Mathematics for Natural Scientists I

Iosif Petrakis

Fakultät für Geowissenschaften, Ludwig-Maximilians-Universität München
Winter term 2019/2020

Contents

CHAPTER 1

Number systems

In this chapter we study some basic properties of the following number systems: the natural numbers $\mathbb{N}$, the integers $\mathbb{Z}$, the rational numbers $\mathbb{Q}$, and the real numbers $\mathbb{R}$. First we need to give a short introduction to the fundamental notions of a set and of a function between sets.

1.1. Sets

Definition 1.1.1. A formula ϕ is a mathematical expression. Let ϕ, ψ be formulas. The *implication* "ϕ implies ψ" is denoted by

$$\phi \Rightarrow \psi.$$

The *conjunction* "ϕ and ψ" is denoted by

$$\phi \wedge \psi, \quad \text{or by} \quad \phi \ \& \ \psi.$$

The *equivalence* "ϕ if and only if ψ" is denoted by

$$\phi \Leftrightarrow \psi$$

and it is the conjunction $\phi \Rightarrow \psi \ \& \ \psi \Rightarrow \phi$. The *negation* "not ϕ" is denoted by $\neg\phi$. The *disjunction* "ϕ or ψ" is denoted by

$$\phi \vee \psi.$$

Let $\phi(x)$ be a formula i.e., the variable x occurs in ϕ. The formula

$$\exists_x \big(\phi(x)\big)$$

is read as "there *exists* x such that $\phi(x)$ holds", and the formula

$$\forall_x \big(\phi(x)\big)$$

is read as "*for all* x we have that $\phi(x)$ holds".

To prove $\phi \Rightarrow \psi$, we suppose ϕ and we prove ψ. If ϕ is false, then the implication $\phi \Rightarrow \psi$ is true, in a trivial way. To prove $\phi \ \& \ \psi$, we prove ϕ and we prove ψ. To prove $\phi \Leftrightarrow \psi$, we prove $\phi \Rightarrow \psi$ and we prove $\psi \Rightarrow \phi$. To prove $\neg\phi$, we suppose ϕ and we reach a contradiction, like $\psi \ \& \ \neg\psi$, for some formula ψ. To prove $\phi \vee \psi$, we prove ϕ, or we prove ψ. Sometimes, to prove $\phi \vee \psi$, we prove $\neg(\neg\phi \ \& \ \neg\psi)$. To prove $\exists_x\big(\phi(x)\big)$, we find x and we prove $\phi(x)$. Sometimes, to prove $\exists_x\big(\phi(x)\big)$, we prove $\neg\neg\big[\exists_x\big(\phi(x)\big)\big]$. To prove $\forall_x\big(\phi(x)\big)$, we suppose an arbitrary x and we prove $\phi(x)$. Some basic examples of formulas appear in the next definition.

Definition 1.1.2. A *set* [Menge] X is a collection of mathematical objects. A mathematical object x that is a member of X is called an *element* of X, and we write

$$x \in X.$$

If y is a mathematical object that is not an element of X, we write

$$y \notin X$$

instead of $\neg(y \in X)$. The set that has no elements is called the *empty* set [leere Menge] and it is denoted by $\emptyset$. Let X and Y be sets. We say that X and Y are *equal*, in symbols $X = Y$, if they have the same elements i.e.,

$$\forall_x\big(x \in X \Leftrightarrow x \in Y\big).$$

We say that X is a *subset* [Teilmenge] of Y, in symbols $X \subseteq Y$, if every element of X is an element of Y i.e.,

$$\forall_x\big(x \in X \Rightarrow x \in Y\big).$$

If $X \subseteq Y$ and there is $y \in Y$ such that $y \notin X$, then we say that X is a *proper* subset [echte Teilmenge] of Y. In this case we write $X \subsetneq Y$. If X is a set, the collection $\mathcal{P}(X)$ of all subsets of X is called the *powerset* [Potenzmenge] of X.

Very often we use the symbols $\{\}$ to denote the elements of a set. E.g., the *set of natural numbers* [natürliche Zahlen] $\mathbb{N}$ is denoted by

$$\mathbb{N} = \{0, 1, 2, 3, \ldots\}.$$

The *set of integers* [ganze Zahlen] $\mathbb{Z}$ is denoted by

$$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\},$$

and

$$\mathbb{N} \subsetneq \mathbb{Z}.$$

If X is a set, we can define a subset X_P of X through a *property* $P(x)$ on X by collecting all elements of X such that $P(x)$ holds. In this case we write

$$X_P = \{x \in X \mid P(x)\}.$$

E.g., the set $\mathtt{Even}$ of even natural numbers [gerade Zahlen] is defined as follows:

$$\mathtt{Even} = \{n \in \mathbb{N} \mid P(n)\}, \quad P(n) :\Leftrightarrow \exists_{m\in\mathbb{N}}\big(n = 2m\big),$$

$$\exists_{m\in\mathbb{N}}\big(n = 2m\big) :\Leftrightarrow \exists_m\big(m \in \mathbb{N} \ \&\ n = 2m\big).$$

Clearly, $\mathtt{Even} \subsetneq \mathbb{N}$. The set $\mathtt{Odd}$ of odd natural numbers [ungerade Zahlen] is defined as follows:

$$\mathtt{Odd} = \{n \in \mathbb{N} \mid Q(n)\}, \quad Q(n) :\Leftrightarrow \exists_{m\in\mathbb{N}}\big(n = 2m + 1\big).$$

Notice that two sets X, Y are equal if they are subsets of each other i.e.,

$$X = Y \Leftrightarrow X \subseteq Y \ \&\ Y \subseteq X.$$

DEFINITION 1.1.3. Let X, Y be sets. The *intersection* [Schnittmenge] $X \cap Y$ of X and Y is the set of all mathematical objects that are elements both of X and Y i.e.,

$$X \cap Y = \{z \mid z \in X \;\&\; z \in Y\}.$$

The *union* [Vereinigungsmenge] $X \cup Y$ of X and Y is the set of all mathematical objects that are elements either of X or of Y i.e.,

$$X \cup Y = \{z \mid z \in X \;\vee\; z \in Y\}.$$

If $A \subseteq X$, the *complement* [Komplement] A' of A in X is the set of all elements of X that do not belong to A i.e.,

$$A' = \{x \in X \mid x \notin A\}.$$

It is easy to see that

$$\texttt{Odd} \cap \texttt{Even} = \emptyset \;\;\&\;\; \texttt{Odd} \cup \texttt{Even} = \mathbb{N} \;\;\&\;\; \texttt{Odd}' = \texttt{Even} \;\;\&\;\; \texttt{Even}' = \texttt{Odd}.$$

PROPOSITION 1.1.4. *If X is a set and A, B, C are subsets of X, the following hold:*

(i) $\emptyset \subseteq X$ *and* $X \subseteq X$.

(ii) $A \cap A = A$ *and* $A \cup A = A$.

(iii) $A \cap B = B \cap A$ *and* $A \cup B = B \cup A$.

(iv) $(A \cap B) \cap C = A \cap (B \cap C)$ *and* $(A \cup B) \cup C = A \cup (B \cup C)$.

(v) $(A \cap B) \cup A = A$ *and* $(A \cup B) \cap A = A$.

(vi) $A \subseteq B \Leftrightarrow A \cap B = A$ *and* $A \subseteq B \Leftrightarrow A \cup B = B$.

(vii) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ *and* $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

PROOF. (i) For $\emptyset \subseteq X$ we need to show that

$$\forall_x \big(x \in \emptyset \Rightarrow x \in X\big).$$

Let x such that $x \in \emptyset$. Since this is impossible by the definition of $\emptyset$, we conclude, in a trivial way, that the implication $x \in \emptyset \Rightarrow x \in X$ holds. Since x is arbitrary, the proof is completed. For $X \subseteq X$ we work similarly, and we use the fact that the implication $x \in X \Rightarrow x \in X$ holds.
(v) We show only $(A \cup B) \cap A = A$. For that we show first that $(A \cup B) \cap A \subseteq A$. If $b \in (A \cup B) \cap A$, we show that $b \in A$. By the definition of intersection we have that $b \in (A \cup B)$ and $b \in A$. Hence we get the required $b \in A$. Next we show that $A \subseteq (A \cup B) \cap A$. If $a \in A$, we show that $a \in (A \cup B) \cap A$ i.e., $a \in A \cup B$ and $a \in A$. Both inclusions follow trivially from the hypothesis $a \in A$.
(ii)-(iv) and (vi) -(vii) is an exercise. $\square$

PROPOSITION 1.1.5. *If X is a set and $A, B \subseteq X$, the following hold:*

(i) $\emptyset' = X$ *and* $X' = \emptyset$.

(ii) $A \cap A' = \emptyset$ *and* $A \cup A' = X$.

(iii) $\big(A'\big)' = A$.

(iv) $(A \cap B)' = A' \cup B'$.

(v) $(A \cup B)' = A' \cap B'$.

(vi) $A \subseteq B \Leftrightarrow B' \subseteq A'$.

PROOF. (i) By the definition of the complement of a subset we have that

$$\emptyset' = \{x \in X \mid x \notin \emptyset\} = X,$$
$$X' = \{x \in X \mid x \notin X\} = \emptyset.$$

For (ii) we use the logical principle $\phi \vee \neg\phi$ (Principle of the Excluded Middle, PEM), and for (iii) the principle $\neg\neg\phi \Rightarrow \phi$ (Double Negation Shift, DNS). □

PROPOSITION 1.1.6. *If X is a set and $A, B \subseteq X$, the difference $A - B$ between A and B is the set of all elements in A that are not in B i.e.,*

$$A - B = \{x \in X \mid x \in A \ \&\ x \notin B\}.$$

If $C \subseteq X$, the following hold:

(i) $A - B = A \cap B'$.

(ii) $(A - B) - C = A - (B \cup C)$.

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

(iv) $(A \cup B) - C = (A - C) \cup (B - C)$.

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

PROOF. Exercise. □

PROPOSITION 1.1.7. *If X is a set and $A, B \subseteq X$, the symmetric difference $A \triangle B$ of A and B is defined by*

$$A \triangle B = (A - B) \cup (B - A).$$

If $C \subseteq X$, the following hold:

(i) $A \triangle \emptyset = A$ *and* $A \triangle A = \emptyset$.

(ii) $A \triangle B = B \triangle A$.

(iii) $A \triangle (B \triangle C) = (A \triangle B) \triangle C$.

(iv) $A \cap (B \triangle C) = (A \cap C) \triangle (A \cap C)$.

PROOF. Exercise. □

DEFINITION 1.1.8. If X, Y are sets, their *product* $X \times Y$ is the set of all pairs (x, y) with $x \in X$ and $y \in Y$ i.e.,

$$X \times Y = \{(x, y) \mid x \in X \ \&\ y \in Y\},$$

where if $(x, y), (x', y') \in X \times Y$, we have that

$$(x, y) = (x', y') \Leftrightarrow x = x' \ \&\ y = y'.$$

1.2. Functions

Definition 1.2.1. If X, Y are sets, a *function* $f : X \to Y$ from X to Y is a rule that associates to every element $x \in X$ a unique element $f(x) \in Y$, which is called the *value* of f at x. To denote that f maps x to $f(x)$ we also write

$$x \mapsto f(x).$$

The set X is called the *domain* [Definitionsmenge] of f, and Y is the *range* of [Zielmenge] of f. The *image* [Wertemenge] $\mathtt{Im}(f)$ of f is the set of values of f i.e.,

$$\mathtt{Im}(f) = \{y \in Y \mid \exists_{x \in X}\big(y = f(x)\big)\}.$$

If $g : X \to Y$, then $f = g$, if f and g are equal on every input $x \in X$ i.e.,

$$f = g \Leftrightarrow \forall_{x \in X}\big(f(x) = g(x)\big).$$

By the uniqueness hypothesis in the Definition 1.2.1 a function $f : X \to Y$ satisfies for every $x, x' \in X$ the implication

$$x = w \Rightarrow f(x) = f(w)$$

i.e., it "maps equal inputs to equal outputs". Clearly, $\mathtt{Im}(f) \subseteq Y$. Let e.g., $f : \mathbb{N} \to \mathbb{N}$ be defined by

$$n \mapsto 2n.$$

By definition, $f(0) = 0$, $f(1) = 2$, and $f(50) = 100$. Clearly, $\mathtt{Im}(f) = \mathtt{Even}$.

Definition 1.2.2. A function $f : X \to Y$ is called an *injection*, or injective [injektiv], if for every $x, x' \in X$ we have that

$$f(x) = f(w) \Rightarrow x = w.$$

Moreover, f is called a *surjection*, or surjective [surjektiv], if $\mathtt{Im}(f) = Y$. A function f is called a *bijection*, or bijective [bijektiv], if it is both an injection and a surjection.

It is easy to see that f is injective, if for every $x, x' \in X$ it satisfies

$$x \neq w \Rightarrow f(x) \neq f(w)$$

i.e., if f "maps unequal inputs to unequal outputs". The function $n \mapsto 2n$ is injective, since $2n = 2m \Rightarrow n = m$, for every $n, m \in \mathbb{N}$, but it is not surjective, since $\mathtt{Im}(f) = \mathtt{Even} \subsetneq \mathbb{N}$. If X is a set, the *identity map* [identische Abbildung] $\mathrm{id}_X : X \to X$ is defined by the rule

$$x \mapsto x.$$

Clearly, id_X is a bijection. Let $g : \mathbb{Z} \to \mathbb{N}$ be defined by

$$g(z) := \begin{cases} z & , z \geq 0 \\ -z & , z < 0. \end{cases}$$

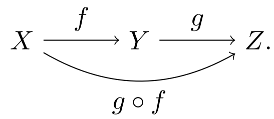
Then g is surjective, since $g(n) = n$, for every $n \in \mathbb{N}$, but g is not injective, since e.g., $g(-1) = g(1) = 1$. If X, Y are sets, and $y_0 \in Y$, let $\widehat{y_0} : X \to Y$ be defined by

$$x \mapsto y_0,$$

for every $x \in X$, is the *constant* function from X to Y with constant value y_0.

Definition 1.2.3. Let X, Y, Z be sets, $f : X \to Y$ and $g : Y \to Z$. The *composition* $g \circ f : X \to Z$ of f and g is defined, for every $x \in X$, by

$$(g \circ f)(x) = g(f(x))$$



Since g and f respect equality, the composition $g \circ f$ also respects equality i.e., if $x, w \in X$, such that $x = w$, then $f(x) = f(w)$ and hence

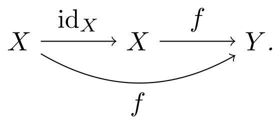
$$(g \circ f)(x) = g(f(x)) = g(f(w)) = (g \circ f)(w).$$

If $f : \mathbb{N} \to \mathbb{N}$ is defined by $f(n) = n+1$, for every $n \in \mathbb{N}$, and if $g : \mathbb{N} \to \mathbb{N}$ is defined by $g(n) = n^2$, for every $n \in \mathbb{N}$, then $g \circ f : \mathbb{N} \to \mathbb{N}$, and for every $n \in \mathbb{N}$

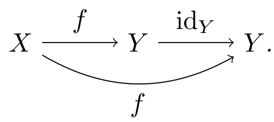
$$(g \circ f)(n) = g(f(n)) = (n+1)^2.$$

Proposition 1.2.4. *Let X, Y, Z, W be sets, and let $f : X \to Y$, $g : Y \to Z$, and $h : Z \to W$. The following hold:*

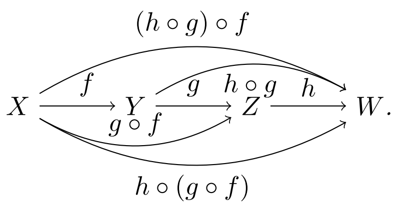
(i) $f \circ \mathrm{id}_X = f$



(ii) $\mathrm{id}_Y \circ f = f$



(iii) $h \circ (g \circ f) = (h \circ g) \circ f$



Proof. (i) By the definition of the equality of functions we need to show that

$$\forall_{x \in X}\big((f \circ \mathrm{id}_X)(x) = f(x)\big).$$

If $x \in X$, then $(f \circ \mathrm{id}_X)(x) = f(\mathrm{id}_X(x)) = f(x)$. Since x is an arbitrary element of X, we conclude that $f \circ \mathrm{id}_X = f$.
(ii) and (iii) Exercise. $\square$

1.3. Induction on $\mathbb{N}$

The induction principle on $\mathbb{N}$ is a fundamental tool in proving properties for *all* natural numbers. All induction principles mentioned in this section are equivalent.

