MATHEMATICS

MATTEO RIZZI

Contents

1. Mathematics in geology

When I was a guy, once I told a friend of mine :"Life is mathematics!"..."Oh, if life were mathematics I would be dead", she desperately said-:) Her disappointed answer is actually true, mathematics is not enough to explain everything in life, but sometimes it can be helpful. Especially for scientists.

In the case of geology, we may be interested in measuring how strong an earthquake is, how acid a solution is, or how old a rock is. For these and many other purposes, geologists absolutely need mathematics.

Assume, for example, that you want to date a rock. For this purpose, you can use the properties of some radioactive isotopes. More precisely, we know that radioactive isotopes are highly unstable materials which tend to evolve to stable states. As a consequence, in the initial configuration of the rock, these materials are unstable. On the contrary, after a while, these materials change and become stable. One possible way to date the rock is to measure the ratio x between the unstable material and the total amount of material. At the beginning, that is at time $t = 0$ this ratio is $x(0) = 1$; after a while it will decrease. The question is: how fast does this ratio

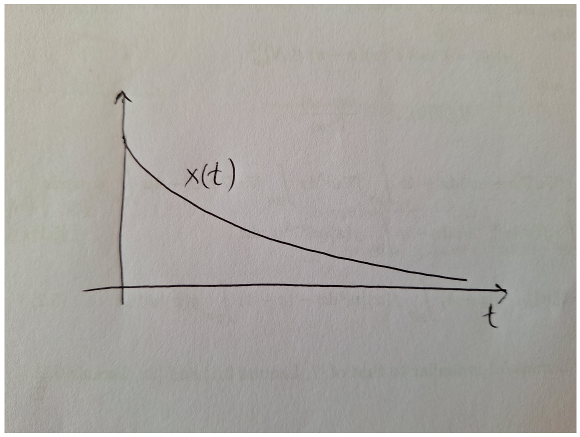

FIGURE 1. The graph of the ratio between the amount of unstable material and the total amount of material in a rock. As we can see, this ratio decreases in time.

decrease?

The answer, which comes from experiments, is the following: the velocity $v(t)$ with which x decreases at time $t > 0$ is proportional to the ratio itself at time t. In other words, we have $v(t) = -\lambda x(t)$, where $\lambda > 0$ is a proportionality constant and the minus sign comes from the fact that this velocity is negative, since $x(t)$ is decreasing in time.

The law $v(t) = -\lambda x(t)$ is called a *differential equation* by mathematicians. This kind of equations will be treated in this course (hopefully). The solution to this equation is an explicit formula which enables us to compute the ratio $x(t)$ at any time $t > 0$ and shows that this ratio decreases in time and tends to vanish after infinite time. The graph of this ratio depending on time is represented in Figure (1). In this sense, mathematics confirms the heuristic prediction that radioactive isotopes tend to become completely stable after a while.

This formula can be inverted, so that, given a value the ratio x_0, we can compute the precise time t_0 such that $x(t_0) = x_0$. As a consequence, by measuring such a ratio x_0 with suitable tools, geologists can deduce that the age of the rock is t_0.

2. Set theory

Definition 2.1. *A set is a collection of objects called* ***elements*** *of the set.*

(1) $A := \{1, 5, 6, 30\}$ is a set.

(2) $B = \{a, e, i, o, u\}$ is a set. That is, B is the set of vowels in the Italian alphabet.
(3) $\mathbb{N} := \{0, 1, 2, \dots\}$ is the set of non-negative integers, or natural numbers.
(4) $C := \{2n : n \in \mathbb{N}\}$ is the set of even numbers. The symbol ":" means "such that".

2.1. **Some useful symbols.** Now we will introduce some useful symbols which often appear when dealing with sets.

- $\in$ means "belongs". For example $10 \in \mathbb{N}$ means "10 belongs to $\mathbb{N}$".
- $\notin$ means "do not belong" or "does not belong". For example $\frac{1}{2} \notin \mathbb{N}$ means "$\frac{1}{2}$ does not belong to $\mathbb{N}$".
- $\emptyset$ is the empty set, that is the set which does not contain any element.
- $A \subseteq B$ means that A is contained in B, that is any element of A belongs to B too. For example $\{1, 2\} \subset \mathbb{N}$.
- $A \nsubseteq B$ means that A is not contained in B, that is there exists an element of A which does not belong to be. For example $\{1, \frac{3}{2}\} \nsubseteq \mathbb{N}$.
- $A \subsetneq B$ means that A is strictly contained in B, or A is a proper subset of B, that is $A \subseteq B$ and $A \neq B$. For example $\{1, 2\} \subsetneq \mathbb{N}$.

Now we introduce some special symbols, known as *quantifiers*, which are crucial to understand the logic structure of a sentence.

- $\forall$ means "for all".
- $\exists$ means "there exist" or "there exists".
- $\nexists$ means "There do not exist" or "there does not exist".
- $\exists!$ means "there exists a unique".

Such quantifiers are often used together with other logic symbols. For example

- $\Rightarrow$ means "implies", or in other words "$A \Rightarrow B$" means "A implies B", that is "if A is true, then B is true".
- If $A \Rightarrow B$, we say that A is a *sufficient condition for* B and B is a *necessary condition for* A.
- $\nRightarrow$ means "does not imply", or in other words "$A \nRightarrow B$" means "A does not imply B", that is "if A is true, then we do not whether B is true or not".
- $\Leftrightarrow$ means "if and only if", or in other words "$A \Leftrightarrow B$" means "A implies B and B implies A" or "A is equivalent to "B". That is "if A is true, then B is true and if B is true, then A is true".
- $\vee$ means "or".
- $\wedge$ means "and".
- $\urcorner$ is used for the negation.

2.1.1. *Examples.*

(1) Let $n \in \mathbb{N}$. If x is not even, then x is odd.
This sentence can be written in the equivalent forms:
- Let $n \in \mathbb{N}$. x is not even $\Rightarrow x$ is odd.
- Let $n \in \mathbb{N}$. $\neg$ (n is even) $\Rightarrow n$ is odd.

(2) $\forall n \in \mathbb{N}$, n is even $\vee$ n is odd.
(3) Every natural number which is a multiple of 4 is even.
This sentence can be written in the form:
$\forall n \in \mathbb{N}$: n is a multiple of 4 $\Rightarrow n$ is even.
(4) Let $n \in \mathbb{N}$. Then n is a multiple of 4 $\Leftrightarrow \exists\, k \in \mathbb{N} : \; n = 4k$.
(5) $\neg(\forall n \in \mathbb{N}$, either n is even or n is a multiple of 3) $\Leftrightarrow$ ($\exists\, n \in \mathbb{N}$: n is not even and n is not a multiple of 3.)
(6) Let a, $b \in \mathbb{N}$. Then
a is even $\wedge$ b is even $\Rightarrow\; a + b$ is even.
a is even $\vee$ b is even $\not\Rightarrow\; a + b$ is even.

2.2. **Operations between sets.** Now we define the notion of subset and we introduce the main operations between sets.

Definition 2.2. *Let A and B be two sets. We say that B is a subset of A if $B \subseteq A$, or equivalently $x \in A \;\Rightarrow\; x \in\; B$.*

Remark 2.1. *Let A and B be two sets. Then*
- $A = B \Leftrightarrow (A \subseteq B \wedge B \subseteq A)$
- $A \neq B \Leftrightarrow (A \not\subseteq B \vee B \not\subseteq A)$

The union, the intersection and the difference of sets are defined as follows

Definition 2.3. *Let A, B be sets. Then*
- *the union of A and B is defined by*
$$A \cup B := \{x : \; x \in A \;\vee\; x \in B\}.$$
- *the intersection between A and B is defined by*
$$A \cap B := \{x : \; x \in A \;\wedge\; x \in B\}$$
- *the difference between A and B is defined by*
$$A \setminus B := \{x : \; x \in A \;\wedge\; x \notin B\}.$$
- *If $B \subseteq A$, the set $A \setminus B$ is knwon as the complement of B in A and it is denoted by $\mathcal{C}_A(B)$.*
- *The cartesian product is defined by*
$$A \times B := \{(a, b) : \; a \in A, \; B \in B\}.$$

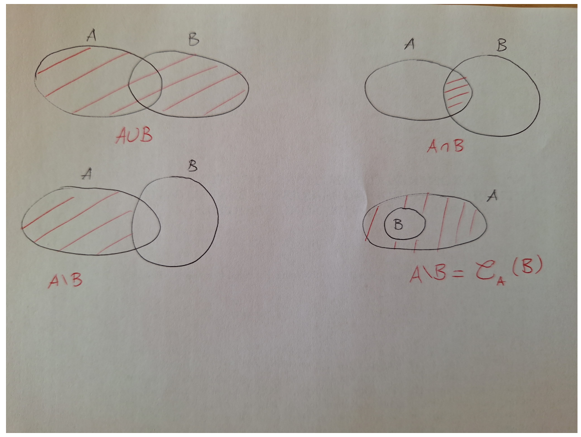

FIGURE 2. Operations between sets

Sometimes, if it is clear from the context what A is, we will use the notation $B^c; = \mathcal{C}_A(B)$. Moreover, we note that

$$A \cap B \subseteq A,\ B \qquad A, B \subseteq A \cup B. \tag{2.1}$$

Moreover, given a set A, we call *cardinality of* A the number of elements of A and we denote it by $\mathrm{Card}(A)$. Note that

$$\mathrm{Card}(A \cup B) = \mathrm{Card}(A) + \mathrm{Card}(B) - \mathrm{Card}(A \cap B). \tag{2.2}$$

Remark 2.2. *Let A, B and C be 3 sets. Then*

- $A \cup B = B \cup A$
- $A \cap B = B \cap A$
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- *If $A \subseteq C$ and $B \subseteq C$, then $\mathcal{C}_C(A \cap B) = \mathcal{C}_C(A) \cup \mathcal{C}_C(B)$ and $\mathcal{C}_C(A \cup B) = \mathcal{C}_C(A) \cap \mathcal{C}_C(B)$.*

2.2.1. *Examples.*

- If $A := \{1, 2, 3\}$ and $B := \{1, 2, 5, 6\}$, then $A \cap B = \{1, 2\}$, $A \cup B = \{1, 2, 3, 5, 6\}$, $A \setminus B = \{3\}$ and $B \setminus A = \{5, 6\}$. In this case, we can see that

$$\mathrm{Card}(A) = 3,\ \mathrm{Card}(B) = 4,\ \mathrm{Card}(A \cap B) = 2,\ \mathrm{Card}(A \cup B) = 5$$

$$\mathrm{Card}(A \setminus B) = 1,\ \mathrm{Card}(B \setminus A) = 2.$$

- $A := \mathbb{N}$, $B := \{4n : n \in \mathbb{N}\}$. Then $\mathcal{C}_A(B) = \{k + 4n : k \in \{1, 2, 3\},\ n \in \mathbb{N}\}$.

- $A := \{2n : n \in \mathbb{N}\}$, $B := \{4n : n \in \mathbb{N}\}$. Then $\mathcal{C}_A(B) = \{2 + 4n : n \in \mathbb{N}\}$. (It follows from the examples that the complement of B in A depends both on A and on B. In other words, if we change A then the complement $\mathcal{C}_A(B)$ also changes.
- Let $\mathbb{R}$ denote the set of real numbers. If $A = B = \mathbb{R}$, then
$$A \times B = \{(x, y) : x \in \mathbb{R},\, y \in \mathbb{R}\} =: \mathbb{R}^2$$
is the cartesian plane.

3. Numeric sets

Here we introduce the main numeric sets that will be used in the course. Each of these sets enables us to make certain operations between its elements, but not all possible operations. For this reason at any step we are lead to consider larger sets.

- $\mathbb{N}$ is the set of non-negative integers, or natural numbers. It is closed with respect to the sum and the product, in the sense that, $\forall\, m,$, we have $m+n \in \mathbb{N}$ and $mn \in \mathbb{N}$. This is not true for the difference or the quotient of natural numbers. For instance $1,\, 2 \in \mathbb{N}$ but $\frac{1}{2} \notin \mathbb{N}$ and $1 - 2 = -1 \notin \mathbb{N}$.
- $\mathbb{Z} = \{-n : n \in \mathbb{N}\} \cup \mathbb{N} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ is the set of integer numbers. It is closed with respect to the sum, the product and the difference, but not with respect to the quotient. Note that $\mathbb{N} \subset \mathbb{Z}$
- $\mathbb{Q} := \{\frac{m}{n} : m \in \mathbb{Z},\, n \in \mathbb{Z}\}$ is the set of rational numbers. It is closed with respect to the all the four operations. Note that $\mathbb{Z} \subset \mathbb{Q}$.

Note that, for any a, b, c, $d \in \mathbb{Z}$, $b \neq 0$, $d \neq 0$, we have

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}, \quad \frac{a}{b} - \frac{c}{d} = \frac{ad - bc}{bd}, \quad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}. \tag{3.1}$$

If, in addition, $c \neq 0$, we have

$$\frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a}{b} \cdot \frac{d}{c}. \tag{3.2}$$

Despite being closed with respect to the four operations, the rational numbers are not enough to deal with elementary geometry. This can be understood by considering the Pitagora's Theorem.

Theorem 3.1 (Pitagora)**.** *Let a and b be the legs and let c be the ipotenuse of a rectangle triangle. Then $a^2 + b^2 = c^2$.*

As a consequence, if $a = b = 1$, then $c^2 = 2$, or equivalently $c = \sqrt{2}$.

Lemma 3.1. *$\sqrt{2}$ is not rational.*

Proof. Assume by contradiction that there exists $m, n \in \mathbb{N}$ such that $\sqrt{2} = \frac{m}{n}$. Without loss of generality, we can assume that m and n are not both even, otherwise we simplify the fraction. However we have $2 = \frac{m^2}{n^2}$, or equivalently $2n^2 = m^2$, so that m^2 is even. As a consequence, m is also even, that is $m = 2k$, for some $k \in \mathbb{N} \setminus \{0\}$. This yields that

$$m^2 = 4k^2 = 2n^2,$$

so that $n^2 = 2k^2$ is also even, hence n is even too. This is a contradiction. □

It is well known that every rational number can be written in decimal form. For example

$$\frac{1}{2} = 0,5, \qquad \frac{1}{3} = 0,3333\cdots = 0,\bar{3}.$$

In the bove examples, we can see that the decimal representation of $\frac{1}{2}$ involves a finite number of digits, while the one of $\frac{1}{3}$ is periodic, in the sense that the same digit is repeated infinitely many times.

Theorem 3.2. *Any rational number can be written in decimal form and its decimal representation is either finite or periodic.*

As a consequence, the number $1,101001000\ldots$ is not rational.

Setting

$$\mathbb{R} := \{n, a_1 a_2 \cdots :\ a_i \in \{0,1,2,3,4,5,6,7,8,9\},\ i \in \mathbb{N} \setminus \{0\}\},$$

the set of rational numbers $\mathbb{Q}$ is a proper subset of $\mathbb{R}$. The elements of $\mathbb{R}$ are known as *real numbers*. The element of $\mathbb{R} \setminus \mathbb{Q}$ are known as irrational numbers. Some famous irrational numbers are $\sqrt{2}$, π, e.

3.1. Properties of the sum.

- Commutative property: $\forall\, a,\ b \in \mathbb{R}$, we have $a + b = b + a$.
- Associative property: $\forall\, a,\ b,\ c \in \mathbb{R}$, we have $a + (b + c) = (a + b) + c =: a + b + c$.
- 0 is the neutral element, that is $\forall\, a \in \mathbb{R}$, we have $a + 0 = a$.
- The existence and uniqueness of the inverse: $\forall\, a \in \mathbb{R}$, $-a$ is the unique real number such that $a + (-a) = 0$.

As a consequence, the following cancellation law holds true:

$$\forall\, a,\ b,\ c \in \mathbb{R},\ a + b = a + c \Leftrightarrow a = b, \tag{3.3}$$

which yields that

$$\forall\, a,\ b,\ c \in \mathbb{R},\ a + b = c \Leftrightarrow a = c - b. \tag{3.4}$$

Properties of the product $a \cdot b = ab$

- Commutative property: $\forall\, a,\ b \in \mathbb{R}$, we have $ab = ba$.

- Associative property: $\forall\, a,\, b,\, c \in \mathbb{R}$, we have $a(bc) = (ab)c =: abc$.
- 1 is the neutral element, that is $\forall\, a \in \mathbb{R}$, we have $a\cdot 1 = a$.
- The existence and uniqueness of the inverse: $\forall\, a \in \mathbb{R}$, $\frac{1}{a}$ is the unique real number such that $a\cdot \frac{1}{a} = 1$.
- Distributive property: $\forall\, a,\, b,\, c \in \mathbb{R}$, we have $a(b + c) = ab + ac$.

As a consequence, the following relations hold:

- $\forall\, a \in \mathbb{R}$, we have $a\cdot 0 = 0$.
- Vanishing: $\forall a,\, b \in \mathbb{R}$, $ab = 0 \Rightarrow a = 0 \vee b = 0$.
- Cancellation law: $\forall\, a,\, b\, c \in \mathbb{R}$, $c \neq 0$, we have $ac = bc \Leftrightarrow a = b$.
- $\forall\, a,\, b\, c \in \mathbb{R}$, $c \neq 0$, we have $ac = b \Leftrightarrow a = \frac{b}{c}$.

Using the above properties we can see that

$$\begin{aligned} &x + 3 = 8 \Leftrightarrow x + 3 = 5 + 3 \Leftrightarrow x = 3 \\ &x + 5 = 5 \Leftrightarrow x = 0 \\ &2x = 1 \Leftrightarrow x = \frac{1}{2}. \end{aligned}$$

3.2. **Ordering in $\mathbb{R}$.** Real numbers can be represented on a straight line, so they can be naturally ordered. The order relations are represented by $>$, $\geq$, $<$, $\leq$. Their meaning is the following

- $a > b$ means that a is on the right of b.
- $a \geq b$ means that $a > b \vee a = b$.
- $a < b$ means that a is on the left of b.
- $a \leq b$ means that $a < b \vee a = b$.

3.3. **Ordering properties.**

(1) $\forall\, a \in \mathbb{R}$, $a \leq a$.
(2) $\forall\, a,\, b \in \mathbb{R}, a = b \Leftrightarrow a \leq b \wedge b \leq a$.
(3) Transitive property: $\forall\, a,\, b,\, c \in \mathbb{R}$, $a \leq b \wedge b \leq c \Rightarrow a \leq c$.
(4) $\forall\, a,\, b \in \mathbb{R}$, $a \leq b \vee b \leq a$. More precisely, we have $a < b \vee a = b \vee a > b$.
(5) $\forall\, a,\, b,\, c \in \mathbb{R}$, $a \leq b \Leftrightarrow a + c \leq b + c$.
(6) $\forall\, a,\, b,\, c \in \mathbb{R}$, $c > 0$, $a \leq b \Leftrightarrow ac \leq bc$.

As a consequence we have

$$\forall\, a,\, b,\, c \in \mathbb{R},\ c < 0,\ a \leq b \Leftrightarrow ac \geq bc.$$

By property (5), we can see that, if $a \leq b$, the inequality is preserved if we add or subtract any real number. Moreover, by the property (6), the inequality is also preserved if we multiply by a positive real number.

Lemma 3.2 (Continuity property of $\mathbb{R}$)**.** *Let A, $B \subset \mathbb{R}$ be two non-empty subsets of $\mathbb{R}$ such that $\forall\, a \in A$, $b \in B$ we have $a \leq b$. Then there exists $c \in \mathbb{R}$ such that $\forall\, a \in A$, $b \in B$ we have $a \leq c \leq b$.*

Lemma 3.3. *Let $a \in \mathbb{R}$ .Then for any $\varepsilon > 0$ there exists $c \in \mathbb{Q}$ such that $a - \varepsilon < c < a + \varepsilon$.*

3.4. **The extrema of a set $A \subseteq \mathbb{R}$.** First, we introduce the notion of interval.

Definition 3.1. *A subset $I \subset \mathbb{R}$ is called an interval if $\forall\, a$, $b \in I$, $a < b$, and $\forall\, x \in \mathbb{R}$ such that $a < x < b$, we have $x \in I$.*

Definition 3.1 says that a set I is an interval if and only if, given two points a, $b \in I$, $a \neq b$, all intermediate points also belong to I.

For example, for a, $b \in \mathbb{R}$, $a < b$, we set $I := (a, b) := \{x \in \mathbb{R} :\ a < x < b\}$. It is possible to see that I is an interval.

Definition 3.2. *Let $A \subseteq \mathbb{R}$ be a subset. Then we say that*

1. *A is open if either $A = \emptyset$ or, for any $x \in A$, there exists $\delta > 0$ such that $(x - \delta, x + \delta) \subseteq A$.*
2. *A is closed if $\mathbb{R} \setminus A$ is open.*
3. *A is bounded if there exists a, $b \in \mathbb{R}$ such that $A \subseteq (a, b)$.*
4. *A is unbounded if it is not bounded.*

Note that, according to Definition 3.2, $\mathbb{R}$ is both open and closed, since it clearly satisfied property (1) and $\mathbb{R} \setminus \mathbb{R} = \emptyset$ is also open. As a consequence $\emptyset$ is both open and closed too.

Moreover, given a set $A \subset \mathbb{R}$, we say that a point $x \in A$ is an interior point if There exists $\delta > 0$ such that $(x - \delta, x + \delta) \subseteq A$ and we denote the set of interior points of A by $\mathring{A}$.

Note that $\mathring{A} \subseteq A$ and equality holds if and only if A is open.

Using the notions introduced in (3.2), intervals can be classified in the following way.

- $\exists\, a, b \in \mathbb{R} :\ I = (a, b) := \{x \in \mathbb{R} :\ a < x < b\}$ (open bounded interval).
- $\exists\, a, b \in \mathbb{R} :\ I = [a, b] := \{x \in \mathbb{R} :\ a \leq x \leq b\}$ (closed bounded interval).
- $\exists\, a, b \in \mathbb{R} :\ I = (a, b] := \{x \in \mathbb{R} :\ a < x \leq b\}$.
- $\exists\, a, b \in \mathbb{R} :\ I = [a, b) := \{x \in \mathbb{R} :\ a \leq x < b\}$.
- $\exists\, a \in \mathbb{R} :\ I = (a, \infty) := \{x \in \mathbb{R} :\ x > a\}$ (open unbounded interval with lower extreme a).

- $\exists\, a \in \mathbb{R} : I = [a, \infty) := \{x \in \mathbb{R} : x \geq a\}$ (closed unbounded interval with lower extreme a).
- $\exists\, a \in \mathbb{R} : I = (-\infty, a) := \{x \in \mathbb{R} : x < a\}$ (open unbounded interval with upper extreme a).
- $\exists\, a \in \mathbb{R} : I = (-\infty, a] := \{x \in \mathbb{R} : x \leq a\}$ (closed unbounded interval with upper extreme a).
- $I = (-\infty, \infty) = \mathbb{R}$ (real line).

Unbounded intervals are also referred to as *half-lines*.

Note that, for example, if either $I = [a, b]$ or $I = (a, b]$ or $I = [a, b)$, for some a, $b \in \mathbb{R}$, $a < b$, we have $\mathring{I} = (a, b)$. If $I = (-\infty, a]$, then $\mathring{I} = (-\infty, a)$ and so on.

Intuition says that any subset $A \subseteq \mathbb{R}$ has a "left extreme" and a "right extreme". This concept can be made precise by introducing the definition of inf (lower extreme) and sup (upper extreme) of a set. In order to do so, we first introduce the notion of **maximum** and **minimum** of a set.

Definition 3.3. *Let $A \subseteq \mathbb{R}$ be a subset of $\mathbb{R}$.*

- *We say that an element $x \in A$ is the maximum of A if $\forall\, a \in A$, we have $a \leq x$.*
- *We say that an element $x \in A$ is the minimum of A if $\forall\, a \in A$, we have $x \leq a$.*

We observe that the maximum and the minimum do not always exist. For example,

- The closed bounded interval $I = [0, 1]$ has both a maximum, that is 1, and a minimum, that is 0.
- The interval $I = [0, 1)$ has a minimum, that is 0, but it has no maximum. In fact, due to Lemma 3.3, for any $x \in [0, 1)$ there exists $c \in \mathbb{R}$ such that $x < c < 1$, thus $c \in I$ and $c > x$. This yields that x is not the maximum of $[0, 1)$.
- The interval $I = (0, 1]$ has a maximum, that is 1, but it has no minimum.
- The open bounded interval $I = (0, 1)$ has neither a maximum nor a minimum.

However, we can introduce a weaker notion, that is the concept of inf and sup, which generalize the notion of minimum and maximum and are defined in such a way that any subset of $\mathbb{R}$ possess both an inf and a sup.

More precisely, given a set $A \subseteq \mathbb{R}$, we introduce the set

$$M(A) := \{x \in \mathbb{R} : x \geq a, \forall\, a \in A\},$$

which is known as the *upper set* of A. Due to the continuity property of $\mathbb{R}$, given by Lemma 3.2, either the set $M(A)$ is empty or it has a minimum. Hence, we are led to the following definition.

