, and therefore that number measures the diagonal.



Hence the geometric interpretation of property R_4 shows us how the set $\mathbb{R}$ is an adequate representation of our intuitive idea of a line, so adequate that **in mathematics the expression “the real line” is often used to denote $\mathbb{R}$** , identifying the number set with its natural geometric representation.

Roots, powers, logarithms and modular arithmetic

4.1 Roots, powers, logarithms

- As a consequence of property R_4 we can perform, in the real field, operations that are only occasionally possible in the rational field, such as extracting roots or raising to a power.

4.1.1 Arithmetic n -th roots

Theorem 4.1

For every $y \in \mathbb{R}$, $y > 0$ and $n \in \mathbb{N}$, $n \geq 1$, there exists one and only one $x \in \mathbb{R}$, $x > 0$, such that $x^n = y$.

- This number is called the arithmetic n -th root of y and is denoted by one of the symbols

$$\sqrt[n]{y} \text{ or } y^{\frac{1}{n}}.$$

Proof. The proof of this Theorem will be given using the properties of continuous functions. □

The arithmetic n -th root is non-negative.

Example 4.1: arithmetic n -th root

For example:

$$\sqrt{4} = 2, \quad \sqrt{9} = 3.$$

Moreover, we have:

$$\sqrt{x^2} = |x| \quad (4.1)$$

Construction of the decimal representation of the n -th root.

- We look for the decimal expansion of

$$\sqrt{2} \approx 1.41421356237$$

this number, not being rational, will be represented by an infinite (non-periodic) expansion.

- We proceed as follows: we build a class of rational numbers of the form:

$$\begin{aligned} 0 &< a_0 \\ &< a_0, a_1 \\ &< a_0, a_1 a_2 \\ &< a_0, a_1 a_2 a_3 \\ &< \dots \\ &< a_0, a_1 a_2 a_3 \dots a_n \\ &< \dots \end{aligned}$$

The rule is: each of these numbers is the largest among those with the same number of digits after the decimal point whose square is less than 2. The first of these numbers are:

$$\begin{array}{lll} 1 & 1^2 & = 1 \\ 1.4 & (1.4)^2 & = 1.96 \\ 1.41 & (1.41)^2 & = 1.9881 \\ 1.414 & (1.414)^2 & = 1.999396 \\ 1.4142 & (1.4142)^2 & = 1.99996164 \\ \dots & & \end{array}$$

- This set of numbers, which we call E_- , is bounded above (each one is < 2); by property R_4 it has a supremum, and this is used to define $\sqrt{2}$.

The number $\sqrt{2}$ is defined precisely as $\sup E_-$

- We could also have built a class of numbers E_+ like the previous one, with the rule that each of them is the smallest among those with the same number of digits after the decimal point whose square is greater than 2; we would have obtained:

$$\begin{array}{rcl}
 2 & 2^2 & = 4 \\
 1.5 & (1.5)^2 & = 2.25 \\
 1.42 & (1.42)^2 & = 2.0164 \\
 1.415 & (1.415)^2 & = 2.002225 \\
 1.4143 & (1.4143)^2 & = 2.00024449 \\
 \dots & &
 \end{array}$$

This set E_+ is bounded below (every element is > 1), therefore it has an infimum; one proves that:

$$\sqrt{2} = \inf E_+ = \sup E_-$$

The numbers of the class E_- approximate $\sqrt{2}$ from below, those of the class E_+ from above.

4.1.2 Powers with real exponent

Extracting the n -th root is the inverse operation of raising to an integer power.

Rational exponent

- The operation of raising to a power can be extended to every rational exponent if the base is positive (using the previous theorem).

$$\text{If } r = \frac{m}{n} \quad \text{and} \quad a > 0 \quad \text{then} \quad a^r = (a^m)^{\frac{1}{n}} = \sqrt[n]{a^m}$$

(we assume $m \in \mathbb{Z}$ and $n \in \mathbb{Z}$ positive)

Real exponent

- If the exponent is real, $b = b_0, b_1 b_2 b_3 \dots b_n \dots$, the number a^b ($a > 0$) will be determined by the class of numbers

$$a^{b_0} \quad a^{b_0, b_1} \quad a^{b_0, b_1 b_2} \quad \dots$$

in a way similar to the case of the root.

If the base a is negative, the operation of raising to a power a^b is defined only in certain cases:

- if the exponent b is an integer, or
- if the exponent $b = \frac{n}{m}$ is rational, *written in lowest terms*, provided that it is not the case that n is odd and m is even.

If $c < 0$ and m is odd, we define

$$\sqrt[m]{c} = -\sqrt[m]{-c}$$

Example 4.2: negative base

For example:

$$(-2)^{\frac{3}{5}} = \sqrt[5]{(-2)^3} = \sqrt[5]{-8} = -\sqrt[5]{8}$$

$$(-2)^{\frac{2}{5}} = \sqrt[5]{(-2)^2} = \sqrt[5]{4}$$

The assumption that the fraction is written in lowest terms is essential: $\frac{1}{3}$ and $\frac{2}{6}$ are the same rational number, but

$$\sqrt[3]{(-8)^1} = \sqrt[3]{-8} = -2 \quad \text{while} \quad \sqrt[6]{(-8)^2} = \sqrt[6]{64} = 2.$$

Only the first expression, with $\frac{n}{m} = \frac{1}{3}$ in lowest terms, defines $(-8)^{\frac{1}{3}}$.

- When we say “does not exist in $\mathbb{R}$ ” we mean that it is not possible to define this operation in such a way that the usual rules of computation remain valid.

Example 4.3: non-existence of a power with real exponent

For example:

$$(-2)^{\frac{3}{4}}$$

does not exist in the real field!

Let

$$a, b \in \mathbb{R}, \quad a > 0 \quad \text{and} \quad b > 0 \quad (\text{positive reals})$$

and

$$c, d \in \mathbb{R} \quad (\text{arbitrary reals}),$$

the main properties of raising to a power are:

$$a^0 = 1 \quad \forall a \neq 0 \quad (4.2a)$$

$$1^c = 1 \quad \forall c \quad (4.2b)$$

$$a^c > 0 \quad \forall c \quad (4.2c)$$

$$a^c > 1 \quad \forall a > 1 \quad \text{and} \quad c > 0 \quad (4.2d)$$

$$a^c < 1 \quad \forall a < 1 \quad \text{and} \quad c > 0 \quad (4.2e)$$

$$a^{c+d} = a^c \cdot a^d \quad (4.3a)$$

$$(a \cdot b)^c = a^c \cdot b^c \quad (4.3b)$$

$$(a^b)^c = a^{b \cdot c} \quad (4.3c)$$

$$c < d \Rightarrow a^c < a^d \text{ if } a > 1 \quad (4.4a)$$

$$c < d \Rightarrow a^c > a^d \text{ if } a < 1 \quad (4.4b)$$

$$0 < a \leq b \Rightarrow a^c \leq b^c \quad \forall c > 0 \quad (4.4c)$$

4.1.3 Logarithms

- Consider the equation

$$a^x = y, \quad a > 0$$

with y given and x unknown. First of all, if $a = 1$, it is solvable only if $y = 1$ (and in that case every real number x is a solution). So let $a \neq 1$.

If $y \leq 0$ it has no solution, see property (4.2c).

The following theorem tells us that it has exactly one solution for every $y > 0$:

Theorem 4.2

Let $a > 0$, $a \neq 1$, $y > 0$. There exists a unique real number x such that $a^x = y$.

- This number is called the logarithm to base a of y and is denoted by the symbol:

$$\log_a y$$

- The logarithm is the inverse operation of raising to a power.
- the logarithm of a number to a given base is the exponent to which the base must be raised to obtain the number itself.

Let

$$x, y, a \in \mathbb{R}, \quad x > 0, y > 0, a > 0 \quad (\text{positive reals}) \quad \text{and} \quad a \neq 1,$$

the main properties of logarithms are (they follow from those of exponentials):

$$\log_a(x \cdot y) = \log_a x + \log_a y \quad (4.5a)$$

$$\log_a\left(\frac{x}{y}\right) = \log_a x - \log_a y \quad (4.5b)$$

$$\log_a\left(\frac{1}{y}\right) = -\log_a y \quad (4.5c)$$

$$\log_a x^\alpha = \alpha \log_a x \quad (\forall \alpha \in \mathbb{R}) \quad (4.6a)$$

$$\log_b x = \frac{\log_a x}{\log_a b} \quad (\forall b > 0, b \neq 1) \quad (4.6b)$$

$$\log_a x = \frac{1}{\log_x a} = -\log_{\frac{1}{a}} x \quad (x \neq 1) \quad (4.6c)$$

Notation:

- $\lg n = \log_2 n \rightarrow$ **binary logarithm**
- $\ln n = \log_e n \rightarrow$ **natural logarithm** (where $e \approx 2.718$ is **Euler's number** (Napier's constant)) (sometimes the natural logarithm $\ln$ is also denoted by $\log$)
- $\lg^k n = (\lg n)^k \rightarrow$ **power**
- $\lg \lg n = \lg(\lg n) \rightarrow$ **composition**

4.1.4 Approximations

- a rational number can always be expressed with absolute precision, using either the fractional or the decimal notation (possibly with repeating digits).
- On the other hand, it is of course not possible to write all the decimal digits of an irrational number, since there are infinitely many of them and they follow one another without periodicity.
- What does it mean, then, to “know” or “specify” an irrational number? It means knowing some algorithm that allows us (at least in theory) to write as many exact decimal digits as we wish.

Example 4.4: irrational numbers

– In the (already considered) example of the irrational number:

0.101001000100001...

built according to the rule: write one digit 1, one digit 0, one digit 1, two digits 0, one digit 1, three digits 0, and so on) it is clear that we could write as many digits as we wish.

- In other cases, such as

$$\sqrt{2} \text{ or } \log_2 3,$$

things are more laborious and require iterative computations to determine each successive decimal digit; nevertheless, these computations can actually be carried out.

- At other times, in Mathematical Analysis, a number is specified in a non-constructive way, by denoting it as the unique number that solves a given problem (once it has been proved that such a solution exists and is unique). This is a less satisfactory approach from an operational point of view, but theoretically flawless.

Practical aspects of computing irrational numbers

- Whenever we perform computations with a **pocket calculator**, it will only display numbers with a fixed number of decimal digits after the decimal point (typically 9). This means that we are working only with rational numbers, indeed with a finite subset of the set $\mathbb{Q}$.
- Working with a **computer** things improve somewhat, but we still remain within finite subsets of $\mathbb{Q}$. Of course, one needs to be aware of this.
- Besides the approximation produced by computing tools, sometimes we ourselves are not interested in too many decimal digits; we then deliberately introduce approximations, writing for example

$$\sqrt{2} \approx 1.414$$

Rounding rule: the last digit written is rounded down (up) if the first digit discarded is from 0 to 4 (respectively, from 5 to 9).

Example 4.5: Rounding

For example:

$$2.4138 \approx 2.41$$

$$2.4152 \approx 2.42$$

4.2 Modular arithmetic

Definition 4.3: of integer part

Given a real number $a \in \mathbb{R}$, we denote by $[a]$ or $\lfloor a \rfloor$ the **integer part** (or “**floor**”) of a :

$$[a] = \text{integer } n \text{ such that } n \leq a < n + 1$$

Equivalently, the integer part of a is the **largest** integer less than or equal to a :

$$\lfloor a \rfloor = \max \{n \in \mathbb{Z} : n \leq a\}.$$

- The two descriptions coincide. If $n \leq a < n + 1$, then n belongs to the set $\{k \in \mathbb{Z} : k \leq a\}$ and every other element k of it satisfies $k \leq a < n + 1$, that is, $k \leq n$ (since k and n are integers): hence n is its maximum. Conversely, if n is the maximum of that set, then $n \leq a$ and moreover $a < n + 1$, since otherwise $n + 1$ would belong to the set and n would not be its maximum.

Example 4.6: Integer part

$$[2.38] = 2; \quad [3] = 3; \quad [-1.8] = -2.$$

- While for positive numbers the integer part is obtained simply by “throwing away the digits after the decimal point”, for negative numbers one must take the largest integer $\leq a$, which is different from what is obtained by throwing away the digits after the decimal point (except when a is already an integer)

Definition 4.4: of fractional part

Given a real number $a \in \mathbb{R}$, we denote by (a) the **fractional part** (or decimal part, in Italian “*mantissa*”) of a :

$$(a) = a - [a]$$

Example 4.7: Fractional part

$$(2.38) = 0.38; \quad (3) = 0; \quad (-1.8) = 0.2.$$

- The fractional part is therefore not an integer but a real number, lying in $[0, 1)$. Indeed, setting $n = [a]$, the definition of integer part gives $n \leq a < n + 1$ and, subtracting n ,

$$0 \leq \underbrace{a - n}_{(a)} < 1.$$

- For positive numbers, it is obtained simply by “throwing away the digits before the decimal point”; for negative numbers the fractional part is the **one’s complement** of the number obtained by throwing away the digits before the decimal point.

Definition 4.5: of upper integer part

Given a real number $a \in \mathbb{R}$, we denote by $\lceil a \rceil$ the **upper integer part** (or “**ceil**”) of a :

$$\lceil a \rceil = \text{integer } n \text{ such that } n - 1 < a \leq n$$

Equivalently, the upper integer part of a is the **smallest** integer greater than or equal to a :

$$\lceil a \rceil = \min \{n \in \mathbb{Z} : n \geq a\}.$$

- Here too the two descriptions coincide. If $n - 1 < a \leq n$, then n belongs to the set $\{k \in \mathbb{Z} : k \geq a\}$ and every other element k of it satisfies $k \geq a > n - 1$, that is, $k \geq n$: hence n is its minimum. Conversely, if n is the minimum of that set, then $a \leq n$ and moreover $n - 1 < a$, since otherwise $n - 1$ would belong to the set and n would not be its minimum.

Example 4.8: of ceil

$$\lceil 2.38 \rceil = 3; \quad \lceil 3 \rceil = 3; \quad \lceil -1.8 \rceil = -1.$$

Some properties:

Given a real number $a \in \mathbb{R}$, we have

$$a - 1 < \lfloor a \rfloor \leq a \leq \lceil a \rceil < a + 1$$

- This follows at once from the two definitions: from $\lfloor a \rfloor \leq a < \lfloor a \rfloor + 1$ we get $\lfloor a \rfloor \leq a$ and $a - 1 < \lfloor a \rfloor$; from $\lceil a \rceil - 1 < a \leq \lceil a \rceil$ we get $a \leq \lceil a \rceil$ and $\lceil a \rceil < a + 1$.

Moreover, the following relations between the two functions hold.

Observation 4.6: integer part, upper integer part and change of sign

For every real number $a \in \mathbb{R}$ we have

$$-\lfloor a \rfloor = \lceil -a \rceil \quad \text{and} \quad -\lceil a \rceil = \lfloor -a \rfloor \quad (4.7)$$

and moreover

$$\lfloor a \rfloor = a \iff a \in \mathbb{Z} \iff \lceil a \rceil = a \quad (4.8)$$

Proof. Set $n = \lfloor a \rfloor$, that is, $n \leq a < n + 1$. Multiplying by -1 (and reversing the inequalities) we obtain

$$-n - 1 < -a \leq -n, \quad \text{that is,} \quad (-n) - 1 < -a \leq -n.$$

Since $-n$ is an integer, the definition of upper integer part gives $\lceil -a \rceil = -n = -\lfloor a \rfloor$: this is the first identity in (4.7). Applying the first identity to the number $-a$ we obtain $-\lfloor -a \rfloor = \lceil a \rceil$, that is, the second one.

As for (4.8): if $\lfloor a \rfloor = a$ then a is an integer, because $\lfloor a \rfloor \in \mathbb{Z}$. Conversely, if $a \in \mathbb{Z}$ then $a \leq a < a + 1$ and hence $\lfloor a \rfloor = a$. In the same way, if $\lceil a \rceil = a$ then $a \in \mathbb{Z}$; and if $a \in \mathbb{Z}$ then $a - 1 < a \leq a$, hence $\lceil a \rceil = a$. $\square$

The two functions are completely characterized by the following inequalities.

Observation 4.7: characterizations of the integer part and of the upper integer part

For every real number $a \in \mathbb{R}$ and every integer $n \in \mathbb{Z}$ we have:

$$\lfloor a \rfloor = n \iff n \leq a < n + 1 \quad (4.9a)$$

$$\lfloor a \rfloor = n \iff a - 1 < n \leq a \quad (4.9b)$$

$$\lceil a \rceil = n \iff n - 1 < a \leq n \quad (4.9c)$$

$$\lceil a \rceil = n \iff a \leq n < a + 1 \quad (4.9d)$$

$$a < n \iff \lfloor a \rfloor < n \quad (4.9e)$$

$$n \leq a \iff n \leq \lfloor a \rfloor \quad (4.9f)$$

$$a \leq n \iff \lceil a \rceil \leq n \quad (4.9g)$$

$$n < a \iff n < \lceil a \rceil \quad (4.9h)$$

and moreover

$$\lfloor a + n \rfloor = \lfloor a \rfloor + n \quad \text{and} \quad \lceil a + n \rceil = \lceil a \rceil + n \quad (4.10)$$

Proof. • Equation (4.9a) is the definition of integer part and (4.9c) is the definition of upper integer part.