Induction principle IND **on** $\mathbb{N}$: Let $\phi(n)$ be a formula on $\mathbb{N}$ such that the following conditions are satisfied:

(i) $\phi(0)$ holds.

(ii) For every $n \in \mathbb{N}$, if $\phi(n)$ holds, then $\phi(n+1)$ holds i.e.,

$$\forall_{n\in\mathbb{N}}\big(\phi(n) \Rightarrow \phi(n+1)\big).$$

Then we can infer that $\phi(n)$ holds, for every $n \in \mathbb{N}$ i.e.,

$$\forall_{n\in\mathbb{N}}\big(\phi(n)\big).$$

Let $\mathbb{N}^+$ be the set of non-zero natural numbers i.e.,

$$\mathbb{N}^+ = \{1, 2, 3, \ldots\}.$$

Induction principle IND$^+$ **on** $\mathbb{N}^+$: Let $\theta(n)$ be a formula on $\mathbb{N}^+$ such that the following conditions are satisfied:

(i) $\theta(1)$ holds.

(ii) For every $n \in \mathbb{N}^+$, if $\theta(n)$ holds, then $\theta(n+1)$ holds i.e.,

$$\forall_{n\in\mathbb{N}^+}\big(\theta(n) \Rightarrow \theta(n+1)\big).$$

Then we can infer that $\theta(n)$ holds, for every $n \in \mathbb{N}^+$ i.e.,

$$\forall_{n\in\mathbb{N}^+}\big(\theta(n)\big).$$

PROPOSITION 1.3.1. *The induction principle* IND *on* $\mathbb{N}$ *implies the induction principle* IND$^+$ *on* $\mathbb{N}^+$.

PROOF. Let $\theta(n)$ be a formula on $\mathbb{N}^+$ such that the conditions of IND$^+$ are satisfied. Let $\phi(n)$ be the following formula on $\mathbb{N}$

$$\phi(n) :\Leftrightarrow \theta(n+1).$$

By definition $\phi(0) :\Leftrightarrow \theta(1)$, which holds by our hypothesis on θ. Let $n \in \mathbb{N}$ such that $\phi(n) :\Leftrightarrow \theta(n+1)$. By our hypothesis on θ we get $\theta((n+1)+1) :\Leftrightarrow \phi(n+1)$, hence by IND we get

$$\forall_{n\in\mathbb{N}}\big(\phi(n)\big) :\Leftrightarrow \forall_{n\in\mathbb{N}}\big(\theta(n+1)\big) \Leftrightarrow \forall_{n\in\mathbb{N}^+}\big(\theta(n)\big).$$

$\square$

PROPOSITION 1.3.2. *The induction principle* IND^+ *on* $\mathbb{N}^+$ *implies the induction principle* IND *on* $\mathbb{N}$.

PROOF. Exercise. □

As an example of using IND^+, let us prove the following formula:

$$\forall_{n\in\mathbb{N}^+}\left(1+2+\ldots+n=\frac{n(n+1)}{2}\right)$$

If $\theta(n)$ is the formula on $\mathbb{N}^+$

$$\theta(n) :\Leftrightarrow 1+2+\ldots+n=\frac{n(n+1)}{2},$$

then by the principle IND^+ it suffices to show

$$\theta(1) :\Leftrightarrow 1=\frac{1(1+1)}{2},$$

which holds trivially, and if $n \in \mathbb{N}^+$ we need to show the following implication:

$$\theta(n) \Rightarrow \theta(n+1) \qquad \text{i.e.,}$$

$$\left[1+2+\ldots+n=\frac{n(n+1)}{2}\right] \Rightarrow \left[1+2+\ldots+n+(n+1)=\frac{(n+1)(n+2)}{2}\right].$$

For that we suppose that

$$1+2+\ldots+n=\frac{n(n+1)}{2}$$

holds, and then we show the equality

$$1+2+\ldots+n+(n+1)=\frac{(n+1)(n+2)}{2}$$

as follows:

$$\begin{aligned}1+2+\ldots+n+(n+1) &= \big[1+2+\ldots+n\big]+(n+1)\\ &= \frac{n(n+1)}{2}+(n+1)\\ &= \frac{n(n+1)}{2}+\frac{2(n+1)}{2}\\ &= \frac{n(n+1)+2(n+1)}{2}\\ &= \frac{(n+1)(n+2)}{2}.\end{aligned}$$

If $a_1, a_2, \ldots a_n \in \mathbb{N}$ we define their sum

$$\sum_{k=1}^{n} a_k = a_1 + a_2 + \ldots a_k.$$

For example, we have that

$$\sum_{k=1}^{5} 2 = 2+2+2+2+2 = 10,$$

$$\sum_{k=1}^{n} m = nm,$$

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

What we showed above is also written as

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

PROPOSITION 1.3.3. *Let* $f : \mathbb{N}^+ \to \mathbb{N}^+$ *with* $a = f(1)$*, and for every* $n, m \in \mathbb{N}^+$

$$f(n+m) = f(n)f(m).$$

Then

$$\forall_{n \in \mathbb{N}^+}\big(f(n) = a^n\big).$$

PROOF. We use the induction principle IND^+ on $\mathbb{N}^+$ for the formula

$$\theta(n) :\Leftrightarrow f(n) = a^n.$$

Clearly, $\theta(1) :\Leftrightarrow f(1) = a^1 = a$, which holds by the definition of a. If $n \in \mathbb{N}^+$, we show the implication

$$\theta(n) \Rightarrow \theta(n+1)$$

i.e., the implication

$$\big[f(n) = a^n\big] \Rightarrow \big[f(n+1) = a^{n+1}\big].$$

By the hypothesis on f we get

$$f(n+1) = f(n)f(1) = a^n f(1) = a^n a = a^{n+1}.$$

□

Induction principle $\mathrm{IND}^<$ **on** $\mathbb{N}$: Let $\phi(n)$ be a formula on $\mathbb{N}$ such that the following conditions are satisfied:

(*i*) $\phi(0)$ holds.

(*ii*) For every $n \in \mathbb{N}^+$, if $\phi(0)$ and $\phi(1)$ and ... and $\phi(n-1)$ hold, then $\phi(n)$ holds:

$$\forall_{n \in \mathbb{N}^+}\Big(\big[\phi(0) \ \& \ \phi(1) \ \& \ \ldots \ \& \ \phi(n-1)\big] \Rightarrow \phi(n)\Big).$$

Then we can infer that $\phi(n)$ holds, for every $n \in \mathbb{N}$ i.e.,

$$\forall_{n \in \mathbb{N}}\big(\phi(n)\big).$$

Proposition 1.3.4. *The induction principle* IND *on* $\mathbb{N}$ *implies the induction principle* $\text{IND}^{<}$ *on* $\mathbb{N}$.

Proof. Let $\phi(n)$ be a formula on $\mathbb{N}$ such that the hypotheses (i) and (ii) of $\text{IND}^{<}$ are satisfied. We show that the hypotheses (i) and (ii) of IND are satisfied, hence the conclusion of IND, which is also the required conclusion of $\text{IND}^{<}$, follows. The hypothesis (i) of IND is the hypothesis (i) of $\text{IND}^{<}$. For the proof of the hypothesis (ii) of IND we suppose $n \in \mathbb{N}$ such that $\phi(n)$, and we show $\phi(n+1)$. Suppose that $\neg(\phi(n+1)$. By the hypothesis (ii) of $\text{IND}^{<}$ there is some $m_1 < n+1$ such that $\neg(\phi(m_1)$ (if for all $m < n+1$ we had that $\phi(m)$ holds, then by the hypothesis (ii) of $\text{IND}^{<}$ we would get $\phi(n+1)$ too). By a similar argument there is some $m_2 < m_1$ such that $\neg(\phi(m_2)$. By repeating this step k number of times, where $k \leq (n+1)$,we get $m_k = 0$, and $\neg(\phi(m_k)$ i.e., $\neg(\phi(0)$. Since we supposed that $\phi(0)$ holds, we reached a contradiction. Hence, our initial hypothesis $\neg(\phi(n+1)$ is false, therefore $\phi(n+1)$ holds. □

Proposition 1.3.5. *The induction principle* $\text{IND}^{<}$ *on* $\mathbb{N}$ *implies the induction principle* IND *on* $\mathbb{N}$.

Proof. Exercise. □

1.4. The algebraic and the ordering axioms for the set of real numbers

We denote by $\mathbb{R}$ the set of real numbers [reele Zahlen] that satisfies the following lists of axioms:

(I) Axioms for addition.

(II) Axioms for multiplication.

(III) Distributivity axiom of multiplication over addition.

(IV) Axioms for order and the Archimedean axiom.

(V) The completeness axiom.

(I) Axioms for addition: There is a function $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$,

$$(x, y) \mapsto x + y$$

such that the following axioms are satisfied:

(A_1) $x + (y + z) = (x + y) + z$, for every $x, y, z \in \mathbb{R}$.

(A_2) There is an element 0 of $\mathbb{R}$ such that $0 + x = x$, for every $x \in \mathbb{R}$.

(A_3) For every $x \in \mathbb{R}$ there is some $y \in \mathbb{R}$ such that $x + y = 0$.

(A_4) $x + y = y + x$, for every $x, y \in \mathbb{R}$.

Notice that the number 0 in (A_2) is uniquely determined. Let $0' \in \mathbb{R}$ such that $0' + x = x$, for every $x \in \mathbb{R}$. If we take $x = 0$, then by (A_2) and (A_4) we get

$$0 = 0' + 0 = 0 + 0' = 0'.$$

The number y in (A_3) is uniquely determined. Let $y' \in \mathbb{R}$ such that $x + y' = 0$. We have that

$$y = 0 + y = (x + y') + y = (y' + x) + y = y' + (x + y) = y' + 0 = y'.$$

We denote this unique element y by $-x$, and we define

$$z - x = z + (-x).$$

We also use the notation

$$\sum_{i=1}^{n} x_i = x_1 + \ldots + x_n.$$

(II) Axioms for multiplication: There is a function $\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$,

$$(x, y) \mapsto x \cdot y$$

such that the following axioms are satisfied:

(M_1) $x \cdot (y \cdot z) = (x \cdot y) \cdot z$, for every $x, y, z \in \mathbb{R}$.

(M_2) There is an element $1 \neq 0$ of $\mathbb{R}$ such that $1 \cdot x = x$, for every $x \in \mathbb{R}$.

(M_3) For every $x \in \mathbb{R}$, such that $x \neq 0$, there is some $y \in \mathbb{R}$ such that $x \cdot y = 1$.

(M_4) $x \cdot y = y \cdot x$, for every $x, y \in \mathbb{R}$.

Notice that the number 1 in (M_2), and the number y in (M_3) are uniquely determined. We denote this unique element y by $x^{-1} = \frac{1}{x}$, and we define

$$\frac{z}{x} = z \cdot \frac{1}{x}.$$

We also use the notation

$$\prod_{i=1}^{n} x_i = x_1 \cdot \ldots \cdot x_n.$$

For simplicity we often write xy instead of $x \cdot y$. If $a \in \mathbb{R}$ and $n \in \mathbb{N}^+$, we define

$$a^n := \begin{cases} a & , n = 1 \\ a^{n-1}a & , n > 1. \end{cases}$$

Hence,

$$a^n = \prod_{i=1}^{n} a.$$

If $a \neq 0$, we define

$$a^0 = 1.$$

It is easy to show by IND that for all $m, n \in \mathbb{N}$ we have that

$$a^{m+n} = a^m a^n.$$

If $n \in \mathbb{N}$, we define

$$a^{-n} = (a^{-1})^n = \left(\frac{1}{a}\right)^n.$$

One can show that for all $m, n \in \mathbb{Z}$ we have that $a^{m+n} = a^m a^n$.

(III) Distributivity axiom of multiplication over addition:

(D) $x \cdot (y + z) = x \cdot y + x \cdot z$, for every $x, y, z \in \mathbb{R}$.

COROLLARY 1.4.1. *If $x, y, z, w \in \mathbb{R}$, the following hold.*

(i) $0 \cdot x = 0$.

(ii) $(-x)y = -(xy)$.

(iii) $(-x)(-y) = xy$.

(iv) $-(x + y) = -x - y$.

(v) If $x, y \neq 0$, then $xy \neq 0$, and $(xy)^{-1} = x^{-1}y^{-1}$.

(vi) If $z, w \neq 0$, then

$$\frac{x}{z}\frac{y}{w} = \frac{xy}{zw} \quad \& \quad \frac{x}{z} + \frac{y}{w} = \frac{xw + yz}{zw}.$$

(vii) If $x \neq 0$ and $xy = xz$, then $y = z$.

PROOF. Exercise. □

By (D) and using the induction principle IND^+ we can show the distributivity of multiplication over an arbitrary sum i.e.,

$$\forall_{n \in \mathbb{N}^+}\left(x \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} x y_i\right).$$

Similarly we can show that

$$\sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j = \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{j=1}^{m} y_j\right) = \left(\sum_{j=1}^{m} y_j\right)\left(\sum_{i=1}^{n} x_i\right) = \sum_{j=1}^{m}\sum_{i=1}^{n} y_j x_i.$$

(IV) Axioms for order: There is a subset P of $\mathbb{R}$, which is called the set of *positive* reals such that the following axioms are satisfied:

(O_1) For every $x \in \mathbb{R}$ we have that

$$x \in P \quad \vee \quad x = 0 \quad \vee \quad -x \in P,$$

and these cases are mutually exclusive i.e., if $x \in P$, then $x \neq 0$ and $-x \notin P$, and if $x = 0$, then $x \notin P$ and $-x \notin P$, while if $-x \in P$, then $x \notin P$ and $x \neq 0$.

(O_2) If $x, y \in P$, then $x + y \in P$ and $x \cdot y \in P$.

DEFINITION 1.4.2. Let $x, y \in \mathbb{R}$. We say that x is *negative* if $-x \in P$. Let

$$x > 0 \Leftrightarrow x \in P,$$

$$x > y \Leftrightarrow x - y > 0,$$

$$y < x \Leftrightarrow x > y,$$

$$x < 0 \Leftrightarrow (-x) > 0,$$

$$x \leq y \Leftrightarrow x < y \ \vee \ x = y.$$

COROLLARY 1.4.3. *If $x, y, z \in \mathbb{R}$, the following hold.*
(i) $1 \in P$.
(ii) For every $n \in \mathbb{N}^+$ we have that $n \cdot 1 \in P$.
(iii) If x, y are negative, then $xy \in P$.
(iv) If $x > 0$ and $y < 0$, then $xy < 0$.
(v) If $x \neq 0$, then $x^2 > 0$.
(vi) If $x > 0$, then $\frac{1}{x} > 0$.
(vii) If $x < y$ and $y < z$, then $x < z$.
(viii) If $x < y$ and $z \in \mathbb{R}$, then $x + z < y + z$.
(ix) If $x < y$ and $z > 0$, then $xz < yz$.
(x) If $x < y$ and $x, y > 0$, then $\frac{1}{y} < \frac{1}{x}$.
(xi) If $xy = 0$, then $x = 0$ or $y = 0$.