Definition 3.4. *If* $M(A) \neq \emptyset$*, we set* $\sup(A) := \min M(A)$*. Otherwise, we set* $\sup(A) = \infty$*.*

Similarly, due to Lemma 3.2, the *lower set* of A given by

$$m(A) := \{x \in \mathbb{R} : \, x \leq a, \, \forall \, a \in A\}$$

has a maximum, at least if it is non-empty. As a consequence, we are led to the definition

Definition 3.5. *If* $m(A) \neq \emptyset$*, we set* $\inf(A) := \max m(A)$*. Otherwise, we set* $\inf(A) = -\infty$*.*

For instance

$$\inf(0, \infty) = \inf(0, 1) = \inf[0, 1) = \inf(0, 1] = \inf[0, 1] = 0, \quad \inf(-\infty, 0) = -\infty$$

and

$$\sup(0, 1) = \sup[0, 1) = \sup(0, 1] = \sup[0, 1] = \sup(-\infty, 1) = 0, \quad \sup(1, \infty) = \infty.$$

Remark 3.1. *Let* $A \subseteq \mathbb{R}$, $A \neq \emptyset$*. Then*

- *A has a minimum if and only if* $\inf(A) \in \mathbb{R}$*. In this case* $\min(A) = \inf(A)$*.*
- *A has a maximum if and only if* $\sup(A) \in \mathbb{R}$*. In this case* $\max(A) = \sup(A)$*.*

For example, the interval $I = (0, 1)$ *fulfills* $\sup I = 1 \notin I$*, hence it has no maximum. On the other hand, the interval* $I = (0, 1]$ *fulfills* $\sup I = 1 \in I$*, so that* $\max I = 1$*.*

Definition 3.6. *Let* $A \subseteq \mathbb{R}$, $A \neq \emptyset$*. We say that*

- *A is bounded below if* $\inf(A) \in \mathbb{R}$*.*
- *A is bounded above if* $\sup(A) \in \mathbb{R}$*.*

Remark 3.2. *We note that $A \subseteq \mathbb{R}$ is bounded if and only if it is both bounded below and bounded above.*

Once again, Remark 3.2 shows that the intervals $(0, 1)$, $[0, 1)$, $(0, 1]$ and $[0, 1]$ are all bounded, while the half-lines and the real line are unbounded. This justifies the names *open bounded interval* and *open unbounded interval* introduced above.

Moreover, any finite set $A := \{a_1, \ldots, a_n\} \subset \mathbb{N}$ is bounded and has both a maximum and a minimum.

The set
$$A := \left\{\frac{1}{n} : n \in \mathbb{N} \setminus \{0\}\right\}$$
fulfills $\sup(A) = \max(A) = 1$ and $\inf(A) = 0 \notin A$, hence A has no minimum. A

3.5. **Powers and roots of a number.** The notions of sup and inf can be used to define the roots of a real number. First, we define integer powers of a given real number.

Definition 3.7 (Integer powers)**.** *Let $x \in \mathbb{R}$ and let $n \in \mathbb{Z}$. Then we set*
$$x^n := \begin{cases} \underbrace{x \cdot \ldots \cdot x}_{n\ times} & if n > 0 \\ 1 & if n = 0,\ x \neq 0 \\ \frac{1}{x^{-n}} & if n < 0, x \neq 0. \end{cases}$$

For example,
$$2^2 = 4, \quad 2^{-3} = \frac{1}{2^3} = \frac{1}{8}, \quad (-2)^{-3} = \frac{1}{(-2)^3} = -\frac{1}{8}.$$

Now we introduce the square root. More precisely, for $y \geq 0$, set
$$\sqrt{y} := \sup\{x \in \mathbb{R} : 0 \leq x^2 \leq y\}. \tag{3.5}$$
Note that the square root is defined for $y \geq 0$ only.

Lemma 3.4. *In the above notation, we have $0 \leq (\sqrt{y})^2 = y$.*

Proof. Set $A_y := \{x \in \mathbb{R} : 0 \leq x^2 \leq y\}$. Since $0 \in A_y$, we have $\sqrt{y} = \sup(A_y) \geq 0$.

Let us first prove that $(\sqrt{y})^2 \leq y$. If we assume by contradiction that $(\sqrt{y})^2 > y$, then by definition of sup there exists $x \in A_y$ such that $x^2 > y$, which is impossible.

Now we will prove that $(\sqrt{y})^2 = y$. In order to do so, let us assume by contradiction that $a := (\sqrt{y})^2 < y$. As a consequence, by definition of sup, we can see that for any $x \in \mathbb{R}$ such that $0 \leq x^2 \leq y$, we have $x^2 \leq a$. This gives $a \geq y$, a contradiction. ☐

Remark 3.3. *Let $y \geq 0$. Then*
- $\sqrt{y} = 0 \Leftrightarrow y = 0$.
- $\sqrt{y} \geq 0,\ \forall\, y \geq 0$.
- $\sqrt{y} > 0,\ \forall\, y > 0$.
- *Let $y \geq 0$ and let $x \in \mathbb{R}$ such that $x^2 = y$. Then $x = \sqrt{y}$ or $x = -\sqrt{y}$.*

For example, $\sqrt{4} = 2$ and equation $x^2 - 4 = 0$ has exactly 2 real solutions, that is $\pm 2 = \pm\sqrt{4}$. Note that, in particular, $2 = \sqrt{(-2)^2}$.

This definition can be generalized to any even exponent as follows.

For $y \geq 0$ and $n \in \mathbb{N} \setminus \{0\}$, we set

$$\sqrt[2n]{y} := \sup\{x \in \mathbb{R} :\ 0 \leq x^{2n} \leq y\}.$$

Arguing as in Lemma 3.4, it is possible to see that $\sqrt[2n]{y} \geq 0$ and $(\sqrt[2n]{y})^{2n} = y$. The number $\sqrt[2n]{y}$ is known as $2n$-th root of y.

Remark 3.4. *We note that the conclusion of Remark 3.3 holds for all even roots. In particular*

Now we will focus on the definition of odd roots. First, we define the absolute value of y by setting

$$|y| := \begin{cases} y & \forall\, y \geq 0, \\ -y & \forall\, y < 0 \end{cases}$$

and the sign of y by setting

$$\mathrm{sgn}(y) := \begin{cases} 1 & \forall\, y \geq 0, \\ 0 & \text{for}\, y = 0, \\ -1 & \forall\, y < 0 \end{cases}$$

We stress that, in these notation, $y = \mathrm{sgn}(y)|y|$. Moreover, for $a > 0$, we have

$$|y| \geq a \Leftrightarrow (y \leq -a \vee y \geq a), \qquad |y| \leq a \Leftrightarrow -a \leq y \leq a.$$

For $y \in \mathbb{R}$ and $n \in \mathbb{N}$, we set

$$\sqrt[2n+1]{y} := \mathrm{sgn}(y) \sup\{x \in \mathbb{R} :\ 0 \leq x^{2n+1} \leq |y|\}. \tag{3.6}$$

Once again, arguing as the proof of Lemma 3.4, we can see that $(\sqrt[2n+1]{y})^{2n+1} = y$. The number $\sqrt[2n+1]{y}$ is known as the $(2n+1)$-th root of y and has the same sign as y. We note that, for any $y \in \mathbb{R}$, equation $x^{2n+1} = y$ has a unique solution given by $x = \sqrt[2n+1]{y}$. To sum up, we have the following facts.

Remark 3.5. *Let $n \in \mathbb{N}$. Then*

- *For any $y \in \mathbb{R}$ and $n \in \mathbb{N}$, the equation $x^{2n+1} = y$ has a unique solution given by $x = \sqrt[2n+1]{y}$. In particular, $\sqrt[2n+1]{0} = 0$.*
- *Assume that $n \neq 0$. Then for any $y \geq 0$, equation $x^{2n} = y$ has exactly 2 solutions given by $x = \pm \sqrt[2n]{y}$.*
- *Assume that $n \neq 0$. Then, for any $y \in \mathbb{R}$, we have $\sqrt[2n]{y^{2n}} = |y|$.*

Some examples: $\sqrt[3]{8} = 2$, $\sqrt[3]{-8} = -2$, $\sqrt{4} = 2$.

Once we have defined integer powers and roots, we can introduce fractional powers as follows.

Definition 3.8.
- *Let $y \geq 0$ be a real number $m \in \mathbb{Z}$ and $n \in \mathbb{N} \setminus \{0\}$. Then we set*

$$y^{\frac{m}{n}} := \begin{cases} \sqrt[n]{y^m} & if y > 0 \\ 0 & if y = 0,\ m > 0. \end{cases}$$

- *Let $y < 0$ be a real number $m \in \mathbb{Z}$ and $n \in \mathbb{N}$. Then we set $y^{\frac{m}{2n+1}} := \sqrt[2n+1]{y^m}$.*

For instance $4^{\frac{1}{2}} = \sqrt{4} = 2$ and $4^{-\frac{1}{2}} = (\sqrt{4})^{-1} = \frac{1}{\sqrt{4}} = \frac{1}{2}$. Using rational powers it is possible to define real powers. More precisely, given a real number $y \geq 0$ and an irrational number $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, $\alpha > 0$, we set

$$y^\alpha := \begin{cases} \sup\{y^q : q \in \mathbb{Q},\ q < \alpha\} & \text{if } y > 1 \\ \inf\{y^q : q \in \mathbb{Q},\ q < \alpha\} & \text{if } 0 < y \leq 1 \end{cases} \tag{3.7}$$

In particular, we note that $0^\alpha = 0$, for any $\alpha > 0$. Moreover, for $x \in \mathbb{R} \setminus \{0\}$ and $\alpha < 0$, we set $x^{-\alpha} := \frac{1}{x^\alpha}$.

3.6. **Properties of the powers.** In this section we will list the main properties of the powers, which are often used in computations.

(1) $\forall\, x, y \in (0, \infty)$, $\alpha \in \mathbb{R}$, $x^\alpha y^\alpha = (xy)^\alpha$.
(2) $\forall\, x \in (0, \infty)$, $\alpha, \beta \in \mathbb{R}$, $x^\alpha x^\beta = x^{\alpha+\beta}$.
(3) $\forall\, x \in (0, \infty)$, $\alpha, \beta \in \mathbb{R}$, $(x^\alpha)^\beta = x^{\alpha\beta}$.
(4) $\forall\, x \in (0, \infty)$, $\alpha \in \mathbb{R}$, $x^{-\alpha} = \frac{1}{x^\alpha}$.
(5) $\forall\, x \in (0, \infty)$, $\alpha, \beta \in \mathbb{R}$, $\frac{x^\alpha}{x^\beta} = x^{\alpha-\beta}$.
(6) $\forall\, x, y \in (0, \infty)$, $\alpha \in \mathbb{R}$, $\frac{x^\alpha}{y^\alpha} = (\frac{x}{y})^\alpha$.
(7) $\forall\, x \in \mathbb{R}$, $x^0 = 1$ and $x^1 = x$.
(8) $\forall\, x \in (0, 1) \cup (1, \infty)$, $\alpha, \beta \in \mathbb{R}, x^\alpha = x^\beta \Leftrightarrow \alpha = \beta$.
(9) $\forall\, x \in (1, \infty)$, $\alpha, \beta \in \mathbb{R}$, we have

$$x^\alpha > x^\beta \Leftrightarrow \alpha > \beta$$
$$x^\alpha \geq x^\beta \Leftrightarrow \alpha \geq \beta$$
$$x^\alpha < x^\beta \Leftrightarrow \alpha < \beta$$
$$x^\alpha \leq x^\beta \Leftrightarrow \alpha \leq \beta.$$

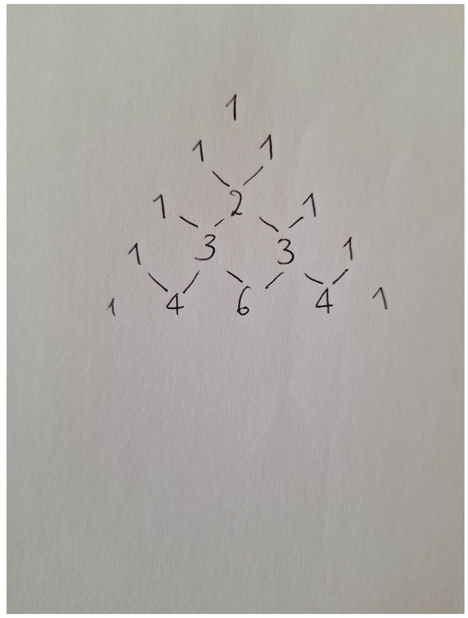

Figure 3. Tartaglia triangle: every element is obtained by taking the sum of the two upper elements

(10) $\forall\, x \in (0,1)$, $\alpha, \beta \in \mathbb{R}$, we have

$$\begin{aligned} x^\alpha > x^\beta &\Leftrightarrow \alpha < \beta \\ x^\alpha \geq x^\beta &\Leftrightarrow \alpha \leq \beta \\ x^\alpha < x^\beta &\Leftrightarrow \alpha > \beta \\ x^\alpha \leq x^\beta &\Leftrightarrow \alpha \geq \beta. \end{aligned}$$

(11) Powers of a binomial: $\forall\, x, y \in \mathbb{R}$, we have

$$\begin{aligned} (x+y)^0 &= 1 \\ (x+y)^1 &= x+y \\ (x+y)^2 &= x^2+2xy+y^2 \\ (x+y)^3 &= x^3+3x^2y+3xy^2+y^3 \end{aligned}$$

(12) Square of a sum of $k \geq 3$ terms: $\forall\, x_1, \dots, x_k \in \mathbb{R}$, we have

$$(\sum_{i=1}^k x_i)^2 = \sum_{i=1}^k x_i^2 + 2 \sum_{1 \leq i < j \leq k} x_i x_j.$$

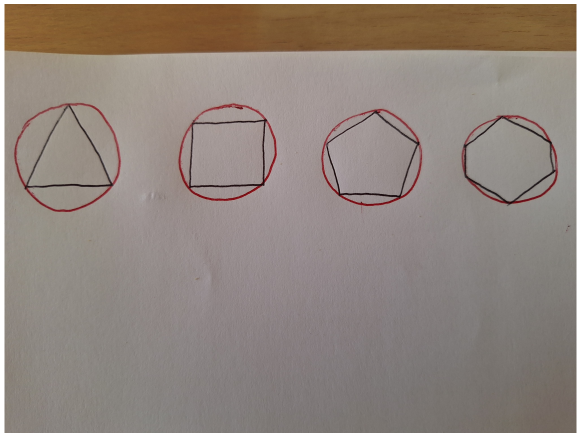

FIGURE 4. Definition of π: the Perimeter of the polygon P_n of n edges approximates, for n large, the length of the circle

For example, for $k = 3$ the latter reduces to

$$(x_1 + x_2 + x_3)^2 = x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3.$$

3.7. **Some famous irrational numbers.** The notion of sup enables us to define some famous irrational numbers such as e and π. More precisely, we have

- The Nepero number e is defined by
$$e := \sup\left\{\left(1 + \frac{1}{n}\right)^{\frac{1}{n}} : n \in \mathbb{N} \setminus \{0\}\right\}.$$
A numeric computation shows that $e = 2,7182....$ This number is well known, since it is the basis of natural logarithms (see below).
- The number π is the ratio between the length of the circle and its diameter. The length of the circle ℓ can be characterized through the notion of sup as follows:
$$\ell = \sup\{\mathrm{Per}(P_n) : n \in \mathbb{N} \setminus \{0, 1, 2\}\},$$
where Per denotes the Perimeter and P_n is a regular polygon of $n \geq 3$ edges contained in the circle with vertices on the circle itself.

3.8. **Logarithms.** Given two real numbers $a, x \in \mathbb{R}$, with $a > 0$, by (3.7) we can define $y := a^x$. Moreover, the following existence and uniqueness result holds.

Theorem 3.3. *Let $a > 0$, $a \neq 1$ be fixed. Then for any $y > 0$, there exists a unique $x \in \mathbb{R}$ such that $y = a^x$.*

In view of this Theorem, we can introduce the following Definition.

Definition 3.9. *Given $a > 0$, $a \neq 1$, $y > 0$, the number $x \in \mathbb{R}$ found in Theorem 3.3 is denoted by $x := \log_a y$ and known as the logarithm in basis a of y.*

From Definition 3.7, it follows that, if $a > 1$, we have

$$x_1 > x_2 \Leftrightarrow a^{x_1} > a^{x_2}, \tag{3.8}$$

as a consequence, for any y_1, $y_2 \in (0, \infty)$, we have

$$y_1 > y_2 \Leftrightarrow \log_a y_1 > \log_a y_2. \tag{3.9}$$

Note that, if $a \in (0, 1)$, in (3.8) and (3.9) the opposite inequality hold.

There are some relevant cases of logarithms, which are very important in the applications. For example, when $a = 10$, we use the notation $\operatorname{Log} y := \log_{10} y$.

If $a = e$ is the Nepero number, we use the notation $\log x := \ln x := \log_e x$, and we call it *natural logarithm.*

Remark 3.6.

- *The Richter scale is often used to measure the magnitude of an earthquake. The formula for computing the magnitude is the following*

$$M := \operatorname{Log} A - \operatorname{Log} A_0,$$

where M the magnitude, A is the amplitude of seismic waves and A_0 is the amplitude of a fixed model wave. This means that the larger the amplitude of the oscillation, the larger is the magnitude. This makes sense since it says that, for example, an earthquake of magnitude 8 is stronger than a one of magnitude 4. In any case, note that the amplitude is not proportional to the magnitude.

- *The* ph *of a solution is defined by* $\mathrm{ph} := -\operatorname{Log}(H_+)$*, where H_+ represents the concentration of hydrogen ions in the solution. Solutions are considered neutral if they have* $\mathrm{ph} = 7$*, acid if they have* $\mathrm{ph} < 7$ *or basic if they have* $\mathrm{ph} > 7$*. Due to the presence of the $-$ sign, we can see that the solutions are acid if the concentration of hydrogen ions is high and basic if such a concentration is law.*

Now we list some properties of the logarithms that are very useful in the applications. Namely, $\forall\, a > 0$, $a \neq 1$, we have

(1) $\log_a 1 = 0$ and $\log_a a = 1$.
(2) $\forall\, x \in \mathbb{R}$, $\log_a a^x = x$.
(3) $\forall\, y > 0$, $a^{\log_a y} = y$.
(4) $\forall\, y_1$, $y_2 > 0$, $\log_a(y_1 y_2) = \log_a y_1 + \log_a y_2$.
(5) $\forall\, y > 0$, $\log_a \frac{1}{y} = -\log_a y$.
(6) $\forall\, y_1$, $y_2 > 0$, $y_2 \neq 0$, $\log_a \frac{y_1}{y_2} = \log_a y_1 - \log_a y_2$.

(7) $\forall\, \alpha > 0,\, y > 0,\, \log_a(y^\alpha) = \alpha \log_a y$.
(8) $\forall\, b > 0,\, b \neq 1,\, \log_a b = \frac{1}{\log_b a}$.
(9) $\forall\, y,\, b > 0,\, b \neq 1,\, \log_a y = \frac{\log_b y}{\log_b a}$

In view of (3.8), for $a > 1$, $x \in \mathbb{R}$ and $y > 0$ we have

$$\begin{aligned} a^x \le y &\Leftrightarrow x \le \log_a y \\ a^x \ge y &\Leftrightarrow x \ge \log_a y \\ a^x < y &\Leftrightarrow x < \log_a y \\ a^x > y &\Leftrightarrow x > \log_a y. \end{aligned} \tag{3.10}$$

If $0 < a < 1$, $x \in \mathbb{R}$, $y > 0$ the inequalities are reversed, in the sense that we have

$$\begin{aligned} a^x \le y &\Leftrightarrow x \ge \log_a y \\ a^x \ge y &\Leftrightarrow x \le \log_a y \\ a^x < y &\Leftrightarrow x > \log_a y \\ a^x > y &\Leftrightarrow x < \log_a y. \end{aligned} \tag{3.11}$$

4. Functions

Definition 4.1. *Let X and Y be two sets. A function $f : X \to Y$ is a law that associates a unique element $y \in Y$ with each element $x \in X$. The set X is known as the domain of f.*

For any $x \in X$, the unique element $y \in Y$ associated to x by f is denoted by $f(x)$ and known as "the value of f at x", or "the image of x under the function f", or simply "f of x".

The set

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

is known as *image* of f, or image of X through f, and is also indicated by $Im(f)$. In other words, the image of f is the set of images of all points $x \in X$.

Note that $f(X) \subseteq Y$, but it is not always true that $f(X) = Y$ (see the examples (1), (2) and (4), (5) above).

On the other hand, given a subset $Z \subseteq Y$, we define the *inverse image* of Z through f by

$$f^{-1}(Z) := \{x \in X : f(x) \in Z\}. \tag{4.1}$$

Remark 4.1.
- *It follows from the definition of inverse image of a set that $f(f^{-1}(Z)) \subseteq Z$, but the equality is not necessarily true (see the examples below).*

- *If $W \subset X$, then $W \subseteq f^{-1}(f(W))$, but the equality is not necessarily true (see the examples below).*

Given a function $f : X \to Y$, the graph of f is the subset of the cartesian product $X \times Y$ given by

$$\Gamma(f) := \{(x, y) : x \in X,\ y = f(x)\}.$$

Examples (see figure (5))

(1) Let $X = \{\text{football players of the Inter}\}$. Then

$$\begin{aligned} f : X &\to \mathbb{N} \\ x &\mapsto \text{number of}\, x \,\text{on the t-shirt} \end{aligned} \tag{4.2}$$

is a function. In fact, every football player has a unique number on the t-shirt.

(2) Let $X := \{\text{Italy, Indonesia, France, Portugal}\}$ and $Y := \{\text{letters of the english alphabet}\}$. Then

$$\begin{aligned} f : X &\to Y \\ x &\mapsto \text{initial letter of}\, x \end{aligned} \tag{4.3}$$

is a function. In fact, every country has a unique initial letter in English.

Moreover, if $Z := \{a, i\}$, then $f^{-1}(Z) = \{\text{Italy, Indonesia}\}$, thus

$$f(f^{-1}(Z)) = \{i\} \subsetneq Z.$$

Furthermore, if $W := \{\text{Italy}\}$, then $f(W) = \{i\}$, hence

$$W \subsetneq f^{-1}(f(W)) = \{\text{Italy, Indonesia}\}.$$

(3) Let $X := \mathbb{R}$, $Y := [0, \infty)$. Then

$$\begin{aligned} f : X &\to Y \\ x &\mapsto x^2 \end{aligned} \tag{4.4}$$

is a function.

(4) Let $X = Y := \mathbb{R}$. Then

$$\begin{aligned} f : X &\to Y \\ x &\mapsto x^3 \end{aligned} \tag{4.5}$$

is a function.

(5) Let $X := \mathbb{R}$ and $Y := \mathbb{R}^2$. Then

$$\begin{aligned} f : X &\to Y \\ x &\mapsto (x^2, x^4 - 1) \end{aligned} \tag{4.6}$$

is a function. Note that

$$f(\mathbb{R}) = \{(x^2, x^4 - 1) : x \in \mathbb{R}\} = \{(t, t^2 - 1) : t \in [0, \infty)\} = \Gamma(g),$$

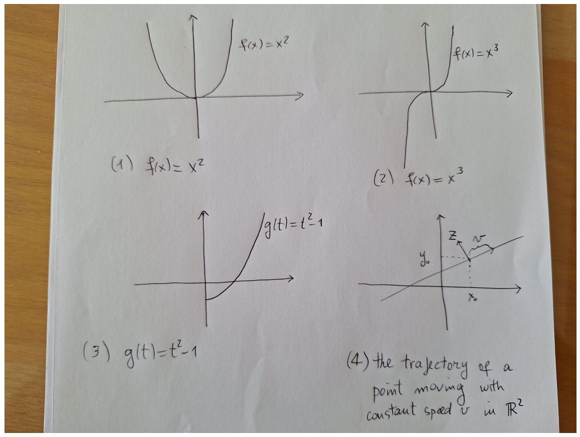

FIGURE 5. Starting from the left we have (1) the graph of the function $f(x) = x^2$, (2) the graph of the function $f(x) = x^3$, (3) the image of the function f defined in 4.6 and (4) the trajectory of a point moving on the plane with constant speed v.

where $g : [0, \infty) \to \mathbb{R}$ is defined by $g(t) := t^2 - 1$.

(6) A car is moving in the plane $\mathbb{R}^2$ with constant speed $v := (v_2, v_2) \in \mathbb{R}^2 \setminus \{(0,0)\}$ and initial position $(x_0, y_0) \in \mathbb{R}^2$. Then the law

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R}^2 \\ t &\mapsto (x_0 + tv_1, y_0 + tv_2) \end{aligned} \tag{4.7}$$

which describes the motion of the point is a function. More precisely, it says that, at time t, we will find the car is at point $f(t) \in \mathbb{R}^2$.

Now we will introduce the notions of *injective* and *surjective* function.

Let X, Y be sets and let $f : X \to Y$ be a function. We recall that the image of f is defined by

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

Definition 4.2. *We say that*

- *f is injective if for any $y \in f(X)$, there exists a unique $x \in X$ such that $y = f(x)$.*
- *f is surjective if $f(X) = Y$.*
- *f is bijective if it is injective and surjective.*

Now we will see which of the functions defined in the above examples are injective or surjective.

(1) The function defined in (4.2) is injective but not surjective. In fact there are not two players with the same number on the t-shirt and not all the positive

integer numbers are assigned to a player of the inter. For example, there is no player number 110089 in the Inter.