- Equation (4.9b) is a rewriting of (4.9a): the condition $n \leq a < n + 1$ is equivalent to “ $n \leq a$ and $a < n + 1$ ”, that is, to “ $n \leq a$ and $a - 1 < n$ ”. In the same way (4.9d) is a rewriting of (4.9c): $n - 1 < a \leq n$ is equivalent to “ $a \leq n$ and $n < a + 1$ ”.
- (4.9e): if $a < n$, then $\lfloor a \rfloor \leq a < n$. Conversely, if $\lfloor a \rfloor < n$, then, both being integers, $\lfloor a \rfloor \leq n - 1$ and hence $a < \lfloor a \rfloor + 1 \leq n$.

- (4.9g): if $a \leq n$, then n belongs to the set $\{k \in \mathbb{Z} : k \geq a\}$, whose minimum is $\lceil a \rceil$, hence $\lceil a \rceil \leq n$. Conversely, if $\lceil a \rceil \leq n$, then $a \leq \lceil a \rceil \leq n$.
- Equations (4.9f) and (4.9h) are obtained by negating both sides of (4.9e) and of (4.9g) respectively: the negation of $a < n$ is $n \leq a$ and the negation of $\lfloor a \rfloor < n$ is $n \leq \lfloor a \rfloor$; the negation of $a \leq n$ is $n < a$ and the negation of $\lceil a \rceil \leq n$ is $n < \lceil a \rceil$.
- (4.10): setting $m = \lfloor a \rfloor$, that is, $m \leq a < m + 1$, and adding n we obtain $m + n \leq a + n < (m + n) + 1$ with $m + n \in \mathbb{Z}$, hence $\lfloor a + n \rfloor = m + n = \lfloor a \rfloor + n$. In the same way, setting $m = \lceil a \rceil$, from $m - 1 < a \leq m$ we obtain $(m + n) - 1 < a + n \leq m + n$, hence $\lceil a + n \rceil = m + n = \lceil a \rceil + n$.

□

Given an integer $n \in \mathbb{Z}$, we have

$$\left\lfloor \frac{n}{2} \right\rfloor + \left\lceil \frac{n}{2} \right\rceil = n$$

Proof. We distinguish the two cases.

- If n is **even**, there exists $m \in \mathbb{Z}$ such that $n = 2m$, and since m is an integer (4.8) gives $\lfloor m \rfloor = \lceil m \rceil = m$:

$$\left\lfloor \frac{n}{2} \right\rfloor + \left\lceil \frac{n}{2} \right\rceil = \left\lfloor \frac{2m}{2} \right\rfloor + \left\lceil \frac{2m}{2} \right\rceil = \lfloor m \rfloor + \lceil m \rceil = m + m = 2m = n.$$

- If n is **odd**, there exists $m \in \mathbb{Z}$ such that $n = 2m + 1$. Since $m \leq m + \frac{1}{2} < m + 1$, (4.9a) gives $\lfloor m + \frac{1}{2} \rfloor = m$; since $(m + 1) - 1 < m + \frac{1}{2} \leq m + 1$, (4.9c) gives $\lceil m + \frac{1}{2} \rceil = m + 1$. Hence

$$\left\lfloor \frac{n}{2} \right\rfloor + \left\lceil \frac{n}{2} \right\rceil = \left\lfloor m + \frac{1}{2} \right\rfloor + \left\lceil m + \frac{1}{2} \right\rceil = m + (m + 1) = 2m + 1 = n.$$

□

Given two positive integers $r, s \in \mathbb{Z}$, $r, s > 0$, we have

$$\left\lceil \frac{r}{s} \right\rceil \leq \frac{r + (s - 1)}{s} \quad (4.11a)$$

$$\left\lfloor \frac{r}{s} \right\rfloor \geq \frac{r - (s - 1)}{s} \quad (4.11b)$$

Given, in addition, a non-negative real number $p \in \mathbb{R}$, $p \geq 0$, we have

$$\left\lceil \frac{\left\lfloor \frac{p}{r} \right\rfloor}{s} \right\rceil = \left\lceil \frac{p}{r s} \right\rceil \quad (4.12a)$$

$$\left\lfloor \frac{\left\lceil \frac{p}{r} \right\rceil}{s} \right\rfloor = \left\lfloor \frac{p}{r s} \right\rfloor \quad (4.12b)$$

Definition 4.8: Divisor

A **divisor** of an integer $n \in \mathbb{Z}$, also called a **factor** of n , is an integer $m \in \mathbb{Z}$ that can be multiplied by some integer $q \in \mathbb{Z}$ to obtain n , i.e., if $n = q \cdot m$.

- If m is a divisor of n , n is a **multiple** of m .
- An integer n is **divisible** by another integer m if m is a *divisor* of n .
- For an integer $a \in \mathbb{Z}$ and a positive integer $n \in \mathbb{Z}, n > 0$, the value $a \bmod n$ is the remainder of the division $\frac{a}{n}$.

Definition 4.9: of modulo (remainder)

Given $a \in \mathbb{Z}$ and $n \in \mathbb{Z}, n > 0$,

$$a \bmod n = a - n \left\lfloor \frac{a}{n} \right\rfloor$$

- It follows that $0 \leq a \bmod n < n$

If $(a \bmod n) = (b \bmod n)$, we write

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

and we say that a is **equivalent** to b , modulo n .

Example 4.9

For example, 23 and 13 are equivalent modulo 5 and we write $23 \equiv 13 \pmod{5}$

- Equivalently, $a \equiv b \pmod{n}$ if a and b have the same remainder when divided by n .
- Equivalently, $a \equiv b \pmod{n}$ if and only if n is a divisor of $|b - a|$.

Summations, geometric and arithmetic progressions

5.1 Summations

Definition 5.1: summation

Let $a_1, a_2, \dots, a_n$ be n real numbers. Their sum

$$a_1 + a_2 + \dots + a_n$$

can be written in compact form with the summation symbol:

$$\sum_{k=1}^n a_k$$

which is read: “sum for k from 1 to n of a_k ”. The symbol k is called the summation index.

- The summation symbol is therefore just a shorthand, which is nevertheless very useful when the terms a_k are defined explicitly as a function of the index k .

Example 5.1: summations

$$\sum_{k=1}^{10} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \frac{1}{9} + \frac{1}{10}$$

$$\sum_{k=3}^n k^2 = 3^2 + 4^2 + 5^2 + \dots + n^2$$

- The summation index is a **dummy index**. This means that if we replace k with i , j or any other index (in all its occurrences), the value of the summation does not change.

Example 5.2: dummy index

We have:

$$\sum_{k=1}^n k^2 = \sum_{i=1}^n i^2$$

since both symbols denote the sum of the squares of the first n natural numbers (without zero).

On the other hand, we have:

$$\sum_{k=1}^n k^2 \neq \sum_{k=1}^m k^2$$

since the two symbols denote the sum of, respectively, the first n or the first m squares (if $n \neq m$ the result will be different).

5.1.1 Main properties of summations

Observation 5.2: product by a constant

Given a summation $\sum_{k=1}^n a_k$ and a real number $c \in \mathbb{R}$, we have:

$$\sum_{k=1}^n (c \cdot a_k) = c \sum_{k=1}^n a_k \quad (5.1)$$

Proof. By the distributive property we have:

$$\underbrace{c a_1 + c a_2 + \cdots + c a_n}_{=\sum_{k=1}^n (c \cdot a_k)} = \underbrace{c (a_1 + a_2 + \cdots + a_n)}_{=c \sum_{k=1}^n a_k}$$

□

Observation 5.3: summation with a constant term

For every natural number $n \geq 1$ and every real number $c \in \mathbb{R}$, we have:

$$\sum_{k=1}^n c = c \cdot n \quad (5.2)$$

Proof. We have:

$$\underbrace{c + c + \cdots + c}_{\text{that is, } c \text{ added } n \text{ times}} = c n$$

□

Observation 5.4: union of summations

Given two summations $\sum_{k=1}^n a_k$ and $\sum_{k=1}^n b_k$, we have:

$$\sum_{k=1}^n a_k + \sum_{k=1}^n b_k = \sum_{k=1}^n (a_k + b_k) \quad (5.3)$$

Proof. We have:

$$\underbrace{a_1 + a_2 + \cdots + a_n + b_1 + b_2 + \cdots + b_n}_{=\sum_{k=1}^n a_k + \sum_{k=1}^n b_k} = \underbrace{a_1 + b_1 + a_2 + b_2 + \cdots + a_n + b_n}_{=\sum_{k=1}^n (a_k + b_k)}$$

□

- The three properties that follow (decomposition, index translation and index reflection) are simply different ways of writing and/or ordering the terms of the summations.

Observation 5.5: decomposition

Given two natural numbers $n \geq 1$ and $m \geq 1$, we have:

$$\sum_{k=1}^{n+m} a_k = \sum_{k=1}^n a_k + \sum_{k=n+1}^{n+m} a_k \quad (5.4)$$

Proof. The terms of the summation $\sum_{k=1}^{n+m} a_k$ can be grouped into the first n and the last m ones:

$$\begin{aligned} & \underbrace{a_1 + a_2 + \cdots + a_n}_{=\sum_{k=1}^n a_k} + \underbrace{a_{n+1} + a_{n+2} + \cdots + a_{n+m}}_{=\sum_{k=n+1}^{n+m} a_k} \\ &= \underbrace{a_1 + a_2 + \cdots + a_n + a_{n+1} + \cdots + a_{n+m}}_{=\sum_{k=1}^{n+m} a_k} \end{aligned}$$

□

Observation 5.6: index translation

Given a summation $\sum_{k=1}^n a_k$ and a natural number $m \geq 1$, we have:

$$\sum_{k=1}^n a_k = \sum_{k=1+m}^{n+m} a_{k-m} = \sum_{k=1-m}^{n-m} a_{k+m} \quad (5.5)$$

Proof. In the summation $\sum_{k=1+m}^{n+m} a_{k-m}$ the index k ranges from $1+m$ to $n+m$, and therefore the index $k-m$ ranges from 1 to n :

$$\underbrace{a_{(1+m)-m} + a_{(2+m)-m} + \cdots + a_{(n+m)-m}}_{=\sum_{k=1+m}^{n+m} a_{k-m}} = \underbrace{a_1 + a_2 + \cdots + a_n}_{=\sum_{k=1}^n a_k}$$

In the same way, in the summation $\sum_{k=1-m}^{n-m} a_{k+m}$ the index k ranges from $1-m$ to $n-m$, and therefore the index $k+m$ ranges from 1 to n :

$$\underbrace{a_{(1-m)+m} + a_{(2-m)+m} + \cdots + a_{(n-m)+m}}_{=\sum_{k=1-m}^{n-m} a_{k+m}} = \underbrace{a_1 + a_2 + \cdots + a_n}_{=\sum_{k=1}^n a_k}$$

□

Observation 5.7: index reflection

Given a summation $\sum_{k=1}^n a_k$, we have:

$$\sum_{k=1}^n a_k = \sum_{k=1}^n a_{n-k+1} = \sum_{k=0}^{n-1} a_{n-k} \quad (5.6)$$

Proof. In the summation $\sum_{k=1}^n a_{n-k+1}$ the index k ranges from 1 to n , and therefore the index $n-k+1$ ranges from n to 1, that is, the terms are the same ones listed in reverse order:

$$\underbrace{a_n + a_{n-1} + \cdots + a_1}_{=\sum_{k=1}^n a_{n-k+1}} = \underbrace{a_1 + a_2 + \cdots + a_n}_{=\sum_{k=1}^n a_k}$$

In the same way, in the summation $\sum_{k=0}^{n-1} a_{n-k}$ the index k ranges from 0 to $n-1$, and therefore the index $n-k$ ranges from n to 1:

$$\underbrace{a_n + a_{n-1} + \cdots + a_1}_{=\sum_{k=0}^{n-1} a_{n-k}} = \underbrace{a_1 + a_2 + \cdots + a_n}_{=\sum_{k=1}^n a_k}$$

□

5.1.2 Some important summations

Observation 5.8: sum of the first n natural numbers (without zero)

For every natural number $n \geq 1$, we have:

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

Proof. Using the index reflection (5.6) we have:

$$\begin{aligned} \sum_{k=1}^n k &= \frac{1}{2} \left(\sum_{k=1}^n k + \sum_{k=1}^n k \right) = \frac{1}{2} \left(\sum_{k=1}^n k + \sum_{k=1}^n (n - k + 1) \right) \\ &= \frac{1}{2} \sum_{k=1}^n (k + n - k + 1) = \frac{1}{2} \sum_{k=1}^n (n + 1) = \frac{n(n+1)}{2} \end{aligned}$$

□

Observation 5.9: sum of the first n odd numbers

For every natural number $n \geq 1$, we have:

$$\sum_{k=1}^n (2k-1) = n^2 \quad \text{or equivalently} \quad \sum_{k=0}^{n-1} (2k+1) = n^2$$

Proof. Using the properties of summations we have:

$$\begin{aligned} \sum_{k=1}^n (2k-1) &= 2 \sum_{k=1}^n k - \sum_{k=1}^n 1 \\ &= 2 \left(\frac{n(n+1)}{2} \right) - n = n^2 + n - n = n^2 \\ \sum_{k=0}^{n-1} (2k+1) &= 2 \sum_{k=0}^{n-1} k + \sum_{k=0}^{n-1} 1 \\ &= 2 \sum_{k=1}^n (k-1) + n = 2 \left(\sum_{k=1}^n k - \sum_{k=1}^n 1 \right) + n \\ &= 2 \left(\frac{n(n+1)}{2} - n \right) + n = n^2 + n - 2n + n = n^2 \end{aligned}$$

□

Observation 5.10: sum of the first n even numbers (without zero)

For every natural number $n \geq 1$, we have:

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

Proof.

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

□

- The three important summations we have just seen can also be proved by induction, in the chapter “[Principle of induction](#)” of Part 1.

5.2 Geometric progressions

Definition 5.11: geometric progression

A sequence of real numbers is a **geometric progression** if the ratio between each term (starting from the second one) and the previous one is constant. This constant is called the common ratio of the progression.