PROOF. (i) By (O_1) we have that $1 \in P$, or $1 = 0$, or $-1 \in P$. Since by (M_2) $1 \neq 0$, we have that $1 \in P$ or $-1 \in P$. Suppose that $-1 \in P$. Then by (O_2) we get $(-1)(-1) = 1^2 = 1 \in P$. Since the cases $1 \in P$ and $-1 \in P$ cannot hold together, we get a contradiction. Hence $1 \in P$ is the only true case.
(ii) It follows with the use of IND^+. The case $n = 1$ is just (i).
(iii) By definition $-x, -y \in P$, and by (O_2) we get $(-x)(-y) = xy \in P$.
(iv) By definition $-y \in P$, hence by (iii) we get $x(-y) = -xy \in P$, hence $xy < 0$.
(v) If $x \neq 0$, then by (O_1) we have that $x \in P$ or $(-x) \in P$. In the first case, by (O_2) we get $xx = x^2 \in P$, and in the second, again by (O_2), we get $xx = x^2 = (-x)(-x) \in P$.
(vi) - (xi) Exercise. □

Let $a \in \mathbb{R}$ such that $a \neq 0$. If there is some $x \in \mathbb{R}$ such that $x^2 = a$, then $a > 0$. If there are y$x, y \in \mathbb{R}$ such that $x^2 = y^2 = a$, then by the Corollary 1.4.3(x) we have that

$$x^2 - y^2 = 0 \Leftrightarrow (x - y)(x + y) = 0 \Leftrightarrow x = y \ \vee \ x = -y.$$

Hence, if there is x such that $x^2 = a$, then the equation $x^2 = a$ has exactly two solutions x and $-x$. In this case we call the unique positive solution to the equation

$x^2 = a$ the square root $\sqrt{a}$ of a. Notice that we cannot prove yet that every positive real number has a square root. We also define

$$\sqrt{0} = 0.$$

If $x \in \mathbb{R}$, then $\sqrt{x^2}$ always exists, and it is either x, if $x \geq 0$, or $-x$, if $x < 0$. Let the function

$$|.| : \mathbb{R} \to \mathbb{R}$$

$$x \mapsto |x| = \sqrt{x^2},$$

where $|x|$ is called the *absolute value* of x.

PROPOSITION 1.4.4. *If $x, y \in \mathbb{R}$, the following hold.*
(i) If $x, y \geq 0$ and $\sqrt{x}, \sqrt{y}$ exist, then $\sqrt{xy}$ exists and

$$\sqrt{xy} = \sqrt{x}\sqrt{y}.$$

Moreover, we have that

$$x \leq y \Rightarrow \sqrt{x} \leq \sqrt{y}.$$

(ii) $|x| \geq 0$.
(iii) $x \leq |x|$.
(iv) $|x| = |-x|$.
(v) $|x| = 0 \Leftrightarrow x = 0$.
(vi) $|xy| = |x||y|$.
(vii) $|x|^2 = x^2$.
(viii) [*Triangle inequality*] $|x + y| \leq |x| + |y|$.

PROOF. We show only (vi) and the rest is an exercise. We have that

$$\begin{aligned} |x + y|^2 &= \big(\sqrt{(x + y)^2}\big)^2 \\ &= (x + y)^2 \\ &= x^2 + 2xy + y^2 \\ &\overset{(iii)}{\leq} x^2 + 2|xy| + y^2 \\ &\overset{(vi)}{=} x^2 + 2|x||y| + y^2 \\ &= \big(|x| + |y|\big)^2, \end{aligned}$$

hence by the second implication of the case (i), and by taking the square roots, we get the required inequality. $\square$

The Archimedean Axiom (Arch): The order relation $<$ of reals satisfies the following axiom

$$\text{(Arch)} \qquad \forall_{x,y\in\mathbb{R}}\Big([x>0 \ \&\ y>0] \Rightarrow \exists_{n\in\mathbb{N}}\big(nx>y\big)\Big).$$

Corollary 1.4.5. $\forall_{x\in\mathbb{R}}\exists_{n\in\mathbb{N}}\exists_{m\in\mathbb{N}}\big(x<n \ \&\ -m<x\big)$.

Proof. If $x=0$, we can take $n=m=1$. If $x>0$, by (Arch) on x and 1, there is $n\in\mathbb{N}$ such that $n>x$. Consequently, if $m=n$, we get $-m=-n<0<x$. If $x<0$, then by the previous case there are $n,m\in\mathbb{N}$ such that $-x<n \ \&\ -m<(-x)$, hence $x<m \ \&\ -n<x$. $\square$

We also write the formula of the previous corollary as follows

$$\forall_{x\in\mathbb{R}}\exists_{n,m\in\mathbb{N}}\big(x<n \ \&\ -m<x\big).$$

The formula

$$\exists_{!x\in X}\big(\phi(x)\big)$$

expresses that there exists a unique $x\in X$ such that $\phi(x)$. I.e.,

$$\exists_{!x\in X}\big(\phi(x)\big) :\Leftrightarrow \exists_{x\in X}\Big(\phi(x) \ \&\ \forall_{y\in X}\big(\phi(y)\Rightarrow y=x\big)\Big).$$

If $x,y,z\in\mathbb{R}$, we use abbreviations like the following:

$$x\leq y<z :\Leftrightarrow x\leq y \ \&\ y<z.$$

Corollary 1.4.6. $\forall_{x\in\mathbb{R}}\exists_{!k\in\mathbb{Z}}\big(k\leq x<k+1\big)$.

Proof. If $x=0$, we take $k=0$. If $x>0$, by the previous corollary there is some $n>x$. Let n_0 be the smallest element of $\mathbb{N}$ such that $x<n$ (we can find n_0 by checking for the predecessors m of n if $m>x$). Since $n_0>x>0$, we have that $n_0\geq 1$, and since by its definition n_0 is the smallest natural number $>x$, we get $n_0-1\leq x$. If $x<0$, then by the previous case there is $k\in\mathbb{Z}$ such that $k\leq(-x)<k+1$. If $k=-x$, then $-k=x<-k+1$. If $k<-x$, then $-(k+1)<x<-k$ i.e., $-k-1\leq x<-k=(-k-1)+1$.

To show the uniqueness of k we work as follows. Let $l\in\mathbb{Z}$ such that $l\leq x<l+1$. Suppose that $l<k$. Then $l+1\leq k$, and

$$l\leq x<l+1\leq k\leq x$$

i.e., we reached the contradiction $x<x$. Hence $l\geq k$. If we suppose $k<l$, we get similarly a contradiction, hence $k\geq l$. By the inequalities $l\geq k$ and $k\geq l$ we conclude that $k=l$. $\square$

We use the symbol $\lfloor x\rfloor$ for this unique $k\in\mathbb{Z}$, and we call $\lfloor x\rfloor$ the *floor* of x.

Corollary 1.4.7. $\forall_{x\in\mathbb{R}}\exists_{!m\in\mathbb{Z}}\big(m-1<x\leq m\big)$.

Proof. We use the Corollary 1.4.6. $\square$

We use the symbol $\lceil x \rceil$ for this unique $m \in \mathbb{Z}$, and we call $\lceil x \rceil$ the *ceiling* of x. We also use an abbreviation of the following form

$$\forall_{\varepsilon>0}\big(\phi(\varepsilon)\big) :\Leftrightarrow \forall_{\varepsilon\in\mathbb{R}}\big(\varepsilon > 0 \Rightarrow \phi(\varepsilon)\big).$$

COROLLARY 1.4.8. $\forall_{\varepsilon>0}\exists_{n\in\mathbb{N}^+}\big(\frac{1}{n} < \varepsilon\big)$.

PROOF. Exercise. □

We can show inductively the *Bernoulli inequality*: if $a > -1$, then

$$\forall_{n\in\mathbb{N}}\big((1+a)^n \geq 1 + na\big).$$

COROLLARY 1.4.9. *Let $a \in \mathbb{R}$. The following hold.*

(i) If $a > 1$, then $\forall_{x\in\mathbb{R}}\exists_{n\in\mathbb{N}}\big(a^n > x\big)$.

(ii) If $0 < a < 1$, then $\forall_{\varepsilon>0}\exists_{n\in\mathbb{N}}\big(a^n < \varepsilon\big)$.

PROOF. Exercise (use the Bernoulli inequality). □

1.5. Sequences of real numbers

DEFINITION 1.5.1. Let X be a set. A *sequence* of elements of X is a function $\alpha : \mathbb{N} \to X$. We also use the notations

$$(\alpha_n)_{n\in\mathbb{N}}, \quad \text{or} \quad (\alpha_n)_{n=0}^{\infty}$$

for α, where

$$\alpha_n = \alpha(n).$$

Sometimes we may also use the notation

$$\big(\alpha_0, \alpha_1, \alpha_2, \alpha_3, \dots\big).$$

An element α_n of a sequence α is called the n-th *term* of α. A sequence of reals is a function $\alpha : \mathbb{N} \to \mathbb{R}$.

(i) If $x \in \mathbb{R}$, the constant sequence with value x is the function $\alpha : \mathbb{N} \to \mathbb{R}$ with

$$\alpha_n = x, \quad \text{for every } n \in \mathbb{N}.$$

This sequence looks as follows:

$$\big(x, x, x, \dots\big).$$

(ii) The sequence $\beta : \mathbb{N} \to \mathbb{R}$, defined by

$$\beta_n = \frac{1}{n+1}, \quad \text{for every } n \in \mathbb{N},$$

looks as follows:

$$\left(1, \frac{1}{2}, \frac{1}{3}, \dots\right).$$

(iii) The sequence $\gamma : \mathbb{N} \to \mathbb{R}$, defined by

$$\gamma_n = (-1)^n, \quad \text{for every } n \in \mathbb{N},$$

looks as follows:

$$\big(1, -1, 1, -1, 1, -1, 1, \dots\big).$$

(iv) The sequence $\delta : \mathbb{N} \to \mathbb{R}$, defined by

$$\delta_n = \frac{n}{n+1}, \quad \text{for every } n \in \mathbb{N},$$

looks as follows:

$$\left(0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots\right).$$

(v) The sequence $\zeta : \mathbb{N} \to \mathbb{R}$, defined by

$$\zeta_n = \frac{n}{2^n}, \quad \text{for every } n \in \mathbb{N},$$

looks as follows:

$$\left(0, \frac{1}{2}, \frac{2}{4} = \frac{1}{2}, \frac{3}{2^3} = \frac{3}{8}, \frac{4}{2^4} = \frac{1}{4}, \dots\right).$$

(vi) The sequence of the *Fibonacci numbers* $\mathtt{Fib} : \mathbb{N} \to \mathbb{R}$ is defined recursively as follows

$$\mathtt{Fib}_n := \begin{cases} 0 & , n = 0 \\ 1 & , n = 1 \\ \mathtt{Fib}_{n-1} + \mathtt{Fib}_{n-2} & , n \geq 2, \end{cases}$$

and it looks as follows:

$$\big(0, 1, 1, 2, 3, 5, 8, 13, 21, \dots\big).$$

Definition 1.5.2. Let $\alpha : \mathbb{N} \to \mathbb{R}$ be a sequence of real numbers, and let $x \in \mathbb{R}$. We say that α *converges* to x, or x is the *limit* of α, if

$$\forall_{\varepsilon > 0} \exists_{N_\varepsilon \in \mathbb{N}} \forall_{n \geq N_\varepsilon} \big(|\alpha_n - x| < \varepsilon\big).$$

In this case we use the notations

$$\alpha_n \xrightarrow{n} x, \quad \text{or} \quad \lim_{n \to \infty} \alpha_n = x, \quad \text{or} \quad \lim \alpha_n = x.$$

A sequence of reals α is called *convergent* if there is some $x \in \mathbb{R}$, such that α converges to x. We say that α is a *divergent* sequence, if there is no $x \in \mathbb{R}$ such that α converges to x. A sequence of reals α is called *bounded*, if

$$\exists_{M > 0} \forall_{n \in \mathbb{N}} \big(|\alpha_n| \leq M\big).$$

In this case we say that M is a *bound* of α, or α is *bounded* by M.

Since

$$|\alpha_n - x| < \varepsilon \Leftrightarrow -\varepsilon < \alpha_n - x < \varepsilon \Leftrightarrow x - \varepsilon < \alpha_n < x + \varepsilon,$$

a sequence α converges to $x \in \mathbb{R}$, if for every ε-interval around x, eventually (i.e., after some index N_ε) all terms α_n of α lie there.

Proposition 1.5.3 (Uniqueness of limit). *If $\alpha : \mathbb{N} \to \mathbb{R}$ is a sequence of real numbers, and $x, y \in \mathbb{R}$, then*

$$\left[\alpha_n \xrightarrow{n} x \;\;\&\;\; \alpha_n \xrightarrow{n} y\right] \Rightarrow x = y.$$

Proof. Let $\varepsilon > 0$. Since $\alpha_n \xrightarrow{n} x$ and $\alpha_n \xrightarrow{n} y$, there are $N_{\frac{\varepsilon}{2}} \in \mathbb{N}$ and $M_{\frac{\varepsilon}{2}} \in \mathbb{N}$, such that

$$\forall_{n \geq N_\varepsilon}\left(|\alpha_n - x| < \frac{\varepsilon}{2}\right) \quad \& \quad \forall_{n \geq M_\varepsilon}\left(|\alpha_n - y| < \frac{\varepsilon}{2}\right).$$

Hence for all $n \geq \max\{N_{\frac{\varepsilon}{2}}, M_{\frac{\varepsilon}{2}}\}$ we have that

$$|x - y| = |x - \alpha_n + \alpha_n - y| \leq |x - \alpha_n| + |\alpha_n - y| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

If $x \neq y$, then if we take $\varepsilon = \frac{|x-y|}{2}$, and since the above holds for every $\varepsilon > 0$, we get the contradiction

$$|x - y| < \frac{|x - y|}{2},$$

hence $x = y$ is the case. □

Proposition 1.5.4. *If $\alpha : \mathbb{N} \to \mathbb{R}$ is a sequence of real numbers that converges to some $x \in \mathbb{R}$, then α is bounded.*

Proof. Since $\alpha_n \xrightarrow{n} x$, there is some $N_1 \in \mathbb{N}$ such that

$$\forall_{n \geq N_1}\big(|\alpha_n - x| < 1\big).$$

Since

$$|\alpha_n| = |\alpha_n - x + x| \leq |\alpha_n - x| + |x| < 1 + |x|,$$

we get

$$\forall_{n \geq N_1}\big(|\alpha_n| < 1 + |x|\big).$$

hence the following real number

$$M = \max\big\{|\alpha_1|, \ldots, |\alpha_{N_1 - 1}|, 1 + |x|\big\}$$

is a bound of the sequence α. □

If $\alpha, \beta, \gamma, \delta$, and ζ are the sequences defined above, the following hold.

(i) $\alpha_n \xrightarrow{n} x$: If $\varepsilon > 0$, let $N_\varepsilon = 0$. Then

$$\forall_{n \geq 0}\big(|\alpha_n - x| = |x - x| = 0 < \varepsilon\big).$$

(ii) $\beta_n \xrightarrow{n} 0$: If $\varepsilon > 0$, then by the Corollary 1.4.8 there exists $N_\varepsilon \in \mathbb{N}^+$, such that $\frac{1}{N_\varepsilon} < \varepsilon$. Then

$$\forall_{n \geq N_\varepsilon - 1}\left(|\beta_n - 0| = |\beta_n| = \frac{1}{n+1} \leq \frac{1}{N_\varepsilon} < \varepsilon\right).$$

(iii) The sequence γ is divergent (although it is bounded by 1). Suppose that there is $x \in \mathbb{R}$, such that $\gamma_n \xrightarrow{n} x$. Hence there is some $N_1 \in \mathbb{N}$ such that

$$\forall_{n \geq N_1} \big(|\gamma_n - x| = |(-1)^n - x| < 1 \big).$$

Hence, for every $n \geq N_1$ we get

$$\begin{aligned} 2 &= |(-1)^{n+1} - (-1)^n| \\ &= |\gamma_{n+1} - \gamma_n| \\ &= |\gamma_{n+1} - x + x - \gamma_n| \\ &\leq |\gamma_{n+1} - x| + |x - \gamma_n| \\ &< 1 + 1 \\ &= 2, \end{aligned}$$

which is a contradiction.

(iv) $\delta_n \xrightarrow{n} 1$: Exercise.

(v) $\zeta_n \xrightarrow{n} 0$: Exercise.