(2) The function defined in (4.3) is neither injective not surjective. In fact $f(\text{Italy}) = f(\text{Indonesia}) = i$ and $f(X) = \{i, f, p\} \subsetneq \{\text{letters of the alphabet}\}$.
(3) Due to Remark (3.5), the function defined in (4.4) surjective but not injective.
(4) Due to Remark (3.5), the function defined in (4.5) is both injective and surjective.
(5) The funcition defined in (4.6) is neither injective nor surjective. In fact we have $f(x) = f(-x)$, for any $x \in \mathbb{R}$, and $(-1, -2) \notin f(\mathbb{R})$, since $x^2 \geq 0$ and $x^4 - 1 \geq -1$, for any $x \in \mathbb{R}$, due to the properties of the powers.
(6) The function defined in (4.7) is injective, in fact

$$\begin{cases} x_0 + tv_1 = x_0 + sv_1 \\ y_0 + tv_2 = y_0 + sv_2 \end{cases} \Rightarrow t = s,$$

but it is not surjective. In fact the point $z := (x_0 + v_2, y_0 - v_1) \notin f(\mathbb{R})$. In order to prove that, we first note that, since $v \neq 0$, we can assume without loss of generality that $v_1 \neq 0$. If we assume by contradiction that there exists $t \in \mathbb{R}$ such that $f(t) = (x_0 + v_2, y_0 - v_1)$, then we have

$$\begin{cases} x_0 + tv_1 = x_0 + v_2 \\ y_0 + tv_2 = y_0 - v_1 \end{cases}$$

This yields that $0 \neq -v_1 = tv_2 = t^2 v_1$, which implies that $t^2 = -1$, a contradiction.

More precisely, the image of f is the straight line passing through (x_0, y_0) of direction v and the point z does not belong to this line.

Let X, Y, Z, W be sets such that $Z \subset Y$. Let $f : X \to Y$ and $g : Z \to W$ be functions. We define the *composed function*

$$g \circ f : f^{-1}(Z) \to W$$

by setting $g \circ f(x) := g(f(x))$. If, for example, $X = Y = W = \mathbb{R}$, $Z = [0, \infty)$, $f(x) = x + 1$ and $g(z) := \sqrt{z}$, then

$$\begin{aligned} g \circ f : [-1, \infty) &\to \mathbb{R} \\ x &\mapsto \sqrt{x+1}. \end{aligned}$$

Given a function $f : X \to Y$ and a subset $A \subseteq X$, we define the restriction of f to A as

$$\begin{aligned} f\big|_A : A &\to Y \\ x &\mapsto f(x). \end{aligned}$$

This notion will be useful in the sequel.

Definition 4.3.
- *Given a set X, we define the identity function*
$$Id_X : X \to X$$
$$x \mapsto x.$$
- *Given a function $f : X \to Y$, we say that f is invertible if there exists a function $g : Y \to X$ such that $g \circ f = Id_X$.*

It is possible to see that, if f is invertible, then the inverse function g is unique. This function g is denoted by f^{-1} and is also invertible with inverse f, that is, it satisfies $f \circ f^{-1} = Id_Y$.

Theorem 4.1. *Let $f : X \to Y$ be a function. Then f is invertible if and only if f is bijective.*

Proof. If f is bijective, then in particular it is surjetive, hence for any $y \in Y$ there exists $x \in X$ such that $y = f(x)$. Moreover, being f injective, this element $x \in X$ is unique. As a consequence, setting $x := g(y)$ we have defined the inverse function.

On the other hand, if f is invertible, then
$$f(x_1) = f(x_2) \Rightarrow f^{-1} \circ f(x_1) = f^{-1} \circ f(x_2) \Leftrightarrow x_1 = x_2,$$
which yields that f is injective. In addtion, for any $y \in Y$, we have $y = f \circ f^{-1}(y) = f(x)$, where we have set $x := f^{-1}(y) \in X$. As a consequence, f is surjective too. □

For example,
- due to Remark 3.5, the function $f(x) := x^3$ defined in (4.5) is invertible with inverse function $f^{-1} : \mathbb{R} \to \mathbb{R}$ given by $f^{-1}(y) = \sqrt[3]{y}$.
- The function $f(x) := x + 1$ is invertible with inverse $f^{-1}(y) = y - 1$. In fact we have
$$y = x + 1 \Leftrightarrow x = y - 1.$$
- The function $f(x) := x^2$ is not invertible since it is not injective. In fact $f(-2) = f(2) = 4$.

Remark 4.2. *Let $f : X \to Y$ be an injective function. Then $f : X \to f(X)$ is invertible, since it is also surjective.*

Given $X, Y \subseteq \mathbb{R}$ and an invertible function $f : X \to Y$, the graph of f^{-1} can be deduced by reflecting the graph of f with respect to the straight line $y = x$ (see for example figures (6, 7)).

For example, the function
$$f : [0, \infty) \to \mathbb{R}$$
$$x \mapsto x^2$$

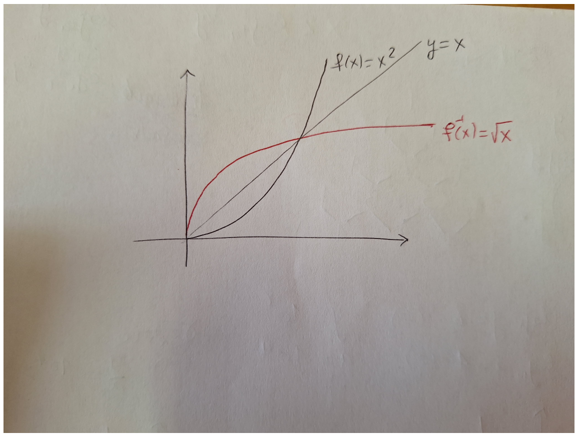

FIGURE 6. The function $f(x) := x^2$ and its inverse $f^{-1}(x) := \sqrt{x}$

is injective but not invertible, since it is not surjective. However, the function

$$\begin{aligned} f : [0, \infty) &\to [0, \infty) \\ x &\mapsto x^2 \end{aligned}$$

is also surjective, hence it is invertible with inverse

$$\begin{aligned} f^{-1} : [0, \infty) &\to [0, \infty) = f([0, \infty)) \\ x &\mapsto \sqrt{x}. \end{aligned}$$

See figure (6).

The function

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto e^x \end{aligned}$$

is injective but not invertible, since it is not surjective. However, the function

$$\begin{aligned} f : \mathbb{R} &\to (0, \infty) \\ x &\mapsto e^x \end{aligned}$$

is invertible, with inverse function given by

$$\begin{aligned} f^{-1} : (0, \infty) &\to \mathbb{R} = f(\mathbb{R}) \\ x &\mapsto \log(x). \end{aligned}$$

See figure (7).

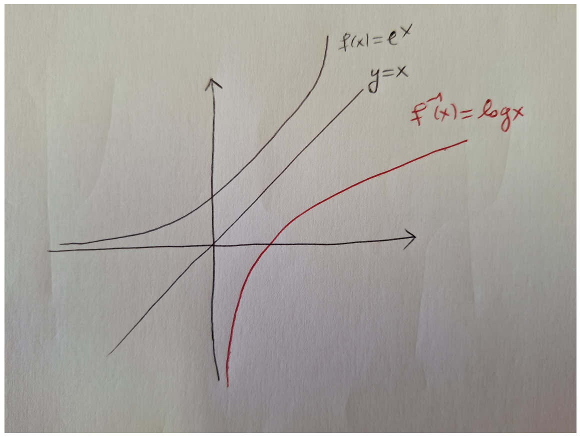

FIGURE 7. The function $f(x) := e^x$ and its inverse $f^{-1}(x) := \log(x)$.

4.1. **Polynomials.** Very relevant examples of functions are polynomials.

Definition 4.4. *Let $n \in \mathbb{N}$. Then the function $p : \mathbb{R} \to \mathbb{R}$ is a polynomial of degree n if there exist $a_0, \dots, a_n \in \mathbb{R}$, $a_n \neq 0$ such that*

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n = \sum_{i=0}^{n} a_i x^i.$$

In particular, the polynomials of degree 0 are constant functions and the polynomials of degree 1 are functions whose graphs are straight lines.

In the sequel, we will be interested to determine the roots of a given polynomial, that the points $x \in \mathbb{R}$ such that $p(x) = 0$. The first step is to to characterize such a set for polynomials of degree 1.

Theorem 4.2. *Let $p(x) := a_0 + a_1 x$ be a polynomial of degree 1. Then $p(x) = 0$ if and only if $x = -\frac{a_0}{a_1}$.*

Proof. Note that, since $a_1 \neq 0$, we can divide by a_1 and deduce that

$$p(x) = 0 \Leftrightarrow a_1 x = -a_0 \Leftrightarrow x = -\frac{a_0}{a_1}.$$

□

The situation becomes more involved when dealing with polynomials of degree 2. In this case, the existence of real roots of the polynomial

$$p(x) := a_0 + a_1 x + a_2 x^2$$

actually depends the sign of $\Delta := a_1^2 - 4a_2 a_0$.

Theorem 4.3. *Let $p(x) := a_0 + a_1x + a_2x^2$ be a polynomial of degree 2.*

(1) *If $\Delta > 0$, then p has exactly 2 real roots given by*
$$x_\pm := \frac{-a_1 \pm \sqrt{a_1^2 - 4a_0a_2}}{2a_2}. \tag{4.8}$$
(2) *If $\Delta = 0$, then p has a unique real root given by $x = -\frac{a_1}{2a_2}$.*
(3) *If $\Delta < 0$, then p has no real roots.*

Proof. Using that $a_2 \neq 0$ and multiplying, if necessary, by -1, we can assume that $a_2 > 0$. Dividing by a_2, it is possible to see that $p(x) = 0$ if and only if
$$\begin{aligned} 0 &= x^2 + \frac{a_1}{a_2}x + \frac{a_0}{a_2} \\ &= x^2 + 2\frac{a_1}{2a_2}x + \frac{a_1^2}{4a_2^2} + \frac{a_0}{a_2} - \frac{a_1^2}{4a_2^2} \\ &= \left(x + \frac{a_1}{2a_2}\right)^2 - \frac{\Delta}{4a_2^2}, \end{aligned} \tag{4.9}$$
or equivalently $p(x) = 0$ if and only if $\left(x + \frac{a_1}{2a_2}\right)^2 = \frac{\Delta}{4a_2^2}$. As a consequence, we see that the equation $p(x) = 0$ has no real solution if $\Delta < 0$. If $\Delta = 0$, we have $x + \frac{a_1}{2a_2} = 0$, that is $x = -\frac{a_1}{2a_2}$. Finally, if $\Delta > 0$, taking the square root of both sides, we have
$$\left|x + \frac{a_1}{2a_2}\right| = \frac{\sqrt{\Delta}}{2a_2} \Leftrightarrow \left(x + \frac{a_1}{2a_2} = \frac{\sqrt{\Delta}}{2a_2} \vee x + \frac{a_1}{2a_2} = -\frac{\sqrt{\Delta}}{2a_2}\right),$$
or equivalently
$$x = \frac{-a_1 + \sqrt{a_1^2 - 4a_0a_2}}{2a_2} \vee x = \frac{-a_1 + \sqrt{a_1^2 - 4a_0a_2}}{2a_2}$$
□

Sometimes relation (4.8) is written in the form
$$x_\pm = \frac{-\frac{a_1}{2} \pm \sqrt{\left(\frac{a_1}{2}\right)^2 - a_0a_2}}{a_2}.$$

For example the solutions to
$$x^2 - 6x + 8 = 0$$
are given by $x_\pm = \frac{3 \pm \sqrt{9-8}}{1}$, that is $x_- = 2$ and $x_+ = 4$.

The equation
$$x^2 - 6x + 10 = 0$$
has no real solutions, since $\frac{\Delta}{4} = 9 - 10 = -1 < 0$.

The equation
$$x^2 - 2x + 1 = 0$$
has a unique solution $x = 1$, since $\frac{\Delta}{4} = 1 - 1 = 0$. In fact, $x^2 - 2x + 1 = (x-1)^2$.

4.2. **Polynomial inequalities.** In this section, given a polynomial p, we are interested in the following questions:

- For which values of x do we have $p(x) > 0$?
- For which values of x do we have $p(x) < 0$?

4.2.1. *Degree* 1. Let us first assume that $p(x) = a_0 + a_1 x$ is a polynomial of degree 1. Let us assume first that $a_1 > 0$. Then we have
$$p(x) > 0 \Leftrightarrow a_1 x > -a_0 \Leftrightarrow x > -\frac{a_0}{a_1}.$$
On the other hand, if $a_1 < 0$, we have
$$p(x) > 0 \Leftrightarrow a_1 x > -a_0 \Leftrightarrow x < -\frac{a_0}{a_1}.$$
In fact, when we multiply both sides of the inequality by a negative number, the inequality itself is reversed. This can be seen by considering the graph of the function $p(x) := a_0 + a_1 x$ in both cases. Similarly, we can treat the inequality $p(x) < 0$ (see figure (8)).

Similarly, if, for example, $a_1 > 0$ we can see that
$$p(x) \geq 0 \Leftrightarrow x \geq -\frac{a_0}{a_1}, \quad p(x) \leq 0 \Leftrightarrow x \leq -\frac{a_0}{a_1} \tag{4.10}$$
and, if $a_1 < 0$, we have
$$p(x) \geq 0 \Leftrightarrow x \leq -\frac{a_0}{a_1}, \quad p(x) \leq 0 \Leftrightarrow x \geq -\frac{a_0}{a_1} \tag{4.11}$$
For example
$$\begin{aligned}
&2x + 3 > 0 \Leftrightarrow x > -\frac{3}{2}, \\
&4x + 8 < 0 \Leftrightarrow x < 2. \\
&-3x + 2 > 0 \Leftrightarrow x < \frac{2}{3}, \\
&-2x + 4 < 0 \Leftrightarrow x > 2.
\end{aligned} \tag{4.12}$$

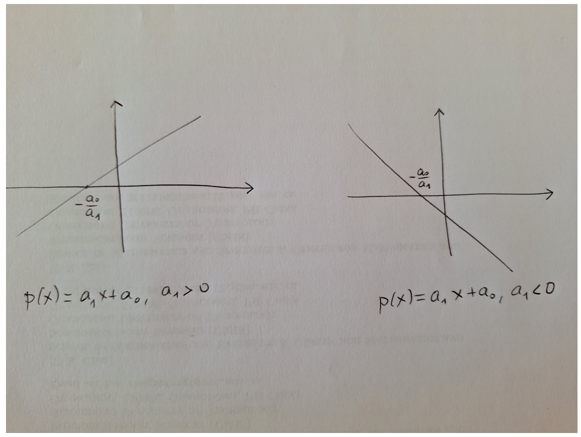

FIGURE 8. The graph of a polynomial of degree 1 is a straight line. It can be used to solve the inequality $p(x) > 0$ and $p(x) < 0$.

Remark 4.3 (Trick to simplify the computations)**.** *In order to simplify the computations, up to changing the inequality one can always assume that* $a_1 > 0$ *before dividing. For example*

$$-3x + 2 > 0 \Leftrightarrow 3x - 2 < 0 \Leftrightarrow x < \frac{2}{3}.$$

4.2.2. *Degree* 2. Given a polynomial $p(x) = a_0 + a_1x + a_2x^2$ of degree 2, by Theorem 4.3, the existence of solutions to the equation $p(x) = 0$ depends on the sign of Δ.

If, for example, $a_2 > 0$, we have:

- if $\Delta > 0$, the polynomial has exactly 2 roots $x_- < x_+$ and
$$p(x) < 0 \Leftrightarrow x_- < x < x_+, \qquad p(x) > 0 \Leftrightarrow (x < x_- \vee x > x_+).$$
- if $\Delta = 0$, then p vanishes if and only if $x = x_+ = x_- = -\frac{a_1}{2a_2}$ and $p(x) = a_2(x + \frac{a_1}{2a_2})^2 > 0$ for $x \neq -\frac{a_1}{2a_2}$.
- if $\Delta < 0$, then $p(x) > 0$ for any $x \in \mathbb{R}$.

This can be deduced from the graphs in figure (9).

If, $a_2 < 0$, then $p(x) > 0$ if and only if $-p(x) < 0$, which is equivalent to say that

- if $\Delta > 0$, the polynomial has exactly 2 roots $x_+ < x_-$ and
$$p(x) > 0 \Leftrightarrow x_+ < x < x_-, \qquad p(x) < 0 \Leftrightarrow (x < x_+ \vee x > x_-).$$
- if $\Delta = 0$, then p vanishes if and only if $x = x_+ = x_- = -\frac{a_1}{2a_2}$ and $p(x) = a_2(x + \frac{a_1}{2a_2})^2 < 0$ for $x \neq -\frac{a_1}{2a_2}$.

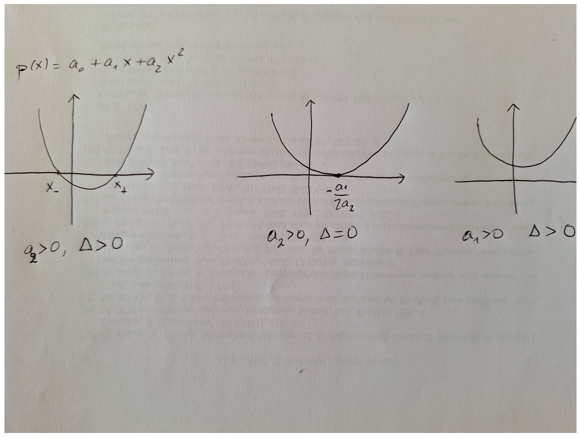

FIGURE 9. Graph of the polynomial $p(x) := a_0+a_1x+a_2x^2$ for $a_2 > 0$. They can be deduce by translating and dilation of the graph of the polynomial $y = x^2$.

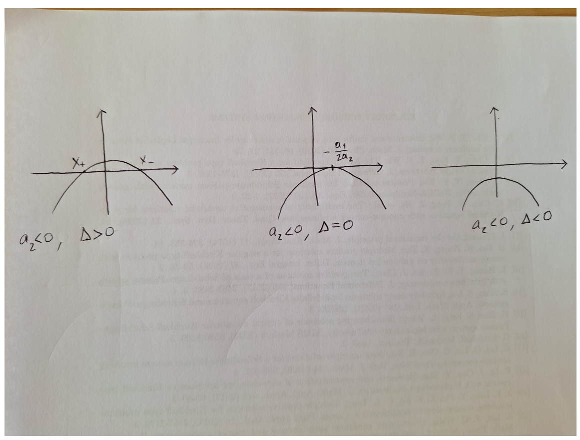

FIGURE 10. Graph of the polynomial $p(x) := a_0 + a_1x + a_2x^2$ for $a_2 < 0$. They can be deduce by translating and dilation of the graph of the polynomial $y = -x^2$.

- if $\Delta < 0$, then $p(x) < 0$ for any $x \in \mathbb{R}$.

This can be deduced from the graphs in figure (10).

4.3. **Higher degree polynomials.** Sometimes it is useful to be able to find the real roots of higher degree polynomials. For this reason, it is useful to keep in mind the following theorem.

Theorem 4.4 (Factorization). *Let p be a polynomial of degree $n \geq 3$ with real coefficients. Then p can be written as the product of at most n polynomials with real coefficients, each of which of degree 1 or 2.*

In any case, it is generally not easy to find such a decomposition for a given polynomial.

Such a factorization is crucial when dealing with polynomial equations or inequalities.

4.3.1. *Polynomial equations of degree $n \geq 3$.* In this section we will give the outlines of the general strategy to solve polynomial equations of degree $n \geq 3$, that is of the form $p(x) = 0$, where p is a polynomial of degree $n \geq 3$. We will see that the first step is to factorize the polynomial, which is always possible thank to Theorem 4.4.

More precisely, by Theorem 4.4, there exist $m, s \in \{1, \ldots, n\}$, $k_1, \ldots, k_m \in \{1, \ldots, n\}$, $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ and polynomials $q_1, \ldots, q_s$ of degree 2 such that the polynomial p can be written in the form

$$p(x) = (x - \alpha_1)^{k_1} \ldots (x - \alpha_m)^{k_m} q_1(x) \ldots q_s(x). \tag{4.13}$$

Without loss of generality, up to changing the exponents k_i, we can assume that $\alpha_i \neq \alpha_j$ if $i \neq j$. Moreover, we can assume that the polynomials q_i have $\Delta < 0$, for any $1 \leq i \leq s$. In fact, if we assume that there exists $i \in \{1, \ldots, s\}$ such that q_i has $\Delta \geq 0$, then there exist β, α_{m+1}, $\alpha_{m+2} \in \mathbb{R}$ such that $q_i(x) = \alpha(x - \alpha_{m+1})(x - \alpha_{m+2})$, hence we reduce ourselves to an expression which is analogue to (4.13) with two additional terms of order 1, that is with m replaced by $m + 2$.

By using (4.13) and recalling that the polynomials of degree 2 with $\Delta < 0$ do not have real roots, we can see that the real roots of our polynomial p are $\alpha_1, \ldots, \alpha_m$. The integer k_i is known as the *algebraic multiplicity* or simply *multiplicity* of the root α_i, for $1 \leq i \leq m$.

The most difficult part is to find a decomposition of the form (4.13) of a given polynomial.

(1) The polynomial $p(x) = x^3 - 2x^2 - x + 2$ can be factorized as

$$p(x) = x^2(x - 2) - (x - 2) = (x^2 - 1)(x - 2) = (x - 1)(x + 1)(x - 2).$$

This yields that

$$p(x) = 0 \Leftrightarrow (x - 1)(x + 1)(x - 2) = 0,$$

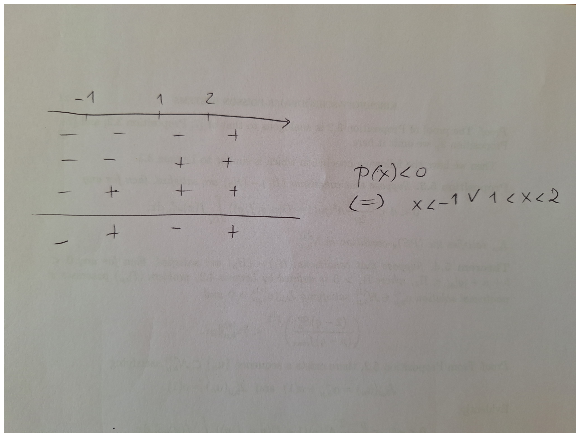

FIGURE 11. Graphic representation of the inequality $x^3 - 2x^2 - x + 2 < 0$

hence the solutions are $x = 1,\ x = -1,\ x = 2$. All the roots have multiplicity 1.

(2) The polynomial $p(x) = x^3 - x^2 - x + 1$ can be factorized as

$$p(x) = x^2(x-1) - (x-1) = (x-1)(x^2-1) = (x-1)^2(x+1),$$

hence the roots are 1 and -1. Note that 1 has multiplicity 2 and -1 has multiplicity 1.

(3) The polynomial $p(x) = x^3 - 2x^2 + x - 2$ can be factorized as

$$p(x) = x^2(x-2) + (x-2) = (x^2+1)(x-2),$$

therefore it has 1 real root only, that is $x = 2$, since the polynomial $x^2 + 1$ has $\Delta < 0$ and hence no real root.

4.3.2. *Polynomial inequalities of degree* $n \geq 3$. Polynomial inequalities are expressions of the form

$$p(x) < 0,\ p(x) \leq 0,\ p(x) > 0,\ p(x) \geq 0,$$

where p is a polynomial of degree n. If $n \leq 2$, we have already seen how to deal with such inequalities. Here we focus on the case $n \geq 3$. Once again we factorize (see Theorem 4.4) the polynomial and observe that $p(x) < 0$ if and only p has an odd number of negative factors.

For example, recalling the above examples we have

(1) $p(x) = x^3 - 2x^2 - x + 2 < 0$ if and only if $(x+1)(x-1)(x-2) < 0$, that is if and only if $x < -1 \vee 1 < x < 2$ (see figure (11)).

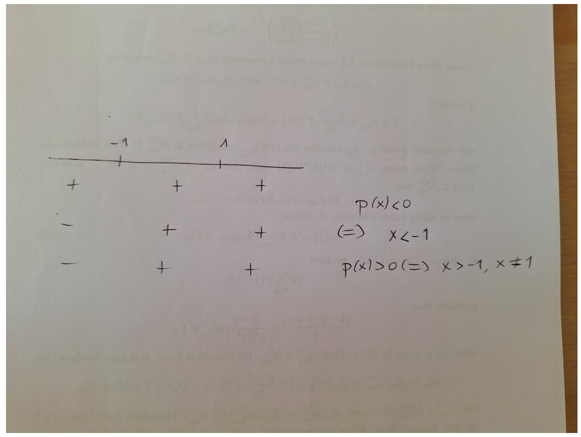

FIGURE 12. Graphic representation of the inequality $x^3 - x^2 - x + 1 < 0$

(2) $p(x) = x^3 - x^2 - x + 1 < 0$ if and only if $(x-1)^2(x+1) < 0$, that is if and only if $x < -1$. On the other hand, $p(x) > 0$ if and only if $x > -1 \wedge x \neq 1$ (see figure (12)).
(3) $p(x) = x^3 - 2x^2 + x - 2 < 0$ if and only if $(x^2+1)(x-2) > 0$, that is if and only if $x < 2$. On the other hand, $p(x) > 0$ if and only if $x > 2$. In fact, we have $x^2 + 1 > 0$ for any $x \in \mathbb{R}$.