Given the first term $a \in \mathbb{R}$ and the common ratio $q \in \mathbb{R}$, the associated geometric progression is:

$$a, \quad aq, \quad aq^2, \quad aq^3, \quad aq^4, \quad \dots$$

Each term (starting from the second one) is obtained from the previous one by multiplying it by q . The k -th term ($k \in \mathbb{N}$, $k \geq 1$) can be written as aq^{k-1} and we have:

$aq^{1-1} = aq^0 = a$	first term, in position $k = 1$
$aq^{2-1} = aq^1 = aq$	second term, in position $k = 2$
$aq^{3-1} = aq^2$	third term, in position $k = 3$
$aq^{4-1} = aq^3$	fourth term, in position $k = 4$
$\dots$	$\dots$

Observation 5.12: k -th term of a geometric progression

Given the first term $a \in \mathbb{R}$ and the common ratio $q \in \mathbb{R}$, let t_k denote the term in position k of the geometric progression. The recursive relation that defines the progression is:

$$t_1 = a \quad \text{and} \quad t_k = t_{k-1} q \quad \text{for every natural number } k \geq 2 \quad (5.7)$$

and the closed formula of the k -th term is:

$$t_k = a q^{k-1} \quad \text{for every natural number } k \geq 1 \quad (5.8)$$

Proof. The closed formula (5.8) is obtained by applying repeatedly the recursive relation (5.7):

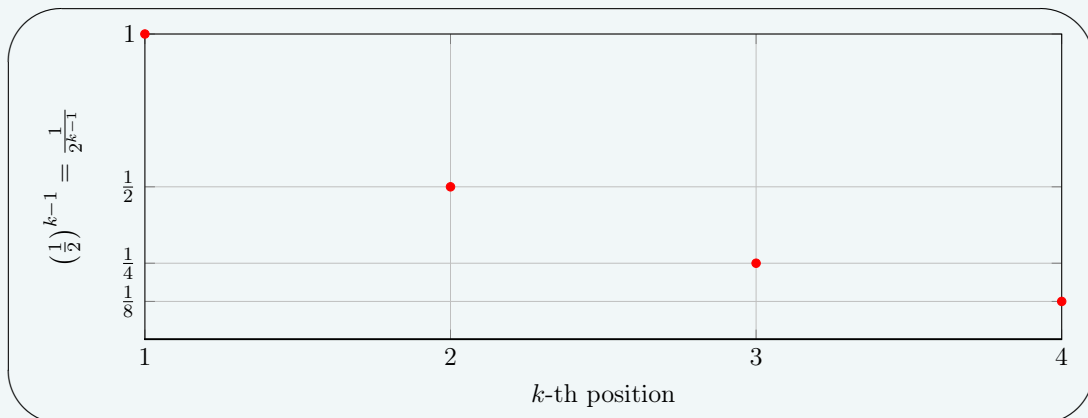
$$\begin{aligned} t_1 &= a = a q^0 \\ t_2 &= t_1 q = a q = a q^1 \\ t_3 &= t_2 q = a q q = a q^2 \\ t_4 &= t_3 q = a q^2 q = a q^3 \\ &\vdots \\ t_k &= t_{k-1} q = a q^{k-2} q = a q^{k-1} \end{aligned}$$

□

Example 5.3: geometric progressions

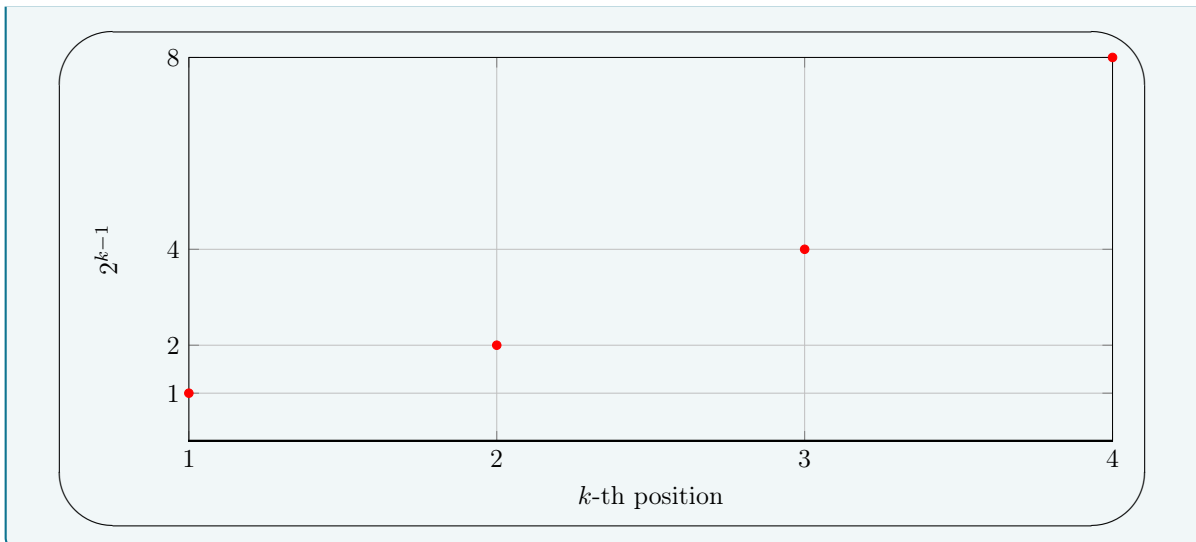
- With $a = 1$ and $q = \frac{1}{2}$, the first 4 terms are:

$$1, \quad \frac{1}{2}, \quad \frac{1}{4}, \quad \frac{1}{8}$$



- With $a = 1$ and $q = 2$, the first 4 terms are:

$$1, \quad 2, \quad 4, \quad 8$$

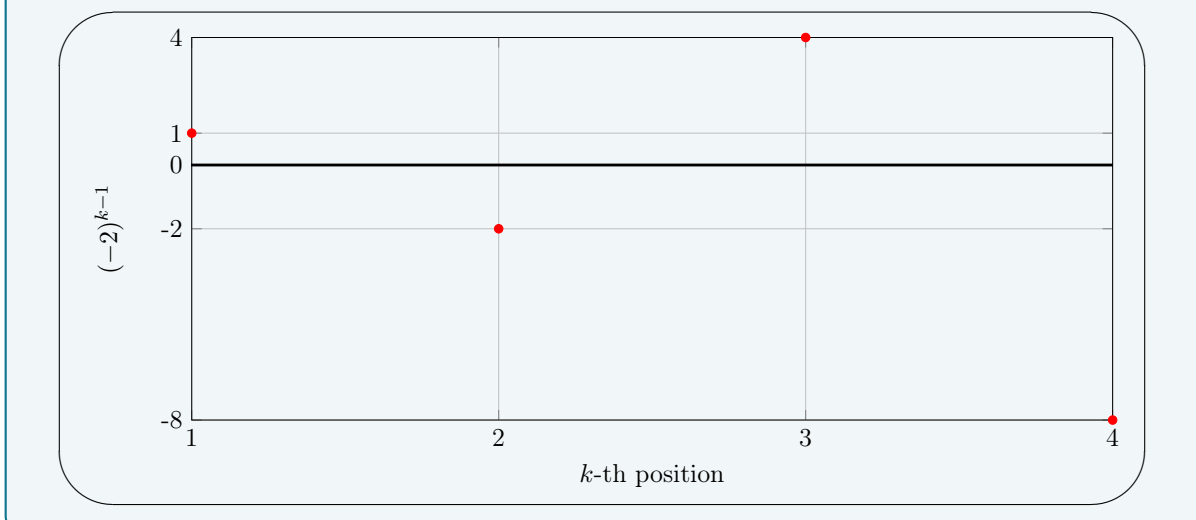


Example 5.4: geometric progression with a negative common ratio

- With $a = 1$ and $q = -2$, the first 4 terms are:

$$1, -2, 4, -8$$

With a negative common ratio the terms change sign at every step: those in odd positions are positive, those in even positions are negative.



5.2.1 Sums of the terms of geometric progressions

Observation 5.13: sum of the first n terms of the geometric progression ($a = 1$)

Given $q \in \mathbb{R}$, for every natural number $n \geq 1$ we have:

$$\sum_{k=1}^n q^{k-1} = \begin{cases} \frac{q^n - 1}{q - 1} & \text{if } q \neq 1 \\ n & \text{otherwise} \end{cases} \quad (5.9)$$

Proof. If $q \neq 1$, we prove the equation in the following equivalent form:

$$(q - 1) \sum_{k=1}^n q^{k-1} = q^n - 1$$

Applying the properties of summations, in particular the product by a constant (5.1) and the index translation (5.5), we get:

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

Dividing by $q - 1 \neq 0$ we obtain:

$$\sum_{k=1}^n q^{k-1} = \frac{q^n - 1}{q - 1}$$

If $q = 1$, we have instead:

$$\sum_{k=1}^n q^{k-1} = \sum_{k=1}^n 1 = n$$

□

- The proof never uses the sign of the common ratio: formula (5.9) holds for every $q \in \mathbb{R}$, positive, zero or negative. The only case to be made precise is $q = 0$: the first term of the summation is $q^{1-1} = 0^0$, which here is understood to be equal to 1 (convention $0^0 = 1$). With this convention, for every $n \geq 1$ we have:

$$\sum_{k=1}^n 0^{k-1} = 1 + 0 + \cdots + 0 = 1 \quad \text{and} \quad \frac{0^n - 1}{0 - 1} = \frac{-1}{-1} = 1$$

and therefore formula (5.9) remains true.

- Given $q \in \mathbb{R}, n \in \mathbb{N}, n \geq 1$ and $a \in \mathbb{R}$, formula (5.9) extends as follows:

$$\sum_{k=1}^n a q^{k-1} = \begin{cases} a \left(\frac{q^n - 1}{q - 1} \right) & \text{if } q \neq 1 \\ a n & \text{otherwise} \end{cases} \quad \text{since} \quad \sum_{k=1}^n a q^{k-1} = a \sum_{k=1}^n q^{k-1} \quad (5.10)$$

clearly:

$$\frac{q^n - 1}{q - 1} = \frac{1 - q^n}{1 - q} \quad \text{and hence we also have} \quad \sum_{k=1}^n a q^{k-1} = \begin{cases} a \left(\frac{1 - q^n}{1 - q} \right) & \text{if } q \neq 1 \\ a n & \text{otherwise} \end{cases}$$

- Formula (5.9) can also be proved by induction, in the chapter “[Principle of induction](#)” of Part 1.

Example 5.5: sum of the first n terms of geometric progressions

- With $a = 2$ and $q = \frac{1}{2}$, the first 4 terms are:

$$2, \quad 1, \quad \frac{1}{2}, \quad \frac{1}{4} \quad \text{and} \quad 2 + 1 + \frac{1}{2} + \frac{1}{4} = \frac{8 + 4 + 2 + 1}{4} = \frac{15}{4}$$

the sum of the first $n = 4$ terms is given by the formula:

$$\sum_{k=1}^4 2 \left(\frac{1}{2} \right)^{k-1} = \sum_{k=1}^4 \frac{2}{2^{k-1}} = 2 \frac{1 - \frac{1}{2^4}}{1 - \frac{1}{2}} = 2 \frac{1 - \frac{1}{16}}{\frac{1}{2}} = \frac{15}{16} \cdot 2 = \frac{15}{8}$$

- With $a = 1$ and $q = \frac{1}{2}$, the first 4 terms are:

$$1, \quad \frac{1}{2}, \quad \frac{1}{4}, \quad \frac{1}{8} \quad \text{and} \quad 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{8 + 4 + 2 + 1}{8} = \frac{15}{8}$$

the sum of the first $n = 4$ terms is given by the formula:

$$\sum_{k=1}^4 \left(\frac{1}{2} \right)^{k-1} = \sum_{k=1}^4 \frac{1}{2^{k-1}} = \frac{1 - \frac{1}{2^4}}{1 - \frac{1}{2}} = \frac{1 - \frac{1}{16}}{\frac{1}{2}} = \frac{15}{16} \cdot 2 = \frac{15}{8}$$

- With $a = 1$ and $q = 2$, the first 4 terms are:

$$1, \quad 2, \quad 4, \quad 8 \quad \text{and} \quad 1 + 2 + 4 + 8 = 15$$

the sum of the first $n = 4$ terms is given by the formula:

$$\sum_{k=1}^4 2^{k-1} = \frac{2^4 - 1}{2 - 1} = 16 - 1 = 15$$

- With $a = 1$ and $q = -2$, the first 4 terms are:

$$1, -2, 4, -8 \quad \text{and} \quad 1 - 2 + 4 - 8 = -5$$

the sum of the first $n = 4$ terms is given by the formula:

$$\sum_{k=1}^4 (-2)^{k-1} = \frac{(-2)^4 - 1}{-2 - 1} = \frac{16 - 1}{-3} = \frac{15}{-3} = -5$$

5.3 Arithmetic progressions

Definition 5.14: arithmetic progression

A sequence of real numbers is an **arithmetic progression** if the difference between each term (starting from the second one) and the previous one is constant. This constant is called the common difference of the progression.

Given the first term $a \in \mathbb{R}$ and the common difference $d \in \mathbb{R}$, the associated arithmetic progression is:

$$a, a + d, a + 2d, a + 3d, a + 4d, \dots$$

Each term (starting from the second one) is obtained from the previous one by adding d to it. The k -th term ($k \in \mathbb{N}, k \geq 1$) can be written as $a + (k - 1)d$ and we have:

$a + (1 - 1)d = a + 0d = a$	first term, in position $k = 1$
$a + (2 - 1)d = a + 1d = a + d$	second term, in position $k = 2$
$a + (3 - 1)d = a + 2d$	third term, in position $k = 3$
$a + (4 - 1)d = a + 3d$	fourth term, in position $k = 4$
$\dots$	$\dots$

- The arithmetic progression is the additive analogue of the geometric progression: in the geometric progression each term is *multiplied* by the common ratio q , in the arithmetic progression the common difference d is *added* to each term.

Observation 5.15: k -th term of an arithmetic progression

Given the first term $a \in \mathbb{R}$ and the common difference $d \in \mathbb{R}$, let t_k denote the term in position k of the arithmetic progression. The recursive relation that defines the progression is:

$$t_1 = a \quad \text{and} \quad t_k = t_{k-1} + d \quad \text{for every natural number } k \geq 2 \quad (5.11)$$

and the closed formula of the k -th term is:

$$t_k = a + (k - 1) d \quad \text{for every natural number } k \geq 1 \quad (5.12)$$

Proof. The closed formula (5.12) is obtained by repeatedly applying the recursive relation (5.11):

$$\begin{aligned} t_1 &= a = a + 0 d \\ t_2 &= t_1 + d = a + d = a + 1 d \\ t_3 &= t_2 + d = a + d + d = a + 2 d \\ t_4 &= t_3 + d = a + 2 d + d = a + 3 d \\ &\vdots \\ t_k &= t_{k-1} + d = a + (k - 2) d + d = a + (k - 1) d \end{aligned}$$

□

- The closed formula (5.12) can also be obtained by the following argument, which does not need the dots. We write the recursive relation (5.11) in the form $t_j - t_{j-1} = d$ for every position j from 2 to k :

$$\begin{aligned} t_2 - t_1 &= d \\ t_3 - t_2 &= d \\ t_4 - t_3 &= d \\ &\vdots \\ t_k - t_{k-1} &= d \end{aligned}$$

We now add all these $k - 1$ equalities. On the right-hand side we add d to itself $k - 1$ times and we obtain $d(k - 1)$. On the left-hand side the sum is **telescopic**: every term appears twice, once with the sign $+$ and once with the sign $-$, except t_k and t_1 :

$$\underbrace{(t_2 - t_1) + (t_3 - t_2) + (t_4 - t_3) + \cdots + (t_k - t_{k-1})}_{=\sum_{j=2}^k (t_j - t_{j-1})} = t_k - t_1 = t_k - a$$

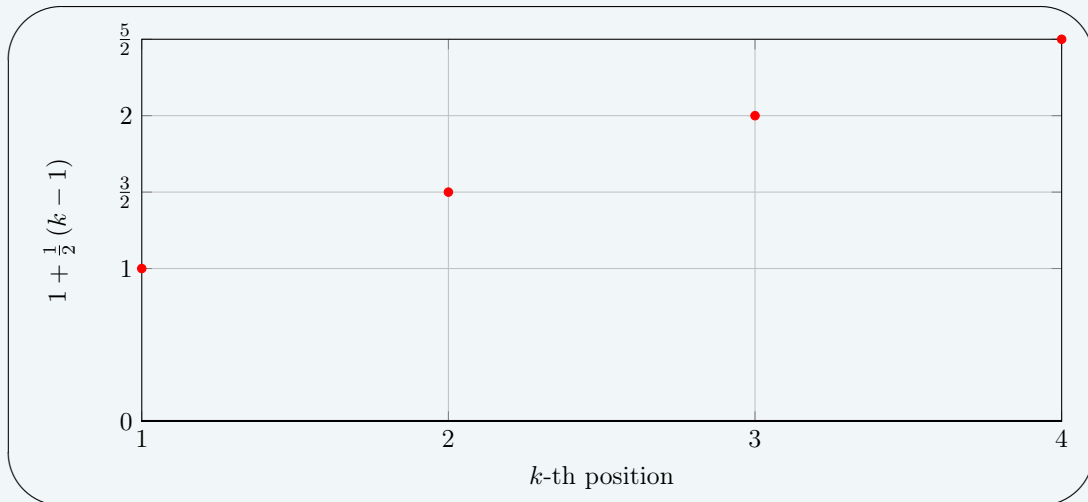
Equating the two sides we get $t_k - a = d(k - 1)$, that is, the closed formula:

$$t_k = a + (k - 1) d \quad \text{for every natural number } k \geq 1$$

Example 5.6: arithmetic progressions

- With $a = 1$ and $d = \frac{1}{2}$, the first 4 terms are:

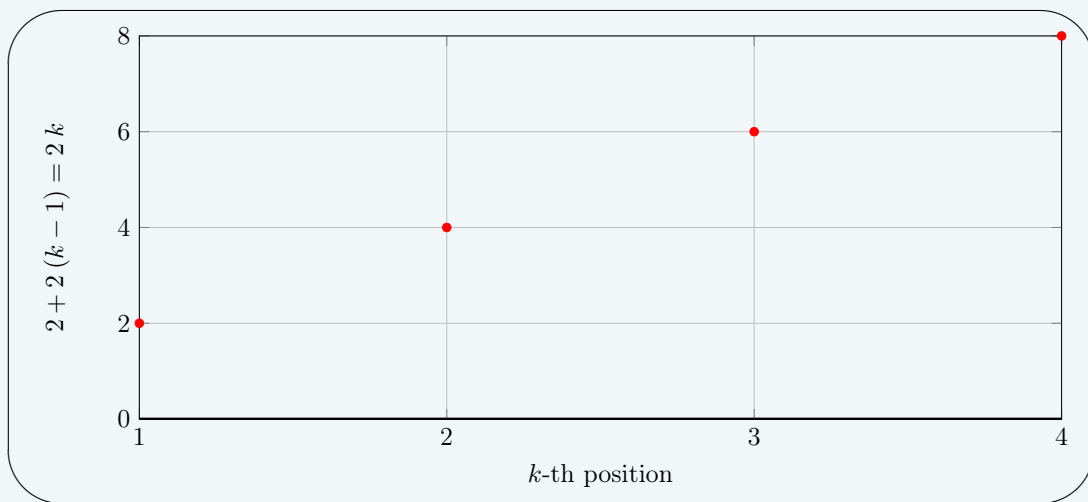
$$1, \frac{3}{2}, 2, \frac{5}{2}$$



- With $a = 2$ and $d = 2$, the first 4 terms are:

$$2, 4, 6, 8$$

that is, the first 4 even numbers (without zero).