PROPOSITION 1.5.5. *Let $(\alpha_n)_{n\in\mathbb{N}}, (\beta_n)_{n\in\mathbb{N}}$ be sequences of reals, and $\lambda, x, y \in \mathbb{R}$. We define the sequences $(\alpha+\beta)_{n\in\mathbb{N}}, (\alpha\cdot\beta)_{n\in\mathbb{N}}, (\lambda\alpha)_{n\in\mathbb{N}}$ and $\big(\frac{1}{\beta}\big)_{n\in\mathbb{N}}$, if $\beta_n \neq 0$, for every $n \in \mathbb{N}$, as follows:*

$$(\alpha+\beta)_n = \alpha_n + \beta_n,$$

$$(\alpha\cdot\beta)_n = \alpha_n \cdot \beta_n,$$

$$(\lambda\alpha)_n = \lambda\alpha_n,$$

$$\left(\frac{1}{\beta}\right)_n = \frac{1}{\beta_n},$$

for every $n \in \mathbb{N}$. If $\alpha_n \xrightarrow{n} x$ and $\beta_n \xrightarrow{n} y$, the following hold:

(i) $(\alpha+\beta)_n \xrightarrow{n} x + y$.

(ii) $(\alpha\cdot\beta)_n \xrightarrow{n} x \cdot y$.

(iii) $(\lambda\alpha)_n \xrightarrow{n} \lambda x$.

(iv) If $y \neq 0$, then there is $n_0 \in \mathbb{N}$ such that $\beta_n \neq 0$, for all $n \geq n_0$, and

$$\left(\frac{1}{\beta}\right)_{n+n_0} \xrightarrow{n} \frac{1}{y},$$

and

$$\left(\frac{\alpha}{\beta}\right)_{n+n_0} \xrightarrow{n} \frac{x}{y}.$$

PROOF. By definition of convergence of a sequence we have that

$$\forall_{\varepsilon>0}\exists_{N_\varepsilon^\alpha\in\mathbb{N}}\forall_{n\geq N_\varepsilon^\alpha}\big(|\alpha_n - x| < \varepsilon\big).$$

$$\forall_{\varepsilon>0}\exists_{N_\varepsilon^\beta\in\mathbb{N}}\forall_{n\geq N_\varepsilon^\beta}\big(|\beta_n - x| < \varepsilon\big).$$

(i) By the triangle inequality we have that

$$\begin{aligned}|(\alpha+\beta)_n - (x+y)| &= |\alpha_n + \beta_n - x - y|\\ &= |(\alpha_n - x) + (\beta_n - y)|\\ &\leq |\alpha_n - x)| + |\beta_n - y|\\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2}\\ &= \varepsilon,\end{aligned}$$

for all $n \geq N_\varepsilon^{\alpha+\beta} = \max\{N_{\frac{\varepsilon}{2}}^\alpha, N_{\frac{\varepsilon}{2}}^\beta\}$.

(ii) If $M > 0$ is a bound of the convergent sequence α, then by the triangle inequality we have that

$$\begin{aligned}|(\alpha\cdot\beta)_n - xy| &= |\alpha_n\beta_n - xy|\\ &= |\alpha_n\beta_n - \alpha_n y + \alpha_n y - xy|\\ &= |(\alpha_n\beta_n - \alpha_n y) + (\alpha_n y - xy)|\\ &\leq |\alpha_n(\beta_n - y)| + |(\alpha_n - x)y|\\ &= |\alpha_n||\beta_n - y| + |\alpha_n - x||y|\\ &\leq M|\beta_n - y| + |\alpha_n - x||y|\\ &< M\frac{\varepsilon}{2M} + |y|\frac{\varepsilon}{2(|y|+1)}\\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2}\\ &= \varepsilon,\end{aligned}$$

for all $n \geq N_\varepsilon^{\alpha\cdot\beta} = \max\{N_{\frac{\varepsilon}{2M}}^\beta, N_{\frac{\varepsilon}{2(|y|+1)}}^\alpha\}$.

(iii) Exercise.

(iv) By the convergence $\beta_n \xrightarrow{n} y$ we have that

$$|\beta_n - y| < \frac{|y|}{2},$$

for every $n \geq n_0 = N_{\frac{|y|}{2}}^\beta$. Hence, for every $n \geq n_0$ we get

$$-|\beta_n - y| > -\frac{|y|}{2}.$$

Since for every $x, y \in \mathbb{R}$ we have shown (Blatt 3, Exercise 4(ii)) that

$$|x - y| \geq \big||x| - |y|\big| \geq |x| - |y|,$$

we get for every $n \geq n_0$

$$\begin{aligned}|\beta_n| &= |y - (y - \beta_n)| \\ &\geq \big||y| - |\beta_n - y|\big| \\ &\geq |y| - |\beta_n - y| \\ &\geq |y| - \frac{|y|}{2} \\ &= \frac{|y|}{2} \\ &> 0.\end{aligned}$$

Moreover, we have that

$$\begin{aligned}\left|\left(\frac{1}{\beta}\right)_n - \frac{1}{y}\right| &= \left|\frac{1}{\beta_n} - \frac{1}{y}\right| \\ &= \left|\frac{y - \beta_n}{\beta_n y}\right| \\ &= \frac{1}{|\beta_n||y|}|\beta_n - y| \\ &\leq \frac{2}{|y|}\frac{1}{|y|}\left(\frac{\varepsilon|y|^2}{2}\right) \\ &= \varepsilon,\end{aligned}$$

for all $n \geq \max\{n_0, N^{\beta}_{\frac{\varepsilon|y|^2}{2}}\}$.
For the convergence

$$\left(\frac{\alpha}{\beta}\right)_{n+n_0} \xrightarrow{n} \frac{x}{y}$$

we use the previous convergence and the case (ii). □

Let the sequence α defined by

$$\alpha_n = \frac{4n^2 + 14n}{n^2 - 2},$$

for every $n \in \mathbb{N}$. Since for $n > 0$ we have that

$$\alpha_n = \frac{n^2(4 + 14\frac{1}{n})}{n^2(1 - 2\frac{1}{n^2})} = \frac{4 + 14\frac{1}{n}}{1 - 2\frac{1}{n^2}},$$

and $\frac{1}{n} \xrightarrow{n} 0$ and hence $\frac{1}{n^2} = \frac{1}{n}\frac{1}{n} \xrightarrow{n} 0$, we get $14\frac{1}{n} \xrightarrow{n} 0$, $-2\frac{1}{n^2} \xrightarrow{n} 0$, and hence

$$\alpha_n \xrightarrow{n} 4.$$

Proposition 1.5.6. *Let* $(\alpha_n)_{n\in\mathbb{N}}, (\beta_n)_{n\in\mathbb{N}}$ *be sequences of reals, and* $x, y \in \mathbb{R}$. *If* $\alpha_n \xrightarrow{n} x$ *and* $\beta_n \xrightarrow{n} y$*, and if*

$$\alpha_n \leq \beta_n,$$

for every $n \in \mathbb{N}$*, then* $x \leq y$.

Proof. Suppose that $\varepsilon = x - y > 0$. For every $n \geq N^{\alpha}_{\frac{\varepsilon}{2}}$ we have that

$$|\alpha_n - x| < \frac{\varepsilon}{2} \Leftrightarrow -\frac{\varepsilon}{2} < \alpha_n - x < \frac{\varepsilon}{2} \Leftrightarrow x - \frac{\varepsilon}{2} < \alpha_n < x + \frac{\varepsilon}{2},$$

hence

$$\alpha_n > x - \frac{x-y}{2} = \frac{x+y}{2}.$$

For every $n \geq N^{\beta}_{\frac{\varepsilon}{2}}$ we have that

$$|\beta_n - y| < \frac{\varepsilon}{2} \Leftrightarrow -\frac{\varepsilon}{2} < \beta_n - y < \frac{\varepsilon}{2} \Leftrightarrow y - \frac{\varepsilon}{2} < \beta_n < y + \frac{\varepsilon}{2},$$

hence

$$\beta_n < y + \frac{x-y}{2} = \frac{x+y}{2}.$$

Hence for every $n \geq \max\{N^{\beta}_{\frac{\varepsilon}{2}}, N^{\beta}_{\frac{\varepsilon}{2}}\}$ we get

$$\beta_n < \frac{x+y}{2} < \alpha_n,$$

which is a contradiction. Hence $x \leq y$ is the case. □

1.6. The completeness axiom

All axioms (I), (II), (III) and (IV) are satisfied also by the set of rational numbers $\mathbb{Q}$. The axiom discussed in this section is the most important axiom for the set of the real numbers $\mathbb{R}$, and, as expected, it is not satisfied by $\mathbb{Q}$.

Lemma 1.6.1. *If* $k, l \in \mathbb{N}$*, the following hold:*

(i) $k \in \texttt{Even} \Rightarrow k^2 \in \texttt{Even}$.

(ii) $k \in \texttt{Odd} \Rightarrow k^2 \in \texttt{Odd}$.

(iii) $k^2 \in \texttt{Even} \Rightarrow k \in \texttt{Even}$.

(iv) $k^2 \in \texttt{Odd} \Rightarrow k \in \texttt{Odd}$.

(v) $k \in \texttt{Even} \Rightarrow kl \in \texttt{Even}$.

(vi) $k, l \in \texttt{Odd} \Rightarrow kl \in \texttt{Odd}$.

PROOF. (i) If $k = 2n$, for some $n \in \mathbb{N}$, then $k^2 = (2n)^2 = 4n^2 = 2(2n^2) \in \mathtt{Even}$.
(ii) If $k = 2n + 1$, for some $n \in \mathbb{N}$, then $k^2 = (2n + 1)^2 = 4n^2 + 4n + 1 = 2[2n^2 + 2n] + 1 \in \mathtt{Odd}$.
(iii) If $k^2 \in \mathtt{Even}$ and $k \in \mathtt{Odd}$, then by (ii) $k^2 \in \mathtt{Odd}$ too, which is a contradiction.
(iv) If $k^2 \in \mathtt{Odd}$ and $k \in \mathtt{Even}$, then by (i) $k^2 \in \mathtt{Even}$ too, which is a contradiction.
(v)-(vi) are left to the reader as a simple exercise. □

LEMMA 1.6.2. *There is no rational number q such that $q^2 = 2$.*

PROOF. Let $p \in \mathbb{Q}$ such that $p^2 = 2$. Moreover, let

$$p = \frac{k}{l},$$

where without loss of generality $p > 0$ and k, l are natural numbers, which are not both of them even (why?). If $k^2 = 2l^2$, then $k^2 \in \mathtt{Even}$, hence $k \in \mathtt{Even}$. Let $k = 2m$, for some $m \in \mathbb{N}^+$. Since $k^2 = 4m^2 = 2l^2$, we get $l^2 = 2m^2$, hence $l^2 \in \mathtt{Even}$, therefore $l \in \mathtt{Even}$, a fact which contradicts our hypothesis on k and l. □

DEFINITION 1.6.3. A sequence $(\alpha_n)_{n\in\mathbb{N}}$ of reals in called a *Cauchy-sequence*, if

$$\forall_{\varepsilon>0}\exists_{C_\varepsilon\in\mathbb{N}^+}\forall_{n,m\geq C_\varepsilon}\big(|\alpha_n - \alpha_m| < \varepsilon\big).$$

PROPOSITION 1.6.4. *If $(\alpha_n)_{n\in\mathbb{N}}$ is a convergent sequence, then $(\alpha_n)_{n\in\mathbb{N}}$ is a Cauchy-sequence.*

PROOF. If $x \in \mathbb{R}$ such that $\alpha_n \xrightarrow{n} x$, we have that

$$|\alpha_n - \alpha_m| = |\alpha_n - x + x - \alpha_m| \leq |\alpha_n - x| + |x - \alpha_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

for all $n, m \geq N_{\frac{\varepsilon}{2}} = C_\varepsilon$. □

Completeness axiom (CA): If $(\alpha_n)_{n\in\mathbb{N}}$ is a Cauchy-sequence of reals, then $(\alpha_n)_{n\in\mathbb{N}}$ is convergent.

Next we use CA to prove the existence of the square root of a positive real number.

THEOREM 1.6.5. *Let $a, b \in \mathbb{R}$ such that $a > 0$ and $b > 0$. Let the sequence $(\alpha_n)_{n\in\mathbb{N}}$ be defined by*

$$\alpha_0 = b,$$
$$\alpha_{n+1} = \frac{1}{2}\left(\alpha_n + \frac{a}{\alpha_n}\right).$$

The following hold:
(i) $\alpha_n > 0$, *for all* $n \in \mathbb{N}$.
(ii) $\alpha_n^2 \geq a$, *for all* $n \geq 1$.
(iii) $\alpha_{n+1} \leq \alpha_n$, *for all* $n \geq 1$.

(iv) If $(\beta_n)_{n\in\mathbb{N}^+}$ is the sequence of reals defined by

$$\beta_n = \frac{a}{\alpha_n}, \quad n \in \mathbb{N}^+,$$

then
(a) $\beta_n^2 \leq a$, for every $n \geq 1$,
(b) $\beta_n \leq \alpha_m$, for every $n, m \geq 1$, and
(c) For every $n \geq 1$ we have that

$$\alpha_n - \beta_n \leq \frac{1}{2^{n-1}}(\alpha_1 - \beta_1).$$

(v) The sequence $(\alpha_n)_{n\in\mathbb{N}}$ is a Cauchy-sequence.
(vi) If $x \in \mathbb{R}$ such that $\alpha_n \xrightarrow{n} x$, then $x \geq 0$ and $x^2 = a$.

Proof. (i) We use the induction principle IND.
(ii) We show that

$$\alpha_n^2 - a = \frac{1}{4}\left(\alpha_{n-1} + \frac{a}{\alpha_{n-1}}\right)^2 - a \geq 0.$$

(iii) Using (i) and (ii) we show that

$$\alpha_n - \alpha_{n+1} \geq 0.$$

(iv)(a) By (ii) we have that

$$\alpha_n^2 \geq a \Rightarrow \frac{a}{\alpha_n^2} \leq 1,$$

hence

$$\beta_n^2 = \frac{a^2}{\alpha_n^2} = \frac{a}{\alpha_n^2}a \leq 1 \cdot a = a.$$

(iv)(b) By (iii) we have that

$$\alpha_{n+1} \leq \alpha_n \Rightarrow \frac{1}{\alpha_{n+1}} \geq \frac{1}{\alpha_n} \overset{a\geq 0}{\Rightarrow} \beta_{n+1} = \frac{a}{\alpha_{n+1}} \geq \frac{a}{\alpha_n} = \beta_n$$

i.e., for every $n \geq 1$ we have that

$$\beta_{n+1} \geq \beta_n. \tag{1.1}$$

Let $n, m \geq 1$. Suppose first that $n \geq m$. By (ii) for every $n \geq 1$ we have that

$$\beta_n = \frac{a}{\alpha_n} \leq \alpha_n. \tag{1.2}$$

By (iii) we have that

$$\alpha_n \leq \alpha_{n-1} \leq \ldots \leq \alpha_m,$$

hence by the Equation 1.2 we get

$$\beta_n \leq \alpha_n \leq \alpha_{n-1} \leq \ldots \leq \alpha_m.$$

Suppose next that $n \leq m$. By the Equations 1.1 and 1.2 we get

$$\beta_n \leq \beta_{n+1} \leq \ldots \leq \beta_m \leq \alpha_m.$$

(iv)(c) We use the induction principle IND^+. If $n = 1$, then

$$\alpha_1 - \beta_1 = \frac{1}{2^0}\big(\alpha_1 - \beta_1\big).$$

If $n > 1$, then by the Equation 1.1 we have that

$$\beta_{n+1} \geq \beta_n \Rightarrow -\beta_{n+1} \leq -\beta_n,$$

hence

$$\begin{aligned}\alpha_{n+1} - \beta_{n+1} &\leq \alpha_{n+1} - \beta_n \\ &= \frac{1}{2}\big(\alpha_n + \beta_n\big) - \beta_n \\ &= \frac{1}{2}\big(\alpha_n - \beta_n\big) \\ &\leq \frac{1}{2}\left[\frac{1}{2^{n-1}}(\alpha_1 - \beta_1)\right] \\ &= \frac{1}{2^n}(\alpha_1 - \beta_1).\end{aligned}$$

(v) We calculate the absolute value $|\alpha_n - \alpha_m|$. Suppose first that $n \leq m$. By (iii) we get $|\alpha_n - \alpha_m| = \alpha_n - \alpha_m$. By the cases (iv)(b) and (iv)(c) we have that

$$|\alpha_n - \alpha_m| = \alpha_n - \alpha_m \leq \alpha_n - \beta_n \leq \frac{1}{2^{n-1}}(\alpha_1 - \beta_1).$$