4.4. **Rational functions.** Let p, $q : \mathbb{R} \to \mathbb{R}$ be polynomials, with q nonconstant. Due to Theorem 4.4, both p and q can be written as the product of polynomials of order 1 or 2. Let us assume that p and q have no common factors. Let

$$Z_q := \{x \in \mathbb{R} : \, q(x) = 0\}$$

be the set of roots of q. Then the function

$$\begin{aligned} f : \mathbb{R} \setminus Z_q &\to \mathbb{R} \\ x &\mapsto \frac{p(x)}{q(x)} \end{aligned} \tag{4.14}$$

is known as a *rational function*. In other words, a function is said to be rational if it is the ratio between two polynomials. Its domain is given by the set where the denominator does not vanish.

For example, the function $f(x) := \frac{x^2+4x+3}{x-2}$ is a rational function and its domain is $\mathbb{R} \setminus \{2\}$. However, the notation can be misleading, because sometimes it is possible to simplify. In fact, the function $f(x) := \frac{x^2+4x+3}{x+1}$ is actually not a rational function, since $x^2 + 4x + 3 = (x+1)(x+3)$, so that $f(x) = x + 3$ for $x \neq -1$.

4.4.1. *Rational equations.* In this section we are interested in finding the zero level set of a given rational function $f(x) = \frac{p(x)}{q(x)}$.

(1) First, we need to decompose both q and p to find common roots.
(2) If there are common roots, then we can simplify the fraction.
(3) Once we have simplified the fraction, we can determine the domain of f and see that f the zeros of f are precisely the points $x \in \mathbb{R}$ such that $p(x) = 0$ but $q(x) \neq 0$.

Some examples

- The solutions to equation $f(x) := \frac{x^2+4x+3}{x-2} = 0$ are $x = -1$ and $x = -3$. In fact $p(x) := x^2 + 4x + 3$ can be decomposed as $p(x) = (x + 1)(x + 3)$, so the numerator and the denominator $q(x) := x - 2$ have no common terms. As a consequence, the domain of $f(x) = \frac{(x+1)(x+3)}{x-2}$ is $X := \mathbb{R} \setminus \{2\}$ and $f(x) = 0$ if and only if $p(x) = 0$, which is the case for $x = -1$ or $x = -3$.
- The unique solution to equation $f(x) := \frac{x^2+4x+3}{x+1}$ is $x = -3$ and the domain of f is $\mathbb{R} \setminus \{-1\}$. In fact $f(x) = \frac{(x+1)(x+3)}{x+1} = x + 3$ for $x \neq -1$.
- The domain of the function $f(x) := \frac{x-1}{x^3+2x^2-x-2}$ is $\mathbb{R} \setminus \{\pm 1, -2\}$ and the equation $f(x) = 0$ has no solutions. In this case, the numerator and the denominator have a common root, that is 1. In fact, simplifying we have

$$f(x) = \frac{x-1}{(x^2-1)(x+2)} = \frac{x-1}{(x-1)(x+1)(x+2)} = \frac{1}{(x+1)(x+2)} \qquad \forall\, x \neq 1.$$

4.4.2. *Rational inequalities.* Let f be a rational function. We want to find the set where $f(x) > 0$ (or $f(x) < 0$). The procedure is the following

- First, we need to decompose both q and p to find common roots.
- If there are common roots, then we can simplify the fraction.
- Once we have simplified the fraction, we can determine the domain of f and see that f the zeros of f are precisely the points $x \in \mathbb{R}$ such that $p(x) = 0$ but $q(x) \neq 0$.
- Finally, note that

$$f(x) > 0 \Leftrightarrow (p(x) > 0 \wedge q(x) > 0) \vee (p(x) < 0 \wedge q(x) < 0)$$

and

$$f(x) < 0 \Leftrightarrow (p(x) > 0 \wedge q(x) < 0) \vee (p(x) < 0 \wedge q(x) > 0).$$

For example, we have

$$\frac{x^2+4x+3}{x-2} > 0 \Leftrightarrow -3 < x < -1 \vee x > 2$$

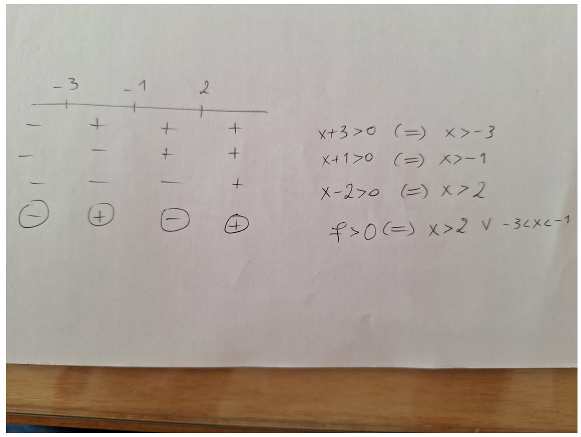

FIGURE 13. Graphical representation of the inequality $\frac{x^2+4x+3}{x-2} > 0$

and

$$\frac{x^2+4x+3}{x-2} < 0 \Leftrightarrow x < -3 \vee -1 < x < 2.$$

See figure 13 for a graphical representation of the inequality

4.5. **Qualitative properties of functions.** First we introduce the notion of monotonicity for a function $f : X \to \mathbb{R}$ with $X \subseteq \mathbb{R}$.

Definition 4.5.
- *We say that a function $f : X \to \mathbb{R}$ is non-decreasing if $x_1, x_2 \in X$, $x_1 \geq x_2 \Rightarrow f(x_1) \geq f(x_2)$.*
- *We say that a function $f : X \to \mathbb{R}$ is non-increasing if $x_1, x_2 \in X$, $x_1 \geq x_2 \Rightarrow f(x_1) \leq f(x_2)$.*
- *We say that a function $f : X \to \mathbb{R}$ is strictly increasing (or increasing) if $x_1, x_2 \in X$, $x_1 > x_2 \Rightarrow f(x_1) > f(x_2)$.*
- *We say that a function $f : X \to \mathbb{R}$ is strictly decreasing (or decreasing) if $x_1, x_2 \in X$, $x_1 > x_2 \Rightarrow f(x_1) < f(x_2)$.*

We say that a function is *monotone* if it is either non-decreasing or non-increasing. We say that a function is *strictly monotone* if it is either increasing or decreasing.

For example, the function $f(x) := x^3$ is strictly increasing in $\mathbb{R}$, while the function $f(x) := x^2$ is decreasing in $(-\infty, 0]$ and increasing in $[0, \infty)$.

Remark 4.4. *Let $X, Y \subseteq \mathbb{R}$ and let $f : X \to Y$ be a strictly monotone function. Then f is injective. As a consequence, the function $f : X \to f(X)$ is invertible.*

For example, for any $n \in \mathbb{N}$, the function $f(x) := x^{2n+1}$ is strictly increasing in $\mathbb{R}$, hence it is injective. Since $f(\mathbb{R}) = \mathbb{R}$, then it is also surjective with inverse function

$f^{-1}(y) := \sqrt[2n+1]{y}$ for $n \geq 1$, $f^{-1}(y) = y$ for $n = 0$.

Now we introduce the notion of symmetry of a function. Let $X \subseteq \mathbb{R}$. We say that X is symmetric if $\forall\, x \in X$, $-x \in X$. For example the interval $[-1, 1]$ is symmetric, while $[-1, 2)$ and $[-1, 1)$ are not.

Definition 4.6. *Let $X, Y \subseteq \mathbb{R}$ and let $f : X \to Y$ be a function. Assume that X is symmetric. Then w say that*

- *f is even if $f(x) = f(-x)$, $\forall\, x \in X$.*
- *f is odd if $f(x) = -f(-x)$, $\forall\, x \in X$.*

Let X be a symmetric domain. Then a function $f : X \subseteq \mathbb{R} \to \mathbb{R}$ is even if and only if its graph is symmetric with respect to the vertical axes, while it is odd if and only if its grph is symmetric with respect to the origin.

For example, the function $f(x) = x^{2n}$ is even for any $n \in \mathbb{N}$, while the function $f(x) := x^{2n+1}$ is odd for any $n \in \mathbb{N}$. However, there are functions which are neither even nor odd, for example $f(x) := e^x$.

Moreover, we introduce the notion of boundedness.

Definition 4.7. *Let $Y \subseteq \mathbb{R}$, $Y \neq \emptyset$. We say that*

- *Y is bounded above if $\sup Y < \infty$.*
- *Y is bounded below if $\inf Y > -\infty$.*
- *Y is bounded if it is both bounded above and bounded below.*
- *Y i unbounded if it is not bounded.*

Accordingly, we define the notion of boundedness for functions.

Definition 4.8. *Given a set X, a set $Y \subseteq \mathbb{R}$ and a function $f : X \to Y$, we say that*

- *f is bounded above if $f(X)$ is bounded above.*
- *f is bounded below if $f(X)$ is bounded below.*
- *f is bounded if $f(X)$ is bounded.*
- *f is unbounded if it is not bounded.*

The infimum and the supremum are defined for functions too. More precisely, we set

$$\inf_{x \in X} f(x) := \inf f(X), \quad \sup_{x \in X} f(x) := \sup f(X).$$

Using this language, we can see that

- f is bounded below if and only if $\inf_{x \in X} f(x) > -\infty$,
- f is bounded above if and only if $\sup_{x \in X} f(x) < \infty$.

- f is bounded if and only if $-\infty < \inf_{x \in X} f(x) \leq \sup_{x \in X} f(x) < \infty$.
- f is unbounded if it is not bounded

For example, the function $f(x) = e^x$ is bounded below but not bounded above. The function $f(x) := \log(x)$ is neither bounded below nor bounded above (see figure (7)). The function

$$\begin{aligned} f :&[0,1] \to \mathbb{R} \\ x &\mapsto x^2 \end{aligned} \tag{4.15}$$

is bounded.

In many applications, functions $f : X \subseteq \mathbb{R} \to \mathbb{R}$ are simply given through their analytical expression, for example $f(x) := x^2$. The domain may be not specified. In these case, it is understood that the domain $X := \mathrm{Dom}(f)$ is the maximal subset of $\mathbb{R}$ such that $f(x)$ is defined for any $x \in X$. For example,

- if $f(x) := x^2$, then $\mathrm{Dom}(f) = \mathbb{R}$.
- if $f(x) := \sqrt{x}$, then $\mathrm{Dom}(f) = [0, \infty)$.
- if $f(x) := \log x$, then $\mathrm{Dom}(f) = (0, \infty)$.
- If $f(x) := \frac{1}{x}$, then $\mathrm{Dom}(f) = (-\infty, 0) \cup (0, \infty)$.

In other words, in order to find the domain of the most relevant functions defined up to now, it is enough to require

- The arguments of logarithms to be positive,
- the arguments of even roots to be nonnegative,
- the denominators to be different form 0.

For example, the domain of the function $\log(x^2 - 1)$ is the set where $x^2 - 1 > 0$, that is

$$\mathrm{Dom}(\log(x^2 - 1)) = (-\infty, -1) \cup (1, \infty).$$

5. Trigonometric functions

5.1. **Sine and cosine.** Let us consider the unit circle

$$S^1 := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}.$$

Then any angle α can be associated to an arc on S^1, simply by drawing a half-line whose angle with the positive x axis is α and taking the arc between the point $(1, 0) \in S^1$ and the intersection p between the aforementioned half-line and S^1 (see figure 14).

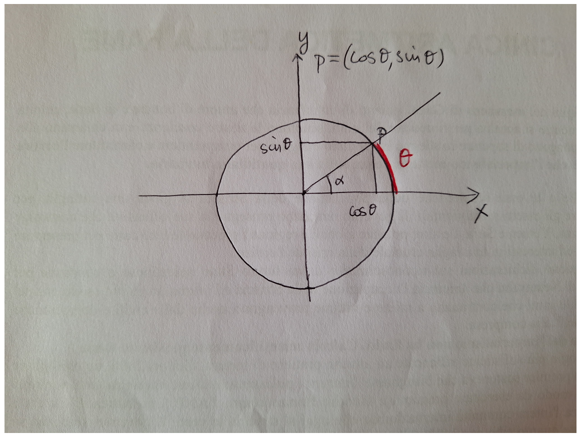

FIGURE 14. The definition of sin and cosine

The length θ of the arc corresponding to α is proportional to α. Considering the fact that the length of the arc corresponding to the angle $\alpha = 360°$ is 2π, we have

$$\frac{\theta}{\alpha} = \frac{2\pi}{360°} \Leftrightarrow \theta = \frac{2\pi}{360°}\alpha.$$

Adopting the convention that angles are positive if measured counterclockwise and negative if measured clockwise, we can define the function

$$\alpha \in \mathbb{R} \mapsto \theta(\alpha) := \frac{2\pi}{360°}\alpha \in \mathbb{R},$$

which associated to any angle α the signed length of the corresponding arc. For example, $\theta(90°) = \pi/2$, $\theta(180°) = \pi$. Since this function is invertible, the length θ can be used to measure angles. This measure is called *radians*. For example the angle $\alpha = 90°$ measures $\pi/2$ radians.

The coordinates of p are known as $\cos\theta$ and $\sin\theta$, or in other words $p = (\sin\theta, \cos\theta)$. In these way we have defined two functions

$$\begin{aligned} \sin : \mathbb{R} &\to \mathbb{R} \\ \theta &\mapsto \sin\theta \end{aligned} \tag{5.1}$$

and

$$\begin{aligned} \cos : \mathbb{R} &\to \mathbb{R} \\ \theta &\mapsto \cos\theta. \end{aligned} \tag{5.2}$$

Using the fact that $p \in S^1$, we can see that

$$|\sin\theta| \le 1, \quad |\cos\theta| \le 1, \quad \cos^2\theta + \sin^2\theta = 1, \quad \forall\,\theta \in \mathbb{R}. \tag{5.3}$$

Definition 5.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be a function, let $T > 0$. Then f is said to be periodic with period T (or T-periodic) if $f(x+T) = f(x)$, for any $x \in \mathbb{R}$.*

Note that, if f is periodic with period T, once we draw the graph of f on $[0, T]$ we know the whole graph of f.

By construction, it is possible to see that the functions sin and cos are periodic of period 2π, that is

$$\sin(\theta + 2\pi) = \sin\theta,\ \cos(\theta + 2\pi) = \cos\theta, \qquad \forall\, \theta \in \mathbb{R}. \tag{5.4}$$

In fact, the values θ and $\theta + 2\pi$ are associated with the same angle in the unit circle.

In particular, using (5.3), (5.4) and the definition of sine and cosine, we can see that the functions

$$\sin, \cos : \mathbb{R} \to [-1, 1]$$

are surjective but not injective (to see that these functions are surjective, it is enough to observe that any horizontal straight line of the form $y = c$ with $c \in [-1, 1]$ intersects the unit circle at least once). As a consequence, they are not invertible. However, if we restict the domains, we get invertible functions. More precisely, the functions

$$\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1, 1], \quad \cos : [0, \pi] \to [-1, 1]$$

are monotone increasing and decreasing respectively; therefore, they are invertible. Their inverses are denoted by

$$\arcsin : [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}], \qquad \arccos : [-1, 1] \to [0, \pi]$$

and known as *arcsine* and *arccosine* respectively. In other words, we have

$$\arcsin(\sin x) = x,\ \forall\, x \in [-\frac{\pi}{2}, \frac{\pi}{2}]; \qquad \arccos(\cos x) = x,\ \forall\, x \in [0, \pi].$$

Moreover, we note that, if f is periodic with period T, any integer multiple of the period still satisfies

$$f(x + kT) = f(x), \qquad \forall\, k \in \mathbb{Z},\ x \in \mathbb{R},$$

so that

$$\sin(\theta + 2\pi k) = \sin\theta,\ \cos(\theta + 2\pi k) = \cos\theta, \qquad \forall\, \theta \in \mathbb{R},\ k \in \mathbb{Z}. \tag{5.5}$$

Remark 5.1.
- *The function* cos *is even, that is,* $\cos(-\theta) = \cos\theta,\ \forall\, \theta \in \mathbb{R}$.
- *The function* sin *is odd, that is,* $\sin(-\theta) = -\sin\theta,\ \forall\, \theta \in \mathbb{R}$.

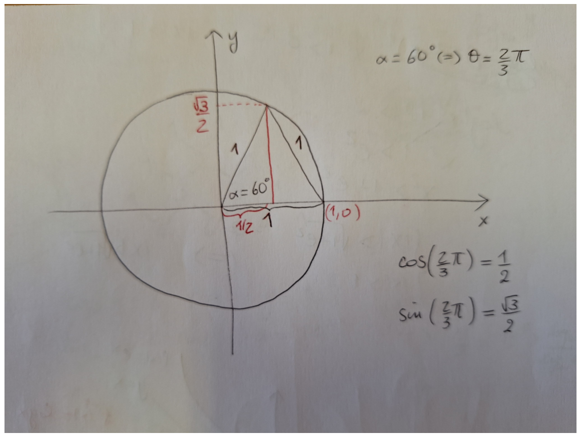

FIGURE 15. Using the properties of the equilater triangle and the Pitagora Theorem, we can see that $\cos(\frac{\pi}{3}) = \frac{1}{2}$ and $\sin(\frac{\pi}{3}) = \frac{\sqrt{3}}{2}$.

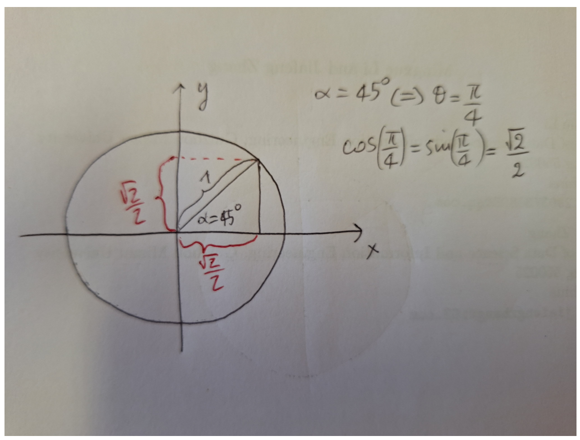

FIGURE 16. Using the property of the isosceles triangle and the Pitagora Theorem, we can see that $\cos(\frac{\pi}{4}) = \sin(\frac{\pi}{4}) = \frac{\sqrt{2}}{2}$.

We are interested in computing the sine and cosine of some known angles. Due to the periodicity, it is enough to consider values $\theta \in [0, 2\pi]$. It is possible to see that

$$\cos 0 = \sin(\frac{\pi}{2}) = 1, \; \cos(\frac{\pi}{2}) = \sin(0) = 0, \; \cos\pi = \sin(\frac{3\pi}{2}) = -1.$$

Some further examples are given in figures (15) and (16), where we show that

$$\cos(\frac{\pi}{3}) = \frac{1}{2}, \; \sin(\frac{\pi}{3}) = \frac{\sqrt{3}}{2}, \; \cos(\frac{\pi}{4}) = \sin(\frac{\pi}{4}) = \frac{\sqrt{2}}{2}.$$

Further interesting examples of angles whose sine and cosine are known are obtained

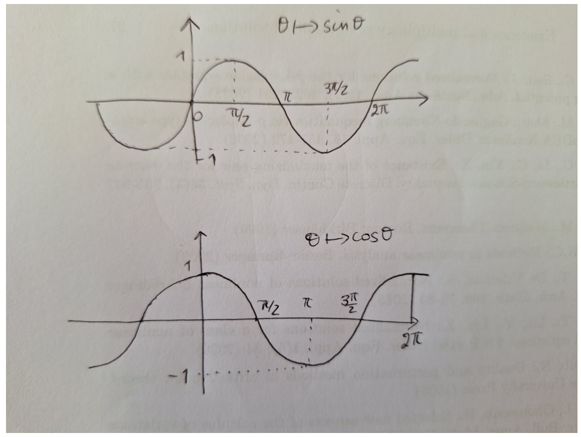

FIGURE 17. The graph of the the functions $\theta \mapsto \sin\theta$ and $\theta \mapsto \cos\theta$. Being these functions periodic of period 2π, it is enough to draw the grph for $\theta \in [0, 2\pi]$.

by reflection from the previous ones.

- By reflection with respect to the straight line $y = x$, we can see that $\sin(\frac{\pi}{6}) = \frac{1}{2}$, $\cos(\frac{\pi}{6}) = \frac{\sqrt{3}}{2}$.
- By reflection with respect to the origin, we have $\cos(\frac{5\pi}{4}) = \sin(\frac{5\pi}{4}) = -\frac{\sqrt{2}}{2}$.

With similar arguments, one can compute the sine and the cosine of

$$\frac{3\pi}{4},\ \frac{7\pi}{4},\ \frac{\pi}{6} + k\pi,\ \frac{\pi}{3} + k\pi,\ k = 1, 2, 3, 4.$$

Recollecting all the above information, we can draw the graph of the sine and cosine functions (see figure 17). fter that, reflecting these graphs with respect to the straight line $y = x$, we can draw the graphs of the functions arcsin, arccos (see figure 18). The sine and cosine can be used to compute the length of the sides of a rectangle triangle. In fact, the length of the hypotenuse $i > 0$ and the length of the legs c_1, $c_2 > 0$ are related through the relations

$$c_1 = i\cos\theta,\ c_2 = i\sin\theta,$$

where $\theta \in (0, \pi/2)$ is the angle between the hypothenus and the leg of length c_1 (see figure 19).

5.1.1. *Addition and subtraction formulas.* For any $x, y \in \mathbb{R}$, we have

- $\cos(x + y) = \cos x \cos y - \sin x \sin y$
- $\sin(x + y) = \sin x \cos y + \cos x \sin y$
- $\cos(x - y) = \cos x \cos y + \sin x \sin y$
- $\sin(x - y) = \sin x \cos y + \cos x \sin y$

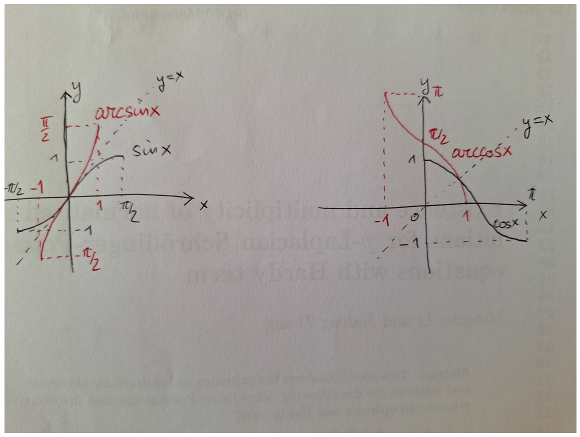

FIGURE 18. The graphs of the functions arcsin and arccos.

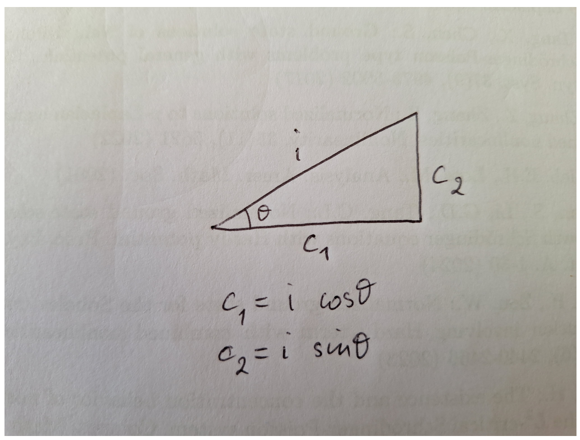

FIGURE 19. The relation between the length of the legs and the hypotenuse of a rectangle triangle. $c_1 = i\cos\theta$, $c_2 = i\sin\theta$.

In particular, we can see that $\sin(x+\frac{\pi}{2}) = \cos x$, for any $x \in \mathbb{R}$. As a consequence, we can obtain the graph of the cosine by translating the one of the sine.

5.1.2. *Duplication and bisection formulas.* For any $x \in \mathbb{R}$, we have

- $\sin(2x) = 2\sin x\cos x$
- $\cos(2x) = \cos^2 x - \sin^2 x$
- $\cos^2(\frac{x}{2}) = \frac{1+\cos x}{2}$
- $\sin^2(\frac{x}{2}) = \frac{1-\cos x}{2}$

6. Tangent and cotangent

In this section, we will define the tangent and the cotangent of an angle.

Definition 6.1. • *For* $\theta \in \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi : k \in \mathbb{N}\}$, *we define the function*

$$\begin{aligned} \tan : \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi : k \in \mathbb{N}\} &\to \mathbb{R} \\ \theta \mapsto \tan\theta &:= \frac{\sin\theta}{\cos\theta}. \end{aligned} \tag{6.1}$$

• *For* $\theta \in \mathbb{R} \setminus \{k\pi : k \in \mathbb{Z}\}$, *we define the function*

$$\begin{aligned} \cot : \mathbb{R} \setminus \{k\pi : k \in \mathbb{N}\} &\to \mathbb{R} \\ \theta \mapsto \cot\theta &:= \frac{\cos\theta}{\sin\theta}. \end{aligned} \tag{6.2}$$

Note that the domain of the tangent is the set where $\cos\theta \neq 0$, while the domain of the cotangent is the set where $\sin\theta \neq 0$.

Moreover, due to (6.1), we have

$$\tan\theta = \frac{1}{\cot\theta}, \qquad \forall\, \theta \in \mathbb{R} \setminus \{\frac{k\pi}{2} : k \in \mathbb{Z}\}.$$

The geometric meaning of the tangent and cotangent of angle is explained in figures (20) and (21).