5.3.1 Sums of the terms of arithmetic progressions

Observation 5.16: sum of the first n terms of the arithmetic progression ($a = 0$)

Given $d \in \mathbb{R}$, for every natural number $n \geq 1$ we have:

$$\sum_{k=1}^n d(k-1) = d \left(\frac{n(n-1)}{2} \right) \quad (5.13)$$

Proof. Using the product by a constant (5.1) and the index reflection (5.6), with the same argument used for the sum of the first n natural numbers, we have:

$$\begin{aligned} \sum_{k=1}^n d(k-1) &= \frac{d}{2} \left(\sum_{k=1}^n (k-1) + \sum_{k=1}^n (k-1) \right) = \frac{d}{2} \left(\sum_{k=1}^n (k-1) + \sum_{k=1}^n (n-k) \right) \\ &= \frac{d}{2} \sum_{k=1}^n (k-1+n-k) = \frac{d}{2} \sum_{k=1}^n (n-1) = d \left(\frac{n(n-1)}{2} \right) \end{aligned}$$

where in the second step the index of the second summation has been reflected: the term in position k is $a_k = k-1$ and therefore $a_{n-k+1} = (n-k+1)-1 = n-k$. $\square$

- Given $d \in \mathbb{R}, n \in \mathbb{N}, n \geq 1$ and $a \in \mathbb{R}$, formula (5.13) extends as follows:

$$\begin{aligned} \sum_{k=1}^n (a + d(k-1)) &= a n + d \left(\frac{n(n-1)}{2} \right) \quad (5.14) \\ \text{since } \sum_{k=1}^n (a + d(k-1)) &= \underbrace{\sum_{k=1}^n a}_{=a n} + \sum_{k=1}^n d(k-1) \end{aligned}$$

where the union of summations (5.3) and the summation with a constant term (5.2) have been used. This formula computes the sum of the first n terms of the arithmetic progression with a generic first term a .

- The same proof by index reflection, applied directly to the summation of the first n terms, gives the equivalent form:

$$\sum_{k=1}^n (a + d(k-1)) = \frac{n}{2} \left(\underbrace{a}_{=t_1} + \underbrace{a + (n-1)d}_{=t_n} \right) = n \left(\frac{t_1 + t_n}{2} \right)$$

since, reflecting the index, the term in position k and the term in position $n-k+1$ have constant sum $t_k + t_{n-k+1} = 2a + (n-1)d = t_1 + t_n$. The sum of the first n terms of an arithmetic progression is therefore n times the average of the first and the last term.

- With $a = 1$ and $d = 1$ the arithmetic progression is $1, 2, 3, \dots$ and formula (5.14) gives $n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2}$, that is, the sum of the first n natural numbers (without zero) already seen in the section “Some important summations”. With $a = 1$ and $d = 2$ the progression is $1, 3, 5, \dots$ and the formula gives $n + 2 \frac{n(n-1)}{2} = n^2$, that is, the sum of the first n odd numbers.

Example 5.7: sum of the first n terms of arithmetic progressions

- With $a = 1$ and $d = \frac{1}{2}$, the first 4 terms are:

$$1, \frac{3}{2}, 2, \frac{5}{2} \quad \text{and} \quad 1 + \frac{3}{2} + 2 + \frac{5}{2} = \frac{2+3+4+5}{2} = \frac{14}{2} = 7$$

the sum of the first $n = 4$ terms is given by the formula:

$$\sum_{k=1}^4 \left(1 + \frac{1}{2} (k-1) \right) = 1 \cdot 4 + \frac{1}{2} \left(\frac{4 \cdot 3}{2} \right) = 4 + 3 = 7$$

- With $a = 2$ and $d = 2$, the first 4 terms are:

$$2, 4, 6, 8 \quad \text{and} \quad 2 + 4 + 6 + 8 = 20$$

the sum of the first $n = 4$ terms is given by the formula:

$$\sum_{k=1}^4 (2 + 2(k-1)) = 2 \cdot 4 + 2 \left(\frac{4 \cdot 3}{2} \right) = 8 + 12 = 20$$

and with the equivalent form: $4 \left(\frac{2+8}{2} \right) = 4 \cdot 5 = 20$.

- With $a = 1$ and $d = -\frac{1}{2}$, the first 4 terms are:

$$1, \frac{1}{2}, 0, -\frac{1}{2} \quad \text{and} \quad 1 + \frac{1}{2} + 0 - \frac{1}{2} = 1$$

the sum of the first $n = 4$ terms is given by the formula:

$$\sum_{k=1}^4 \left(1 - \frac{1}{2} (k-1) \right) = 1 \cdot 4 - \frac{1}{2} \left(\frac{4 \cdot 3}{2} \right) = 4 - 3 = 1$$

Factorials, binomial coefficients and triangle inequality

6.1 Factorials

Definition 6.1: factorial of n

The factorial of n is the product of the first n positive integers. It is denoted by $n!$ and is read “ n factorial”. In formulas:

$$n! = \prod_{k=1}^n k = 1 \cdot 2 \cdot 3 \cdot \dots \cdot (n-1) \cdot n$$

By definition, we set $0! = 1$.

- The number $n!$ grows very rapidly as n increases. The first values are:

n	0	1	2	3	4	5	6	7	8	9	10
$n!$	1	1	2	6	24	120	720	5,040	40,320	362,880	3,628,800

- Some properties of the factorial, which are immediate to verify, are:

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

$$\frac{n!}{(n-k)!} = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-k+1), \quad \text{with } 1 \leq k \leq n \quad (6.1)$$

With $1 \leq k \leq n$, it becomes the product of k factors, starting from n and decreasing by one unit at a time.

Example 6.1: Computing the factorial

$$\frac{100!}{95!} = \frac{100!}{(100-5)!} = 100 \cdot 99 \cdot 98 \cdot 97 \cdot 96 = 9,034,502,400$$

It is always convenient to simplify expressions containing the factorial as much as possible, before computing them!

6.2 Binomial coefficients

Definition 6.2: binomial coefficient

The **binomial coefficient** is defined as the number:

$$c_{n,k} = \frac{n!}{k! (n-k)!} \quad \text{with } 0 \leq k \leq n. \quad (6.2)$$

The binomial coefficient $c_{n,k}$ is usually denoted by the symbol: $\binom{n}{k}$ which is read “ n choose k ”.

- Given rule (6.1), we have:

$$c_{n,k} = \frac{n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-k+1)}{k!} \quad (6.3)$$

with $k \geq 1$, an expression that is more convenient for computing the binomial coefficient.

We have:

$$\binom{n}{n-k} = \binom{n}{k} \quad (6.4)$$

Since:

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

We have:

$$\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k} \quad \text{for every } n \geq 2 \text{ and } 1 \leq k \leq n-1 \quad (6.5)$$

(the proof below needs $(k-1)!$ and $(n-k-1)!$, hence $k \geq 1$ and $k \leq n-1$). Since:

$$\begin{aligned}
 \binom{n-1}{k-1} + \binom{n-1}{k} &= \frac{(n-1)!}{(k-1)! \underbrace{(n-1-(k-1))!}_{=(n-k)!(n-k-1)!}} + \frac{(n-1)!}{k!(n-k-1)!} \\
 &= \frac{(n-1)!}{(k-1)!(n-k)(n-k-1)!} + \frac{(n-1)!}{k(k-1)!(n-k-1)!} \\
 &= \frac{k(n-1)! + (n-k)(n-1)!}{k(k-1)!(n-k)(n-k-1)!} = \frac{\overbrace{(n-1)!}^{=n!} n}{\underbrace{k(k-1)!}_{k!} \underbrace{(n-k)(n-k-1)!}_{(n-k)!}} \\
 &= \frac{n!}{k!(n-k)!} = \binom{n}{k}
 \end{aligned}$$

It also follows that:

$$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k} \quad \text{for every } n \geq 1 \text{ and } 1 \leq k \leq n$$

6.2.1 Newton's formula

The n -th power of a binomial $(a+b)$ can be computed with the following formula (from which the name binomial coefficient derives):

Observation 6.3: Newton's formula

For every integer $n \geq 0$, with $a, b \in \mathbb{R}$, we have:

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k \quad (6.6)$$

Proof. By induction on n .

- **Base case of the induction**

Let $n = 0$. Then the statement becomes: $(a+b)^0 = \binom{0}{0} a^0 b^0$ i.e. $1 = 1$ which is clearly true (with the convention $0^0 = 1$, which also covers the case $a = b = 0$).

- **Inductive step**

Suppose it is true for n , and let us prove it for $(n+1)$. By the inductive hypothesis,

we have: $(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k$. Then:

$$\begin{aligned}
 (a + b)^{n+1} &= (a + b) \cdot (a + b)^n = (a + b) \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k \\
 &= \sum_{k=0}^n \binom{n}{k} a^{n-k+1} b^k + \underbrace{\sum_{k=0}^n \binom{n}{k} a^{n-k} b^{k+1}}_{= \sum_{k=1}^{n+1} \binom{n}{k-1} a^{n-k+1} b^k} \\
 &= a^{n+1} + \sum_{k=1}^n \binom{n}{k} a^{n-k+1} b^k + b^{n+1} + \sum_{k=1}^n \binom{n}{k-1} a^{n-k+1} b^k \\
 &= a^{n+1} + b^{n+1} + \sum_{k=1}^n \underbrace{\left(\binom{n}{k} + \binom{n}{k-1} \right)}_{= \binom{n+1}{k}} a^{n-k+1} b^k \\
 &= \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k} b^k
 \end{aligned}$$

which is exactly the desired statement, for $n + 1$. □

Observation 6.4

For every integer $n \geq 0$ we have:

$$\sum_{k=0}^n \binom{n}{k} = 2^n \quad (6.7)$$

Proof. We write:

$$2^n = (1 + 1)^n$$

We now apply Newton's formula (6.6):

$$(1 + 1)^n = \sum_{k=0}^n \binom{n}{k} 1^{n-k} 1^k = \sum_{k=0}^n \binom{n}{k}$$

□

6.2.2 Recursive computation of binomial coefficients

Relation (6.5) allows us to compute the binomial coefficients $\binom{n}{k}$ by means of the so-called **Pascal's triangle** (or Tartaglia's triangle).

The rules for building the triangle are:

1. At the top of the triangle we place the number $\binom{0}{0} = 1$ (by definition).
2. On the sides we place the numbers $\binom{n}{0} = \binom{n}{n} = 1$ for every $n \geq 1$.
3. For $0 < k < n$, the number $\binom{n}{k}$ is written at the intersection of the n -th row and the k -th column.
4. Again for $0 < k < n$, the number $\binom{n}{k}$ is the sum of the two numbers located in the previous row, the one in the same column and the one in the previous column.

Example 6.2: Pascal's triangle (or Tartaglia's triangle)

Pascal's triangle with $n \leq 10$ and $k \leq 10$ is:

	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$	$k = 8$	$k = 9$	$k = 10$
$n = 0$	1										
$n = 1$	1	1									
$n = 2$	1	2	1								
$n = 3$	1	3	3	1							
$n = 4$	1	4	6	4	1						
$n = 5$	1	5	10	10	5	1					
$n = 6$	1	6	15	20	15	6	1				
$n = 7$	1	7	21	35	35	21	7	1			
$n = 8$	1	8	28	56	70	56	28	8	1		
$n = 9$	1	9	36	84	126	126	84	36	9	1	
$n = 10$	1	10	45	120	210	252	210	120	45	10	1

To compute the binomial coefficient $\binom{5}{3}$, corresponding to the blue cell, we can use relation (6.5) and add the two binomial coefficients $\binom{4}{2}$ and $\binom{4}{3}$, corresponding to the red cells:

$$\binom{5}{3} = \binom{4}{2} + \binom{4}{3} = 6 + 4 = 10.$$

Example 6.3: computing the power of a binomial using Pascal's triangle

Thanks to the previous Pascal's triangle and to Newton's formula (6.6) we can compute:

$$(a + b)^5 = a^5 + 5 a^4 b + 10 a^3 b^2 + 10 a^2 b^3 + 5 a b^4 + b^5$$

6.3 Absolute value

Definition 6.5: absolute value

The **absolute value** of a real number $a \in \mathbb{R}$ (or **modulus** of a) is the non-negative number defined as follows:

$$|a| = \begin{cases} a & \text{if } a \geq 0 \\ -a & \text{if } a < 0 \end{cases} \quad (6.8)$$

From the definition of absolute value it immediately follows that:

$$\forall \varepsilon \geq 0, a \in \mathbb{R}, \quad |a| \leq \varepsilon \iff -\varepsilon \leq a \leq \varepsilon \quad (6.9)$$

6.3.1 Triangle inequality in $\mathbb{R}$

Observation 6.6: triangle inequality in $\mathbb{R}$

$$|b + c| \leq |b| + |c| \quad \forall b, c \in \mathbb{R} \quad (6.10)$$

Proof. We write the two relations:

$$-|b| \leq b \leq |b|, \quad -|c| \leq c \leq |c|$$

and add them side by side:

$$-(|b| + |c|) \leq b + c \leq |b| + |c|$$

Hence, by (6.9), with $\varepsilon = |b| + |c| \geq 0$ and $a = b + c$, (6.10) follows. $\square$

- The triangle inequality is also used in the following form:

$$|d - e| \leq |d - f| + |e - f| \quad \forall d, e, f \in \mathbb{R} \quad (6.11)$$

To obtain it, it suffices to set in (6.10):

$$b = d - f, \quad c = f - e$$

we obtain:

$$|d - f + f - e| = |d - e| \leq |d - f| + |f - e| = |d - f| + |e - f|$$

since:

$$|f - e| = |e - f| \quad \forall e, f \in \mathbb{R}$$

- Moreover, the triangle inequality can also be written in the following form:

$$|g| \leq |g - h| + |h| \quad \text{i.e.} \quad |g| - |h| \leq |g - h| \quad \forall g, h \in \mathbb{R} \quad (6.12)$$

To obtain it, it suffices to set in (6.10):

$$b = g - h, \quad c = h$$

Observation 6.7: reverse triangle inequality in $\mathbb{R}$

$$||g| - |h|| \leq |g - h|, \quad \forall g, h \in \mathbb{R}. \quad (6.13)$$

Proof. From (6.12) we have

$$|g| - |h| \leq |g - h| \quad \forall g, h \in \mathbb{R}$$

Similarly, swapping g and h in (6.12) we obtain:

$$|h| - |g| \leq |h - g| = |g - h| \quad \text{that is} \quad |g| - |h| \geq -|g - h|$$

hence we have

$$-(|g - h|) \leq |g| - |h| \leq |g - h| \quad \forall g, h \in \mathbb{R}$$

Hence, by (6.9), with $\varepsilon = |g - h| \geq 0$ and $a = |g| - |h|$, the reverse triangle inequality follows. $\square$

- Inequality (6.10) can easily be extended to the case of k terms:

$$\left| \sum_{i=1}^k b_i \right| \leq \sum_{i=1}^k |b_i|. \quad (6.14)$$

- The following immediate properties also hold:

$$\begin{aligned} |bc| &= |b| |c|, & |-b| &= |b| \quad \forall b, c \in \mathbb{R}, \\ \left| \frac{b}{c} \right| &= \frac{|b|}{|c|} & & \forall b, c \in \mathbb{R}, \quad c \neq 0. \end{aligned} \quad (6.15)$$

The Fibonacci sequence

7.1 The Fibonacci sequence

Definition 7.1: Fibonacci sequence

$$\begin{aligned} F_0 &= 0, \\ F_1 &= 1, \\ F_i &= F_{i-1} + F_{i-2}, \quad \text{with } i \geq 2, i \in \mathbb{N} \end{aligned} \quad (7.1)$$

- Therefore, each Fibonacci number is the sum of the two previous ones, which gives the sequence

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, \dots$$

- The Fibonacci sequence is related to the **golden ratio** ϕ and to the **conjugate of the golden ratio** $\hat{\phi}$,