Suppose next that $n \geq m$. By (iii) we get $|\alpha_n - \alpha_m| = \alpha_m - \alpha_n$. By the cases (iv)(b) and (iv)(c) we have that

$$|\alpha_n - \alpha_m| = \alpha_m - \alpha_n \leq \alpha_m - \beta_m \leq \frac{1}{2^{m-1}}(\alpha_1 - \beta_1).$$

Suppose that $\alpha_1 \neq \beta_1 \Leftrightarrow \alpha_1 - \beta_1 > 0$, since $\alpha_1 \geq \beta_1$. If $\zeta_n = \frac{1}{2^{n-1}}$, for every $n \geq 1$, then $\zeta_n \xrightarrow{n} 0$, and for every $n, m \geq N^{\zeta}_{\frac{\varepsilon}{\alpha_1 - \beta_1}} = C_\varepsilon$ we have that

$$|\alpha_n - \alpha_m| < \varepsilon$$

i.e., $(\alpha_n)_{n \in \mathbb{N}}$ is a Cauchy-sequence. Notice that if $\alpha_1 - \beta_1 = 0$, then what we want follows trivially. In this case we have that

$$\alpha_1 = \frac{a}{\alpha_1} \Leftrightarrow \alpha_1^2 = a,$$

and by the case (iv)(c) we have that

$$0 \leq |\alpha_n - \beta_n| \leq \frac{1}{2^{n-1}}(\alpha_1 - \beta_1) = 0$$

i.e., $\alpha_n = \beta_n$, for every $n \geq 1$, hence $\alpha_n^2 = a$, for every $n \geq 1$, and by (i) $(\alpha_n)_{n\geq 1}$ is the constant sequence $\sqrt{a}$.
(vi) We show that $\beta_n \xrightarrow{n} x$. Since

$$|\beta_n - x| \leq |\beta_n - \alpha_n| + |\alpha_n - x|,$$

and since by (iv)(c) $|\beta_n - \alpha_n| \xrightarrow{n} 0$, we get $\beta_n \xrightarrow{n} x$. Hence

$$\begin{aligned} x^2 &= \big(\lim_{n\longrightarrow\infty} \beta_n\big) \cdot \big(\lim_{n\longrightarrow\infty} \beta_n\big) \\ &= \lim_{n\longrightarrow\infty} \beta_n^2 \\ &\overset{(iv)(a)}{\leq} a \\ &\overset{(ii)}{\leq} \lim_{n\longrightarrow\infty} \alpha_n^2 \\ &= \big(\lim_{n\longrightarrow\infty} \alpha_n\big) \cdot \big(\lim_{n\longrightarrow\infty} \alpha_n\big) \\ &= x \cdot x \\ &= x^2. \end{aligned}$$

From the inequalities $x^2 \leq a \leq x^2$ we conclude that $x^2 = a$. □

As a consequence of the previous theorem, if we define the sequence

$$\alpha_0 = 1,$$

$$\alpha_{n+1} = \frac{1}{2}\bigg(\alpha_n + \frac{2}{\alpha_n}\bigg),$$

then

$$\alpha_n \xrightarrow{n} \sqrt{2}.$$

Using this sequence we can show that the set $\mathbb{Q}$ of rational numbers does not satisfy CA (Exercise). As a generalization of the previous theorem, CA implies the existence of the k-th root of a positive real, for every $k \geq 2$.

Theorem 1.6.6. *Let $k \in \mathbb{N}$ such that $k \geq 2$, and let $a, b \in \mathbb{R}$ such that $a > 0$ and $b > 0$. Let the sequence $(\alpha_n)_{n\in\mathbb{N}}$ be defined by*

$$\alpha_0 = b,$$

$$\alpha_{n+1} = \frac{1}{k}\bigg((k-1)\alpha_n + \frac{a}{\alpha_n^{k-1}}\bigg).$$

The following hold:
(i) $\alpha_n > 0$, for all $n \in \mathbb{N}$.
(v) The sequence $(\alpha_n)_{n\in\mathbb{N}}$ is a Cauchy-sequence.
(vi) If $x \in \mathbb{R}$ such that $\alpha_n \xrightarrow{n} x$, then $x \geq 0$ and $x^k = a$.

DEFINITION 1.6.7. The set $\mathbb{I}$ of *irrational* real numbers is defined by

$$\mathbb{I} = \{x \in \mathbb{R} \mid x \notin \mathbb{Q}\}$$

i.e., $\mathbb{I}$ is the complement of $\mathbb{Q}$ in $\mathbb{R}$.

Clearly, $\sqrt{2}, \sqrt{3} \in \mathbb{I}$.

1.7. Infinite series of real numbers

DEFINITION 1.7.1. Let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence of real numbers. The sequence $(\sigma_n)_{n\in\mathbb{N}}$ of *partial sums* of $(\alpha_n)_{n\in\mathbb{N}}$ is defined by

$$\sigma_n = \sum_{k=0}^{n} \alpha_k = \alpha_0 + \alpha_1 + \ldots + \alpha_n,$$

for every $n \in \mathbb{N}$. If $(\sigma_n)_{n\in\mathbb{N}}$ converges to a real number x, we write

$$x = \lim_{n\longrightarrow\infty} \sigma_n = \sum_{n=0}^{\infty} \alpha_n.$$

If $(\sigma_n)_{n\in\mathbb{N}}$ converges, we write

$$\sum_{n=0}^{\infty} \alpha_n \in \mathbb{R}.$$

If $(\sigma_n)_{n\in\mathbb{N}}$ is divergent, we write

$$\sum_{n=0}^{\infty} \alpha_n \notin \mathbb{R}.$$

Note that if $n \geq m$, then

$$\begin{aligned}
\sigma_n - \sigma_m &= \left(\sum_{k=0}^{n} \alpha_k\right) - \left(\sum_{k=0}^{m} \alpha_k\right) \\
&= \big(\alpha_0 + \alpha_1 + \ldots + \alpha_m + \alpha_{m+1} + \ldots \alpha_n\big) - \big(\alpha_0 + \alpha_1 + \ldots + \alpha_m\big) \\
&= \alpha_{m+1} + \ldots + \alpha_n \\
&= \sum_{k=m+1}^{n} \alpha_k.
\end{aligned}$$

As a special case we get

$$\begin{aligned}
\sigma_n - \sigma_{n-1} &= \left(\sum_{k=0}^{n} \alpha_k\right) - \left(\sum_{k=0}^{n-1} \alpha_k\right) \\
&= \big(\alpha_0 + \alpha_1 + \ldots + \alpha_{n-1} + \alpha_n\big) - \big(\alpha_0 + \alpha_1 + \ldots + \alpha_{n-1}\big) \\
&= \alpha_n.
\end{aligned}$$

If $\alpha_n = 0$, for every $n \in \mathbb{N}$, then for the corresponding sequence of partial sums we have that

$$\sigma_n = \sum_{k=0}^{n} \alpha_k = 0 + 0 + \ldots + 0 = 0,$$

hence

$$\sum_{n=0}^{\infty} \alpha_n = 0.$$

If $x \neq 0$, and $\alpha_n = x$, for every $n \in \mathbb{N}$, then for the corresponding sequence of partial sums we have that

$$\sigma_n = \sum_{k=0}^{n} \alpha_k = x + x + \ldots + x = (n+1)x.$$

By the Archimedean axiom we get that the sequence $(\sigma_n)_{n\in\mathbb{N}}$ is unbounded, hence

$$\sum_{n=0}^{\infty} \alpha_n \notin \mathbb{R}.$$

If $(\alpha_n)_{n\in\mathbb{N}}$ is a sequence of real numbers, each term α_n can be written as a *telescoping sum*:

$$\begin{aligned}\alpha_n &= \alpha_0 + (\alpha_1 - \alpha_0) + (\alpha_2 - \alpha_1) + \ldots + (\alpha_n - \alpha_{n-1})\\ &= \alpha_0 + \sum_{k=1}^{n}(\alpha_k - \alpha_{k-1})\\ &= \alpha_0 + \sum_{k=0}^{n-1}(\alpha_{k+1} - \alpha_k).\end{aligned}$$

We can use this writing of α_n to calculate an infinite series as follows. Suppose that we need to calculate

$$\sum_{n=1}^{\infty} \gamma_n,$$

for some sequence $(\gamma_n)_{n\in\mathbb{N}}$ of real numbers.

Step 1. Let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence of real numbers such that for every $k \geq 1$

$$\gamma_k = \alpha_k - \alpha_{k-1}.$$

Step 2. By the above writing of α_n as a telescoping sum we get

$$\sum_{k=1}^{n} \gamma_k = \sum_{k=1}^{n}(\alpha_k - \alpha_{k-1}) = \alpha_n - \alpha_0.$$

Step 3. If $x \in \mathbb{R}$ such that $\lim_{n\longrightarrow\infty} \alpha_n = x$, then

$$\sum_{n=1}^{\infty} \gamma_n = \lim_{n\longrightarrow\infty} \left(\sum_{k=0}^{n} \gamma_k\right)$$

$$
\begin{aligned}
&= \lim_{n\longrightarrow\infty} \big(\alpha_n - \alpha_0\big) \\
&= \lim_{n\longrightarrow\infty} \alpha_n - \lim_{n\longrightarrow\infty} \alpha_0 \\
&= x - \alpha_0.
\end{aligned}
$$

Example. Suppose that we need to find

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)}.$$

Since

$$\gamma_k = \frac{1}{k(k+1)} = \frac{k}{k+1} - \frac{k-1}{k} = \alpha_k - \alpha_{k-1},$$

where

$$\alpha_n = \frac{n}{n+1}, \quad n \in \mathbb{N},$$

and since $\alpha_0 = 0$ and $x = \lim_{n\longrightarrow\infty} \alpha_n = 1$, we get

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1 - 0 = 1.$$

The following result is an immediate consequence of the Proposition 1.5.5.

PROPOSITION 1.7.2. *Let $(\alpha_n)_{n\in\mathbb{N}}$, $(\beta_n)_{n\in\mathbb{N}}$ be sequences of real numbers, and let $\lambda, \mu \in \mathbb{R}$. If*

$$\sum_{n=0}^{\infty} \alpha_n \in \mathbb{R} \quad \& \quad \sum_{n=0}^{\infty} \beta_n \in \mathbb{R},$$

then

$$\sum_{n=0}^{\infty} \big(\lambda\alpha_n + \mu\beta_n\big) \in \mathbb{R}, \quad \textit{and}$$

$$\sum_{n=0}^{\infty} \big(\lambda\alpha_n + \mu\beta_n\big) = \lambda\bigg(\sum_{n=0}^{\infty} \alpha_n\bigg) + \mu\bigg(\sum_{n=0}^{\infty} \beta_n\bigg).$$

As a corollary of the above proposition, from the previous example we get

$$\sum_{n=1}^{\infty} \frac{5}{n(n+1)} = 5\bigg(\sum_{n=1}^{\infty} \frac{1}{n(n+1)}\bigg) = 5 \cdot 1 = 5.$$

PROPOSITION 1.7.3 (Infinite geometric series). *If $x \in \mathbb{R}$ such that $|x| < 1$, then*

$$\sum_{n=0}^{\infty} x^n \in \mathbb{R} \quad \& \quad \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

Proof. If $n \in \mathbb{N}$, then

$$(1-x)\bigg(\sum_{k=0}^{n} x^k\bigg) = 1 - x^{n+1}.$$

Since $x \neq 1$, we get

$$\sigma_n = \sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}.$$

Since $|x| < 1$, by the Exercise 4(ii)(a) of Sheet 4 we have that $\lim_{n\longrightarrow\infty} x^{n+1} = 0$, hence

$$\lim_{n\longrightarrow\infty} \sigma_n = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

□

As a corollary of the previous proposition, we get

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = 1,$$

since

$$\sum_{n=0}^{\infty} \frac{1}{2^n} = \sum_{n=0}^{\infty} \bigg(\frac{1}{2}\bigg)^n = \frac{1}{1-\frac{1}{2}} = 2,$$

hence

$$2 = \bigg(\frac{1}{2}\bigg)^0 + \sum_{n=1}^{\infty} \frac{1}{2^n} = 1 + \sum_{n=1}^{\infty} \frac{1}{2^n}.$$

Proposition 1.7.4 (Cauchy-criterion of convergence). *Let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence of real numbers. The sequence $(\sigma_n)_{n\in\mathbb{N}}$ of partial sums of $(\alpha_n)_{n\in\mathbb{N}}$ converges if and only if*

$$\forall_{\varepsilon>0}\exists_{C_\varepsilon\in\mathbb{N}}\forall_{n\geq m\geq C_\varepsilon}\bigg(\bigg|\sum_{k=m+1}^{n} \alpha_k\bigg| < \varepsilon\bigg).$$

Proof. By the Proposition 1.6.4 and the Completeness axiom the sequence of partial sums $(\sigma_n)_{n\in\mathbb{N}}$ converges if and only of $(\sigma_n)_{n\in\mathbb{N}}$ is a Cauchy-sequence. By definition this means that for every $\varepsilon > 0$ there is $C_\varepsilon \in \mathbb{N}$ such that for all $n \geq m \geq C_\varepsilon$ we have that

$$|\sigma_n - \sigma_m| = \bigg|\sum_{k=m+1}^{n} \alpha_k\bigg| < \varepsilon.$$

□

Proposition 1.7.5 (Criterion of non-convergence of an infinite series). *Let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence of real numbers. If the sequence $(\sigma_n)_{n\in\mathbb{N}}$ of partial sums of $(\alpha_n)_{n\in\mathbb{N}}$ converges, then $\lim_{n\longrightarrow\infty} \alpha_n = 0$.*

PROOF. By the Cauchy-criterion of convergence we have that

$$\forall_{\varepsilon>0}\exists_{C_\varepsilon\in\mathbb{N}}\forall_{n\geq C_\varepsilon+1}\bigg(\bigg|\sum_{k=n-1}^{n}\alpha_k\bigg| = |\alpha_n| < \varepsilon\bigg).$$

□

Now we have one more explanation, why for $x \neq 0$

$$\sum_{n=0}^{\infty} x \notin \mathbb{R},$$

since the constant sequence x does not converge to 0. The converse to the Proposition 1.7.5 does not hold, in general. One can show that

$$\sum_{n=0}^{\infty}\frac{1}{n} = \infty, \quad \text{although} \quad \lim_{n\longrightarrow\infty}\frac{1}{n} = 0.$$

DEFINITION 1.7.6. Let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence of real numbers. The sequence $(\sigma_n)_{n\in\mathbb{N}}$ of partial sums of $(\alpha_n)_{n\in\mathbb{N}}$ converges *absolutely* if

$$\sum_{n=0}^{\infty}|\alpha_n| \in \mathbb{R}.$$

PROPOSITION 1.7.7. *If $(\alpha_n)_{n\in\mathbb{N}}$ is a sequence of real numbers, then*

$$\sum_{n=0}^{\infty}|\alpha_n| \in \mathbb{R} \ \Rightarrow \ \sum_{n=0}^{\infty}\alpha_n \in \mathbb{R}.$$

PROOF. Exercise. □

PROPOSITION 1.7.8 (Comparison test). *If $(\alpha_n)_{n\in\mathbb{N}}$ and $(\beta_n)_{n\in\mathbb{N}}$ are sequences of real numbers, such that*

$$\forall_{n\in\mathbb{N}}\big(|\alpha_n| \leq \beta_n\big), \quad \textit{and} \quad \sum_{n=0}^{\infty}\beta_n \in \mathbb{R},$$

then

$$\sum_{n=0}^{\infty}|\alpha_n| \in \mathbb{R}.$$

PROOF. By the Cauchy-criterion of convergence we have that

$$\forall_{\varepsilon>0}\exists_{C_\varepsilon^\beta\in\mathbb{N}}\forall_{n\geq m\geq C_\varepsilon^\beta}\bigg(\bigg|\sum_{k=m+1}^{n}\beta_k\bigg| < \varepsilon\bigg).$$