The graphs of the tangent and cotangent are shown in figure (22). It is clear from the graphs that the tangent is invertible in $(-\frac{\pi}{2}, \frac{\pi}{2})$ and the cotangent is invertible in $(0, \pi)$. The inverse functions are known as *arctangent* and arccotangent and denoted by

$$\arctan : \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2}), \qquad \text{arccot} : \mathbb{R} \to (0, \pi)$$

respectively. Their graphs are drawn in figure

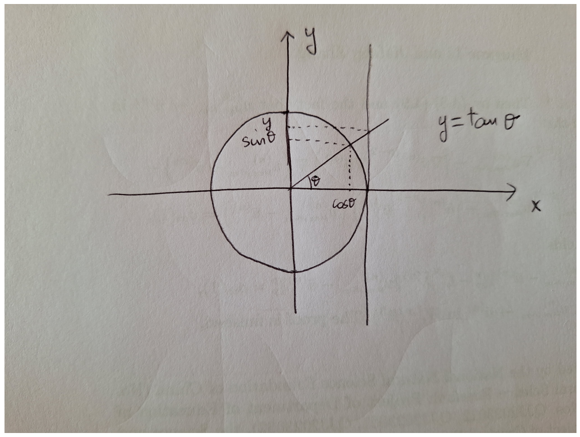

FIGURE 20. The tangent of θ is the y- coordinate of the intersection between the half-line of inclination θ and the vertical line $x = 1$. In fact, due to the proportionality between the edges of similar triangles, we have $\frac{\sin\theta}{\cos\theta} = \frac{y}{1}$.

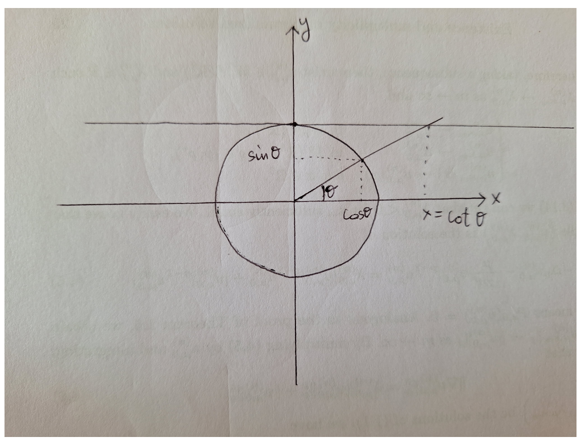

FIGURE 21. The cotangent of θ is the x- coordinate of the intersection between the half-line of inclination θ and the horizontal line $y = 1$. In fact, due to the proportionality between the edges of similar triangles, we have $\frac{\cos\theta}{\sin\theta} = \frac{x}{1}$.

7. LIMITS OF A FUNCTION

Consider the function $f(x) := \frac{1}{x^2}$. The domain of f is $\mathbb{R} \setminus \{0\}$. Despite the fact that the function is not defined in $x = 0$, we may be interested in the behavior of f

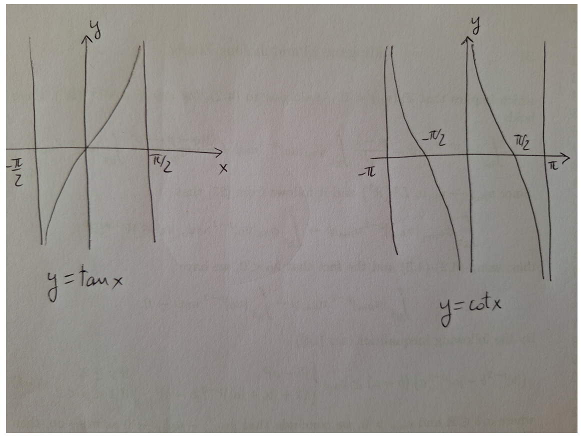

FIGURE 22. The graphs of the tangent and cotangent

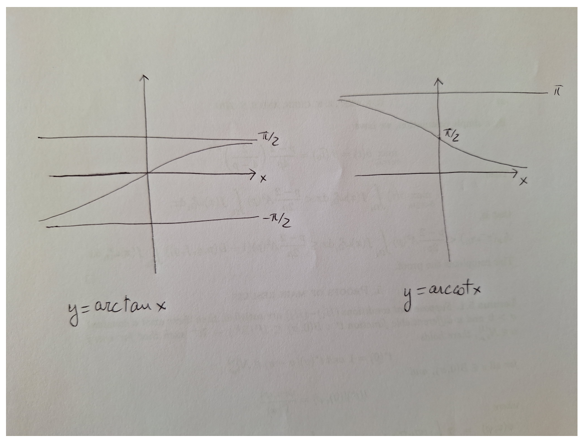

FIGURE 23. The graph of the arctan and *arccot*.

when x approaches 0. Looking at the graph, it is clear that as x approaches 0, $f(x)$ becomes very large (see figure 24).

This can be expressed analytically by saying that

$$\lim_{x \to 0} \frac{1}{x^2} = \infty.$$

In the sequel, we will introduce the precise definition of limit, in order to give a precise meaning to the above expressions.

7.1. **Neighborhoods and accumulation points.** For any $x \in \bar{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$, we introduce the notion of neighborhood.

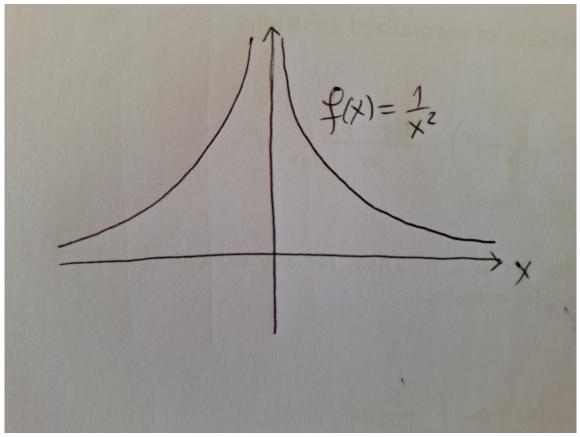

FIGURE 24. The graph of the function $f(x) := \frac{1}{x^2}$.

Definition 7.1.
- *Given a point $x \in \mathbb{R}$ and an open bounded interval $I \subseteq \mathbb{R}$, we say that I is a neighborhood of x if $x \in I$.*
- *We say that an open interval $I \subseteq \mathbb{R}$ is a neighborhood of ∞ if there exists $a \in \mathbb{R}$ such that $I = (a, \infty)$.*
- *We say that an open interval $I \subseteq \mathbb{R}$ is a neighborhood of $-\infty$ if there exists $a \in \mathbb{R}$ such that $I = (-\infty, a)$.*

For example, $I := (0, 1)$ is a neighborhood of $x = 1/2$, $I := (0, \infty)$ is a neighborhood of ∞ and $I := (-\infty, 1)$ is a neighborhood of $-\infty$. Thanks to the notion of neighborhood, we can define accumulation points.

Definition 7.2. *Let $X \subseteq \mathbb{R}$ be a set and let $x_0 \in \bar{\mathbb{R}}$. We say that x_0 is an accumulation point of X if for any neighborhood $I \subseteq \mathbb{R}$ of x_0, we have $(I \cap X) \setminus \{x_0\} \neq \emptyset$.*

In other words, x_0 is an accumulation point of X if any neighborhood $I \subseteq \mathbb{R}$ of x_0 contains at least a point of X which is different from x_0.

Let $X \subseteq \mathbb{R}$ and let

$$\mathcal{A}(X) := \{x_0 \in \bar{\mathbb{R}} : x_0 \text{ is an accumulation point of } X\}$$

be the set of accumulation points of X.

From the definition it follows that, if $\subseteq \mathbb{R}$ is an interval, any point $x_0 \in I$ is also an accumulation point of X, since for any neighborhood I of x_0 we can find a point $x \neq x_0$ such that $x \in X \cap I$. As a consequence, $I \subseteq \mathcal{A}(I)$. However, the opposite inclusion is not always true, even if I is an interval. For example, ∞ and 1 are an accumulation points of the interval $I := (1, \infty)$, even though $1, \infty \notin I$.

Such an inclusion is true if, for example, $I = [a, b]$ is a closed bounded interval (a, $b \in \mathbb{R}$). In this case we have $I = \mathcal{A}(I)$. If the interval is open, that is $I := (a, b)$ with a, $b \in \mathbb{R}$, then $\mathcal{A}(I) = [a, b]$.

On the other hand, the inclusion $X \subseteq \mathcal{A}(X)$ is not true for any $X \subset \mathbb{R}$. For example, if $X = \{0\} \cap (1, 2)$, then $0 \in X \setminus \mathcal{A}(X)$, since the interval $I := (-\frac{1}{2}, \frac{1}{2})$ fulfills $I \cap X = \{0\}$, hence $(I \cap X) \setminus \{0\} = \emptyset$. As a consequence, we are led to the following definition.

Definition 7.3. *Let $X \subseteq \mathbb{R}$ and let $x_0 \in X$. Then x_0 is an isolated point of X if x_0 is not an accumulation point of X.*

From the definition it follows that $x_0 \in X$ is an isolated point if and only if there exists a neighborhood $I \subseteq \mathbb{R}$ of x_0 such that $I \cap X = \{x_0\}$. For example, 0 is an isolated point of $\{0\} \cap (1, 2)$.

Finally, we can see that any point $x \in X$ can be isolated or an accumulation point of X, and accumulation points of X can belong to X or not.

7.2. **Definition of limit.**

Definition 7.4. *Let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be a function, let $y_0 \in \bar{\mathbb{R}}$ and let $x_0 \in \bar{\mathbb{R}}$ be an accumulation point of X. We say that*

$$\lim_{x \to x_0} f(x) = y_0 \tag{7.1}$$

if for any neighborhood $J \subseteq \mathbb{R}$ of y_0, there exists a neighborhood $I \subseteq \mathbb{R}$ of x_0 such that, $\forall\, x \in (I \cap X) \setminus \{x_0\}$, we have $f(x) \in J$.

In Definition 7.4, x_0 is an accumulation point of the domain. It may belong to it or not. Moreover, both x_0 and the limit y_0 can be finite or not.

The idea of the definition is the following: for any neighborhood I of the limit y_0, no matter how small it is, we can find a small neighborhood J which depends on I such that, if $x \in (I \cap X) \setminus \{x_0\}$, then $f(x) \in J$. Note that, being x_0 an accumulation point of the domain, f is not necessarily defined at x_0. For this reason we only need to check the values of f in $I \setminus \{x_0\}$.

Remark 7.1. *The fact that $x_0 \in \bar{\mathbb{R}}$ is an accumulation point of the domain X of a function does not imply that f admits a limit as $x \to x_0$. For example,*

- *periodic functions $f : \mathbb{R} \to \mathbb{R}$ do not admit limits as $x \to \pm\infty$.*

- *Let $x_0 \in \mathbb{R}$. Then the limit as $x \to x_0$ of the function*

$$f : x \in \mathbb{R} \mapsto \begin{cases} 1 & \forall\, x \in \mathbb{Q} \\ 0 & \forall\, x \in \mathbb{R} \setminus \mathbb{Q} \end{cases} \tag{7.2}$$

does not exist.

The definition 7.4 is quite general but a bit abstract. In the application, it is more convenient to have some equivalent formulations that are more intuitive to handle. For example, if $x \in \mathbb{R}$, any neighborhood of x can be written in the form $(x-a, x+b)$, for some a, $b > 0$. As a consequence, if x_0, $y_0 \in \mathbb{R}$, Definition 7.4 reduces to

$$\forall\, \varepsilon > 0,\, \exists\, \delta = \delta(\varepsilon) > 0 :\, \forall\, x \in X \setminus \{x_0\} \text{ with } |x - x_0| < \delta, \text{ we have } |f(x) - y_0| < \varepsilon. \tag{7.3}$$

Here $\varepsilon > 0$ is arbitrarily small and $\delta > 0$ depends on ε. For example, if

$$f(x) := \begin{cases} 2x + 3 & \text{for } x \neq 0 \\ 1 & \text{for } x = 0, \end{cases}$$

then it is possible to see that

$$\lim_{x \to 0} f(x) = 3 \neq f(0) = 1.$$

In fact, for any $\varepsilon > 0$ (small) we have $3 - \varepsilon < f(x) = 2x + 3 < 3 + \varepsilon$ if $0 < |x| < \delta := \frac{\varepsilon}{2}$. Moreover, it is possible to see that, if we change the definition of f at 0, the limit does not change. For example, setting $f(0) = 3 = 2 \cdot 0 + 3$, which is the natural value of f at $x = 0$, we still have

$$\lim_{x \to 0} f(x) = 3.$$

Even if we do not define f at 0, that is we restrict the function to $X := (-\infty, 0) \cup (0, \infty)$, 0 is an accumulation point of x, hence we can calculate the limit as $x \to 0$, which still has the same value.

The case in which both x_0 and y_0 do not belong to $\mathbb{R}$ can be treated similarly. For example, if $x_0 = y_0 = \infty$, Definition 7.4 reduces to

$$\forall\, M > 0,\, \exists\, K = K(M) > 0 :\, \forall\, x \in X \text{ with } x > K, \text{ we have } f(x) > M. \tag{7.4}$$

Here both M and K are large. For example, if $f(x) = x^3$, we have $f(x) > M$ if $x > K := \sqrt[3]{M}$. Hence we have

$$\lim_{x \to \infty} x^3 = \infty.$$

7.3. **Horizontal asymptotes.** Similarly, using that any neighborhood of ∞ can be written in the form (M, ∞), where $M > 0$ is an arbitrary (large) number, in case $y_0 \in \mathbb{R}$ and $x_0 = \infty$, Definition 7.4 reduces to

$$\forall \varepsilon > 0, \exists M > 0 : \forall x \in X \text{ with } x > M, \text{ we have } |f(x) - y_0| < \varepsilon. \tag{7.5}$$

Similarly, if $x_0 = -\infty$, Definition 7.4 reduces to

$$\forall \varepsilon > 0, \exists M > 0 : \forall x \in X \text{ with } x < -M, \text{ we have } |f(x) - y_0| < \varepsilon. \tag{7.6}$$

Definition 7.5. *Let $f : X \subset \mathbb{R} \to \mathbb{R}$ be a function. Assume that either $\infty \in \mathcal{A}(X)$ and there exists $y_0 := \lim_{x\to\infty} f(x) \in \mathbb{R}$ or $-\infty \in \mathcal{A}(X)$ and $y_0 := \lim_{x\to\infty} f(x) \in \mathbb{R}$. Then we say that the straight line of equation $y = y_0$ is a horizontal asymptote of f.*

The idea is that the graph of f approaches the straight line $y = y_0$ as $x \to \infty$ (or $x \to -\infty$).

For example, it is possible to see from figure 23 that

$$\lim_{x\to-\infty} \arctan x = -\frac{\pi}{2} \qquad \lim_{x\to\infty} \arctan x = \frac{\pi}{2}.$$

In particular, the function $f(x) := \arctan x$ has two horizontal asymptotes given by the straight lines $y = \frac{\pi}{2}$ and $y = -\frac{\pi}{2}$.

7.4. **Finite limits as $x \to \infty$ or $x \to -\infty$.** When, for example, $x_0 \in \mathbb{R}$ and $y_0 = \infty$, Definition 7.4 reduces to

$$\forall M > 0, \exists \delta = \delta(M) > 0 : \forall x \in X \setminus \{x_0\} \text{ with } |x - x_0| < \delta, \text{ we have } f(x) > M. \tag{7.7}$$

Similarly, if $y_0 = -\infty$, the definition of limit can be expressed in the following way

$$\forall M > 0, \exists \delta = \delta(M) > 0 : \forall x \in X \setminus \{x_0\} \text{ with } |x - x_0| < \delta, \text{ we have } f(x) < -M. \tag{7.8}$$

Let us consider, for example, the case $f(x) := \frac{1}{x^2}$, whose graph is drawn in figure 7.4. If $M > 0$ is a large number, we have $\frac{1}{x^2} > M$ if $0 < |x| < \delta := \frac{1}{\sqrt{M}}$. This is the precise proof of the fact that

$$\lim_{x\to 0} \frac{1}{x^2} = \infty.$$

This agrees with intuition. Note that 0 does not belong to the domain of the function $\frac{1}{x^2}$, that is $X := (-\infty, 0) \cup (0, \infty)$, but it is an accumulation point of X.

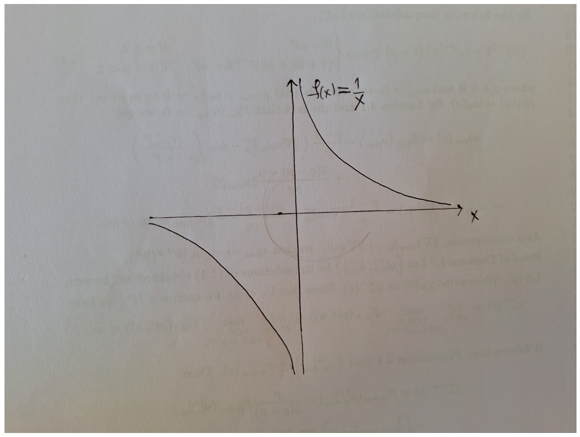

FIGURE 25. The graph of the function $f(x) := \frac{1}{x}$.

7.5. **Vartical asymptotes.** In order to define vertical asymptotes, we need to define the notion of *right* and *left neighborhood* of a point.

Definition 7.6 (Right and left neighborhood)**.** *Let $x \in \mathbb{R}$. Then we say that*

- *$I \subseteq \mathbb{R}$ is a right neighborhood of x if there exists a neighborhood $J \subseteq \mathbb{R}$ of x such that $I = J \cap (x_0, \infty)$.*
- *$I \subseteq \mathbb{R}$ is a left neighborhood of x if there exists a neighborhood $J \subseteq \mathbb{R}$ of x such that $I = J \cap (-\infty, x_0)$.*

Using the notion of left and right neighborhood, it is possible to define the notions of limit from the right and from the left

Definition 7.7 (Limit from the right and from the left)**.** *Let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be a function, let $y_0 \in \bar{\mathbb{R}}$ and let $x_0 \in \mathbb{R}$ be an accumulation point of X. We say that*

$$\lim_{x \to x_0^+} f(x) = y_0, \qquad \lim_{x \to x_0^-} f(x) = y_0 \tag{7.9}$$

if for any neighborhood $J \subseteq \mathbb{R}$ of y_0, there exists a right (or left respectively) neighborhood $I \subseteq \mathbb{R}$ of x_0 such that, $\forall\, x \in (I \cap X) \setminus \{x_0\}$, we have $f(x) \in J$.

It is possible to see that the limit from the right and the limit from the left may be different. For example, we have

$$\lim_{x \to 0^+} \frac{1}{x} = \infty, \qquad \lim_{x \to 0^-} \frac{1}{x} = -\infty \tag{7.10}$$

In fact, we have $\frac{1}{x} > M$ if $0 < x < \delta := \frac{1}{M}$ and $\frac{1}{x} < -M$ if $-\frac{1}{M} =: -\delta < x < 0$ (see figure (25)).

Remark 7.2. *Let $f : X \to \mathbb{R}$ be a function and let x_0 be an accumulation point of the domain such that, for any neighborhood I of x_0, we have $I \cap X \cap (x_0, \infty) \neq \emptyset$ and $I \cap X \cap (-\infty, x_0) \neq \emptyset$. Then there exists $\lim_{x\to x_0} f(x)$ if and only if both $\lim_{x\to x_0^+} f(x)$ and $\lim_{x\to x_0^-} f(x)$ exist and are equal.*

As a consequence, the function

$$\begin{aligned} f : \mathbb{R} \setminus \{0\} &\to \mathbb{R} \\ x &\mapsto \frac{1}{x} \end{aligned} \tag{7.11}$$

admits no limit as $x \to 0$, since the limit from the right and the limit from the left are different.

Definition 7.8 (Vertical asymptotes)**.** *Let $f : X \subset \mathbb{R} \to \mathbb{R}$ be a function. Assume that $x_0 \in \mathbb{R} \cap \mathcal{A}(X)$ and assume that either there exists $\lim_{x\to x_0^-} f(x) \in \bar{\mathbb{R}} \setminus \mathbb{R}$ or there exists $\lim_{x\to x_0^+} f(x) \in \bar{\mathbb{R}} \setminus \mathbb{R}$. Then we say that the straight line $x = x_0$ is a vertical asymptote of f.*

The idea is that the graph of f approaches the straight line $x = x_0$ as $x \to x_0$.

For example, the functions $\frac{1}{x^2}$ and $\frac{1}{x}$ both have a vertical asymptote of equation $x = 0$. Note that the existence of the limit as $x \to x_0$ is not a necessary condition for the existence of a vertical asymptote. In fact, in the case of the function $\frac{1}{x}$, the limit from the right and from the left both exist, but they are different.

7.6. **Limits from above and from below.** Moreover, it is possible to define the *limit from above* and *from below.*

Definition 7.9. *Let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be a function, let $y_0 \in \mathbb{R}$ and let $x_0 \in \bar{\mathbb{R}}$ be an accumulation point of X. We say that*

$$\lim_{x\to x_0^+} f(x) = y_0^\pm, \qquad \lim_{x\to x_0^-} f(x) = y_0^\pm, \qquad \lim_{x\to x_0} f(x) = y_0^\pm \tag{7.12}$$

if for any right (or left) neighborhood $J \subseteq \mathbb{R}$ of y_0, there exists a (right or left) neighborhood $I \subseteq \mathbb{R}$ of x_0 such that, $\forall\, x \in (I \cap X) \setminus \{x_0\}$, we have $f(x) \in J$.

The above definitions can be reformulated using ε and δ (or M).

By figure 7, we can see that

$$\lim_{x\to\infty} e^x = \infty,\ \lim_{x\to\infty} \log x = \infty,\ \lim_{x\to-\infty} e^x = 0^+,\ \lim_{x\to 0^+} \log x = -\infty. \tag{7.13}$$

As a consequence, the function e^x has a horizontal asymptote of equation $y = 0$ and the function $\log x$ has a vertical asymptote of equation $x = 0$ (see Figure 7).

7.7. **Operations with limits.** We are often interested in computing the limit of a sum or product of functions f and g. Most of the time, these limits can be deduced from the ones of f and g. Roughly speaking, one expects to be allowed to exchange the operations of sum and product with the limit. This is possible in the majority of the cases; however there are some special cases in which these rules fail and, if the limit exists, we need to find a different strategy to compute it.

We adopt the convention that

$$\alpha+\infty := \infty,\ \alpha-\infty := -\infty,\ \infty+\infty := \infty,\ -\infty-\infty := -\infty, \qquad \forall\, \alpha \in \mathbb{R}. \tag{7.14}$$

Theorem 7.1 (The limit of a sum)**.** *Let $f, g : X \to \mathbb{R}$ be functions. Let x_0 be an accumulation point of X and assume that $\lim_{x\to x_0} f(x) =: \alpha \in \bar{\mathbb{R}}$ and $\lim_{x\to x_0} g(x) =: \beta \in \bar{\mathbb{R}}$ exist. If $(\alpha, \beta) \notin \{(\infty, -\infty), (-\infty, \infty)\}$, then $\lim_{x\to x_0}(f(x) + g(x)) = \alpha + \beta$.*

Let us explain Definitions (7.15). First, we prove Theorem 7.1. We see that, for example, if $\lim_{x\to x_0} f(x) = \infty$ and $\lim_{x\to x_0} g(x) = \infty$, then $\lim_{x\to x_0} f(x)+g(x) = \infty$. This shows that it makes sense to define $\infty + \infty = \infty$.

If $\lim_{x\to x_0} f(x) = \infty$ and $\lim_{x\to x_0} g(x) = -\infty$, we can say nothing about $\lim_{x\to x_0}(f(x)+ g(x))$. For example, if $f(x) = x + \sin x$ and $g(x) = -x$, then

$$\lim_{x\to\infty} f(x) = \infty, \qquad \lim_{x\to\infty} g(x) = -\infty,$$

but $f(x) + g(x) = \sin x$ admits no limit as $x \to \infty$, since it is periodic.

If we take $f(x) := x^2$ and $g(x) := -2x$, then we have

$$\lim_{x\to\infty} f(x) = \infty, \qquad \lim_{x\to\infty} g(x) = -\infty$$

and

$$\lim_{x\to\infty} (f(x) + g(x)) = \lim_{x\to\infty} (x^2 - 2x) = \infty.$$

In fact, we have $x^2 - 2x > M$ if and only if $x < 1 - \sqrt{1 + M} \vee x > 1 + \sqrt{1 + M}$. So in particular $x^2 - 2x > M$ for any $x > K := 1 + \sqrt{M + 1}$. The same argument shows that

$$\lim_{x\to-\infty} (2x - x^2) = -\infty.$$

As a consequence, it does not make sense to define $\infty-\infty$. In fact, the result changes according to the situation. For this reason, we say that $\infty - \infty$ and $-\infty + \infty$ are *undefined forms.*

Now we will deal with the limit of a product. For $\alpha > 0$, we set $\alpha\cdot(\pm\infty) := \pm\infty$, while for $\alpha < 0$, we set $\alpha\cdot(\pm\infty) := \mp\infty$. Moreover, we set

$$\infty\cdot\infty = (-\infty)\cdot(-\infty) := \infty;\ \infty\cdot(-\infty) = (-\infty)\cdot\infty := -\infty. \tag{7.15}$$

These definitions are justified by the following result.