Definition 7.2: golden ratio

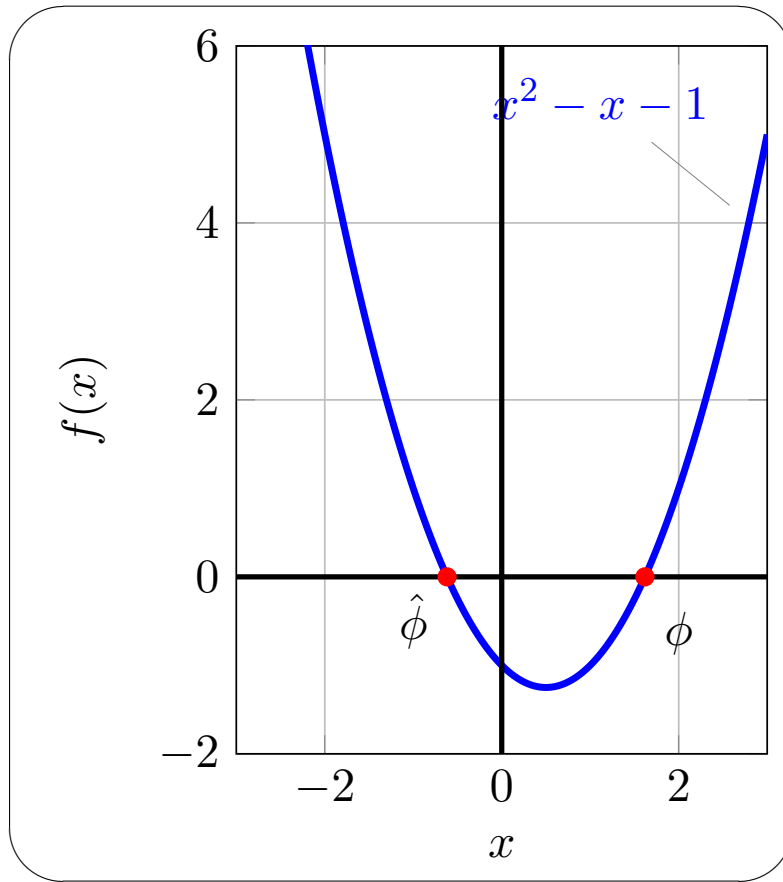
The **golden ratio** ϕ and the **conjugate of the golden ratio** $\hat{\phi}$ are the two roots of the equation:

$$x^2 = x + 1$$

and are given by the formulas:

$$\begin{aligned} \phi &= \frac{1 + \sqrt{5}}{2} & \hat{\phi} &= \frac{1 - \sqrt{5}}{2} \\ &\approx 1.61803 & &\approx -0.61803 \end{aligned}$$

The following figure shows the graph of the function $f(x) = x^2 - x - 1$ on the interval $[-3, 3]$; the red dots correspond to the roots of the equation $x^2 = x + 1$.



The golden ratio ϕ and the conjugate of the golden ratio $\hat{\phi}$ clearly satisfy the equation $x^2 = x + 1$ since:

$$\phi^2 = \left(\frac{1 + \sqrt{5}}{2} \right)^2 = \frac{1 + 2\sqrt{5} + 5}{4} = \frac{3 + \sqrt{5}}{2} = \frac{1 + \sqrt{5}}{2} + 1 = \phi + 1$$

$$\hat{\phi}^2 = \left(\frac{1 - \sqrt{5}}{2} \right)^2 = \frac{1 - 2\sqrt{5} + 5}{4} = \frac{3 - \sqrt{5}}{2} = \frac{1 - \sqrt{5}}{2} + 1 = \hat{\phi} + 1$$

Observation 7.3: closed formula of the Fibonacci sequence

$$F_i = \frac{\phi^i - \hat{\phi}^i}{\sqrt{5}}, \quad i = 0, 1, 2, \dots$$

Proof. By induction on i .

- **Base case of the induction**

We prove that the formula holds for $i = 0$ and $i = 1$:

$$F_0 = \frac{\phi^0 - \hat{\phi}^0}{\sqrt{5}} = \frac{1 - 1}{\sqrt{5}} = 0, \quad F_1 = \frac{\phi^1 - \hat{\phi}^1}{\sqrt{5}} = \frac{\sqrt{5}}{\sqrt{5}} = 1.$$

- **Inductive step**

Suppose it is true for $i = k$ and $i = k - 1$ with $k \geq 1$, and let us prove it for $i = k + 1$. By the definition of the Fibonacci sequence, we have $F_{k+1} = F_k + F_{k-1}$, while by the inductive hypothesis we have:

$$F_k = \frac{\phi^k - \hat{\phi}^k}{\sqrt{5}} \quad \text{and} \quad F_{k-1} = \frac{\phi^{k-1} - \hat{\phi}^{k-1}}{\sqrt{5}}$$

Hence we can write

$$\begin{aligned} F_{k+1} &= F_k + F_{k-1} \\ &= \frac{\phi^k - \hat{\phi}^k}{\sqrt{5}} + \frac{\phi^{k-1} - \hat{\phi}^{k-1}}{\sqrt{5}} = \frac{(\phi^k - \hat{\phi}^k) + (\phi^{k-1} - \hat{\phi}^{k-1})}{\sqrt{5}} \\ &= \frac{(\phi^k + \phi^{k-1}) - (\hat{\phi}^k + \hat{\phi}^{k-1})}{\sqrt{5}} = \frac{\phi^{k-1}(\phi + 1) - \hat{\phi}^{k-1}(\hat{\phi} + 1)}{\sqrt{5}} \\ &= \frac{\phi^{k-1}(\phi^2) - \hat{\phi}^{k-1}(\hat{\phi}^2)}{\sqrt{5}} = \frac{\phi^{k+1} - \hat{\phi}^{k+1}}{\sqrt{5}} \end{aligned}$$

which is exactly the desired statement, for $i = k + 1$. □

Observation 7.4: computing F_i by rounding

$$F_i = \left\lfloor \frac{\phi^i}{\sqrt{5}} + \frac{1}{2} \right\rfloor, \quad i = 0, 1, 2, \dots$$

Proof. Since $|\hat{\phi}| < 1$, we have

$$\frac{|\hat{\phi}^i|}{\sqrt{5}} \leq \frac{1}{\sqrt{5}} < \frac{1}{2} \quad \text{and} \quad -\frac{1}{2} < \frac{\hat{\phi}^i}{\sqrt{5}} < \frac{1}{2}, \quad i = 0, 1, 2, \dots$$

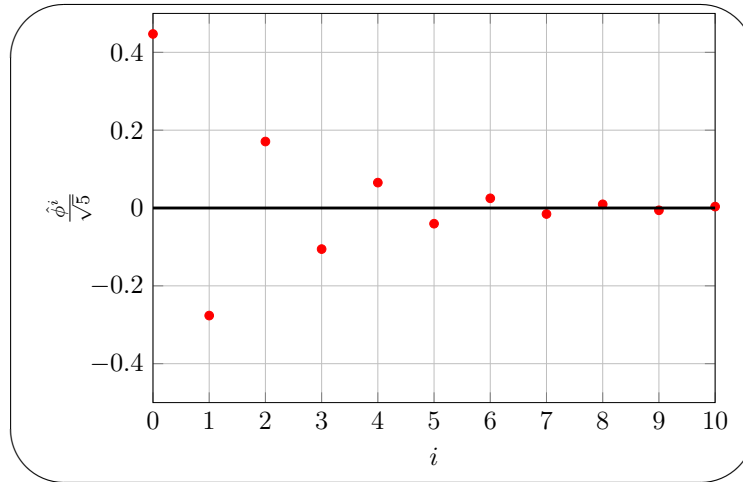
Since $F_i \in \mathbb{N}$ and $F_i = \frac{\phi^i}{\sqrt{5}} - \frac{\hat{\phi}^i}{\sqrt{5}}$, the i -th Fibonacci number F_i is equal to $\frac{\phi^i}{\sqrt{5}}$ rounded to the nearest integer:

$$F_i = \left\lfloor \frac{\phi^i}{\sqrt{5}} + \frac{1}{2} \right\rfloor, \quad i = 0, 1, 2, \dots$$



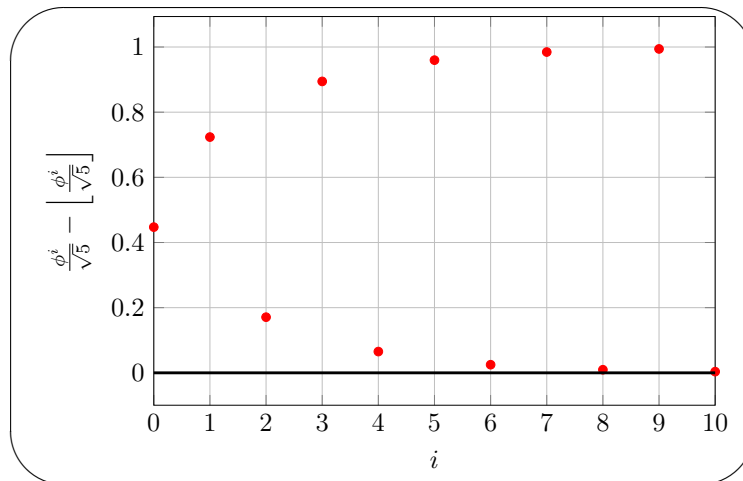
- The values of

$$\frac{\hat{\phi}^i}{\sqrt{5}} \text{ with } i = 0, 1, \dots, 10 \text{ are :}$$



- The values of

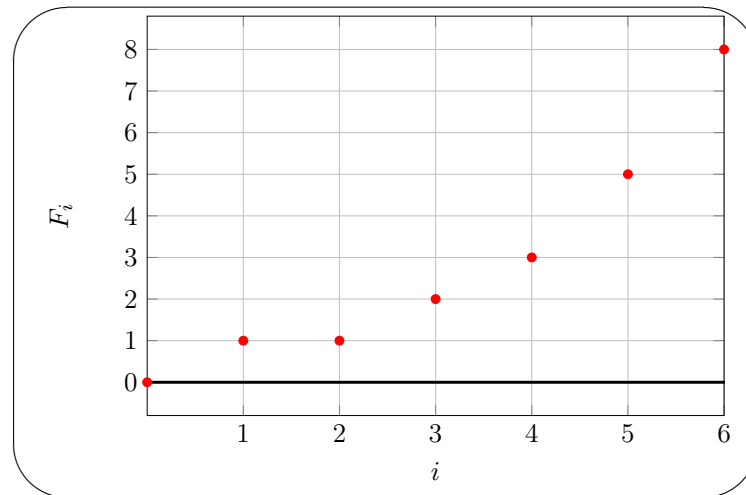
$$\frac{\phi^i}{\sqrt{5}} - \left\lfloor \frac{\phi^i}{\sqrt{5}} \right\rfloor \text{ with } i = 0, 1, \dots, 10 \text{ are :}$$



- Since $F_i \in \mathbb{N}$, we moreover have:

$$\left(\frac{\phi^i}{\sqrt{5}} - \left\lfloor \frac{\phi^i}{\sqrt{5}} \right\rfloor \right) - \frac{\hat{\phi}^i}{\sqrt{5}} \in \{0, 1\}, \quad i = 0, 1, 2, \dots$$

- The values of F_i , with $i = 0, 1, \dots, 10$, are:



Cardinality of infinite sets

8.1 Countable cardinality

Is it possible to compare the “size” of infinite sets?

- Let us approach the question starting from the important number sets that we have introduced: $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and ask ourselves: how many elements does each of these sets have?
- Intuitively, the answer seems obvious: each of these sets has infinitely many elements; however, the rational numbers are more numerous than the integers, since $\mathbb{Z} \subset \mathbb{Q}$, and for the same reason the real numbers are more numerous than the rational numbers, since $\mathbb{Q} \subset \mathbb{R}$.
- How can we state, at the same time, that two sets are both infinite, but one is more numerous than the other?

How is it possible, in general, to compare the size of infinite sets?

- To give meaning to these questions, even before being able to answer them, we must define what we mean by two sets having the same size.
- Abstracting from the experience of counting the elements of a finite set, the idea of **same size** has been identified with that of **one-to-one correspondence** (bijection):

Definition 8.1: equal cardinality of two sets

Two sets A , B are said to have **equal cardinality** (or **power**) if they can be put in one-to-one correspondence with each other, i.e., if there exists a rule that associates with each element of A one and only one element of B , and vice versa.

- The cardinality (or power) of a set expresses the intuitive idea of size

Observation 8.2

The set of integers $\mathbb{Z}$ and the set of natural numbers $\mathbb{N}$ have the same cardinality

Proof. The following rule puts the integers and the natural numbers in one-to-one correspondence:

$\mathbb{Z}$	0	1	-1	2	-2	...	n	$-n$	...
	$\updownarrow$	$\updownarrow$	$\updownarrow$	$\updownarrow$	$\updownarrow$		$\updownarrow$	$\updownarrow$	
$\mathbb{N}$	0	1	2	3	4	...	$2n-1$	$2n$	...

□

- Therefore, even though from the point of view of inclusion $\mathbb{Z}$ has “more elements” than $\mathbb{N}$ (in the sense that it has all the elements of $\mathbb{N}$ plus others), the two sets have the same cardinality.

Two sets that have the same cardinality are also called **equipotent** and should be thought of as equally numerous.

Definition 8.3: countable set

A set that has the same cardinality as $\mathbb{N}$ is called **countable**.

- The term countable indicates that the elements of the set can be enumerated, that is, arranged in a numbered list (place 1, place 2, place 3, ...)

The cardinality of $\mathbb{N}$ is called the **cardinality of the countable** (countable infinity).

Theorem 8.4

The set $\mathbb{Q}$ is countable

Proof. • We begin by proving that the set of positive rational numbers is countable.

- To do this, we represent the positive rational numbers as fractions n/m , with n, m positive integers, and we arrange these fractions in an **infinite triangular table**, in the following way:

$\frac{n}{m}$, with $n + m = \dots$				
2	$\frac{1}{1}$			
3	$\frac{1}{2}$	$\frac{2}{1}$		
4	$\frac{1}{3}$	$\frac{2}{2}$	$\frac{3}{1}$	
5	$\frac{1}{4}$	$\frac{2}{3}$	$\frac{3}{2}$	$\frac{4}{1}$
...	...			

Note that: all the fractions representing positive rational numbers appear in this table at least once; some numbers are repeated several times (e.g., $\frac{1}{1} = \frac{2}{2}$). Each row has finite length.

- Hence we can put $\mathbb{N}$ in one-to-one correspondence with the set of positive rationals, by going through the table row by row (and skipping an element when it is equal to one already encountered). For example:

1	2	3	4	5	6	7	8	9	...
$\frac{1}{1}$	$\frac{1}{2}$	$\frac{2}{1}$	$\frac{1}{3}$	$\frac{3}{1}$	$\frac{1}{4}$	$\frac{2}{3}$	$\frac{3}{2}$	$\frac{4}{1}$	...

(note that we skipped $\frac{2}{2}$ because it is equal to $\frac{1}{1}$, already encountered).

- This proves that the set of positive rationals is countable.
- Then $\mathbb{Q}$ is also countable, and this can be proved with an argument similar to the one we used to prove that $\mathbb{Z}$ is countable: denoting by $q_1, q_2, q_3, \dots$ the positive rationals, we set:

$\mathbb{N}$	0	1	2	3	4	5	6	...
	$\updownarrow$	$\updownarrow$	$\updownarrow$	$\updownarrow$	$\updownarrow$	$\updownarrow$	$\updownarrow$	
$\mathbb{Q}$	0	q_1	$-q_1$	q_2	$-q_2$	q_3	$-q_3$	...

which gives a one-to-one correspondence between $\mathbb{N}$ and $\mathbb{Q}$.

□

- Having discovered that several infinite sets, each properly contained in the next ($\mathbb{N}, \mathbb{Z}, \mathbb{Q}$), have the same cardinality, might suggest that this is true for all infinite sets. This is not true, as shown in the next section.

8.2 Cardinality of the continuum

Theorem 8.5

The set of real numbers $\mathbb{R}$ is not countable

- The proof technique is called **Cantor's diagonal argument**

Proof. • We will prove that the interval $[0, 1]$ has uncountable cardinality, from which the uncountability of $\mathbb{R}$ obviously follows

- Suppose therefore, **by contradiction**, that $[0, 1]$ is countable, and arrange **all** the real numbers of the interval $[0, 1]$ in a list $r_1, r_2, r_3, \dots$
- We write each number r_i in decimal form:

$$0.a_1a_2a_3\dots$$

where the a_i are digits from 0 to 9; if the digits are all zero we have $r_i = 0$ and if they are all nine we have $r_i = 0.\overline{9} = 1$:

$$r_1 = 0.a_{11}a_{12}a_{13}\dots$$

$$r_2 = 0.a_{21}a_{22}a_{23}\dots$$

$$r_3 = 0.a_{31}a_{32}a_{33}\dots$$

...