If we define $C_\varepsilon^{|\alpha|} = C_\varepsilon^\beta$, then for every $n \geq m \geq C_\varepsilon^{|\alpha|}$ we get

$$\sum_{k=m+1}^{n}|\alpha_k| \leq \sum_{k=m+1}^{n}\beta_k \leq \bigg|\sum_{k=m+1}^{n}\beta_k\bigg| < \varepsilon,$$

hence by the Cauchy-criterion of convergence again we get $\sum_{n=0}^{\infty} |\alpha_n| \in \mathbb{R}$. □

As a corollary of the comparison test we show (exercise) that

$$\sum_{n=1}^{\infty} \frac{1}{n^k} \in \mathbb{R}, \qquad k \geq 2.$$

PROPOSITION 1.7.9 (Quotient-criterion). *Let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence of real numbers, such that $\alpha_n \neq 0$, for every $n \geq n_0$, and some $n_0 \in \mathbb{N}$. Let $\theta \in \mathbb{R}$ such that*
(i) $0 < \theta < 1$, and
(ii) for every $n \geq n_0$, it holds

$$\left|\frac{\alpha_{n+1}}{\alpha_n}\right| \leq \theta.$$

Then

$$\sum_{n=0}^{\infty} |\alpha_n| \in \mathbb{R}.$$

PROOF. Since

$$\sum_{n=0}^{\infty} |\alpha_n| = \sum_{k=0}^{n_0-1} |\alpha_k| + \sum_{k=n_0}^{\infty} |\alpha_k|,$$

it suffices to show that

$$\sum_{k=n_0}^{\infty} |\alpha_k| \in \mathbb{R}.$$

Because of this, we suppose without loss of generality that $\alpha_n \neq 0$, for every $n \in \mathbb{N}$ i.e., $n_0 = 0$. Because of (ii), a simple induction shows that

$$\forall_{n\in\mathbb{N}}\big(|\alpha_n| \leq |\alpha_0|\theta^n\big).$$

Since

$$\sum_{n=0}^{\infty} |\alpha_0|\theta^n = |\alpha_0|\bigg(\sum_{n=0}^{\infty} \theta^n\bigg) = |\alpha_0|\frac{1}{1-\theta},$$

what we want follows from the comparison test. □

CHAPTER 2

Real-valued functions of a real variable

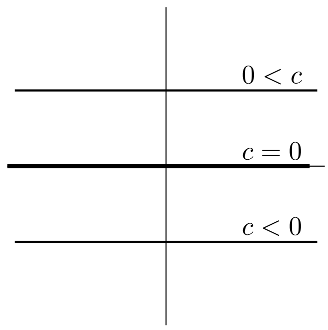
In this chapter we study the continuity, the differentiability and the integration of functions defined on a subset of $\mathbb{R}$ with values in $\mathbb{R}$. We can picture these functions through the representation of their graph in the Euclidean plane $\mathbb{R}^2$. First we study the notion of a continuous function $f : D \subseteq \mathbb{R} \to \mathbb{R}$. As it is indicated by the term continuous, the graph of a continuous function is a continuous curve in the plane $\mathbb{R}^2$.

2.1. The graph of a real-valued function of a real variable

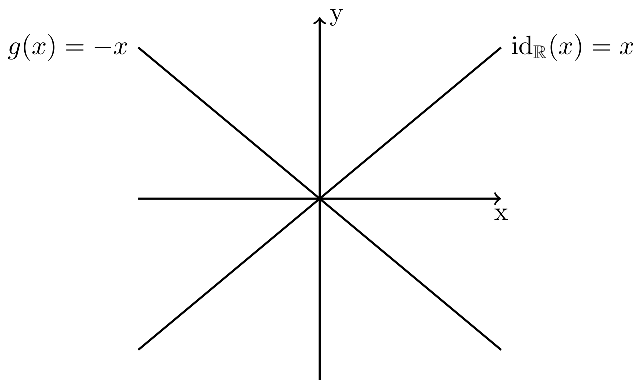
Definition 2.1.1. A *real-valued function of a real variable* is a function $f : D \to \mathbb{R}$, where D is a subset of $\mathbb{R}$. The *graph* $\mathrm{Gr}(f)$ of f is defined by

$$\mathrm{Gr}(f) = \{(x, y) \in D \times \mathbb{R} \mid y = f(x)\}.$$

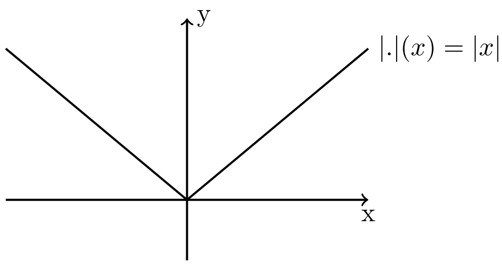
Example 1. If $c \in \mathbb{R}$, let the *constant function* c is the function $f_c : D \to \mathbb{R}$, defined by $f_c(x) = c$, for every $x \in \mathbb{R}$. If $D = \mathbb{R}$, the graph of f_c is a straight line parallel to the axis of x's, the position of which depends on the value of c, as it is shown in the following figure.



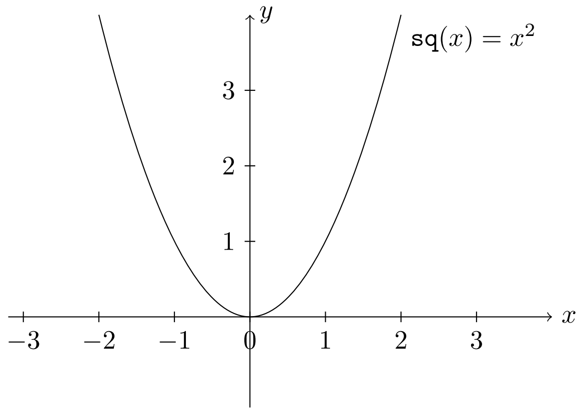
Example 2. The *identity function* $\mathrm{id}_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$, where $x \mapsto x$, for every $x \in \mathbb{R}$, has as graph the following diagonal line, while the graph of the function $g : \mathbb{R} \to \mathbb{R}$, defined by $g(x) = -x$, for every $x \in \mathbb{R}$, is the line symmetric to the graph of $\mathrm{id}_{\mathbb{R}}$, with respect to the horizontal axis.



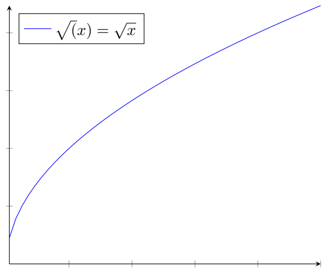
Example 3. The *absolute value function* $|.| : \mathbb{R} \to \mathbb{R}$, defined by $x \mapsto |x|$, for every $x \in \mathbb{R}$, has as graph the following curve



Example 4. The *square function* $\mathtt{sq} : \mathbb{R} \to \mathbb{R}$, defined by $\mathtt{sq}(x) = x^2$, for every $x \in \mathbb{R}$, has as graph the following curve



Example 5. The function $\sqrt{} : \mathbb{R}^+ \to \mathbb{R}$, defined by $\sqrt{}(x) = \sqrt{x}$, for every $x \in \mathbb{R}^+$, has as graph the following curve



Example 6. The *Dirichlet function* $\texttt{Dir} : \mathbb{R} \to \mathbb{R}$ is defined by

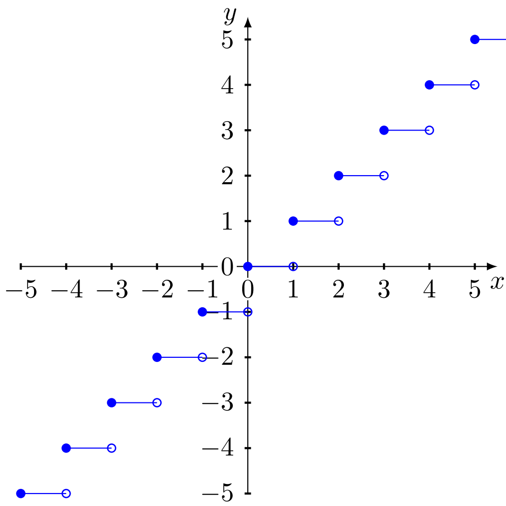
$$\texttt{Dir}(x) := \begin{cases} 1 & , x \in \mathbb{Q} \\ 0 & , x \in \mathbb{I}, \end{cases}$$

and its graph cannot be represented by a continuous curve in the plane.

Example 7. The *floor function* $\lfloor . \rfloor : \mathbb{R} \to \mathbb{R}$ is defined by $x \mapsto \lfloor x \rfloor$, for every $x \in \mathbb{R}$, where $\lfloor x \rfloor$ is the unique integer such that

$$\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1.$$

The graph of the floor function is pictured in the following figure.



Definition 2.1.2. Let $D \subseteq \mathbb{R}$, $f, g : D \to \mathbb{R}$, and $\lambda \in \mathbb{R}$. Let the functions $f + g, \lambda f, f \cdot g : D \to \mathbb{R}$, defined by

$$(f + g)(x) = f(x) + g(x),$$

$$(\lambda f)(x) = \lambda f(x),$$

$$(f \cdot g)(x) = f(x) \cdot g(x),$$

for every $x \in D$, respectively. If

$$D_g^* = \{x \in D \mid g(x) \neq 0\},$$

we define the function $\frac{f}{g} : D_g^* \to \mathbb{R}$, where

$$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)},$$

for every $x \in D_g^*$.

Example 8. The function $\frac{1}{\mathtt{sq}} : \mathbb{R}^* \to \mathbb{R}$, where $\mathbb{R}^* = \{x \in \mathbb{R} \mid x \neq 0\}$, is defined by

$$\left(\frac{1}{\mathtt{sq}}\right)(x) = \frac{1}{x^2},$$

for every $x \in \mathbb{R}^*$.

Example 9. A *polynomial* p is a function $p : \mathbb{R} \to \mathbb{R}$, where

$$\begin{aligned} p &= \sum_{k=0}^{n} a_k \mathrm{id}_{\mathbb{R}}^k \\ &= a_0 \mathrm{id}_{\mathbb{R}}^0 + a_1 \mathrm{id}_{\mathbb{R}}^1 + \ldots a_n \mathrm{id}_{\mathbb{R}}^n \\ &= a_0 + a_1 \mathrm{id}_{\mathbb{R}} + a_2 \mathrm{id}_{\mathbb{R}}^2 + \ldots + a_n \mathrm{id}_{\mathbb{R}}^n, \end{aligned}$$

where $a_0, a_1, \ldots, a_n \in \mathbb{R}$. If $a_n \neq 0$, the number n is called the *degree* of p. If $x \in \mathbb{R}$, then by definition we get

$$\begin{aligned} p(x) &= \left(\sum_{k=0}^{n} a_k \mathrm{id}_{\mathbb{R}}^k\right)(x) \\ &= \big(a_0 + a_1 \mathrm{id}_{\mathbb{R}} + a_2 \mathrm{id}_{\mathbb{R}}^2 + \ldots + a_n \mathrm{id}_{\mathbb{R}}^n\big)(x) \\ &= a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n. \end{aligned}$$

The identity function $\mathrm{id}_{\mathbb{R}}$ is a polynomial of degree 1 ($a_0 = 0$ and $a_1 = 1$), while the function $h(x) = x^2$ is a polynomial of degree 2 ($a_0 = a_1 = 0$ and $a_2 = 1$).

Example 10. If $p = \sum_{k=0}^{n} a_k \mathrm{id}_{\mathbb{R}}^k$ and $q = \sum_{k=0}^{m} b_k \mathrm{id}_{\mathbb{R}}^k$ are polynomials, the *rational function* R_{pq} is a function $R_{pq} : D_q^* \to \mathbb{R}$, defined by

$$\begin{aligned} R_{pq}(x) &= \frac{p(x)}{q(x)} \\ &= \frac{\left(\sum_{k=0}^{n} a_k \mathrm{id}_{\mathbb{R}}^k\right)(x)}{\left(\sum_{k=0}^{m} b_k \mathrm{id}_{\mathbb{R}}^k\right)(x)} \end{aligned}$$

$$= \frac{a_0 + a_1x + a_2x^2 + \ldots + a_nx^n}{b_0 + b_1x + b_2x^2 + \ldots + b_mx^m},$$

where

$$D_q^* = \{x \in \mathbb{R} \mid b_0 + b_1x + b_2x^2 + \ldots + b_mx^m \neq 0\}.$$

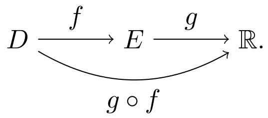
The next definition is a special case of the Definition 1.2.3.

DEFINITION 2.1.3. Let $D, E \subseteq \mathbb{R}$, and let $f : D \to \mathbb{R}$ and $g : E \to \mathbb{R}$, such that

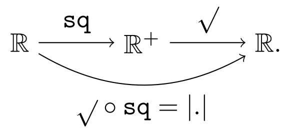
$$\mathtt{Im}(f) = \{f(x) \mid x \in D\} \subseteq E.$$

The *composition* $g \circ f : D \to \mathbb{R}$ of f and g is defined, for every $x \in D$, by

$$(g \circ f)(x) = g(f(x))$$



Example 11. If $\mathtt{sq}(x) = x^2$, then $(\sqrt{} \circ \mathtt{sq})(x) = \sqrt{}(\mathtt{sq}(x)) = \sqrt{x^2} = |x|$



2.2. Continuity

DEFINITION 2.2.1. Let $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$, and $x_0, l \in \mathbb{R} \cup \{-\infty, +\infty\}$. Let also the set

$$\mathbb{F}(\mathbb{N}, \mathbb{R}) = \{\alpha : \mathbb{N} \to \mathbb{R}\}.$$

(i) Let $D(x_0)$ be the set of all sequences in D that converge to x_0 i.e.,

$$D(x_0) = \{(\alpha_n)_{n\in\mathbb{N}} \in \mathbb{F}(\mathbb{N}, \mathbb{R}) \mid \forall_{n\in\mathbb{N}}(\alpha_n \in D) \ \& \ \lim_{n\longrightarrow\infty} \alpha_n = x_0\}.$$

If the set $D(x_0)$ is non-empty, we say that x_0 is an *accumulation-point* of D. If x_0 is an accumulation-point of D, we define

$$\lim_{x\longrightarrow x_0} f(x) = l :\Leftrightarrow \forall_{(\alpha_n)_{n\in\mathbb{N}}\in D(x_0)} \left(\lim_{n\longrightarrow\infty} f(\alpha_n) = l \right)$$

i.e.,

$$\left[\lim_{n\longrightarrow\infty} \alpha_n = x_0 \right] \Rightarrow \left[\lim_{n\longrightarrow\infty} f(\alpha_n) = l \right],$$

for every sequence of real numbers $(\alpha_n)_{n\in\mathbb{N}}$ in D.