Theorem 7.2 (The limit of a product)**.** *Let $f, g : X \to \mathbb{R}$ be functions. Let x_0 be an accumulation point of X and assume that $\lim_{x\to x_0} f(x) =: \alpha \in \bar{\mathbb{R}}$ and $\lim_{x\to x_0} g(x) =: \beta \in \bar{\mathbb{R}}$ exist. If $\alpha, \beta \notin \{(0, \pm\infty), (\pm\infty, 0)\}$, then $\lim_{x\to x_0}(f(x)g(x)) = \alpha\beta$.*

If $\alpha = 0$ and $\beta = \pm\infty$, the product is not defined. In fact, if we take $f(x) := \frac{1}{x}$ and $g(x) := x^2$, we can see that $\lim_{x\to\infty} f(x) = 0$, $\lim_{x\to\infty} g(x) = \infty$ and

$$\lim_{x\to\infty} f(x)g(x) = \lim_{x\to\infty} \frac{1}{x}x^2 = \lim_{x\to\infty} x = \infty.$$

On the other hand, taking $f(x) := \frac{1}{x^2}$ and $g(x) := x$, we have $\lim_{x\to\infty} f(x) = 0$, $\lim_{x\to\infty} g(x) = \infty$ but

$$\lim_{x\to\infty} f(x)g(x) = \lim_{x\to\infty} \frac{1}{x^2}x = \lim_{x\to\infty} \frac{1}{x} = 0.$$

If we set $f(x) = x^2$ and $g(x) := \frac{1}{x^2+1}$, then $\lim_{x\to\infty} f(x) = \infty$, $\lim_{x\to\infty} g(x) = 0$ and

$$\lim_{x\to\infty} f(x)g(x) = \lim_{x\to\infty} \frac{x^2}{x^2+1} = 1.$$

In fact we have

$$\left|\frac{x^2}{x^2+1} - 1\right| < \varepsilon \Leftrightarrow -\varepsilon < \frac{1}{x^2-1} < \varepsilon \Leftrightarrow x < -\sqrt{1+\frac{1}{\varepsilon}} \vee x > \sqrt{1+\frac{1}{\varepsilon}}.$$

So in particular, for $x > M := \sqrt{1+\frac{1}{\varepsilon}}$, we have $|f(x)g(x) - 1| < \varepsilon$.

As a consequence, $0\cdot(\pm\infty)$ and $(\pm\infty)\cdot 0$ are undefined forms.

Remark 7.3. *Applying Theorem 7.2, we can see that*

$$\lim_{x\to\pm\infty} x^{2k} = \infty,\ \forall\, k \in \mathbb{N}\setminus\{0\}; \qquad \lim_{x\to\pm\infty} x^{2k+1} = \pm\infty,\ \forall\, k \in \mathbb{N}.$$

Moreover, we have

$$\lim_{x\to\pm\infty} \frac{1}{x^k} = 0, \qquad \forall\, k \in \mathbb{N}\setminus\{0\}.$$

Now we will deal with the limit of a composition. Roughly speaking, we expect that, when computing the limit of $f \circ g$, we can first calculate the limit of g and replace it in the expression of $f \circ g$. The rule is true with quite mild assumptions.

Theorem 7.3. *Let $g : A \to \mathbb{R}$ and $f : B \to \mathbb{R}$ be functions. Assume that $g(A) \subseteq B$, $x_0 \in \mathcal{A}(A)$ and*

- $\lim_{x\to x_0} g(x) = y_0$
- *$y_0 \in \mathcal{A}(B)$ and there exists a neighborhood $I \subseteq \mathbb{R}$ of x_0 such that $g(x) \neq y_0$, $\forall\, x \in (I \cap X) \setminus \{x_0\}$.*
- $\lim_{y\to y_0} f(y) = z_0$.

Then $\lim_{x\to x_0} f \circ g(x) = z_0$.

Remark 7.4.
- *Let $f : X \to \mathbb{R}$ be a function and let $x_0 \in \mathcal{A}(X)$. Assume that $\lim_{x\to x_0} f(x_0) = 0^+$. Then, by Theorem 7.3, we have $\lim_{x\to x_0} \frac{1}{f(x_0)} = \infty$. Roughly speaking, 1 divided by a very small positive number is very large. Similarly, if $\lim_{x\to x_0} f(x_0) = 0^-$, then we have $\lim_{x\to x_0} \frac{1}{f(x)} = -\infty$.*
- *In the spirit of Theorem 7.3, saying that $0\cdot\infty$ is an undefined form is the same as saying that $\frac{\infty}{\infty}$ is an undefined form.*

Here we list some special cases in which we want to compute the limit of $\frac{f(x)}{g(x)}$ and both f and g tend to ∞, or equivalently $f(x) \to \infty$ and $\frac{1}{g(x)} \to 0$.

- $\lim_{x\to\infty} \frac{\log_a x}{x^k} = 0$, $\forall\, a > 1$, $k \geq 1$.
- $\lim_{x\to\infty} \frac{x^k}{a^x} = 0$, $\forall\, a > 1$, $k \geq 1$.

Roughly speaking, logarithms go to infinity slower than powers and powers go to ∞ slower than the exponential.

Remark 7.5. *Once again, saying that $0\cdot\infty$ is not defined is the same as saying that $\frac{0}{0}$ is not defined. Here we list some special cases in which we want to compute the limit of $\frac{f(x)}{g(x)}$ and both f and g tend to 0, or equivalently $f(x) \to 0$ and $\frac{1}{g(x)} \to \infty$.*

- $\lim_{x\to 0} \frac{\sin x}{x} = 1$
- $\lim_{x\to 0} \frac{1-\cos x}{x^2} = \frac{1}{2}$
- $\lim_{x\to 0} \frac{\arctan x}{x} = 1$
- $\lim_{x\to 0} \frac{a^x-1}{x} = \log a$
- $\lim_{x\to 0} \frac{\tan x}{x} = 1$
- $\lim_{x\to 0} \frac{\log_a(1+x)}{x} = \frac{1}{\log a}$
- $\lim_{x\to 0} \frac{\arcsin x}{x} = 1$
- $\lim_{x\to 0} \frac{(x+1)^\alpha-1}{x} = \alpha$.

Now we will deal with the case of the limit of a function of the form $f(x)^{g(x)}$. In order to do so, we first set $0^\infty := 0$, $\infty^\infty := \infty$.

Theorem 7.4 (Limits of powers)**.** *Let f, $g : X \subseteq \mathbb{R} \to \mathbb{R}$. Let x_0 be an accumulation point of X and assume that $\lim_{x\to x_0} f(x) =: \alpha \in [0,\infty]$ and $\lim_{x\to x_0} g(x) =: \beta \in \bar{\mathbb{R}}$*

exist. If $(\alpha, \beta) \notin \{(1, \pm\infty), (0,0), (\infty, 0)\}$*, then*

$$\lim_{x \to x_0} f(x)^{g(x)} = \alpha^\beta.$$

Proof. We use Theorem 7.3 and the properties of logarithms to conclude that

$$f(x)^{g(x)} = e^{\log(f(x)^{g(x)})} = e^{g(x)\log(f(x))} \to e^{\beta \log \alpha} = \alpha^\beta$$

as $x \to x_0$. □

Once again, 1^∞ is not defined, in fact

$$(1+x)^{\frac{1}{x^2}} = e^{\frac{\log(1+x)}{x^2}} = e^{\frac{\log(1+x)}{x}\frac{1}{x}} \to \infty \qquad \text{as } x \to 0,$$

while

$$(1+x)^{\frac{1}{\sqrt{x}}} = e^{\frac{\log(1+x)}{\sqrt{x}}} = e^{\frac{\log(1+x)}{x}\sqrt{x}} \to 1 \qquad \text{as } x \to 0.$$

The same technique shows that

$$(1+x)^{\frac{1}{x}} = e^{\frac{\log(1+x)}{x}} \to e \qquad \text{as } x \to 0. \tag{7.16}$$

Remark 7.6 (A trick to remember about undefined forms)**.** *It is enough to remember about the following undefined forms*

(1) $\infty - \infty$
(2) $0 \cdot \infty$
(3) 1^∞
(4) 0^0

The other ones can be deduced from the above cases, by the commutative property of the sum and the product and Theorem 7.3. More explicitely, we can see that $0 \cdot (-\infty)$, $(\pm\infty) \cdot 0$, $-\infty + \infty$, $1^{-\infty}$ and ∞^0 are also undefined forms. In fact, if, for instance, ∞^0 were defined, then

$$\infty^0 = e^{\log(\infty^0)} = e^{0 \log \infty} = e^{0 \cdot \infty}$$

would be defined too. Hence $0 \times \infty$ would be defined too, which is not the case.

7.8. **Limits of polynomials and rational functions.**

Theorem 7.5 (Limits of polynomials)**.** *Let* $p : \mathbb{R} \to \mathbb{R}$ *be a polynomial of degree* $n \geq 1$ *of the form*

$$p(x) := a_0 + a_1 x + \cdots + a_n x^n, \qquad a_n \neq 0$$

- *If* $a_n > 0$ *and* n *is even, then* $\lim_{x \to \pm\infty} p(x) = \infty$.
- *If* $a_n > 0$ *and* n *is odd, then* $\lim_{x \to \infty} p(x) = \infty$ *and* $\lim_{x \to -\infty} p(x) = -\infty$.
- *If* $a_n < 0$ *and* n *is even, then* $\lim_{x \to \pm\infty} p(x) = -\infty$.
- *If* $a_n < 0$ *and* n *is odd, then* $\lim_{x \to \infty} p(x) = -\infty$ *and* $\lim_{x \to -\infty} p(x) = \infty$.

Proof. The proof is based on Theorem 7.1 and 7.2. More precisely, we use the decomposition

$$p(x) = a_n x^n \left(\frac{a_0}{a_n x^x} + \frac{a_1 x}{a_n x^n} + \cdots + 1 \right) =: a_n x^n g(x),$$

Remark 7.3 and the fact that $\lim_{x\to\pm\infty} g(x) = 1$. □

Similarly, we can treat the case of rational functions.

Theorem 7.6 (Limits of rational functions)**.** *Assume that*

$$p(x) := a_0 + a_1 x + \cdots + a_n x^n, \qquad q(x) := b_0 + b_1 x + \cdots + b_m x^m$$

are polynomials of degree n, $m \geq 1$ *without common factors.*

- *If* $m > n$, *then* $\lim_{x\to\pm\infty} \frac{p(x)}{q(x)} = 0$.
- *If* $m = n$, *then* $\lim_{x\to\pm\infty} \frac{p(x)}{q(x)} = \frac{a_n}{b_m}$.
- *If* $m < n$ *and* $n - m$ *is even, then*

$$\lim_{x\to\pm\infty} \frac{p(x)}{q(x)} = \operatorname{sgn}\left(\frac{a_n}{b_m}\right)\infty$$

- *If* $m < n$ *and* $n - m$ *is odd, then*

$$\lim_{x\to\infty} \frac{p(x)}{q(x)} = \operatorname{sgn}\left(\frac{a_n}{b_m}\right)\infty, \qquad \lim_{x\to-\infty} \frac{p(x)}{q(x)} = -\operatorname{sgn}\left(\frac{a_n}{b_m}\right)\infty.$$

Proof. It is enough to write the fraction as

$$f(x) = \frac{a_n x^n \left(\frac{a_0}{a_n x^x} + \frac{a_1 x}{a_n x^n} + \cdots + 1 \right)}{b_m x^m \left(\frac{b_0}{b_m x^m} + \frac{b_1 x}{b_m x^m} + \cdots + 1 \right)}$$

and apply the known results about the limits of x^k, for $k \in \mathbb{Z}$. □

Limits can be used to draw graphs of functions. For example, let $f(x) := \frac{x^2+4x+3}{x-2} = \frac{(x+3)(x+1)}{x-2}$.

- As we saw above, the domain of this function is given by $\mathbb{R} \setminus \{2\}$.
- The equation $f(x) = 0$ is solved if and only if $X = -1$ or $x = -3$. This means that the intersection of the graph with the x-axis is given by the points $(-3, 0)$ and $(-1, 0)$.
- Inequality $f(x) > 0$ is satisfied for $-3 < x < -1$ and $x > 2$.
- Computing the right and left limits as $x \to 2$. More precisely, we can see that

$$\lim_{x\to 2^-} \frac{(x+3)(x+1)}{x-2} = -\infty, \qquad \lim_{x\to 2^+} \frac{(x+3)(x+1)}{x-2} = \infty. \tag{7.17}$$

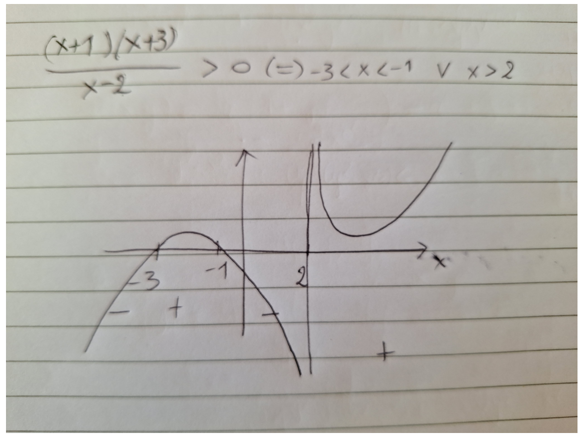

FIGURE 26. The graph of the function $f(x) := \frac{(x+3)(x+1)}{x-2}$.

This yields that the straight line $x = 2$ is a horizontal asymptote of f.

Moreover, using Theorem 7.2 we have

$$\lim_{x \to -\infty} \frac{(x+3)(x+1)}{x-2} = -\infty, \qquad \lim_{x \to -\infty} \frac{(x+3)(x+1)}{x-2} = \infty. \tag{7.18}$$

Using this information, we can draw the graph of $f(x) := \frac{(x+3)(x+1)}{x-2}$ (see Figure 26).

7.9. **Some useful Theorems about limits.** Some limits may be hard to compute. This kind of difficulty can be overcome by finding an upper bound and a lower bound of our given function for which the limits are known. If these limits coincide, then our original function admits a limit too, and this limit coincides with the common value of the limits of the upper and lower bounds.

Theorem 7.7 (Comparison or policemen Theorem)**.** *Let f, g, $h : X \subseteq \mathbb{R} \to \mathbb{R}$ be functions and let $x_0 \in \mathcal{A}(X)$. Assume that there exists a neighborhood $I \subseteq \mathbb{R}$ of x_0 such that*

$$f(x) \leq h(x) \leq g(x) \qquad \forall\, x \in (X \cap I) \setminus \{x_0\}.$$

If $\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = \alpha \in \bar{\mathbb{R}}$, then $\lim_{x\to x_0} h(x) = \alpha$.

For example, taking $h(x) := \frac{\sin x}{x}$, we can see that

$$-\frac{1}{x} \leq h(x) \leq \frac{1}{x} \qquad \forall\, x \in \mathbb{R} \setminus \{0\},$$

therefore $\lim_{x\to\infty} h(x) = 0$.

Theorem 7.7 can be used to compute the limit of products or quotients when we know, for example, that one of the two terms is bounded.

Corollary 7.1. *Let* f, $g : X \subseteq \mathbb{R} \to \mathbb{R}$ *be functions and let* $x_0 \in \mathcal{A}(X)$.

- *If there exists a neighborhood* $I \subseteq \mathbb{R}$ *of* x_0 *such that* f *is bounded in* $(X \cap I) \setminus \{x_0\}$ *and* $\lim_{x \to x_0} g(x) = \infty$, *then* $\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0$.
- *If there exists a neighborhood* $I \subseteq \mathbb{R}$ *of* x_0 *such that* f *is bounded in* $(X \cap I) \setminus \{x_0\}$ *and* $\lim_{x \to x_0} g(x) = 0$, *then* $\lim_{x \to x_0} f(x)g(x) = 0$.

Moreover, if the limit as $x \to x_0$ exists and it is not zero, then we can deduce that the sign of f close to x_0 coincides with the sign of the limit.

Theorem 7.8 (Sign permanence Theorem)**.** *Let* $f : X \subseteq \mathbb{R} \to \mathbb{R}$ *be a function and let* $x_0 \in \mathcal{A}(X)$. *Assume that there exists* $\lim_{x \to x_0} f(x) =: \alpha \in (0, \infty]$. *Then there exists a neighborhood* $I \subseteq \mathbb{R}$ *of* x_0 *such that* $f(x) > 0$ *for any* $x \in (X \cap I) \setminus \{x_0\}$.

On the other hand, applying Theorem 7.8 to $-f$, we can see that, under the assumptions of Theorem 7.8, if $\lim_{x \to x_0} f(x) =: \alpha \in [-\infty, 0)$, then there exists a neighborhood $I \subseteq \mathbb{R}$ of x_0 such that $f(x) < 0$ for any $x \in (X \cap I) \setminus \{x_0\}$. As a consequence, we deduce the following corollary.

Corollary 7.2. *Let* $f : X \subseteq \mathbb{R} \to \mathbb{R}$ *be a function and let* $x_0 \in \mathcal{A}(X)$. *Assume that there exists* $\lim_{x \to x_0} f(x) =: \alpha \in \bar{\mathbb{R}}$.

- *If there exists a neighborhood* I *of* x_0 *such that* $f(x) \geq 0$ *for any* $x \in (X \cap I) \setminus \{x_0\}$, *then* $\alpha \in [0, \infty]$.
- *If there exists a neighborhood* I *of* x_0 *such that* $f(x) \leq 0$ *for any* $x \in (X \cap I) \setminus \{x_0\}$, *then* $\alpha \in [-\infty, 0]$.

8. Continuous functions

In this section we will introduce the notion of continuous function. Roughly speaking, a function is continuous if, when drawing the graph, we never lift the pen from the sheet. It is the case if, for example, f is a polynomial, or the sine or cosine. This does not happen if $f(x) := \frac{1}{x}$. In fact, in this case, we need to lift the pen from the sheet when we approach the point $x = 0$. This can be translated into a precise mathematical language by saying that the function $\frac{1}{x}$ is not continuous at $x = 0$.

Now we give a precise definition of continuity.

Definition 8.1. *Let* $f : X \subseteq \mathbb{R} \to \mathbb{R}$ *be a function and let* $x_0 \in X$.

- *We say that* f *is continuous at* x_0 *if*

$$\forall \varepsilon > 0, \exists \delta > 0 : \forall x \in X \text{ with} |x - x_0| < \delta, \text{ we have} |f(x) - f(x_0)| < \varepsilon. \tag{8.1}$$

- *We say that f is continuous in X if f is continuous at x_0 for any $x_0 \in X$.*

If x_0 is an accumulation point of the domain X, Definition 8.1 can be expressed in terms of the notion of limit.

Remark 8.1.
- *If $x_0 \in X \cap \mathcal{A}(X)$ is an accumulation point of X, then f is continuous at x_0 if and only if $\lim_{x\to x_0} f(x) = f(x_0)$.*
- *If x_0 is an isolated point of X, then f is continuous at x_0. In fact, for any $\varepsilon > 0$, we can choose $\delta > 0$ such that $X \cap (x_0 - \delta, x_0 + \delta) = \{x_0\}$, so that*

$$|f(x) - f(x_0)| = 0 < \varepsilon \qquad \forall\, x \in X \cap (x_0 - \delta, x_0 + \delta).$$

Note that the number $\delta > 0$ appearing in definition 8.1 actually depends both on ε and on x_0. If it is possible to find such a δ independent of x_0, we say that f is *uniformly continuous* in X. More precisely, we have the following definition.

Definition 8.2. *Let $f : X \subset \mathbb{R} \to \mathbb{R}$ be a function. We say that f is uniformly continuous in X if*

$$\forall\, \varepsilon > 0,\, \exists\, \delta > 0 : \forall\, x,\, y \in X \text{ with } |x - y| < \delta, \text{ we have } |f(x) - f(y)| < \varepsilon.$$

We can see that uniform continuity implies continuity, but the converse is not true.

An example of uniformly continuous functions is given by functions that are continuous on a closed bounded interval $[a, b]$.

Theorem 8.1. *Let $a, b \in \mathbb{R}$ with $a < b$ and let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then f is uniformly continuous in $[a, b]$.*

It is possible to see that the sum, the product and the composition of continuous functions is still a continuous function.

Theorem 8.2. *Let $f, g : X \subset \mathbb{R} \to \mathbb{R}$ be continuous at $x_0 \in X$. Then $f + g$, $f \cdot g$ are also continuous at x_0.*

Proof. If x_0 is an isolated point, then any function is continuous at x_0, so in particular this is true for $f + g$ and $f \cdot g$. If $x_0 \in \mathcal{A}(X)$, the statement follows from the definition of limit. $\square$

As a consequence, If f and g are continuous in X, then $f + g$ and $f \cdot g$ are also continuous in X.

Theorem 8.3. *Let $X, Y \subseteq \mathbb{R}$, let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be continuous at $x_0 \in X$ and assume that $f(x_0) \in Y$. Let $g : Y \to \mathbb{R} \to \mathbb{R}$ be continuous at $f(x_0)$. Then $f \circ g$ is continuous at x_0.*

As a consequence, if $f(X) \subset Y$, f is continuous in X and g is continuous in Y, then $f \circ g$ is continuous in X.

Some well-known examples of continuous functions are polynomials. In fact, the identity function $f(x) = x$ is clearly continuous in $\mathbb{R}$, and the same is true for constant functions. As a consequence, Theorem 8.2 yields the result that integer powers x^k, $k \geq 2$ are also continuous and the same is true for polynomials. Other examples of continuous functions are trigonometric functions such as $\sin x$ or $\cos x$.

8.1. **Classification of discontinuity points.** Let $f : X \subset \mathbb{R} \to \mathbb{R}$ be a continuous function and let $x_0 \in \mathcal{A}(X) \cup X$. Then only 3 cases are possible.

(1) $x_0 \in X$ and f is continuous at x_0.
(2) $x_0 \in X$ and f is not continuous at x_0.
(3) $x_0 \notin X$.

If either case (2) or case (3) occur, then we say that x_0 is a *discontinuity point*.

We are interested in classifying discontinuity points, that is, to understand how a function can fail to be continuous. In fact, many situations are possible: the right and left limit as $x \to x_0$ may both exist and be equal, or different, or not exist.

Definition 8.3. (1) *We say that a discontinuity point* $x_0 \in \mathcal{A}(X) \cup X$ *is a removable discontinuity point (or a removable singularity) if* $\lim_{x\to x_0^-} f(x) = \lim_{x\to x_0^+} f(x)$ *(both limits exist and are equal).*

(2) *We say that a discontinuity point* $x_0 \in \mathcal{A}(X) \cup X$ *is a jump discontinuity point (or jump singularity) if* $\lim_{x\to x_0^-} f(x), \lim_{x\to x_0^+} f(x) \in \mathbb{R}$ *and* $\lim_{x\to x_0^-} f(x) \neq \lim_{x\to x_0^+} f(x)$ *(both limits exist and are finite but different).*

(3) *We say that a discontinuity point* $x_0 \in \mathcal{A}(X) \cup X$ *is a second type discontinuity point (or second type singularity) if none of the previous two conditions is satisfied.*

Note that f is not necessarily defined at discontinuity points.

The removable singularity has this name because it can be removed just by defining $f(x_0) := \lim_{x\to x_0} f(x)$.

In case f has a jump singularity at $x = x_0$, then the *amplitude* of the jump is given by $A := |\lim_{x\to x_0^-} f(x) - \lim_{x\to x_0^+} f(x)|$.

For example, the function

$$f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$$
$$x \mapsto x^3$$

has a removable discontinuity point at $x = 0$. This can be removed by defining $f(0) := 0$. The same is true for the function

$$f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$$
$$x \mapsto \begin{cases} x^3 & \text{if } x \neq 0 \\ 1 & \text{if } x = 0. \end{cases}$$

If either f is defined at $x = 0$ but $f(0) \neq 0$ or f is not defined at $x = 0$, the point $x = 0$ is a removable singularity. More in general, if f is continuous in X and we change its definition at a point $x_0 \in X$, then x_0 becomes a removable singularity.

The function

$$f : \mathbb{R} \to \mathbb{R}$$
$$x \mapsto \begin{cases} x^3 & \forall\, x \leq 0 \\ 1 & \forall\, x > 0 \end{cases}$$

has a jump discontinuity at $x = 1$. The amplitude of the jump is 1. A similar situation occurs for functions $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = a$ for $x \leq x_0$ and $f(x) = b$ for $x > x_0$ if $b \neq a$. The same holds true if the function is not defined at $x = x_0$.

The function $f(x) := \frac{1}{x}$ has a second type singularity at $x = 0$, since $\lim_{x \to 0^-} \frac{1}{x} = -\infty$ and $\lim_{x \to 0^+} \frac{1}{x} = \infty$.

The function

$$f : \mathbb{R} \to \mathbb{R}$$
$$x \mapsto \begin{cases} x^3 & \forall\, x \leq 0 \\ \frac{1}{x} & \forall\, x > 0 \end{cases}$$

has a second type singularity at $x = 0$.

8.2. **Some useful results about continuous functions.** Here we state some crucial results about continuous functions.