- We now define the following decimal number:

$$r = 0.b_1b_2b_3\dots$$

where the digits b_i are defined by the following rule:

$$b_i = \begin{cases} 5 & \text{if } a_{ii} \text{ is one of the digits } 0, 1, 2, 3 \text{ or } 4 \\ 4 & \text{if } a_{ii} \text{ is one of the digits } 5, 6, 7, 8 \text{ or } 9 \end{cases}$$

With this definition we have $b_i \neq a_{ii}$ for every i . Note that the number r has been constructed by reasoning on the diagonal of the infinite table whose rows are the numbers r_i , hence the name of the argument. □

Proof. • We now observe that r is a real number belonging to the interval $[0, 1]$ (because it is of the form $0.b_1b_2b_3\dots$ with all digits equal to 4 or 5) and on the other hand it is not equal to any of the numbers r_i in the list, since:

$\rightarrow r \neq r_1$: the first digit of r is different from the first digit of r_1 ($b_1 \neq a_{11}$);

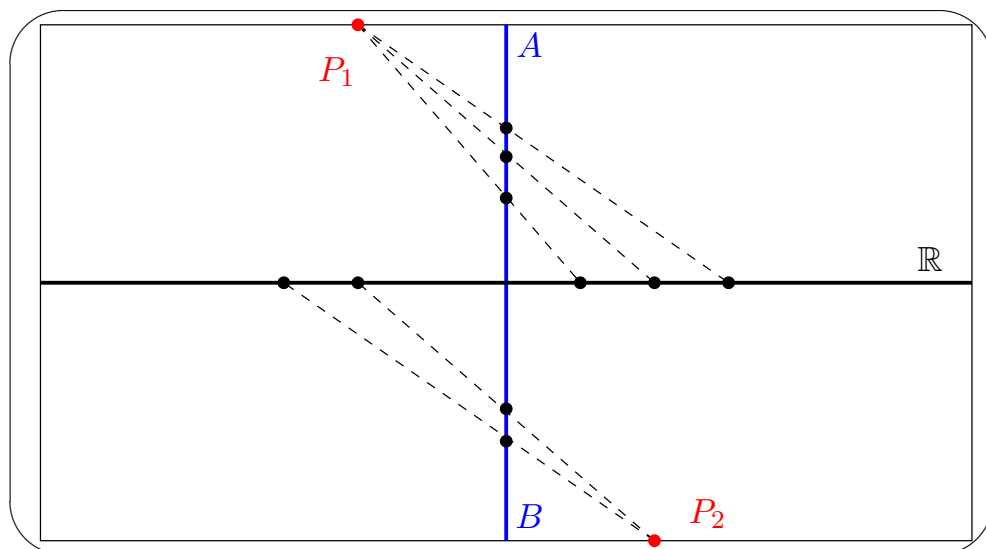
$\rightarrow r \neq r_2$: the second digit of r is different from the second digit of r_2 ($b_2 \neq a_{22}$);
 $\rightarrow \dots$

- But this leads to a **contradiction**, since we had assumed that the r_i completely exhausted the set of real numbers of the interval $[0, 1]$. Therefore $[0, 1]$ is not countable.

□

Since $\mathbb{N}$ cannot be put in one-to-one correspondence with $\mathbb{R}$, but it can be put in one-to-one correspondence with a proper subset of $\mathbb{R}$ ($\mathbb{N}$ itself!), we say that $\mathbb{N}$ has **smaller cardinality** than $\mathbb{R}$.

- We also note that the interval $[0, 1]$, and any interval of $\mathbb{R}$ (open or closed, bounded or unbounded), has the same cardinality as $\mathbb{R}$.
- This fact can be proved, for example, with an elementary geometric construction that gives a one-to-one correspondence between the points of the line and those of a segment $[A, B]$; as shown in the following figure, by drawing suitable segments from the two fixed points P_1, P_2 to the line itself:



- This implies that, for example, the points of a line have the same cardinality as the points of a segment.

The cardinality of $\mathbb{R}$ is called the **cardinality of the continuum**.

- It is not only every interval of $\mathbb{R}$ that has this same cardinality. It can be proved that the same holds for the plane, for three-dimensional space and for their “continuous” subsets: for example, a plane and a sphere both have the cardinality of the continuum.

- So far we have encountered only two hierarchical levels of infinity: the cardinality of the countable and that of the continuum. One should not believe that only these two exist! For example, the set of all subsets of $\mathbb{R}$ (its power set) is even more numerous than $\mathbb{R}$; and with this procedure we can always construct a set more numerous than a given set.
- Therefore the hierarchical levels of infinite cardinalities are themselves infinite.

Complex numbers

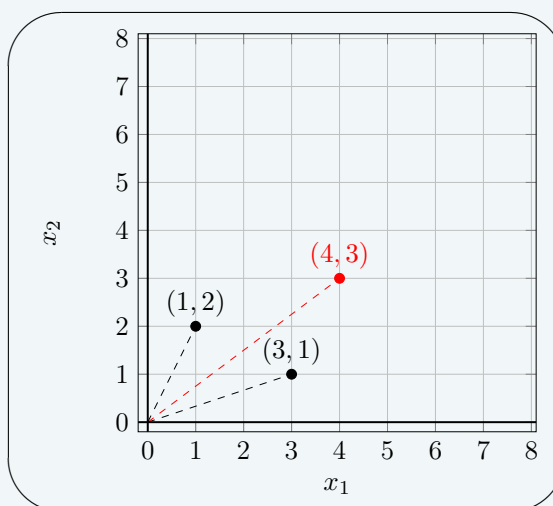
9.1 Definition of $\mathbb{C}$ and field structure

- We have denoted by $\mathbb{R}^2$ (short for $\mathbb{R} \times \mathbb{R}$) the set of ordered pairs (a, b) of real numbers.
- On these pairs we directly define the operations of sum and product with the following rules:

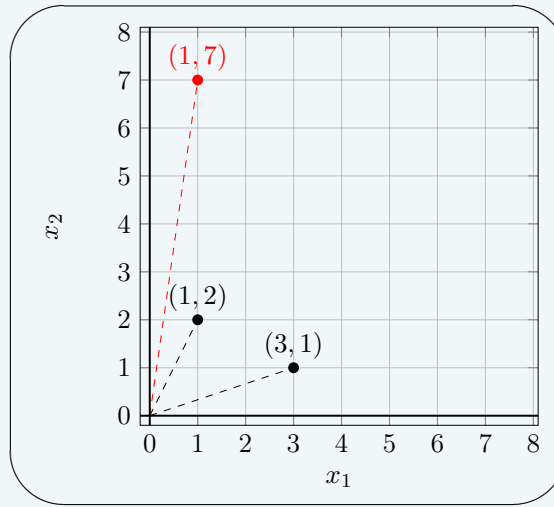
$$(a, b) + (c, d) = (a + c, b + d) \quad (9.1)$$

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc) \quad (9.2)$$

Example 9.1: sum



Example 9.2: product



- This “sum” and this “product” satisfy the **commutative**, **associative** and **distributive** properties
- We also observe that, $\forall (a, b) \in \mathbb{R}^2$:

$$(a, b) + (0, 0) = (0, 0) + (a, b) = (a, b)$$

hence the pair $(0, 0)$ is the **identity element for the sum**. Moreover:

$$(a, b) \cdot (1, 0) = (1, 0) \cdot (a, b) = (a, b)$$

hence the pair $(1, 0)$ is the **identity element for the product**. We have:

$$(a, b) + (-a, -b) = (0, 0)$$

hence $(-a, -b)$ is the **opposite** of (a, b) . Moreover, if $(a, b) \neq (0, 0)$ then:

$$(a, b) \cdot \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right) = (1, 0)$$

hence the pair $(a/(a^2 + b^2), -b/(a^2 + b^2))$ is the **reciprocal** of (a, b) .

Definition 9.1: field of complex numbers

Properties R_1, R_2 are satisfied by the sum and the product defined above, and therefore the set $\mathbb{R}^2$ with this structure is a field, which we will call the **field of complex numbers** and denote by $\mathbb{C}$

- We now observe that $\mathbb{C}$ contains the subset $\mathbb{C}_0$ of pairs of the form $(a, 0)$; it is a subfield of $\mathbb{C}$, since the sum and product of pairs of this form are again pairs of the same form; indeed we have:

$$(a, 0) + (b, 0) = (a + b, 0) \quad \text{and} \quad (a, 0) \cdot (b, 0) = (a \cdot b, 0)$$

Moreover, $\mathbb{C}_0$ can be ordered by setting $(a, 0) < (b, 0)$ if $a < b$.

- If we then put the set of real numbers $\mathbb{R}$ in one-to-one correspondence with $\mathbb{C}_0$, by setting

$$(a, 0) \longleftrightarrow a$$

we can identify the real numbers $a \in \mathbb{R}$ with the complex numbers of the form $(a, 0) \in \mathbb{R}^2$. In this sense the field of complex numbers $\mathbb{C}$ is an **extension** of the field of real numbers $\mathbb{R}$.

Let us now consider the number $(0, 1) \in \mathbb{C}$. It has the remarkable property that:

$$(0, 1) \cdot (0, 1) = (-1, 0)$$

i.e., its square coincides with the real number -1

Definition 9.2: imaginary unit

The pair $(0, 1) \in \mathbb{C}$ is denoted by the letter “ i ” and is called the **imaginary unit**

9.2 Algebraic form of complex numbers

We observe that, if we write any complex number $(c, 0)$ simply as c , we have:

$$(a, b) = (a, 0) + \underbrace{(0, 1)}_{=i} \cdot (b, 0) = a + ib$$

With this notation, rules (9.1) and (9.2) are the ordinary rules of algebraic calculation, keeping in mind that $i^2 = -1$:

$$(a + ib) + (c + id) = (a + c) + i(b + d) \quad (9.3)$$

$$(a + ib) \cdot (c + id) = (ac - bd) + i(ad + bc) \quad (9.4)$$

Definition 9.3: algebraic form, real part and imaginary part

The expression:

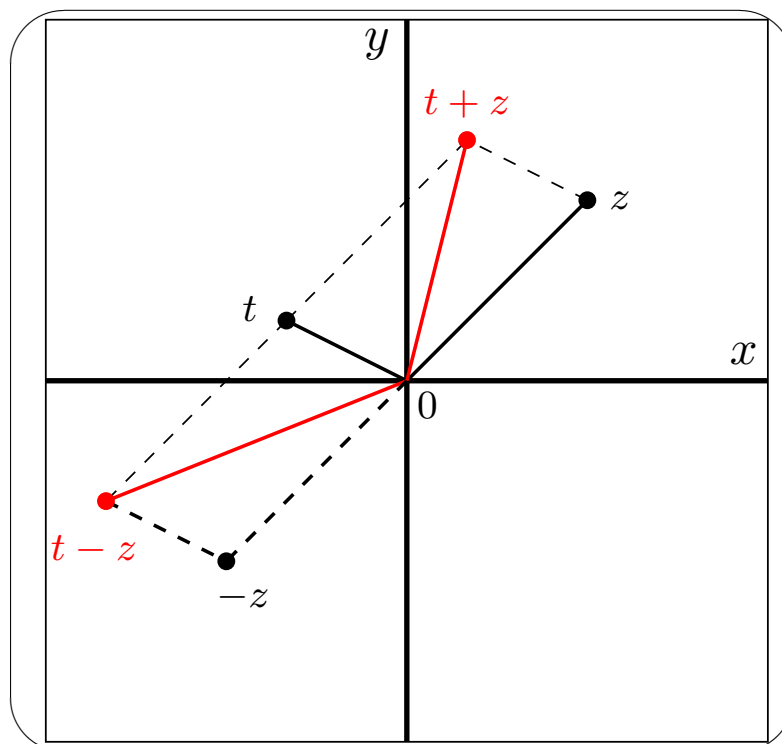
$$z = a + ib \quad (9.5)$$

is called the **algebraic form of complex numbers**; a is called the **real part** of z and is denoted by $\Re(z)$ (or $\text{Re}(z)$) while b is called the **imaginary part** and is denoted by $\Im(z)$ (or $\text{Im}(z)$).

Complex plane

- In a Cartesian plane, the complex numbers $a + ib$ can be represented as points with coordinates (a, b) . In this context:

- the plane is called the **complex plane** or **Gauss plane**
- the x , y axes are called the **real axis** and the **imaginary axis**
- the points on the real axis are the **real numbers**
- the points on the imaginary axis are the **purely imaginary numbers** (i.e., of the form ib)
- The **sum of two complex numbers** is the complex number whose coordinates are the sums of the coordinates: the geometric meaning of this fact is that the point $z + t$ is constructed from the points z , t according to the “**parallelogram rule**”, illustrated in the following figure:



Observation 9.4

The set of complex numbers $\mathbb{C}$ is not an ordered field

Proof. • We have seen that $\mathbb{C}$ satisfies the field axioms; however, it does not satisfy those of an ordered field, that is, it is not possible to define a relation $\leq$ between complex numbers in such a way that properties $R3$ hold

- It can be proved that properties $R3$ imply that the square of any number is never negative, and on the other hand, if a number is positive its opposite is negative. Now, in $\mathbb{C}$ we have:

$$1^2 = 1 \quad \text{and} \quad i^2 = -1$$

We therefore have two squares, each the opposite of the other. Neither of them, however, can be negative (because they are squares), and this is absurd (because

between a and $-a$ one must be negative, if $a \neq 0$). We conclude that $\mathbb{C}$ is not an ordered field. $\square$

Definition 9.5: conjugate

The complex number $a - ib$ is called the **complex conjugate** of $z = a + ib$ and is denoted by $\bar{z}$.

We have:

$$z + \bar{z} = 2a = (a + ib) + (a - ib) = 2a = 2 \Re(z) \quad (9.6)$$

$$z - \bar{z} = 2ib = (a + ib) - (a - ib) = 2ib = 2i \Im(z) \quad (9.7)$$

$$z \cdot \bar{z} = (a + ib) \cdot (a - ib) = a^2 - aib + iba - \underbrace{i^2}_{=-1} b^2 = a^2 + b^2 \geq 0 \quad (9.8)$$

- The conjugation operation has the following elementary properties with respect to sum and product:

$$\overline{(z_1 + z_2)} = \bar{z}_1 + \bar{z}_2 \quad (9.9)$$

$$\overline{(z_1 \cdot z_2)} = \bar{z}_1 \cdot \bar{z}_2 \quad (9.10)$$

$$\overline{\left(\frac{1}{z}\right)} = \frac{1}{\bar{z}} \quad (9.11)$$

Definition 9.6: modulus

The **modulus** of $z = a + ib$ is the non-negative real number $\sqrt{a^2 + b^2}$; it is denoted by $|z|$.

- If $z = a$ is real, its modulus is called the absolute value and is still denoted by $|a|$. The following properties hold:
 - $|z| = 0 \iff z = 0$, moreover $|z| \geq 0$
 - $|z| = |\bar{z}|$
 - $\Re(z) \leq |z|$ $\Im(z) \leq |z|$ $|z| \leq |\Re(z)| + |\Im(z)|$
 - $|z_1 + z_2| \leq |z_1| + |z_2|$ **triangle inequality**
 - $|z_1 + z_2| \geq ||z_1| - |z_2||$
- Properties a), b), c) are immediately verified.