(ii) If $x_0 \in \mathbb{R}$, let the set

$$D^+(x_0) = \{(x_n)_{n\in\mathbb{N}} \in \mathbb{F}(\mathbb{N}, \mathbb{R}) \mid \forall_{n\in\mathbb{N}}(\alpha_n \in D \ \& \ \alpha_n > x_0) \ \& \ \lim_{n\longrightarrow\infty} \alpha_n = x_0\}.$$

If the set $D^+(x_0)$ is non-empty, we define

$$\lim_{x\longrightarrow x_0^+} f(x) = l :\Leftrightarrow \forall_{(\alpha_n)_{n\in\mathbb{N}}\in D^+(x_0)}\Big(\lim_{n\longrightarrow\infty} f(\alpha_n) = l\Big).$$

(iii) If $x_0 \in \mathbb{R}$, let the set

$$D^-(x_0) = \{(\alpha_n)_{n\in\mathbb{N}} \in \mathbb{F}(\mathbb{N},\mathbb{R}) \mid \forall_{n\in\mathbb{N}}(\alpha_n \in D \;\&\; \alpha_n < x_0) \;\&\; \lim_{n\longrightarrow\infty} \alpha_n = x_0\}.$$

If the set $D^-(x_0)$ is non-empty, we define

$$\lim_{x\longrightarrow x_0^-} f(x) = l :\Leftrightarrow \forall_{(\alpha_n)_{n\in\mathbb{N}}\in D^-(x_0)}\Big(\lim_{n\longrightarrow\infty} f(\alpha_n) = l\Big).$$

(iv) Let D be *unbounded above* i.e.,

$$\forall_{n\in\mathbb{N}}\exists_{x\in D}(x \geq n).$$

For such a set D we define

$$D(+\infty) = \{(\alpha_n)_{n\in\mathbb{N}} \in \mathbb{F}(\mathbb{N},\mathbb{R}) \mid \forall_{n\in\mathbb{N}}(\alpha_n \in D) \;\&\; \lim_{n\longrightarrow\infty} \alpha_n = +\infty\},$$

and

$$\lim_{x\longrightarrow +\infty} f(x) = l :\Leftrightarrow \forall_{(\alpha_n)_{n\in\mathbb{N}}\in D(+\infty)}\Big(\lim_{n\longrightarrow\infty} f(\alpha_n) = l\Big).$$

(v) Let D be *unbounded below* i.e.,

$$\forall_{n\in\mathbb{N}}\exists_{x\in D}(x \leq -n).$$

For such a set D we define

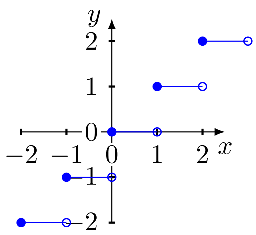
$$D(-\infty) = \{(\alpha_n)_{n\in\mathbb{N}} \in \mathbb{F}(\mathbb{N},\mathbb{R}) \mid \forall_{n\in\mathbb{N}}(\alpha_n \in D) \;\&\; \lim_{n\longrightarrow\infty} \alpha_n = -\infty\},$$

and

$$\lim_{x\longrightarrow -\infty} f(x) = l :\Leftrightarrow \forall_{(\alpha_n)_{n\in\mathbb{N}}\in D(-\infty)}\Big(\lim_{n\longrightarrow\infty} f(\alpha_n) = l\Big).$$

If $x_0 \in D$, then x_0 is an accumulation-point of D, since the constant sequence x_0 is in $D(x_0)$.

Example 1. If we consider the floor function (Example 7 in the previous section),



then

$$\lim_{x \longrightarrow 0^+} \lfloor x \rfloor = 0 \quad \& \quad \lim_{x \longrightarrow 0^-} \lfloor x \rfloor = -1,$$

since if $(\alpha_n)_{n\in\mathbb{N}} \in \mathbb{R}^+(0)$, then by definition $\alpha_n > 0$, for every $n \in \mathbb{N}$, and since $\lim_{n\longrightarrow\infty} \alpha_n = 0$, then for every $n \geq n_0$, for some $n_0 \in \mathbb{N}$, we have that $\alpha_n \in (0, 1]$, hence $\lfloor \alpha_n \rfloor = 0$, for every $n \geq n_0$. Hence

$$\lim_{n\longrightarrow\infty} \lfloor \alpha_n \rfloor = \lim_{n\longrightarrow\infty} 0 = 0.$$

Similarly, if $(\alpha_n)_{n\in\mathbb{N}} \in \mathbb{R}^-(0)$, then by definition $\alpha_n < 0$, for every $n \in \mathbb{N}$, and since $\lim_{n\longrightarrow\infty} \alpha_n = 0$, then for every $n \geq n_0$, for some $n_0 \in \mathbb{N}$, we have that $\alpha_n \in (-1, 0]$, hence $\lfloor \alpha_n \rfloor = -1$, for every $n \geq n_0$. Hence

$$\lim_{n\longrightarrow\infty} \lfloor \alpha_n \rfloor = \lim_{n\longrightarrow\infty} -1 = -1.$$

DEFINITION 2.2.2. Let D be a subset of $\mathbb{R}$ and $x_0 \in D$. A function $f : D \to \mathbb{R}$ is *continuous at* x_0, if

$$\lim_{x\longrightarrow x_0} f(x) = f(x_0).$$

The function f is called *continuous on* D, if it is continuous at very point in D.

By the Definition 2.2.1 $f : D \to \mathbb{R}$ is continuous at $x_0 \in D$ if and only if

$$\big[\lim_{n\longrightarrow\infty} \alpha_n = x_0\big] \Rightarrow \big[\lim_{n\longrightarrow\infty} f(\alpha_n) = f(x_0)\big],$$

for every sequence of real numbers $(\alpha_n)_{n\in\mathbb{N}}$ in D. The Examples 1-5 of real functions in the previous section are continuous functions on their domain of definition, while the Examples 6 and 7 are not.

PROPOSITION 2.2.3. *Let* $D \subseteq \mathbb{R}$, $x_0 \in D$, $f, g : D \to \mathbb{R}$, *and* $\lambda \in \mathbb{R}$.

(I) Suppose that f, g *are continuous at* x_0.

(i) The functions $f + g, \lambda f, f \cdot g : D \to \mathbb{R}$, *defined in the Definition 2.1.2, are continuous at* x_0.

(ii) If $g(x_0) \neq 0 \Leftrightarrow x_0 \in D_g^*$, *the function* $\frac{f}{g} : D_g^* \to \mathbb{R}$, *defined also in the Definition 2.1.2, is continuous at* x_0.

(II) (i) If f, g *are continuous on* D, *then the functions* $f + g, \lambda f, f \cdot g : D \to \mathbb{R}$ *are also continuous on* D.

(ii) If f, g *are continuous on* D_g^*, *the function* $\frac{f}{g} : D_g^* \to \mathbb{R}$ *is continuous on* D_g^*.

PROOF. (I)(i) Let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence of reals in D such that $\lim_{n\longrightarrow\infty} \alpha_n = x_0$. By the Proposition 1.5.5 we get

$$\begin{aligned}
\lim_{n\longrightarrow\infty} \big(f + g\big)(\alpha_n) &= \lim_{n\longrightarrow\infty} \big[f(\alpha_n) + g(\alpha_n)\big] \\
&= \lim_{n\longrightarrow\infty} f(\alpha_n) + \lim_{n\longrightarrow\infty} g(\alpha_n) \\
&= f(x_0) + g(x_0)
\end{aligned}$$

$$= \big(f + g\big)(x_0),$$

$$\lim_{n\longrightarrow\infty} \big(\lambda f\big)(\alpha_n) = \lim_{n\longrightarrow\infty} \lambda f(\alpha_n) = \lambda \lim_{n\longrightarrow\infty} f(\alpha_n) = \lambda f(x_0) = \big(\lambda f\big)(x_0),$$

and

$$\begin{aligned}\lim_{n\longrightarrow\infty} \big(f \cdot g\big)(\alpha_n) &= \lim_{n\longrightarrow\infty} \big[f(\alpha_n) \cdot g(\alpha_n)\big] \\ &= \lim_{n\longrightarrow\infty} f(\alpha_n) \cdot \lim_{n\longrightarrow\infty} g(\alpha_n) \\ &= f(x_0) \cdot g(x_0) \\ &= \big(f \cdot g\big)(x_0).\end{aligned}$$

(I)(ii) Let $(\beta_n)_{n\in\mathbb{N}}$ be a sequence of reals in D^*_g such that $\lim_{n\longrightarrow\infty} \beta_n = x_0$. By the Proposition 1.5.5 we get

$$\lim_{n\longrightarrow\infty} \left(\frac{f}{g}\right)(\beta_n) = \lim_{n\longrightarrow\infty} \frac{f(\beta_n)}{g(\beta_n)} = \frac{\lim_{n\longrightarrow\infty} f(\beta_n)}{\lim_{n\longrightarrow\infty} g(\beta_n)} = \frac{f(x_0)}{g(x_0)} = \left(\frac{f}{g}\right)(x_0).$$

(II)(i) and (II)(ii) follow immediately from (I)(i) and (I)(ii), respectively. □

By the previous proposition the real functions in the Examples 8-10 of the previous section are continuous functions on their domain of definition.

Proposition 2.2.4. *Let $D, E \subseteq \mathbb{R}$, $x_0 \in D$ and $y_0 \in E$, and let $f : D \to \mathbb{R}$ and $g : E \to \mathbb{R}$, such that $\mathtt{Im}(f) = \{f(x) \mid x \in D\} \subseteq E$ and $y_0 = f(x_0)$. The following hold:*

(i) If f is continuous at x_0 and g is continuous at y_0, the composition $g \circ f : D \to \mathbb{R}$ is continuous at x_0.

(ii) If f is continuous on D and g is continuous on E, the composition $g \circ f$ is continuous on D.

Proof. (i) Let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence of reals in D such that $\lim_{n\longrightarrow\infty} \alpha_n = x_0$. By the definition of continuity of f at x_0 and of g at y_0 we have that

$$\lim_{n\longrightarrow\infty} f(\alpha_n) = f(x_0) = y_0 \quad \& \quad \lim_{n\longrightarrow\infty} g\big(f(\alpha_n)\big) = g(y_0) = g(f(x_0)).$$

Hence,

$$\begin{aligned}\lim_{n\longrightarrow\infty} \big(g \circ f\big)(\alpha_n) &= \lim_{n\longrightarrow\infty} g\big(f(\alpha_n)\big) \\ &= g(f(x_0)) \\ &= \big(g \circ f\big)(x_0).\end{aligned}$$

(ii) It follows immediately from (i). □

Example. Since the function $\mathtt{sq}(x) = x^2$ is continuous on $\mathbb{R}$ and the function $\sqrt{\ }$ is continuous on $\mathbb{R}^+$ (Exercise), by the previous proposition we have that the absolute value-function is continuous on $\mathbb{R}$ [recall that $(\sqrt{\ } \circ \mathtt{sq})(x) = \sqrt{(\mathtt{sq}(x))} = \sqrt{x^2} = |x|$]

$$\mathbb{R} \xrightarrow{\mathtt{sq}} \mathbb{R}^+ \xrightarrow{\sqrt{\ }} \mathbb{R}.$$
$$\sqrt{\ } \circ \mathtt{sq} = |.|$$

THEOREM 2.2.5 (Intermediate value theorem). *Let $a, b \in \mathbb{R}$ such that $a < b$, and let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$. If $f(a)f(b) < 0$, then there exists $x_0 \in [a, b]$ such that $f(x_0) = 0$.*

PROOF. See [**1**], p. 106. □

Notice that the condition $f(a)f(b) < 0$ above is equivalent to

$$\big[f(a) < 0 \ \& \ f(b) > 0\big] \quad \text{or} \quad \big[f(a) > 0 \ \& \ f(b) < 0\big].$$

COROLLARY 2.2.6. *Let $a, b \in \mathbb{R}$ such that $a < b$, and let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$. If $c \in \mathbb{R}$ such that $f(a) < c < f(b)$, then there exists $x_0 \in [a, b]$ such that $f(x_0) = c$.*

PROOF. Let the function $g : [a, b] \to \mathbb{R}$, defined by

$$g(x) = f(x) - c,$$

for every $x \in [a, b]$. Since $g(a) = f(a) - c < 0$ and $g(b) = f(b) - c > 0$, by the Theorem 2.2.5 there exists $x_0 \in [a, b]$ such that

$$g(x_0) = 0 \Leftrightarrow f(x_0) - c = 0 \Leftrightarrow f(x_0) = c.$$

□

COROLLARY 2.2.7. *Let $p : \mathbb{R} \to \mathbb{R}$ a polynomial function of odd degree i.e., there is $n \in \mathbb{N}$ such that*

$$p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_{2n} x^{2n} + x^{2n+1},$$

for every $x \in \mathbb{R}$. Then there exists $x_0 \in \mathbb{R}$ such that $p(x_0) = 0$.

PROOF. If $x \neq 0$, we have that

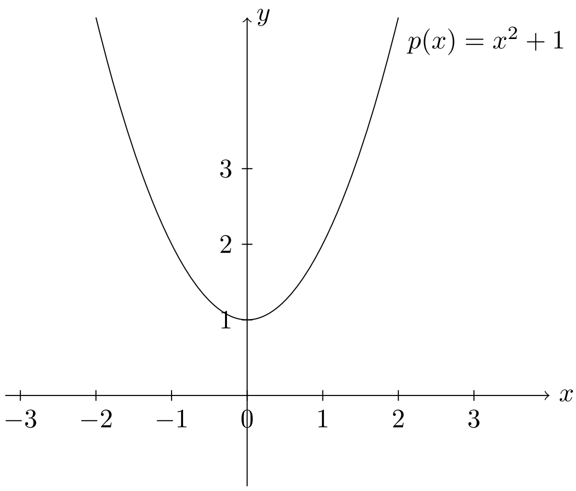
$$p(x) = x^{2n+1}\left(\frac{a_0}{x^{2n+1}} + \frac{a_1}{x^{2n}} + \frac{a_2}{x^{2n-1}} + \ldots + \frac{a_{2n}}{x} + 1\right).$$

Then we get

$$\lim_{x \longrightarrow +\infty} p(x) = +\infty \quad \& \quad \lim_{x \longrightarrow -\infty} p(x) = -\infty.$$

Hence we can find $a < 0 < b$ such that $p(a) < 0 < p(b)$. Since p is continuous on $\mathbb{R}$, it is also continuous on $a, b]$, hence by the Theorem 2.2.5 there is $x_0 \in \mathbb{R}$ such that $p(x_0) = 0$. □

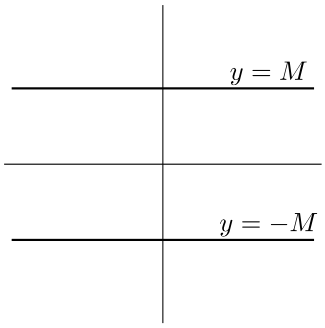
Notice that there is a polynomial function of even degree without any roots i.e., $p(x) \neq 0$, for every $x \in \mathbb{R}$. Consider for example the polynomial function $p(x) = x^2 + 1$, where $x \in \mathbb{R}$. This is also evident by the graph of p.



DEFINITION 2.2.8. A real function $f : D \to \mathbb{R}$ is called *bounded*, if there is $M \in \mathbb{R}$ such that $M > 0$ and

$$\forall_{x \in D}\big(|f(x)| \leq M\big).$$

The geometric interpretation of a bounded function f with bound $M > 0$ is that its graph $\mathrm{Gr}(f)$ is between the horizontal lines $y = M$ and $y = -M$.



A continuous function defined on an unbounded interval can be unbounded. E.g., the above function $p(x) = x^2 + 1$ is defined on $\mathbb{R}$ and its graph cannot be between any two horizontal lines. If a continuous function though, is defined on a bounded set, it is always a bounded function.

THEOREM 2.2.9. *Let $a, b \in \mathbb{R}$ such that $a < b$, and let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$. Then there exists $x_0, x_1 \in [a, b]$ such that $f(x_0) = m$, $f(x_1) = M$, and*

$$\forall_{x \in [a,b]}\big(m \leq f(x) \leq M\big).$$

PROOF. See [**1**], p. 110. □

2.3. Elementary Functions

PROPOSITION 2.3.1. *For every $x \in \mathbb{R}$ the exponential series*

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \frac{x^0}{0!} + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

converges absolutely.

PROOF. We use the quotient-criterion (Proposition 1.7.9). Let

$$\alpha_n = \frac{x^n}{n!}.$$

For every $x \neq 0$ and $n \geq 2|x|$ we have that

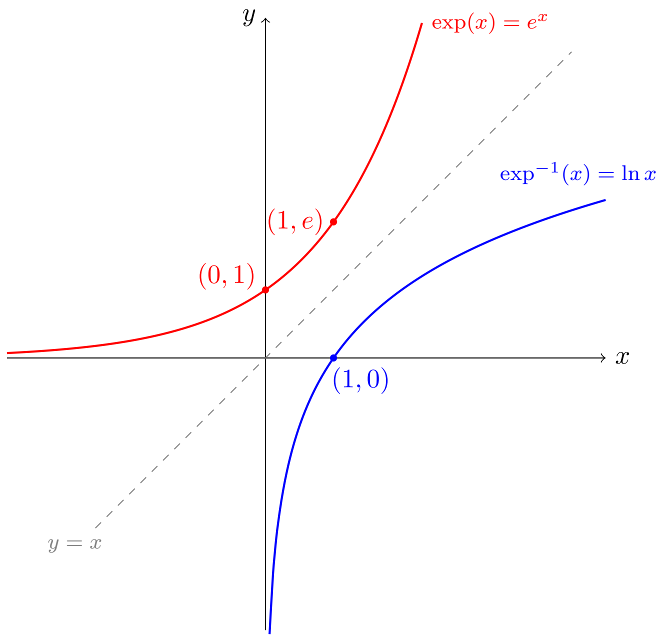
$$\left|\frac{\alpha_{n+1}}{\alpha_n}\right| = \left|\frac{x^{n+1} n!}{x^n (n+1)!}\right| = \frac{|x|}{n+1} \leq \frac{1}{2}.$$

□

With the help of the exponential series we define the famous number e of Euler

$$e = \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!} = 1 + 1 + \frac{1}{2} + \frac{1}{3!} + \dots.$$

and the *exponential function* $\exp : \mathbb{R} \to \mathbb{R}$ with $\exp(x) = e^x$, for every $x \in \mathbb{R}$.