Let $f : X \to \mathbb{R}$ be a function. We say that f has a maximum on X if $\sup f(X)$ is achieved, or in other words if there exists $x_0 \in X$ such that

$$f(x_0) = \sup f(X) =: \sup_{x \in X} f(x).$$

If it is the case, we set

$$\max_{x\in X} f(x) := \sup_{x\in X} f(x)$$

and we say that x_0 is a *maximum point* or a *global maximum point.*

Similarly, we say that a function $f : X \to \mathbb{R}$ has a minimum on X if $\inf f(X)$ is achieved, or in other words if there exists $x_0 \in X$ such that

$$f(x_0) = \inf f(X) =: \inf_{x\in X} f(x).$$

If it is the case, we set

$$\min_{x\in X} f(x) := \inf_{x\in X} f(x)$$

and we say that x_0 is a *minimum point* or a *global minimum point.*

Sometimes, the words *minimum point* and *maximum point* are replaced by *minimiser* and *maximiser* respectively.

For example, the function $f(x) := \arctan x$ has neither a maximum nor a minimum on $\mathbb{R}$ (see figure (23)), while the functions $g(x) := \sin x$ and $h(x) := \cos x$ both a maximum and a minimum, that is 1 and -1 respectively. Both g and h have infinitely many maximum and minimum points (see figure 17).

Theorem 8.4 (Weierstrass)**.** *Let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be continuous on a closed and bounded domain X. Then f has a maximum and a minimum in X.*

Note that the hypothesis of the Weierstrass Theorem are sharp. In fact, the function $f(x) := 3x + 1$ is continuous on $\mathbb{R}$, which is closed but unbounded, and it has neither a maximum nor a minimum. The same is true for the function $f(x) := \arctan x$, which fulfills

$$\inf_{x\in\mathbb{R}} \arctan x = -\frac{\pi}{2}, \qquad \sup_{x\in\mathbb{R}} \arctan x = \frac{\pi}{2},$$

but these extrema are not fulfilled.

The function

$$\begin{aligned} f : (0, 1] &\to \mathbb{R} \\ x &\mapsto \frac{1}{x} \end{aligned}$$

has no maximum. Note that f is continuous on the bounded interval $(0, 1]$, but $(0, 1]$ is not closed.

Moreover, the interval $[-\pi/2, \pi/2]$ is closed and bounded but the function

$$\begin{aligned} f : [-\pi/2, \pi/2] &\to \mathbb{R} \\ x &\mapsto \begin{cases} \tan x & \forall\, x \in (-\pi/2, \pi/2) \\ 0 & \text{for}\, x = -\pi/2 \,\text{or}\, x = \pi/2 \end{cases} \end{aligned} \tag{8.2}$$

has neither a maximum nor a minimum. In fact, it is not continuous.

Now we state another important result about continuous functions, which enables us to prove the existence of at least one solution to equation $f(x) = 0$ if f is continuous.

Theorem 8.5. *Let $I \subseteq \mathbb{R}$ be an interval and let $f : I \to \mathbb{R}$ be a continuous function. Then $f(I)$ is either a point or an interval.*

We note that $f(I)$ is a point if and only if f is constant. As a consequence of Theorem 8.5, we deduce the following result.

Theorem 8.6 (Intermediate value Theorem)**.** *Let a, $b \in \mathbb{R}$, $a < b$, let $f : [a, b] \to \mathbb{R}$ be a continuous function and let $y \in \mathbb{R}$ be such that*

$$\min_{x \in [a,b]} f(x) \leq y \leq \max_{x \in [a,b]} f(x)$$

Then there exists $x_0 \in [a, b]$ such that $f(x_0) = y$.

Proof. The result is trivial if f is constant, hence we can assume that f is not constant. By Theorem 8.5, $I =: f([a, b])$ is an interval. As a consequence, any real number $y \in [\min_{x \in [a,b]} f(x), \max_{x \in [a,b]} f(x)]$, belongs to $f([a, b])$. This yields that there exists $x_0 \in [a, b]$ such that $f(x_0) = y$. □

Applying Theorem 8.6 to the case in which $f(a)f(b) < 0$, we deduce the following corollary.

Corollary 8.1. *Let a, $b \in \mathbb{R}$, $a < b$, let $f : [a, b] \to \mathbb{R}$ be a continuous function such that $f(a)f(b) < 0$. Then there exists $x_0 \in (a, b)$ such that $f(x_0) = 0$.*

Proof. It is enough to apply Theorem 8.6 with $y = 0$. □

9. Differentiable functions

In this section we introduce the notion of *differentiable function* and *derivative* of a function. Roughly, the derivative of a function f at x_0 is the slope of the graph of f at the point x_0. It is clear from intuition that such a slope is not always defined. This happens, for instance, for the function $|x|$ at $x = 0$ or the function $\frac{1}{x}$ at the point $x = 0$.

9.1. **The definition of derivative.** Let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be a function and let $x_0 \in \mathring{X}$. Then there exists $\delta > 0$ such that $(x_0 - \delta, x_0 + \delta) \subseteq X$, so that the function

$$h \in (-\delta, \delta) \setminus \{0\} \mapsto \frac{f(x_0 + h) - f(x_0)}{h} \in \mathbb{R}$$

is well defined. This function is known as the *incremental ratio*. Since 0 is an accumulation point of $(-\delta, \delta) \setminus \{0\}$, it makes sense to wonder whether the limit as $h \to 0$ of the incremental ratio exists or not. As a consequence, we are led to the following definition.

Definition 9.1. *Let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be a function and let $x_0 \in \mathring{X}$. Then we say that f is differentiable at x_0 if there exists*

$$f'(x_0) := \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} \tag{9.1}$$

and $f'(x_0) \in \mathbb{R}$. If it is the case, $f'(x_0)$ is known as the derivative (or first derivative) of f at x_0.

Sometimes, the derivative is denoted by $\frac{df}{dx}(x_0)$.

Moreover, in the above notations, we say that f is differentiable in $\mathring{X}$ if it is differentiable at x_0 for any $x_0 \in \mathring{X}$. If it is the case, then the function

$$\begin{aligned} f' : \mathring{X} &\to \mathbb{R} \\ x &\mapsto f'(x) \end{aligned} \tag{9.2}$$

is known as the *derivative* of f.

The rough idea is the following. Let us consider a point x_0 and a small increment $h > 0$. Then the straight line

$$y = \frac{f(x_0 + h) - f(x_0)}{h}(x - x_0) + f(x_0)$$

intersects the graph of f at least twice, that is at $(x_0, f(x_0))$ and $(x_0 + h, f(x_0 + h))$. As h becomes smaller, these two points become closer and closer and the straight line tends to intersect the graph of at $(x_0, f(x_0))$ only and to have the same slope as the graph at $(x_0, f(x_0))$ (see Figure (27)).

This intuitive idea leads us to the give the following definition.

Definition 9.2. *Let $X \subseteq \mathbb{R}$, $x_0 \in \mathring{X}$ and $f : X \to \mathbb{R}$ be differentiable at x_0. Then the tangent line to the graph of f at x_0 is given by the straight line of equation*

$$y = f'(x_0)(x - x_0) + f(x_0). \tag{9.3}$$

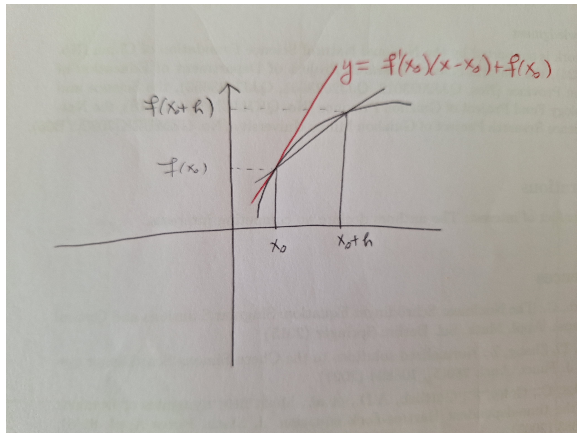

FIGURE 27. The incremental ratio approaches the derivative, that is the slope of the graph, as $x \to x_0$.

It is possible to see from the definition that constant functions are differentiable on $\mathbb{R}$ and have derivative 0, since the incremental ratio is identically equal to 0.

Let a, $b \in \mathbb{R}$, $a \neq 0$ and let $f(x) := ax + b$. This kind of function is referred to as *affine functions*. Then its derivative is given by

$$f'(x) = \lim_{h \to 0} \frac{a(x+h) + b - (ax+b)}{h} = \lim_{h \to 0} \frac{ah}{h} = a.$$

As a consequence, the tangent line to the graph of f is the straight line of equation

$$y = a(x - x_0) + (ax_0 + b) = ax + b = f(x).$$

In other words, affine functions are differentiable on $\mathbb{R}$ and their tangent line coincides with f itself at every point.

For more general differentiable functions, the idea is that the tangent line at x_0 is the "best affine approximation of f in a neighborhood of x_0".

Computing the derivative of a given function is crucial for applications, but most of the time it is quite long and not so easy, at least if one tries to do that applying the definition. As a consequence, it is useful to remember basic rules which enable us to make the job easier. The first result in this direction is the following.

Theorem 9.1. *Let $X \subseteq \mathbb{R}$, α, $\beta \in \mathbb{R}$ and let f, $g : X \to \mathbb{R}$ be differentiable functions at $x_0 \in \mathring{X}$. Then*

(1) ***Linearity of the derivative.*** *$\alpha f + \beta g$ is differentiable at x_0 and*

$$(\alpha f + \beta g)'(x_0) = \alpha f'(x_0) + \beta g'(x_0).$$

(2) ***Leibnitz rule.*** *$f \cdot g$ is differentiable at x_0 and $(f \cdot g)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$.*

(3) ***Derivative of the inverse function.*** *Assume that there exists $\delta > 0$ such that f is invertible in $(x_0 - \delta, x_0 + \delta)$. Then f^{-1} is differentiable at $y_0 := f(x_0)$ and $(f^{-1})'(y_0) = \frac{1}{f'(f^{-1}(y_0))}$.*

Using, for example, the Leibnitz rule, we can see that

$$(x^k)' = kx^{k-1}, \qquad \forall\, x \in \mathbb{R},\, k \in \mathbb{N} \setminus \{0\}. \tag{9.4}$$

Using the linearity of the derivative, we can compute the derivative of a polynomial. More precisely, if $p(x) := a_0 + a_1 x + \cdots + a_n x^n$, we can see that

$$p'(x) := a_1 + 2a_2 x + \cdots + n a_n x^{n-1}, \qquad \forall\, x \in \mathbb{R},\, n \in \mathbb{N} \setminus \{0\}.$$

For example, the derivative of x^2 is $2x$, the derivative of x^3 is $3x^2$ and so on.

Using Remark 7.5, we can see that

$$(e^x)' = \lim_{h \to 0} \frac{e^{x+h} - e^x}{h} = \lim_{h \to 0} \left(e^x \frac{e^h - 1}{h} \right) = e^x, \qquad \forall\, x \in \mathbb{R}, \tag{9.5}$$

as a consequence, by Theorem 9.1, we have

$$(\log x)' = \frac{1}{e^{\log x}} = \frac{1}{x} \qquad \forall\, x \in (0, \infty). \tag{9.6}$$

Moreover, by the addition formulas of the sine and Remark 7.5, we have

$$\begin{aligned} (\sin x)' &= \lim_{h \to 0} \frac{\sin(x+h) - \sin x}{h} \\ &= \lim_{h \to 0} \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} \\ &= \lim_{h \to 0} \left(\sin x \frac{\cos h - 1}{h} + \cos x \frac{\sin h}{h} \right) = \cos x, \qquad \forall\, x \in \mathbb{R}. \end{aligned} \tag{9.7}$$

Similarly, we can see that

$$(\cos x)' = -\sin x \tag{9.8}$$

As a consequence, applying Theorem 9.1 to the product between $\sin x$ and $\frac{1}{\cos x}$, we can see that

$$(\tan x)' = \frac{(\sin x)'}{\cos x} + \sin x \left(-\frac{(\cos x)'}{\cos^2 x} \right) = 1 + \tan^2 x = \frac{1}{\cos^2 x}, \qquad \forall\, x \in \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi \,:\, k \in \mathbb{Z}\}. \tag{9.9}$$

Similarly, we have

$$(\cot x)' = -1 - \cot^2 x = -\frac{1}{\sin^2 x}, \qquad \forall\, x \in \mathbb{R} \setminus \{k\pi :\ k \in \mathbb{Z}\}. \tag{9.10}$$

Using Theorem 9.1, we can compute the derivatives of the inverse functions. For example

$$(\arctan x)' = \frac{1}{(\tan)'(\arctan x)} = \frac{1}{1 + \tan^2(\arctan x)} = \frac{1}{1 + x^2}, \qquad \forall\, x \in \mathbb{R}. \tag{9.11}$$

Similarly, we have

$$\begin{aligned} (\arcsin x)' &= \frac{1}{\sqrt{1 - x^2}}, \qquad \forall\, x \in (-1, 1) \\ (\arccos x)' &= -\frac{1}{\sqrt{1 - x^2}}, \qquad \forall\, x \in (-1, 1) \\ (\operatorname{arccot} x)' &= -\frac{1}{1 + x^2}, \qquad \forall\, x \in \mathbb{R}. \end{aligned} \tag{9.12}$$

It is also very useful to be able to compute the derivative of a composed function.

Theorem 9.2 (Chain rule)**.** *Let $f : X \subset \mathbb{R} \to \mathbb{R}$ and $g : Y \subset \mathbb{R} \to \mathbb{R}$ be functions. Assume that $x_0 \in \mathring{X}$, $f(x_0) \in \mathring{Y}$, f is differentiable at x_0 and g is differentiable at y_0. Then $g \circ f$ is differentiable at x_0 and $(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0)$.*

For example, if $f : X \subseteq \mathbb{R} \to (0, \infty)$, we have

$$(\log f(x))' = \frac{f'(x)}{f(x)}.$$

Now we will see the relation between differentiability and continuity of a function.

Theorem 9.3. *Let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be a function and let $x_0 \in \mathring{X}$. If f is differentiable at x_0, then it is continuous at x_0.*

As a consequence, we deduce the following corollary.

Corollary 9.1. *Let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be a function. Assume that f is differentiable in $\mathring{X}$. Then f is continuous in $\mathring{X}$.*

Taking, for example, $X = [0, \infty)$ and

$$\begin{aligned} f : [0, \infty) &\to \mathbb{R} \\ x &\mapsto \begin{cases} x^2 & \forall\, x > 0 \\ 1 & \text{for } x = 0, \end{cases} \end{aligned} \tag{9.13}$$

we can see that the differentiability in $\mathring{X}$ does not imply the continuity in X, but only in $\mathring{X}$.

9.2. **Differentiability from the right and from the left.** Given a, $b \in \mathbb{R}$, $a < b$ and a function $f : [a, b] \to \mathbb{R}$, according to Definition 9.1, we can wonder whether f is differentiable in the open interval (a, b) only. Such a Definition does not say anything about the extrema a and b of the interval, since they are not interior points. However, we may be interested in computing the slope of the graph at the extrema a and b. For this purpose, we define the notion of derivative *from the right* and *from the left.*

Definition 9.3. *Let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be a function and let $x_0 \in X \cap \mathcal{A}(X)$. Then f is*

- *differentiable from the right if there exists*
$$f'_+(x_0) := \lim_{h \to 0^+} \frac{f(x_0 + h) - f(x_0)}{h}$$
and $f'_+(x_0) \in \mathbb{R}$. In this case, $f'_+(x_0)$ is known as the right derivative of f at x_0.
- *differentiable from the right if there exists*
$$f'_-(x_0) := \lim_{h \to 0^-} \frac{f(x_0 + h) - f(x_0)}{h}$$
and $f'_+(x_0) \in \mathbb{R}$. In this case, $f'_-(x_0)$ is known as the left derivative of f at x_0.

From Definition 9.3, we can see that f is differentiable at $x_0 \in X \cap \mathcal{A}(X)$ if and only if it is differentiable both from the right and from the left at x_0 and $f'_+(x_0) = f'_-(x_0)$. In this case, we have $f'(x_0) = f'_+(x_0) = f'_-(x_0)$.

The notion of differentiability from the right or from the left enables us to extend the definition of derivative to the extrema of an interval. More precisely, given a, $b \in \mathbb{R}$, $a < b$ and a function $f : [a, b] \to \mathbb{R}$, we say that

- f is differentiable at a if there exists $f'_+(a)$ and $f'_+(a) \in \mathbb{R}$. In this case, we set $f'(a) := f'_+(a)$.
- f is differentiable at b if there exists $f'_-(b)$ and $f'_-(b) \in \mathbb{R}$. In this case, we set $f'(b) := f'_-(b)$.

9.3. **Classification of non-differentiability points.** If f is not differentiable at a point $x_0 \in X$, many different behaviors are possible. Here we will list some of the most common cases.

Definition 9.4. *Let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be a function. Assume that both $f'_+(x_0)$ and $f'_-(x_0)$ exist. A point $x_0 \in X \cap \mathcal{A}(X)$ is said to be*

- *an angular point if either $f'_+(x_0) \in \mathbb{R}$ or $f'_-(x_0) \in \mathbb{R}$.*
- *a cusp if $f'_+(x_0) = \infty \wedge f'_-(x_0) = -\infty$ or $f'_+(x_0) = -\infty \wedge f'_-(x_0) = \infty$.*

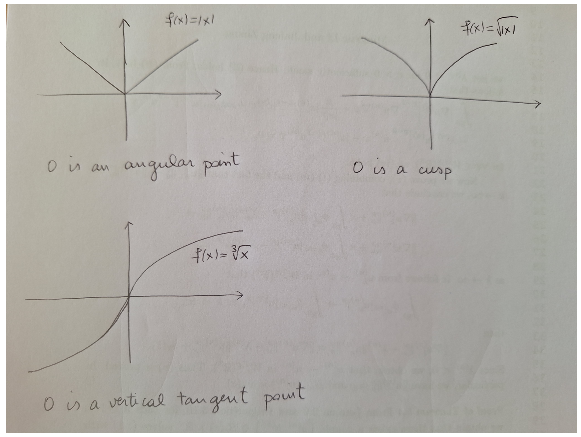

FIGURE 28. Some examples of non-differentiability points

- *a vertical tangent point if* $f'_+(x_0) = f'_-(x_0) = \infty$ *or* $f'_+(x_0) = f'_-(x_0) = -\infty$.

For example,

- 0 is an angular point of $f(x) := |x|$. In fact
$$f'_+(0) = \lim_{h\to 0^+} \frac{|h|}{h} = 1 \neq -1 = \lim_{h\to 0^-} \frac{|h|}{h} = f'_-(0).$$
- 0 is a cusp of $f(x) = \sqrt{|x|}$, since
$$f'_+(0) = \lim_{h\to 0^+} \frac{\sqrt{|h|}}{h} = \lim_{h\to 0^+} \frac{1}{\sqrt{h}} = \infty,$$
while
$$f'_-(0) = \lim_{h\to 0^-} \frac{\sqrt{|h|}}{h} = \lim_{h\to 0^-} \frac{\sqrt{-h}}{-(\sqrt{-h})^2} = \lim_{h\to 0^-} \frac{1}{-\sqrt{-h}} = -\infty.$$
- 0 is a vertical tangent point of the function $f(x) := \sqrt[3]{x}$. In fact
$$\lim_{h\to 0^+} \frac{\sqrt[3]{h}}{h} = \lim_{h\to 0^-} \frac{\sqrt[3]{h}}{h} = \infty.$$

These examples are shown in Figure 28. In particular, if $a, b \in \mathbb{R}$, $a < b$ and $I := [a, b]$ is a closed and bounded interval and $f : I \to \mathbb{R}$ is a function, we say

9.4. **Critical points.** In this section, we introduce the notion of *critical points* of a function. This notion is crucial in mathematical analysis, in general. In particular, for our purposes, we need to determine the critical points of a function to compute its maximum and minimum, if they are achieved.

Definition 9.5. *Let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be a function. A point $x_0 \in \mathring{X}$ is said to be a critical point of f if f is differentiable at x_0 and $f'(x_0) = 0$.*

Note that, since the equation of the tangent line to a graph is given $y = f'(x_0)(x - x_0) + f(x_0)$, $x_0 \in \mathring{X}$ is a critical point of f if and only if such a tangent line is horizontal.

There is an important link between critical points and the extrema of a function, that is maximum and minimum points. In Section 8, we introduced the notion of maximum and minimum points of a function. However, the extrema are not necessarily global; indeed, they can be only local, in the sense of the following definition.

Definition 9.6. *Let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be a function. We say that $x_0 \in X$ is*

- *a local minimum point if there exists $\delta > 0$ such that $f(x) \geq f(x_0)$ for any $x \in (x_0 - \delta, x_0 + \delta) \cap X$.*
- *a strict local minimum point if there exists $\delta > 0$ such that $f(x) > f(x_0)$ for any $x \in (x_0 - \delta, x_0 + \delta) \cap X$.*
- *a local maximum point if there exists $\delta > 0$ such that $f(x) \leq f(x_0)$ for any $x \in (x_0 - \delta, x_0 + \delta) \cap X$.*
- *a strict local maximum point if there exists $\delta > 0$ such that $f(x) < f(x_0)$ for any $x \in (x_0 - \delta, x_0 + \delta) \cap X$.*

It follows from the definition that a strict local minimum point is also a local minimum point and a global minimum point is also a local minimum point, but the opposite implications are not necessarily true. A similar statement holds for maximum points.

Some examples of critical points are shown in Figure (29).

- 0 is a global minimum point of $f(x) := x^2$.
- 1 is a global maximum point of $f(x) = x(2 - x)$.
- $\pm 1/\sqrt{3}$ are a strict local maximum point and a strict local minimum point of $f(x) := x(1 - x^2)$ respectively.
- 0 is neither a local minimum point nor a local maximum point of x^3.

We look for criteria to establish the existence of critical points of a given function $f : X \subset \mathbb{R} \to \mathbb{R}$. The first result in this direction is the Rolle Theorem 9.5, which follows from the more general Lagrange 9.4 Theorem as a corollary.

Theorem 9.4 (Lagrange)**.** *Let $a, b \in \mathbb{R}$, $a < b$ and $f : [a, b] \to \mathbb{R}$ be a continuous function. Assume that f is differentiable in (a, b). Then there exists $\xi \in (a, b)$ such that*

$$f(b) - f(a) = f'(\xi)(b - a).$$

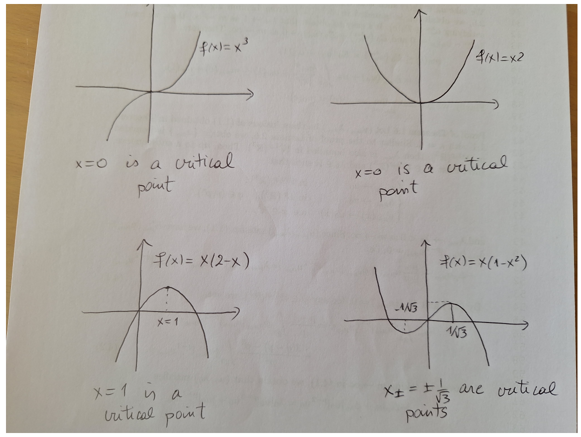

FIGURE 29. Some examples of critical points

In particular, if $f(a) = f(b)$, we obtain the following corollary, known as the Rolle Theorem.

Theorem 9.5 (Rolle)**.** *Let $a, b \in \mathbb{R}$, $a < b$ and $f : [a, b] \to \mathbb{R}$ be a continuous function. Assume that f is differentiable in (a, b) and $f(a) = f(b)$. Then there exists $\xi \in (a, b)$ such that $f'(\xi) = 0$.*

However, the Rolle Theorem only establishes the existence of at least one critical point, but it does not enable us to compute it exactly. For this purpose, we introduce the following necessary condition.

Proposition 9.1. *Let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be a function and let $x_0 \in \mathring{X}$ be a local minimum point or a local maximum point. If f is differentiable at x_0, then $f'(x_0) = 0$.*

Proof. Assume, for example, that x_0 is a local minimum point. Hence there exists $\delta_1 > 0$ such that $f(x) \geq f(x_0)$ for any $x \in (x_0 - \delta_1, x_0 + \delta_1) \cap X$. Since $x \in \mathring{X}$, there exists $\delta_2 > 0$ such that $(x_0 - \delta_2, x_0 + \delta_2) \subseteq X$. Taking $\delta := \min\{\delta_1, \delta_2\} > 0$, we have $(x_0 - \delta, x_0 + \delta) \subseteq X$ and $f(x) \geq f(x_0)$, for any $x \in (x_0 - \delta, x_0 + \delta)$. As a consequence,

$$f(x_0 + h) - f(x) \geq 0 \qquad \forall\, h \in (-\delta, \delta).$$

This yileds that

$$f'_+(x_0) = \lim_{h \to 0^+} \frac{f(x_0 + h) - f(x)}{h} \geq 0, \qquad f'_-(x_0) = \lim_{h \to 0^-} \frac{f(x_0 + h) - f(x)}{h} \leq 0.$$

Moreover, since f is differentiable at x_0, we have $f'_+(x_0) = f'_-(x_0) = f'(x_0)$. In conclusion, we have $f'(x_0) = 0$.

A similar proof can be done in the case of local maximisers. □

We note that Proposition 9.1 also holds for strict local minimisers (or maximisers) and global minimisers (or maximisers), since they are also local minimisers (or maximisers).

Note that the hypothesis that f is differentiable at x_0 is crucial, both in Theorem 9.5 and in Proposition 9.1. In fact, if we consider the function $f(x) := |x|$ in the interval $[-1, 1]$, we have $f(1) = f(-1)$ but f has no critical points in $(-1, 1)$.