Proof. • Let us prove properties d) and e). They are equivalent to the following:

$$(|z_1| - |z_2|)^2 \leq |z_1 + z_2|^2 \leq (|z_1| + |z_2|)^2$$

- Setting $z_1 = a + ib$, $z_2 = c + id$ we obtain:

$$\begin{aligned} |z_1 + z_2|^2 &= |(a + ib) + (c + id)|^2 \\ &= |(a + c) + i(b + d)|^2 \\ &= \left(\sqrt{(a + c)^2 + (b + d)^2} \right)^2 = (a + c)^2 + (b + d)^2 \end{aligned}$$

and hence we have:

$$\begin{aligned} \underbrace{\left(\sqrt{a^2 + b^2} - \sqrt{c^2 + d^2} \right)^2}_{=(a^2+b^2)+(c^2+d^2)-2\sqrt{a^2+b^2}\cdot\sqrt{c^2+d^2}} &\leq \underbrace{(a + c)^2 + (b + d)^2}_{=a^2+2ac+c^2+b^2+2bd+d^2} \\ &\leq \underbrace{\left(\sqrt{a^2 + b^2} + \sqrt{c^2 + d^2} \right)^2}_{=(a^2+b^2)+(c^2+d^2)+2\sqrt{a^2+b^2}\cdot\sqrt{c^2+d^2}} \end{aligned}$$

After simplification, this double inequality reduces to:

$$-\sqrt{a^2 + b^2} \cdot \sqrt{c^2 + d^2} \leq ac + bd \leq \sqrt{a^2 + b^2} \cdot \sqrt{c^2 + d^2}$$

which is equivalent to the following:

$$|ac + db| \leq \sqrt{a^2 + b^2} \cdot \sqrt{c^2 + d^2}$$

Squaring both sides we arrive at:

$$\underbrace{(ac + db)^2}_{=a^2c^2+d^2b^2+2acbd} \leq \underbrace{(a^2 + b^2) \cdot (c^2 + d^2)}_{=a^2c^2+a^2d^2+b^2c^2+b^2d^2}$$

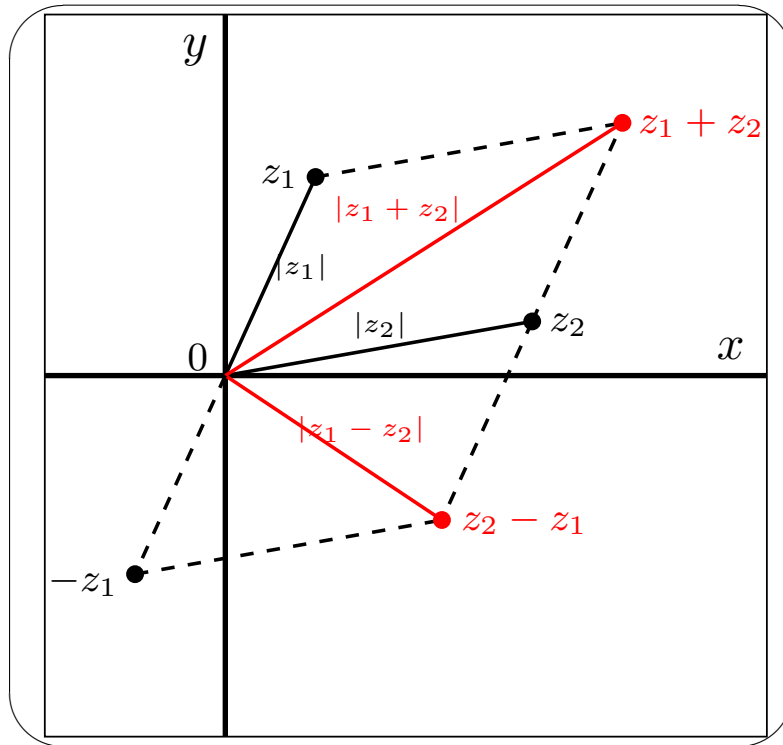
that is,

$$0 \leq -2acbd + a^2d^2 + b^2c^2 = (ad - bc)^2$$

which is true for every $a, b, c, d \in \mathbb{R}$. □

- Geometrically, $|z|$ represents the distance of the point (or complex number) z from the origin; $|z_1 - z_2|$ represents the distance between the two points z_1 and z_2 ; inequalities d) and e) express the well-known theorem on the **lengths of the sides of a triangle**¹:

¹In a non-degenerate triangle, the sum of the lengths of two sides is greater than the length of the third one.



- Using the concepts just introduced, we can write the quotient of two complex numbers in algebraic form:

$$\frac{a + ib}{c + id}$$

it suffices to multiply the numerator and the denominator by $c - id$, and we obtain:

$$\frac{a + ib}{c + id} = \frac{(a + ib)(c - id)}{\underbrace{(c + id)(c - id)}_{=c^2+d^2}} = \frac{ac - aid + ibc - i^2bd}{c^2 + d^2} = \frac{(ac + bd)}{c^2 + d^2} + i \frac{(bc - ad)}{c^2 + d^2}$$

9.2.1 Equations in the complex field

- Let us see how to solve an equation in the complex field, when it involves the unknown $z = x + iy$ also through $\Re(z)$, $\Im(z)$, z , $|z|$.
- We illustrate, with the following example, the procedure of transforming the equation in one complex unknown into a system of two equations in two real unknowns. The method consists in taking the real and imaginary parts of the equation.

Example 9.3: equations in the complex field (algebraic method)

- We want to solve:

$$z^2 + i \Im z + 2 \bar{z} = 0$$

We set $z = x + iy$, with x, y real unknowns, and rewrite the equation:

$$\begin{cases} z^2 = (x + iy)^2 = x^2 - y^2 + 2ixy \\ i \Im z = iy \\ 2\bar{z} = 2(x - iy) = 2x - 2iy \end{cases}$$

Substituting, we therefore obtain:

$$(x^2 - y^2 + 2ixy) + (iy) + (2x - 2iy) = 0$$

- Now, a complex number is zero if and only if its real part and imaginary part are zero. Therefore we separate the real part and the imaginary part of the left-hand side:

$$(x^2 - y^2 + 2x) + i(2xy + y - 2y) = 0$$

and set both equal to zero:

$$\begin{cases} x^2 - y^2 + 2x = 0 \\ 2xy - y = 0 \end{cases}$$

We have thus transformed the equation in one complex unknown into a system of two equations in two real unknowns.

- We solve the system. The second equation gives:

$$y = 0 \quad \text{or} \quad x = \frac{1}{2}$$

1. For $y = 0$ the first equation becomes

$$x^2 + 2x = 0 \quad \text{which gives} \quad x = 0 \quad \text{or} \quad x = -2$$

2. For $x = \frac{1}{2}$ the first equation becomes

$$-y^2 + \frac{5}{4} = 0 \quad \text{which gives} \quad y = \pm \frac{\sqrt{5}}{2}$$

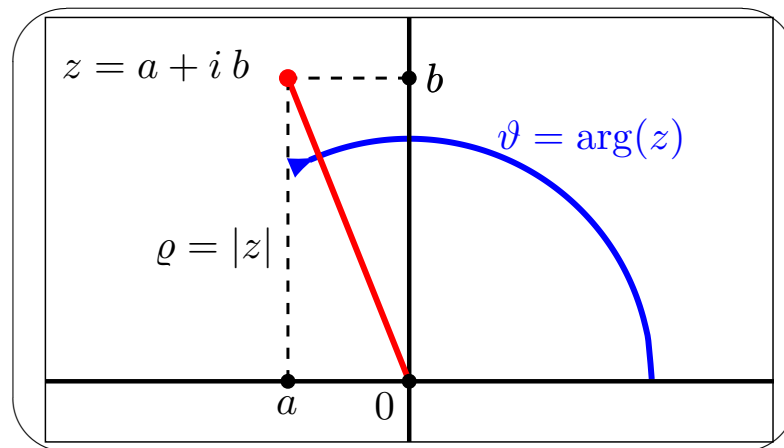
- Hence the equation has the following 4 solutions:

$$z = 0, \quad z = -2, \quad z = \frac{1}{2} + i \frac{\sqrt{5}}{2}, \quad z = \frac{1}{2} - i \frac{\sqrt{5}}{2}$$

The method seen in this example can in principle be applied to any equation in $\mathbb{C}$, but a generic system of two equations in two unknowns is almost always impossible to solve algebraically.

9.3 Trigonometric form of complex numbers

- As is known from Geometry, the points of the plane can be identified not only by their Cartesian coordinates, but also by their **polar coordinates**:
 - $\varrho \rightarrow$ **polar radius**, i.e., the distance of the point from the origin
 - $\vartheta \rightarrow$ **polar angle**, i.e., the angle that the line joining the point to the origin forms with the positive x -axis, measured counterclockwise.
- Clearly, a pair ϱ, ϑ , with $\varrho > 0$, identifies a well-defined point of the plane; conversely, a point of the plane uniquely determines the coordinate ϱ , but the angle ϑ , measured in radians, is determined only up to multiples of 2π .



- Given a complex number z , its **modulus** $|z|$ coincides with the **polar radius** ϱ of the point representing it in the complex plane.
- We call **argument** of z , and denote by $\arg(z)$, any of the angles ϑ associated with the point z . In this way the argument of z is not uniquely determined. Often this indeterminacy does not cause any problem. At other times, however, it is preferable to assign a well-defined argument to a complex number. This can be done in infinitely many ways, by fixing any interval of length 2π within which the angle ϑ is allowed to vary
- The intervals most commonly used for this purpose are $[0, 2\pi)$ and $(-\pi, \pi]$; the argument of z is then called the **principal argument**.

Example 9.4: arguments of complex numbers

- the number $-i$ has argument $-\pi/2$ or $3\pi/2$ or any other value of the form $-\pi/2 + 2k\pi$ with $k \in \mathbb{Z}$. Its principal argument will be $3\pi/2$ if we adopt the convention $\vartheta \in [0, 2\pi)$, and $-\pi/2$ with the convention $\vartheta \in (-\pi, \pi]$.
- Positive real numbers have principal argument 0 and negative ones π , with both conventions.

Given the number $z = a + ib$, from trigonometry we immediately obtain the relations between the Cartesian coordinates a, b and the polar ones ϱ, ϑ :

$$a = \varrho \cos \vartheta \quad \text{and} \quad b = \varrho \sin \vartheta \quad (9.12)$$

The inverse relations are:

$$\varrho = \sqrt{a^2 + b^2}, \quad \cos \vartheta = \frac{a}{\sqrt{a^2 + b^2}} \quad \text{and} \quad \sin \vartheta = \frac{b}{\sqrt{a^2 + b^2}} \quad (9.13)$$

Definition 9.7: trigonometric form

A complex number $z = a + ib$ can also be written in the form

$$z = \varrho (\cos \vartheta + i \sin \vartheta) \quad (9.14)$$

which is called the **trigonometric form** of complex numbers.

Example 9.5: trigonometric form

Let us write the following complex number in trigonometric form:

$$z = \sqrt{3} + i$$

We have $\varrho = \sqrt{a^2 + b^2} = \sqrt{3 + 1} = 2$, hence:

$$\sqrt{3} + i = 2 \left(\frac{\sqrt{3}}{2} + i \frac{1}{2} \right) = 2 \left(\cos \frac{\pi}{6} + i \sin \frac{\pi}{6} \right)$$

9.3.1 De Moivre's formulas

- The trigonometric form is convenient for expressing products and quotients of complex numbers. Indeed, if we have:

$$z_1 = \varrho_1 (\cos \vartheta_1 + i \sin \vartheta_1) \quad z_2 = \varrho_2 (\cos \vartheta_2 + i \sin \vartheta_2)$$

for the **product** we obtain:

$$\begin{aligned} z_1 z_2 &= \varrho_1 \varrho_2 \cdot \left\{ \cos \vartheta_1 \cos \vartheta_2 - \sin \vartheta_1 \sin \vartheta_2 + i (\sin \vartheta_1 \cos \vartheta_2 + \cos \vartheta_1 \sin \vartheta_2) \right\} \\ &= \varrho_1 \varrho_2 \cdot \left\{ \cos (\vartheta_1 + \vartheta_2) + i \sin (\vartheta_1 + \vartheta_2) \right\} \end{aligned} \quad (9.15)$$

for the **quotient**, if $z_2 \neq 0$, we have:

$$\frac{z_1}{z_2} = \frac{\varrho_1}{\varrho_2} \cdot \frac{\cos \vartheta_1 + i \sin \vartheta_1}{\cos \vartheta_2 + i \sin \vartheta_2},$$

multiplying numerator and denominator by $(\cos \vartheta_2 - i \sin \vartheta_2)$ and, taking into account that $(\cos \vartheta_2)^2 + (\sin \vartheta_2)^2 = 1$, we obtain:

$$\begin{aligned}\frac{z_1}{z_2} &= \frac{\varrho_1}{\varrho_2} \cdot \left\{ (\cos \vartheta_1 + i \sin \vartheta_1) \cdot (\cos \vartheta_2 - i \sin \vartheta_2) \right\} \\ &= \frac{\varrho_1}{\varrho_2} \cdot \left\{ \cos(\vartheta_1 - \vartheta_2) + i \sin(\vartheta_1 - \vartheta_2) \right\}\end{aligned}\quad (9.16)$$

- Therefore the modulus of the product and of the quotient of two complex numbers is, respectively, the product and the quotient of the moduli; the argument is, respectively, the sum and the difference of the arguments:

$$|z_1 \cdot z_2| = |z_1| \cdot |z_2| \quad \text{and} \quad \arg(z_1 \cdot z_2) = \arg(z_1) + \arg(z_2) \quad (9.17)$$

$$\left| \frac{z_1}{z_2} \right| = \frac{|z_1|}{|z_2|} \quad \text{and} \quad \arg\left(\frac{z_1}{z_2}\right) = \arg(z_1) - \arg(z_2) \quad (9.18)$$

- Formula (9.15) generalizes to the case of any number of factors $z_1, z_2, \dots, z_n$:

$$z_1 z_2 \dots z_n = \varrho_1 \varrho_2 \dots \varrho_n \cdot \left\{ \cos(\vartheta_1 + \vartheta_2 + \dots + \vartheta_n) + i \sin(\vartheta_1 + \vartheta_2 + \dots + \vartheta_n) \right\} \quad (9.19)$$

If, moreover, all the factors are equal, we obtain:

$$z^n = \varrho^n \cdot \left\{ \cos(n \vartheta) + i \sin(n \vartheta) \right\} \quad (9.20)$$

- These relations on products and quotients of complex numbers are known as De Moivre's formulas.

Example 9.6: powers of complex numbers with De Moivre's formulas

Write in algebraic form:

$$z = (1 + i)^7$$

We determine the modulus and argument of $(1 + i)$ and then apply De Moivre's formula.

$$|1 + i| = \sqrt{2} \quad \text{and} \quad \arg(1 + i) = \frac{\pi}{4} \quad \left(\cos \vartheta = \frac{1}{\sqrt{2}}, \quad \sin \vartheta = \frac{1}{\sqrt{2}} \right)$$

Hence we have:

$$|(1 + i)^7| = (\sqrt{2})^7 = 2^{\frac{7}{2}} = 2^{3+\frac{1}{2}} = 8\sqrt{2} \quad \text{and} \quad \arg(1 + i)^7 = \frac{7}{4} \pi$$

Consequently:

$$(1 + i)^7 = 8\sqrt{2} \left(\underbrace{\cos \frac{7}{4} \pi}_{=\frac{1}{\sqrt{2}}} + i \underbrace{\sin \frac{7}{4} \pi}_{=-\frac{1}{\sqrt{2}}} \right) = 8\sqrt{2} \left(\frac{1}{\sqrt{2}} - i \frac{1}{\sqrt{2}} \right) = 8 - 8i$$

- De Moivre's formulas allow us to give a geometric interpretation of the product of complex numbers.
- Let z , to begin with, be a complex number with modulus 1, hence of the form $(\cos \vartheta + i \sin \vartheta)$. Then, multiplying a number by z means adding ϑ to its argument, i.e., performing a **rotation by the angle ϑ** .
- If z has modulus ϱ instead of 1, in addition to a rotation we perform a **dilation by the factor ϱ** .

Example 9.7: geometric interpretation of the product of complex numbers

- multiplying by i means performing a rotation by $\frac{\pi}{2}$;
- multiplying by -1 means performing a rotation by π ;
- multiplying by $(1 + i)$ means performing a dilation by the factor $\sqrt{2}$ and a rotation by $\frac{\pi}{4}$

Example 9.8: equations in the complex field (trigonometric method)

- We want to solve:

$$z^3 - |z| = 0$$

We rewrite the equation in the form

$$z^3 = |z|$$

and we set $z = \varrho (\cos \vartheta + i \sin \vartheta)$ (trigonometric form) and rewrite the equation:

$$\begin{cases} z^3 = \varrho^3 (\cos 3\vartheta + i \sin 3\vartheta) \\ |z| = \varrho \end{cases}$$

Substituting, we therefore obtain:

$$\varrho^3 (\cos 3\vartheta + i \sin 3\vartheta) = \varrho$$

The equation is satisfied if and only if the two sides have equal moduli and arguments that differ by multiples of 2π (the right-hand side has argument 0), that is:

$$\begin{cases} \varrho^3 = \varrho \\ 3\vartheta = 2k\pi \quad \text{with } k \in \mathbb{Z} \end{cases}$$

The first equation gives $\varrho = 0$ and $\varrho = 1$ (careful: ϱ must be ≥ 0 because it is the modulus of the complex number; therefore $\varrho = -1$ is not acceptable); the second one gives $\vartheta = \frac{2k\pi}{3}, k \in \mathbb{Z}$. Hence:

$$z = 0, \quad z = \cos \frac{2k\pi}{3} + i \sin \frac{2k\pi}{3} \quad \text{with } k \in \mathbb{Z}$$

Explicitly:

$$z = 0, \quad z = 1, \quad z = -\frac{1}{2} + i \frac{\sqrt{3}}{2}, \quad z = -\frac{1}{2} - i \frac{\sqrt{3}}{2}$$

9.3.2 n -th roots of complex numbers

Definition 9.8: n -th root of a complex number

Given a complex number w , we say that z is an n -th (complex) root of w if $z^n = w$.

Theorem 9.9

Let $w \in \mathbb{C}$, $w \neq 0$, and n an integer ≥ 1 . There exist exactly n complex n -th roots $z_0, z_1, \dots, z_{n-1}$ of w ; setting

$$w = r (\cos \varphi + i \sin \varphi) \quad \text{and} \quad z_k = \varrho_k (\cos \vartheta_k + i \sin \vartheta_k)$$

we have

$$\begin{cases} \varrho_k = r^{1/n} \\ \vartheta_k = \frac{\varphi + 2k\pi}{n} \end{cases} \quad k = 0, 1, 2, \dots, n-1 \quad (9.21)$$

Proof. • The numbers z_k are clearly roots of w , as can be seen by computing z_k^n with De Moivre's formula. Let us show that there are no others.

- If a number $R(\cos \psi + i \sin \psi)$ is an n -th root of w , we must have

$$R^n = r \quad \text{and} \quad n\psi = \varphi + 2h\pi \quad \text{with} \quad h \in \mathbb{Z}$$

or equivalently:

$$R = r^{1/n} \quad \text{and} \quad \psi = \varphi/n + 2h\pi/n \quad \text{with} \quad h \in \mathbb{Z}$$

- Giving h the values $0, 1, \dots, n-1$ we find exactly the numbers z_k .
- Giving h any other value $\bar{h}$ different from the previous ones, it can be written in the form $\bar{h} = k + mn$ ($m \in \mathbb{Z}$ is the quotient and k is the remainder of the division of $\bar{h}$ by n), so that we would have

$$\psi = \frac{\varphi}{n} + \frac{2k\pi}{n} + 2m\pi = \vartheta_k + 2m\pi$$

and we would find again the same z_k as before. □

Example 9.9: fifth root of a complex number

- Let us compute $\sqrt[5]{1+i}$: the complex number $(1+i)$ has modulus $\sqrt{2}$ and argument $\pi/4$. The fifth roots will therefore be:

$$\sqrt[5]{1+i} = \sqrt[5]{\sqrt{2}} \left[\cos \left(\frac{\frac{\pi}{4} + 2k\pi}{5} \right) + i \sin \left(\frac{\frac{\pi}{4} + 2k\pi}{5} \right) \right] =$$

$$= \sqrt[10]{2} \left[\cos \left(\frac{\pi}{20} + \frac{2}{5}k\pi \right) + i \sin \left(\frac{\pi}{20} + \frac{2}{5}k\pi \right) \right] \quad \text{with } k = 0, 1, 2, 3, 4$$

- The angles found are not standard angles, but if desired, their sines and cosines can be computed approximately using a calculator.

- Unfortunately, a somewhat ambiguous notation is used for complex roots, the same one used to denote the arithmetic root; that is, $\sqrt[n]{z}$ or $z^{1/n}$ denotes the set of the n complex roots of z .
- This can create confusion when z is real. Indeed, the symbol $\sqrt{4}$, understood as the arithmetic root of 4, is 2; understood as the complex root of 4, it is the set of the two numbers $+2$ and -2 .

Example 9.10: cube root of a complex number

- Let us compute $\sqrt[3]{-1}$: the complex number (-1) has modulus 1 and argument π . The cube roots will therefore be:

$$\sqrt[3]{-1} = \underbrace{\sqrt[3]{1}}_{=1} \left[\cos \left(\frac{\pi}{3} + \frac{2}{3}k\pi \right) + i \sin \left(\frac{\pi}{3} + \frac{2}{3}k\pi \right) \right] \quad \text{with } k = 0, 1, 2$$

that is, the numbers z_k of the form:

$$\cos \vartheta_k + i \sin \vartheta_k \quad \text{with } \vartheta_k = \frac{\pi}{3} + \frac{2}{3}k\pi \quad \text{and } k = 0, 1, 2$$

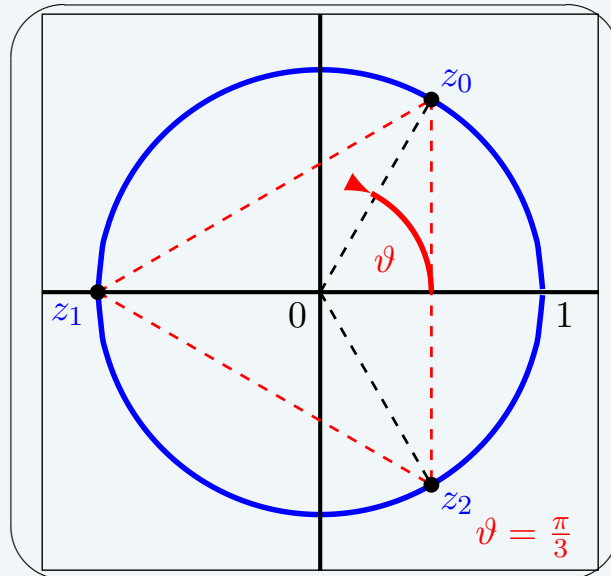
- Explicitly we have:

$$\begin{cases} z_0 = \cos \left(\frac{\pi}{3} \right) + i \sin \left(\frac{\pi}{3} \right) = \frac{1}{2} + i \frac{\sqrt{3}}{2} = \frac{1}{2}(1 + i\sqrt{3}) \\ z_1 = \cos \left(\frac{\pi}{3} + \frac{2\pi}{3} \right) + i \sin \left(\frac{\pi}{3} + \frac{2\pi}{3} \right) = \cos(\pi) + i \sin(\pi) = -1 \\ z_2 = \cos \left(\frac{\pi}{3} + \frac{4\pi}{3} \right) + i \sin \left(\frac{\pi}{3} + \frac{4\pi}{3} \right) \\ \quad = \cos \left(\frac{5\pi}{3} \right) + i \sin \left(\frac{5\pi}{3} \right) = \frac{1}{2} - i \frac{\sqrt{3}}{2} = \frac{1}{2}(1 - i\sqrt{3}) \end{cases}$$

- The arrangement of the roots of complex numbers in the Gauss plane is not random.
- Indeed, if $w = r(\cos \varphi + i \sin \varphi)$, the n -th roots $z_0, z_1, \dots, z_{n-1}$ of w lie at the vertices of the regular polygon with n sides inscribed in the circle with center 0 and radius $r^{1/n}$, with the vertex z_0 located at the point with argument $\vartheta = \varphi/n$.

Example 9.11: roots in the Gauss plane

- The figure shows the cube roots of -1 : z_0, z_1, z_2 from the previous exercise:

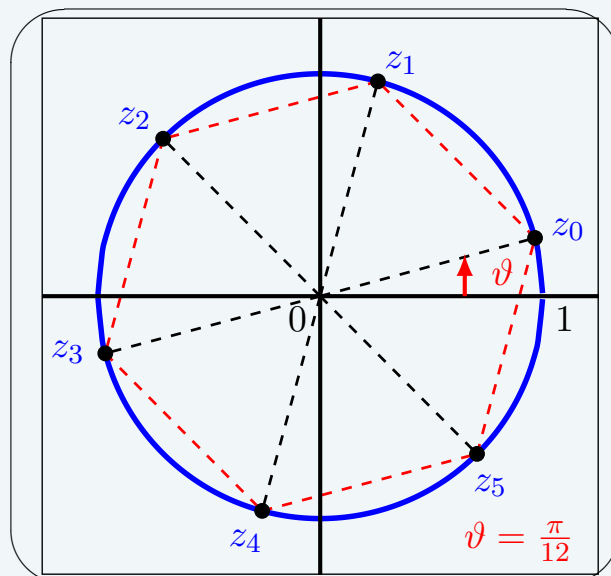


- The figure shows the sixth roots of i : $z_0, z_1, z_2, z_3, z_4, z_5$,

$$\sqrt[6]{i} = \underbrace{\sqrt[6]{1}}_{=1} \left[\cos \left(\frac{\pi}{2 \cdot 6} + \frac{2}{6} k \pi \right) + i \sin \left(\frac{\pi}{2 \cdot 6} + \frac{2}{6} k \pi \right) \right] \quad \text{with } k = 0, 1, 2, 3, 4, 5$$

that is, the numbers z_k of the form:

$$\cos \vartheta_k + i \sin \vartheta_k \quad \text{with } \vartheta_k = \frac{\pi}{12} + \frac{1}{3} k \pi \quad \text{and } k = 0, 1, 2, 3, 4, 5$$

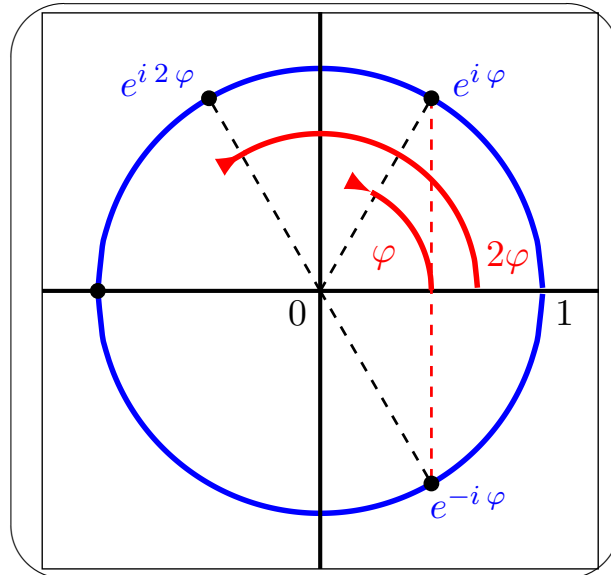


9.3.3 Exponential form of complex numbers

- The following notation is useful:

$$e^{i\varphi} = \cos \varphi + i \sin \varphi \quad \text{with } \varphi \in \mathbb{R}$$

where e is Euler's number (Napier's constant).



With this notation, the *trigonometric form* can be equivalently rewritten as:

$$z = |z| (\cos \varphi + i \sin \varphi) = |z| e^{i\varphi}, \quad \varphi = \arg(z) \quad \text{with } z \neq 0$$

This notation is called the **exponential form** of complex numbers.

- This form is useful since we have:

$$e^{i\varphi_1} \cdot e^{i\varphi_2} = e^{i(\varphi_1 + \varphi_2)}$$

that is, it is very easy to compute products of complex numbers.

De Moivre's formulas become:

$$(e^{i\varphi})^n = e^{in\varphi}$$

The exponential notation is widely used in applications because it greatly simplifies manipulations involving trigonometric quantities.

We have:

$$e^{2k\pi i} = 1 \quad \text{and} \quad e^{i(\varphi + 2k\pi)} = e^{i\varphi} \quad \forall k \in \mathbb{Z}, \quad \text{moreover} \quad |e^{i\varphi}| = 1$$

We also have:

$$e^{i\pi} = -1, \quad e^{i\frac{\pi}{2}} = i, \quad e^{i\frac{3\pi}{2}} = -i, \quad e^{i\frac{3\pi}{4}} = \frac{-1+i}{\sqrt{2}}$$

Rewriting, we obtain **Euler's formula**:

$$e^{i\pi} + 1 = 0$$

which links in a simple way five of the most important constants: 0, 1, i , π , e .

The exponential notation can be generalized consistently with the properties of powers, by defining the **complex exponential**:

$$e^z = e^{x+iy} = e^x e^{iy} = e^x (\cos y + i \sin y), \quad \forall z = x + iy \in \mathbb{C}$$

- From the definition of exponential form it follows that

$$e^{-i\varphi} = \cos(-\varphi) + i \sin(-\varphi) = \cos \varphi - i \sin \varphi$$

and we easily obtain:

$$\cos \varphi = \frac{e^{i\varphi} + e^{-i\varphi}}{2}, \quad \sin \varphi = \frac{e^{i\varphi} - e^{-i\varphi}}{2i} \quad \text{with } \varphi \in \mathbb{R}$$

These formulas allow us to express sine and cosine as combinations of complex exponentials.

9.4 Quadratic equations

- A **second-degree equation** or **quadratic equation** in one **real** unknown x is an algebraic equation in which the highest power of the unknown is 2, and it can always be written in the form:

$$a x^2 + b x + c = 0 \quad (a \neq 0) \quad (9.22)$$

where a, b and c are real numbers.

The solutions are also called **roots** or **zeros** of the equation.

Observation 9.10: quadratic formula

Given a quadratic equation $a x^2 + b x + c = 0$ ($a \neq 0$), its zeros or roots are:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (9.23)$$

- A quadratic equation is called a “complete quadratic equation” when all its coefficients are different from 0. It is solved with the so-called **method of completing the square**, so called because the equation is modified until its left-hand side becomes the square of a binomial.

Proof. Isolating the constant term, we obtain:

$$\begin{aligned}
 a x^2 + b x &= -c \\
 \underbrace{4 a^2 x^2}_{=(2 a x)^2} + \underbrace{4 a b x}_{=2 (2 a x) b} &= -4 a c \\
 (2 a x)^2 + 2 (2 a x) b + b^2 &= b^2 - 4 a c \\
 (2 a x + b)^2 &= \underbrace{b^2 - 4 a c}_{:=\Delta \text{ (discriminant)}} \\
 2 a x + b &= \pm \sqrt{b^2 - 4 a c} \\
 x &= \frac{-b \pm \sqrt{b^2 - 4 a c}}{2 a}
 \end{aligned}$$

□

- If the discriminant Δ is negative there are no real solutions.
- If $\Delta = 0$, the quadratic formula becomes:

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

therefore there is only one root, with *multiplicity two*.

- Finally, if $\Delta < 0$, the equation has no real solutions. In particular, there are always two solutions, but they belong to the field of complex numbers: they are **two complex conjugate numbers** and are computed with the two formulas:

$$\begin{aligned}
 x_+ &= \frac{-b}{2a} + i \left(\frac{\sqrt{4ac - b^2}}{2a} \right) \\
 x_- &= \frac{-b}{2a} - i \left(\frac{\sqrt{4ac - b^2}}{2a} \right)
 \end{aligned}$$

where i is the **imaginary unit** ($i^2 = -1$).

A quadratic equation in one **complex** unknown z has the following form

$$a z^2 + b z + c = 0 \quad (a \neq 0) \quad (9.24)$$

where a, b and c are complex numbers. It is solved with the same formula as in the case of a real unknown:

$$z = \frac{-b \pm \sqrt{b^2 - 4 a c}}{2 a} \quad (9.25)$$

where the square root is understood in the complex sense (the $\pm$ sign is superfluous, because in the complex field the root denotes two numbers, each the opposite of the other)

Example 9.12: quadratic equations in the complex field

- We want to solve:

$$z^2 + 2i z - \sqrt{3}i = 0 \quad a = 1, b = 2i, c = -\sqrt{3}i$$

we have (switching to the trigonometric form to compute the root):

$$\begin{aligned} z &= \frac{-2i \pm \sqrt{2^2 i^2 + 2^2 \sqrt{3}i}}{2} = \frac{-2i \pm 2\sqrt{-1 + \sqrt{3}i}}{2} = -i \pm \sqrt{-1 + \sqrt{3}i} \\ &= -i \pm \sqrt{2 \left(-\frac{1}{2} + i \frac{\sqrt{3}}{2} \right)} = -i \pm \sqrt{2 \left(\cos \frac{2}{3}\pi + i \sin \frac{2}{3}\pi \right)} \\ &= -i \pm \sqrt{2} \left(\underbrace{\cos \frac{\pi}{3}}_{=\frac{1}{2}} + i \underbrace{\sin \frac{\pi}{3}}_{=\frac{\sqrt{3}}{2}} \right) = \pm \frac{\sqrt{2}}{2} + i \left(-1 \pm \frac{\sqrt{6}}{2} \right) \end{aligned}$$

- The previous theorem tells us that a polynomial of the form $z^n + a$ (with a complex) has exactly n roots in $\mathbb{C}$; in the real field, instead, the equation $x^n + a = 0$ may have two, one, or no roots (examples: $x^2 - 1 = 0, x^3 - 1 = 0, x^2 + 1 = 0$); we now know that such an equation always has n roots in $\mathbb{C}$, but only occasionally one or two of them lie in $\mathbb{R}$.
- The result is much more general, as stated by the following theorem, whose proof we do not report.

Theorem 9.11: fundamental theorem of algebra

A polynomial equation of the form

$$a_0 + a_1 z + \cdots + a_n z^n = 0 \quad (a_n \neq 0)$$

with arbitrary complex coefficients has exactly n roots in $\mathbb{C}$, if each of them is counted with its multiplicity ^a.

^aIf $P(z)$ is a polynomial in z of degree n and z_0 is one of its roots, we say that z_0 has multiplicity k (k integer, ≥ 1) if the formula $P(z) = (z - z_0)^k Q(z)$ holds, where Q is a polynomial such that $Q(z_0) \neq 0$.

Cosines and sines of the main angles