It is also crucial to assume that x_0 is an interior point. In fact, if $f(x) = 2x + 1$ and we consider its restriction $g := f\big|_{[0,1]}$ to the closed interval $[0, 1]$, we have

$$g(0) = 1 = \min_{x\in[0,1]} g(x) < \max_{x\in[0,1]} g(x) = g(1) = 3$$

but $g'(0) = g'(1) = 2 \neq 0$.

Proposition 9.1 can be used as follows: if we want to find and classify the extrema (maximisers and minimisers) of a given differentiable function f, we first compute its derivative and solve equation $f'(x) = 0$, in order to find the critical points. However, not all the critical points are extrema. See for example the function $f(x) = x^3$ represented in Figure (29). As a consequence, computing the derivative is not enough to find and classify the extrema of a function. For this purpose we need some additional strategy.

9.5. **Derivatives and monotonicity.** In this Section we will investigate the relation between the sign of the derivative of a function and its monotonicity. First we consider functions which are defined on an interval.

Theorem 9.6. *Let $I \subseteq \mathbb{R}$ be an interval and let $f : I \to \mathbb{R}$ be a differentiable function.*

(1) *If f monotone non-decreasing if and only if $f'(x) \geq 0$ for any $x \in I$.*
(2) *If f monotone decreasing if and only if $f'(x) \leq 0$ for any $x \in I$.*

Proof. (1) On the one hand, if f is non-increasing and $x \in \mathring{I}$, then there exists $\delta > 0$ such that

$$\frac{f(x+h) - f(x)}{h} \geq 0, \qquad \forall h \in (-\delta, \delta) \setminus \{0\}.$$

As a consequence, taking the limit as $h \to 0$, we have $f'(x) \geq 0$. A similar proof can be done in case $x \in I \setminus \mathring{I}$.

The opposite implication is a consequence of the Lagrange Theorem 9.4. In fact, if x, $y \in I$ are such that $y > x$, then by the Lagrange Theorem there exists $\xi \in (x, y)$ such that

$$f(y) - f(x) = f'(\xi)(y - x) \geq 0,$$

since $f'(\xi) \geq 0$ for any $\xi \in I$.

(2) The proof is the same as above.

□

Using once again the Lagrange Theorem, we can prove the following result.

Theorem 9.7. *Let $I \subseteq \mathbb{R}$ be an interval and let $f : I \to \mathbb{R}$ be a differentiable function.*

(1) *If $f'(x) > 0$ for any $x \in I$, then f is strictly increasing.*
(2) *If $f'(x) < 0$ for any $x \in I$, then f is strictly decreasing.*

Proof. The proof of relies on the Lagrange Theorem and the fact that f' has a constant sign in I. □

The opposite implications are in general not true in Theorem 9.7. A counterexample is provided by the function $f(x) := x^3$ which is strictly monotone in $\mathbb{R}$ but its derivative $f'(x) = 3x^2$ vanishes at $x = 0$ (see Figure 29).

The assumption that f is defined on an interval is crucial in Theorems 9.6 and 9.7. For example, the function $f(x) := \frac{1}{x}$ is defined on $X := \mathbb{R} \setminus \{0\}$, which is not an interval. Its derivative fulfills $f'(x) = -\frac{1}{x^2} < 0$ for any $x \in X$, but f is not non-increasing in X. In fact $f(-1) = -1 < 1 = f(1)$.

As a consequence, if we want to classify the critical points of a function, we need to study the sign of the derivative in a neighborhood of such critical points. In particular, the following result is true.

Proposition 9.2. *Let $f : X \subset \mathbb{R} \to \mathbb{R}$ be a function. Assume that f is differentiable in $\mathring{X}$ and $x_0 \in \mathring{X}$ is such that $f'(x_0) = 0$.*

- *If there exists $\delta > 0$ such that $f'(x) \leq 0$ for any $x \in (x_0 - \delta, x_0)$ and $f'(x) \geq 0$ for any $x \in (x_0, x_0 + \delta)$, then x_0 is a local minimiser.*
- *If there exists $\delta > 0$ such that $f'(x) < 0$ for any $x \in (x_0 - \delta, x_0)$ and $f'(x) > 0$ for any $x \in (x_0, x_0 + \delta)$, then x_0 is a strict local minimiser.*
- *If there exists $\delta > 0$ such that $f'(x) \geq 0$ for any $x \in (x_0 - \delta, x_0)$ and $f'(x) \leq 0$ for any $x \in (x_0, x_0 + \delta)$, then x_0 is a local maximiser.*
- *If there exists $\delta > 0$ such that $f'(x) > 0$ for any $x \in (x_0 - \delta, x_0)$ and $f'(x) < 0$ for any $x \in (x_0, x_0 + \delta)$, then x_0 is a strict local maximiser.*

9.6. **Examples.** We want to draw a precise graph of the function $f(x) := x(1-x^2) = x - x^3$ (see Figure (29)).

- The domain is $X := \mathbb{R}$.
- $X \setminus \mathcal{A}(X) = \{\pm\infty\}$ and $\lim_{x\to\pm\infty}(x - x^3) = \lim_{x\to\pm\infty}(\frac{1}{x^2} - 1)(-x^3) = \mp\infty$.
- $f'(x) = 1 - 3x^2$ satisfies $f'(x) = 0$ if and only if $x = \pm\frac{1}{\sqrt{3}}$, $f'(x) > 0$ if and only if $-\frac{1}{\sqrt{3}} < x < \frac{1}{\sqrt{3}}$ and $f'(x) < 0$ if and only if $x < -\frac{1}{\sqrt{3}} \vee x > \frac{1}{\sqrt{3}}$.
- Finally, $\frac{1}{\sqrt{3}}$ is a local maximiser, $-\frac{1}{\sqrt{3}}$ is a local minimiser but both extrema are not global, since $\sup_{x\in\mathbb{R}} f(x) = \infty$ and $\inf_{x\in\mathbb{R}} f(x) = -\infty$.

9.7. **Concavity and convexity of a function.** Let $\Omega \subset \mathbb{R}^2$ be a subset. We say that Ω is convex if for any p, $q \in \Omega$, the segment joining p and q is contained in Ω.

Let $I \subseteq \mathbb{R}$ be an interval and let $f : I \to \mathbb{R}$ be a function. We say that f is *convex* if the set

$$\Omega^+(f) := \{(x, y) \in \mathbb{R}^2 : y > f(x)\}$$

is convex.

On the other hand, we say that f is *concave* if the set

$$\Omega^-(f) := \{(x, y) \in \mathbb{R}^2 : y < f(x)\}$$

is convex.

The situation is represented in Figure (30).

We want to have a criterium to establish if a given function is concave, convex or none of the two. In order to have such a criterium, we need to introduce the second derivative.

Definition 9.7. *Let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be a function. Assume that f is differentiable in $\mathring{X}$. If f' is differentiable in $x_0 \in \mathring{X}$, then we set $f''(x_0) := (f')'(x_0)$. If it is the case, we say that f is twice differentiable at x_0 and $f''(x_0)$ is its second derivative at x_0.*

Moreover, if f' is differentiable at x_0 for any $x_0 \in \mathring{X}$, we say that f is twice differentiable in $\mathring{X}$ with second derivative f''.

In case X is an interval, the second derivative can be defined at the extrema too, similarly to the derivative (or first derivative).

The second derivative is related to the convexity (or concavity) of the function in this way.

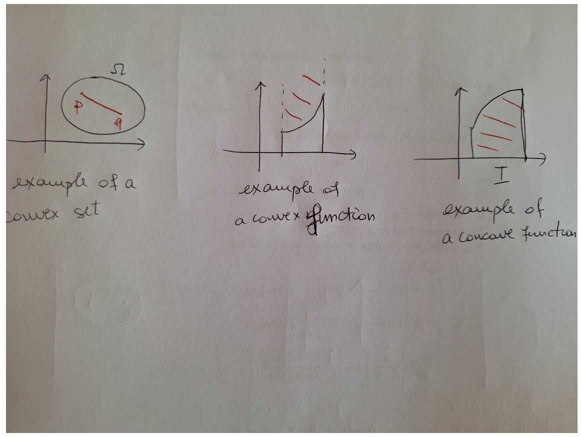

FIGURE 30. Examples of a convex set, a convex function and a concave function

Proposition 9.3. *Let $I \subset \mathbb{R}$ be an interval and let $f : I \to \mathbb{R}$ be a twice differentiable function. Then*

- *f is convex if and only if $f''(x) \geq 0$, for any $x \in I$.*
- *f is concave if and only if $f''(x) \leq 0$, for any $x \in I$.*

For example,

- the function $f(x) = x^2$ satisfies $f'(x) = 2x$ and $f''(x) = 2$, hence it is convex;
- the function $f(x) = \log x$ fulfills $f'(x) = \frac{1}{x}$ and $f''(x) = -\frac{1}{x^2} < 0$, hence it is concave;
- the function $f(x) := x - x^3$ satisfies $f'(x) = 1 - 3x^2$ and $f''(x) = -6x$. Hence it is convex in $(-\infty, 0)$ and concave in $(0, \infty)$.

9.8. **The classification of critical points and the second derivative.** The second derivative can be used to classify the critical points of a given twice differentiable function.

Definition 9.8. *Let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be a function. Assume that f is twice differentiable in $\mathring{X}$ and $x_0 \in \mathring{X}$ is a critical point of f. Then we say that x_0 is nondegenerate if $f''(x_0) \neq 0$.*

Nondegenerate critical points can be either local minimisers or local maximisers.

Proposition 9.4. *Let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be a function. Assume that f is twice differentiable in $\mathring{X}$ and $x_0 \in \mathring{X}$ is a nondegenerate critical point of f.*

- *If $f''(x_0) > 0$, then x_0 is a strict local minimum point.*
- *If $f''(x_0) < 0$, then x_0 is a strict maximum point.*

As a consequence, a possible criterium to understand if a critical point is a strict local minimiser or a strict local maximiser is to compute the second derivative at x_0.

However, this criterium does not say whether the maximum (or minimum) is global or not.

Note, if the critical point is degenerate, that is $f''(x_0) = 0$, we cannot say anything about the nature of the critical point itself, in the sense that we cannot classify it. In fact,

- if $f(x) = x^3$, we have $f''(0) = 0$ but 0 is neither a local minimiser nor a local maximiser;
- if $f(x) = x^4$, we still have $f''(0) = 0$ and 0 is a strict local (actually global) minimiser;
- if $f(x) = -x^4$, we still have $f''(0) = 0$ and 0 is a strict local (actually global) maximiser;
- if
$$f(x) = \begin{cases} x^4 & \forall\, x < 0 \\ 0 & \forall\, 0 \le x \le 1 \\ (x-1)^4 & \forall\, x > 1, \end{cases}$$
then any $x \in [0,1]$ fulfills $f''(x) = 0$ is a local minimiser (not strict) and.

10. Riemann integrals

Given a, $b \in \mathbb{R}$ and a function $f : [a,b] \to \mathbb{R}$, we are interested in computing the area below the graph of f between a and b. If f is either constant or a straight line, this can be done by using elementary Euclidean geometry, which allows us to calculate the area of triangles and rectangles. For more complicated functions, this cannot be done. As a consequence, we need to introduce more sophisticated tools, that is the integral calculus.

10.1. **The primitive of a function.** Before computing the area below a graph, we need to introduce the "inverse differentiation" operation. In this section we introduce the concept of the *primitive of a function.*

Definition 10.1. *Let $X \subseteq \mathbb{R}$ be a set with $\mathring{X} \neq \emptyset$ and let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be a function. Then a function $F : X \subset \mathbb{R}$ is said to be a primitive of f in X if F is differentiable in $\mathring{X}$ and $F' = f$.*

As an immediate consequence of the definition, we have the following result.

Proposition 10.1. *Let $I \subseteq \mathbb{R}$ be an interval and let $f : I \to \mathbb{R}$ be a function. F_1 and F_2 be two primitives of f in $\mathring{I}$. Then $F_1 - F_2$ is constant.*

Proof. It is enough to observe that $(F_1 - F_2)' = 0$, which yields that $F_1 - F_2 \equiv c$, for some constant $c \in \mathbb{R}$. $\square$

If F is a primitive of f in X and $c \in \mathbb{R}$ is a constant, we say that the *integral* or *indefinite integral* of f is

$$\int f(x)\,dx := F(x) + c, \qquad \forall\, x \in \mathring{X}. \tag{10.1}$$

Note that the integral of a given function is not unique, since it defined up to a constant.

Roughly speaking, the integration is the inverse operation of differentiation.

Some relevant examples are given by

$$\begin{aligned}
\int x^\alpha dx &= \frac{x^{\alpha+1}}{\alpha+1} + c, \qquad \forall\, x \in D(x^\alpha),\ \alpha \neq -1\\
\int \frac{1}{x} dx &= \log x + c, \qquad \forall\, x \in (0,\infty)\\
\int e^x dx &= e^x + c, \qquad \forall\, x \in \mathbb{R}\\
\int \sin x\,dx &= -\cos x + c, \qquad \forall\, x \in \mathbb{R}\\
\int \cos x\,dx &= -\sin x + c, \qquad \forall\, x \in \mathbb{R}\\
\int \frac{1}{\cos^2 x} dx &= \tan x + c, \qquad \forall\, x \in \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi\ :\ k \in \mathbb{Z}\}\\
\int \frac{1}{\sin^2 x} dx &= -\cot x + c, \qquad \forall\, x \in \mathbb{R} \setminus \{k\pi\ :\ k \in \mathbb{Z}\}\\
\int \frac{1}{1+x^2} dx &= \arctan x + c, \qquad \forall\, x \in \mathbb{R}\\
\int \frac{-1}{1+x^2} dx &= \mathrm{arccot} x + c, \qquad \forall\, x \in \mathbb{R}\\
\int \frac{1}{\sqrt{1-x^2}} dx &= \arcsin x + c \qquad \forall x \in (-1,1)\\
\int \frac{-1}{\sqrt{1-x^2}} dx &= \arccos x + c \qquad \forall x \in (-1,1)
\end{aligned} \tag{10.2}$$

Some useful calculus techniques are the following.

Theorem 10.1 (Linearity of the integral)**.** *Let $X \subset \mathbb{R}$ be a set with $\mathring{X} \neq \emptyset$ and let $f, g : X \to \mathbb{R}$ be functions. Let $\alpha, \beta \in \mathbb{R}$. Assume that F is a primitive of f and G is a primitive of G. Then $\alpha F + \beta G$ is a primitive of $\alpha f + \beta g$. In other words*

$$\alpha \int f(x)dx + \beta \int g(x)dx = \int (\alpha f(x) + \beta g(x))dx. \tag{10.3}$$

Proof. This can be proved by using the linearity of the derivative. □

Theorem 10.2 (Integration by parts)**.** *Let $X \subset \mathbb{R}$ be a set with $\mathring{X} \neq \emptyset$ and let $f, g : X \to \mathbb{R}$ be functions. Assume that both f and g are differentiable in $\mathring{X}$. Then*

$$\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx. \tag{10.4}$$

Proof. This is a consequence of the Leibnitz rule $(fg)' = f'g + fg'$. □

10.2. **Riemann integral.** Now we can define the *Riemann integral*, which is a tool to calculate the area below a given graph.

Definition 10.2. *Let $a, b \in \mathbb{R}$ with $a < b$ and let $I := [a, b]$. A partition of I is any finite set $P \subseteq [a, b]$ such that $a, b \in P$.*

Let $f : [a, b] \to \mathbb{R}$ be a bounded function, let $n \in \mathbb{N} \setminus \{0\}$ and let $P := \{x_0, \dots, x_n\}$ be a partition of $[a, b]$ of cardinality $n + 1$ such that

$$a = x_0 < x - 1 < \cdots < x_n = b.$$

We define the *upper sums* and the *lower sums* of f relative to P as

$$L(f, P) := \sum_{i=1}^{n} (x_i - x_{i-1}) \left(\inf_{x_{i-1} \le x \le x_i} f(x) \right), \qquad U(f, P) := \sum_{i=1}^{n} (x_i - x_{i-1}) \left(\sup_{x_{i-1} \le x \le x_i} f(x) \right) \tag{10.5}$$

For any partition $P \subset [a, b]$, we have

$$(b - a) \left(\inf_{x \in [a,b]} f(x) \right) \le L(f, P) \le U(f, P) \le (b - a) \left(\sup_{x \in [a,b]} f(x) \right).$$

As a consequence, defining the *lower integral* of f on $[a, b]$ as

$$\underline{\int_a^b f(x)dx} := \sup\{L(f, P) : \, P \text{ is a partition of } [a, b]\}$$

and the *upper integral* of f on $[a, b]$ as

$$\overline{\int_a^b f(x)dx} := \inf\{U(f, P) : \, P \text{ is a partition of } [a, b]\},$$

we have

$$(b-a)\left(\inf_{x\in[a,b]} f(x)\right) \le \underline{\int_a^b} f(x)dx \le \overline{\int_a^b} f(x)dx \le (b-a)\left(\inf_{x\in[a,b]} f(x)\right).$$

In view of these inequalities, we can define *Riemann-integrable functions* as follows.

Definition 10.3. *Let a, $b \in \mathbb{R}$ and let $f : [a,b] \to \mathbb{R}$ be a bounded function. We say that f is Riemann-integrable if*

$$\underline{\int_a^b} f(x)dx = \overline{\int_a^b} f(x)dx.$$

If f is Riemann integrable, we set

$$\int_a^b f(x)dx := \underline{\int_a^b} f(x)dx = \overline{\int_a^b} f(x)dx.$$

This number is known as the integral of f over $[a,b]$.

The integral of a function over $[a,b]$ represents the area of the region between the graph of f over $[a,b]$ and the x-axis, or more precisely the are signed area. In fact, from the definition it follows that, if f is positive, then its integral is also positive, while, if f is negative, its integral is also negative.

Not all bounded functions defined on a closed and bounded interval are Riemann-integrable. For example, the function $f : [0,1] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1 & \text{if } x \in [0,1] \cap \mathbb{Q} \\ 0 & \text{if } x \in [0,1] \setminus \mathbb{Q} \end{cases}$$

fulfills

$$0 = \underline{\int_0^1} f(x)dx < \overline{\int_0^1} f(x)dx = 1,$$

which yields that f is not Riemann-integrable.

A first example of Riemann-integrable functions is given by continuous functions.

Theorem 10.3. *Let a, $b \in \mathbb{R}$ with $a < b$. Let $f : [a,b] \to \mathbb{R}$ be a continuous function. Then it is Riemann-integrable.*

Proof. By Theorem 8.1, f is uniformly continuous in $[a,b]$. As a consequence, for any $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ for any x, $y \in [a,b]$ such that

$|x-y|>\delta$. Taking a partition $P=\{x_0,\dots,x_n\}$ of $[a,b]$ such that $x_0=a$, $x_n=b$ and $|x_i-x_{i-1}|<\delta$ for any $1\le i\le n$, we have

$$0\le U(f,P)-L(f,P)\le\sum_{i=1}^n\left(\max_{x_{i-1}\le x\le x_i}f(x)-\min_{x_{i-1}\le x\le x_i}f(x)\right)(x_i-x_{i-1})\le\varepsilon(b-a).$$

As a consequence, we can see that

$$0\le\overline{\int_a^b f(x)dx}-\underline{\int_a^b f(x)dx}\le\varepsilon(b-a).$$

Since $\varepsilon>0$ is arbitrary, we conclude the proof. □

In view of Theorem 10.3, given a, $b\in\mathbb{R}$ with $a<b$ and a continuous function $f:[a,b]\to\mathbb{R}$, we can always define the *integral function*

$$\begin{aligned}&F:[a,b]\to\mathbb{R}\\&F(x):=\int_a^x f(t)dt.\end{aligned}\tag{10.6}$$

We will see that F is a primitive of f, or, in other words, $F'=f$. In order to do so, we need a preliminary result.

Theorem 10.4 (Mean value Theorem for integrals)**.** *Let a, $b\in\mathbb{R}$ with $a<b$. Let $f:[a,b]\to\mathbb{R}$ be a continuous function. Then there exists $c\in[a,b]$ such that*

$$f(c)=\frac{1}{b-a}\int_a^b f(x)dx.$$

Proof. Due to Definition 10.3, we have

$$y_0:=\frac{1}{b-a}\int_a^b f(x)dx\in[\min_{x\in[a,b]}f(x),\max_{x\in[a,b]}f(x)].$$

Since f is continuous, the conclusion follows from the intermediate value Theorem 8.6. □

Moreover, we have the following results, which follow from the definition of integral.

Theorem 10.5 (Linearity of the integral)**.** *Let a, $b\in\mathbb{R}$ with $a<b$. Let f, $g:[a,b]\to\mathbb{R}$ be continuous functions and let λ, $\mu\in\mathbb{R}$. Then*

$$\int_a^b(\lambda f(x)+\mu g(x))dx=\lambda\int_a^b f(x)dx+\mu\int_a^b g(x)dx.$$

Theorem 10.6 (Additivity)**.** *Let a, b, $c \in \mathbb{R}$ with $a < b < c$. Let $f : [a, c] \to \mathbb{R}$ be a continuous functions. Then*

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx.$$

Now we ready to prove that the function F defined in 10.6 is actually a primitive of f.

Theorem 10.7 (Existence of a primitive)**.** *Let a, $b \in \mathbb{R}$ with $a < b$ and let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then the function F defined in 10.6 is differentiable in $[a, b]$ and satisfies $F' = f$ in $[a, b]$.*

Proof. Since f is continuous on $[a, b]$, for any $\varepsilon > 0$ and for any $x \in [a, b]$, there exists $\delta > 0$ such that, for any $h \in (-\delta, \delta)$, we have $|f(x + h) - f(x)| < \varepsilon$.

As a consequence, due to the additivity of the integral (see Theorem 10.6) and the mean value Theorem for integrals (see Theorem 10.4), taking $h \in (-\delta, \delta) \setminus \{0\}$, for any $x \in [a, b]$ there exists $c \in [x - h, x + h]$ such that

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h}\left(\int_a^{x+h} f(t)dt - \int_a^x f(t)dt\right) = \frac{1}{h}\int_x^{x+h} f(t)dt = f(c).$$

Finally, due to the uniform continuity of f, we have

$$\left|\frac{F(x+h) - F(x)}{h} - f(x)\right| = |f(c) - f(x)| < \varepsilon$$

if $0 < |h| < \delta$. This concludes the proof. □

Remark 10.1. *In particular, Theorem 10.7 shoes that any continuous function $f : [a, b] \to \mathbb{R}$ has a primitive.*

The aim of the Riemann integral is to compute the signed area of the region between the graph of a function $f : [a, b] \to \mathbb{R}$ and the x-axis. The next result, known as the *fundamental Theorem of calculus*, enables us to make this computation for a given function f. Such a result can be proved by applying Theorem 10.7.

Theorem 10.8. *Let a, $b \in \mathbb{R}$ with $a < b$ and let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then*

$$\int_a^b f(x)dx = F(b) - F(a),$$

where F is a primitive of f in $[a, b]$.

Proof. For any $\varepsilon > 0$, we can consider a partition P_1 of $[a, b]$ such that
$$\int_a^b f(t)dt - \varepsilon < L(f, P_1)$$
and a partition P_2 of $[a, b]$ such that
$$U(f, P_2) < \int_a^b f(t)dt + \varepsilon.$$
This follows from the definition of sup and inf. Taking $P := P_1 \cup P_2$, we have
$$\int_a^b f(t)dt - \varepsilon < L(f, P_1) \leq L(f, P) \leq U(f, P) \leq U(f, P_2) < \int_a^b f(t)dt + \varepsilon. \quad (10.7)$$
Writing $P = \{x_0, \dots, x_n\}$, with $x_0 = a$, $x_n = b$, and using the continuity of f, we can apply Theorem 10.7 and the Lagrange Theorem, which yield that, for any $1 \leq i \leq n$, there exists $c_i \in (x_{i-1}, x_i)$ such that
$$F(b) - F(a) = \sum_{i=1}^{n} (F(x_i) - F(x_{i-1})) = \sum_{i=1}^{n} f(c_i)(x_i - x_{i-1}).$$
As a consequence, we can see that
$$\sum_{i=1}^{n} \left(\inf_{x_{i-1} \leq x \leq x_i} f(x) \right) (x_i - x_{i-1}) \leq F(b) - F(a) \leq \sum_{i=1}^{n} \left(\sup_{x_{i-1} \leq x \leq x_i} f(x) \right) (x_i - x_{i-1}),$$
or in other words
$$L(f, P) \leq F(b) - F(a) \leq U(f, P).$$
By 10.7, we can see that
$$\int_a^b f(t)dt - \varepsilon < L(f, P) \leq F(b) - F(a) \leq U(f, P) < \int_a^b f(t)dt + \varepsilon.$$
Since $\varepsilon > 0$ is arbitrary, we conclude the proof. □

10.3. **Some useful tools for computations.**

Theorem 10.9 (Integration by parts)**.** *Let f, $g : [a, b] \to \mathbb{R}$ be differentiable functions. Then*
$$\int_a^b f(x)g'(x)dx = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)g(x)dx.$$

Theorem 10.10. *Let $a, b, c, , d \in \mathbb{R}$ with $a < b$, $c < d$. Let $f : [a, b] \to \mathbb{R}$ be a Riemann-integrable function and let $g : [c, d] \to [a, b]$ be a differentiable function. Then*
$$\int_c^d f(g(x))g'(x)dx = \int_a^b f(y)dy.$$

Universitá degli studi di Bari Aldo Moro
Email address: matteo.rizzi@uniba.it, mrizzi1988@gmail.com