PROPOSITION 2.3.2. *For every $x, y \in \mathbb{R}$ the following hold:*
(i) $\exp(x+y) = \exp(x)\exp(y)$.
(ii) $\exp(x) > 0$.
(iii) $\exp(-x) = \frac{1}{\exp(x)}$.
(iv) $\exp(k) = e^k$, *for every* $k \in \mathbb{Z}$.

PROOF. See [**1**], p. 80. □

Basic limit 1. Next we explain why

$$\lim_{x \longrightarrow 0} \frac{\exp(x) - 1}{x} = 1.$$

From the definition of $\exp(x)$ we have that

$$\exp(x) - 1 = \left(\sum_{n=0}^{\infty} \frac{x^n}{n!} \right) - 1 = \sum_{n=1}^{\infty} \frac{x^n}{n!},$$

hence, if $x \neq 0$ we have that

$$\begin{aligned} \frac{\exp(x) - 1}{x} &= \frac{1}{x} \sum_{n=1}^{\infty} \frac{x^n}{n!} \\ &= \frac{1}{x} \left(\frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \right) \\ &= 1 + \left(\frac{x^1}{2!} + \frac{x^2}{3!} + \frac{x^3}{4!} + \dots \right), \end{aligned}$$

which converges to 1, as x converges to 0.

The exponential function exp is continuous and *strictly increasing* ($x < y \Rightarrow \exp(x) < \exp(y)$), and maps $\mathbb{R}$ bijectively onto

$$\mathbb{R}^{+*} = \{x \in \mathbb{R} \mid x > 0\}.$$

Its inverse function

$$\ln : \mathbb{R}^{+*} \to \mathbb{R} \qquad x \mapsto \ln(x)$$

is also continuous and strictly monotone, and it is called the *natural logarithmic function*. By definition we have that

$$\exp(\ln(x)) = e^{\ln(x)} = x,$$
$$\ln(\exp(x)) = \ln e^x = x,$$

and since exp and ln are injective functions we have that

$$\exp(x) = \exp(y) \Rightarrow x = y, \quad x, y \in \mathbb{R},$$
$$\ln(x) = \ln(y) \Rightarrow x = y, \quad x, y \in \mathbb{R}^+.$$

It is then easy to show (Exercise) that

$$\ln(x \cdot y) = \ln(x) + \ln(y),$$

for every $x, y \in \mathbb{R}^{+*}$. Instead of $\ln(x)$, one also writes $\log(x)$.

For every $x \in \mathbb{R}$ the infinite series

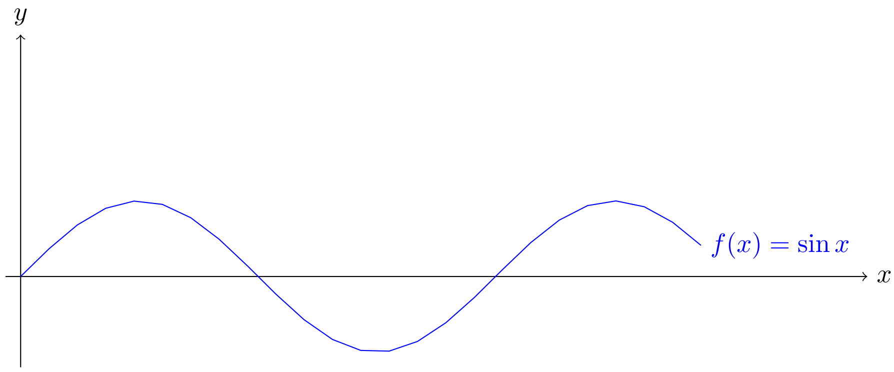
$$\sin(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = \frac{x}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} \mp \ldots,$$

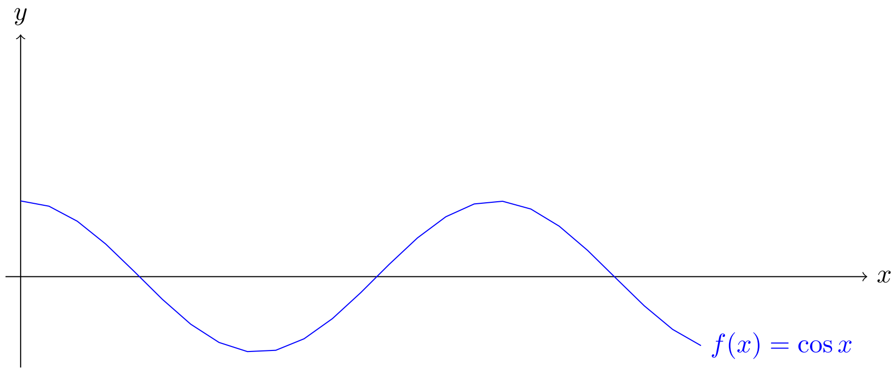
$$\cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = \frac{x^0}{0!} - \frac{x^2}{2!} + \frac{x^4}{4!} \mp \ldots$$

are absolutely convergent (see [**1**], p. 140). Their absolute convergence follows from the absolute convergence of the infinite exponential series. The functions

$$\sin : \mathbb{R} \to \mathbb{R} \quad x \mapsto \sin(x), \quad \& \quad \cos : \mathbb{R} \to \mathbb{R} \quad x \mapsto \cos(x)$$

are shown to be continuous on $\mathbb{R}$.





The real number $\frac{\pi}{2}$ is the unique root of the function cos in the interval $[0, 2]$ (see [**1**], pp. 142-143). Based on the previous definitions it is not trivial to show the fundamental equality

$$\sin(x)^2 + \cos(x)^2 = 1.$$

Basic limit 2. Next we explain why

$$\lim_{x \longrightarrow 0} \frac{\sin(x)}{x} = 1.$$

From the definition of $\sin(x)$ we have that

$$\begin{aligned}\frac{\sin(x)}{x} &= \frac{1}{x} \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \\ &= \frac{1}{x} \left(\frac{x}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} \mp \dots \right) \\ &= 1 - \frac{x^2}{3!} + \frac{x^4}{5!} \mp \dots,\end{aligned}$$

which converges to 1, as x converges to 0.

The *tangent function* is defined on the set

$$D_{\tan} = \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi \mid k \in \mathbb{Z} \right\}$$

through the rule

$$\tan(x) = \frac{\sin(x)}{\cos(x)}.$$

The *cotangent function* is defined on the set

$$D_{\cot} = \mathbb{R} \setminus \{k\pi \mid k \in \mathbb{Z}\}$$

through the rule

$$\cot(x) = \frac{\cos(x)}{\sin(x)}.$$

One can show that

$$\cot(x) = \tan\left(\frac{\pi}{2} - x\right).$$

2.4. Differentiation

DEFINITION 2.4.1. Let $D \subseteq \mathbb{R}$ and $f : D \to \mathbb{R}$ a real function. We say that f is *differentiable at* $x_0 \in D$, is the limit

$$f'(x_0) = \lim_{h \longrightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

exists. For the calculation of this limit we consider sequences $(\alpha_n)_{n \in \mathbb{N}}$ of real numbers such that $\lim_{n \longrightarrow \infty} h_n = 0$ and

$$[h_n \neq 0 \ \ \& \ \ x_0 + h_n \in D], \qquad \text{for all } n \in \mathbb{N}.$$

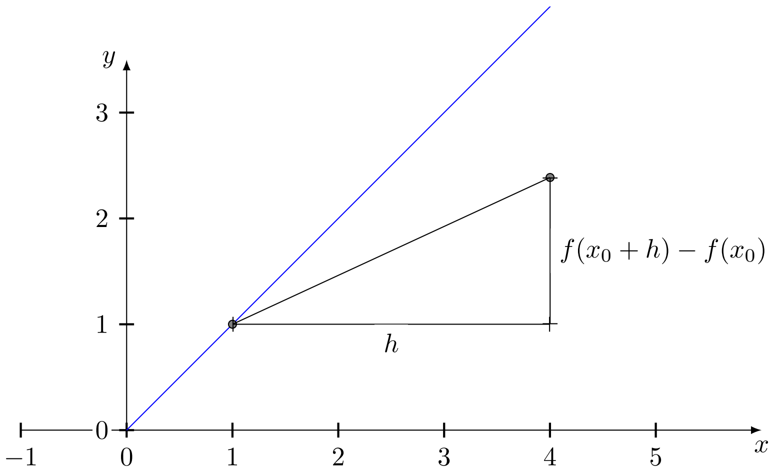
The limit $f'(x_0)$ is called the *derivative* of f at x_0. The function f is called *differentiable in* D, if f is differentiable at every point $x \in D$. We also use the notations

$$\frac{df(x_0)}{dx}, \text{ or } \frac{df}{dx}(x_0) \text{ for } f'(x_0).$$

The ratio

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

is the tangent of the following angle in the triangle $(x_0, f(x_0))$, $(x_0+h, f(x_0+h))$, and $(x_0+h, f(x_0))$.



By taking the limit

$$h \longrightarrow 0 \Leftrightarrow x_0 + h \longrightarrow x_0,$$

the derivative $f'(x_0)$ of f at x_0 is the slope of the line that is tangent to the graph of f at the point $(x_0, f(x_0))$.

Example 1. For the constant function $f : \mathbb{R} \to \mathbb{R}, f(x) = c$, for all $x \in \mathbb{R}$, we get

$$f'(x_0) = \lim_{h\longrightarrow 0} \frac{f(x_0+h)-f(x_0)}{h} = \lim_{h\longrightarrow 0} \frac{c-c}{h} = \lim_{h\longrightarrow 0} 0 = 0.$$

Example 2. For the identity map $\mathrm{id}_\mathbb{R} : \mathbb{R} \to \mathbb{R}, \mathrm{id}_\mathbb{R}(x) = x$, for all $x \in \mathbb{R}$, we get

$${\mathrm{id}_\mathbb{R}}'(x_0) = \lim_{h\longrightarrow 0} \frac{\mathrm{id}_\mathbb{R}(x_0+h)-\mathrm{id}_\mathbb{R}(x_0)}{h} = \lim_{h\longrightarrow 0} \frac{x_0+h-x_0}{h} = \lim_{h\longrightarrow 0} 1 = 1.$$

Example 3. For the function $g : \mathbb{R} \to \mathbb{R}, g(x) = \lambda x$, for all $x \in \mathbb{R}$, where $\lambda \in \mathbb{R}$, we get

$$g'(x_0) = \lim_{h\longrightarrow 0} \frac{\lambda(x_0+h)-\lambda x_0}{h} = \lim_{h\longrightarrow 0} \frac{\lambda x_0+\lambda h-\lambda x_0}{h} = \lim_{h\longrightarrow 0} \lambda = \lambda.$$

Example 4. For the function $\mathtt{sq} : \mathbb{R} \to \mathbb{R}$, where $\mathtt{sq}(x) = x^2$, for all $x \in \mathbb{R}$, we get

$$\begin{aligned}
\mathtt{sq}'(x_0) &= \lim_{h \longrightarrow 0} \frac{\mathtt{sq}(x_0 + h) - \mathtt{sq}(x_0)}{h} \\
&= \lim_{h \longrightarrow 0} \frac{(x_0 + h)^2 - x_0^2}{h} \\
&= \lim_{h \longrightarrow 0} \frac{x_0^2 + 2x_0 h + h^2 - x_0^2}{h} \\
&= \lim_{h \longrightarrow 0} \frac{2x_0 h + h^2}{h} \\
&= \lim_{h \longrightarrow 0} \frac{h(2x_0 + h)}{h} \\
&= \lim_{h \longrightarrow 0} (2x_0 + h) \\
&= \lim_{h \longrightarrow 0} 2x_0 + \lim_{h \longrightarrow 0} h \\
&= 2x_0 + 0 \\
&= 2x_0.
\end{aligned}$$

Example 5. For the inverse function $\mathtt{inv} : \mathbb{R}^* \to \mathbb{R}$, where

$$\mathtt{inv}(x) = \frac{1}{x},$$

for all $x \in \mathbb{R}^*$, we get

$$\begin{aligned}
\mathtt{inv}'(x_0) &= \lim_{h \longrightarrow 0} \frac{\mathtt{inv}(x_0 + h) - \mathtt{inv}(x_0)}{h} \\
&= \lim_{h \longrightarrow 0} \frac{\frac{1}{x_0 + h} - \frac{1}{x_0}}{h} \\
&= \lim_{h \longrightarrow 0} \frac{\frac{x_0 - x_0 - h}{(x_0 + h)x_0}}{h} \\
&= \lim_{h \longrightarrow 0} \frac{-h}{h \cdot x_0(x_0 + h)} \\
&= \lim_{h \longrightarrow 0} \frac{-1}{x_0(x_0 + h)} \\
&= -\frac{1}{\lim_{h \longrightarrow 0} x_0(x_0 + h)} \\
&= -\frac{1}{\lim_{h \longrightarrow 0} (x_0^2 + x_0 h)} \\
&= -\frac{1}{\lim_{h \longrightarrow 0} x_0^2 + \lim_{h \longrightarrow 0} x_0 h} \\
&= -\frac{1}{x_0^2 + 0}
\end{aligned}$$

$$= -\frac{1}{x_0^2}.$$

Example 6. For the exponential function $\exp : \mathbb{R} \to \mathbb{R}$ we get (Exercise)

$$\exp'(x_0) = x_0,$$

for all $x_0 \in \mathbb{R}$.

Example 7. For the sinus function $\sin : \mathbb{R} \to \mathbb{R}$ we get (Exercise)

$$\sin'(x_0) = \cos(x_0),$$

for all $x_0 \in \mathbb{R}$.

Example 8. For the cosinus function $\cos : \mathbb{R} \to \mathbb{R}$ we get (Exercise)

$$\cos'(x_0) = -\sin(x_0),$$

for all $x_0 \in \mathbb{R}$.

Example 9. The absolute-value function $|.| : \mathbb{R} \to \mathbb{R}$, where $|.|(x) = |x|$, for all $x \in \mathbb{R}$, is not differentiable at $x_0 = 0$. Suppose that

$$\lim_{h \longrightarrow 0} \frac{|x_0 + h| - |x_0|}{h} = l \in \mathbb{R}.$$

Let the following sequences of real numbers:

$$\alpha_n = \frac{1}{n+1}, \quad \beta_n = -\frac{1}{n+1}, \qquad n \in \mathbb{N}.$$

We get

$$\begin{aligned}
1 &= \lim_{n \longrightarrow \infty} \frac{\frac{1}{n+1}}{\frac{1}{n+1}} \\
&= \lim_{n \longrightarrow \infty} \frac{\left|0 + \frac{1}{n+1}\right| - |0|}{\frac{1}{n+1}} \\
&= \lim_{n \longrightarrow \infty} \frac{|0 + \alpha_n| - |\alpha_n|}{\alpha_n} \\
&= l \\
&= \lim_{n \longrightarrow \infty} \frac{|0 + \beta_n| - |\beta_n|}{\beta_n} \\
&= \lim_{n \longrightarrow \infty} \frac{\left|0 - \frac{1}{n+1}\right| - |0|}{-\frac{1}{n+1}} \\
&= \lim_{n \longrightarrow \infty} -\frac{\frac{1}{n+1}}{\frac{1}{n+1}} \\
&= -1.
\end{aligned}$$

PROPOSITION 2.4.2. *If the function $f : D \to \mathbb{R}$ is differentiable at $x_0 \in D$, then f is continuous at x_0.*

PROOF. It suffices to show (Exercise 2(i), Sheet 10)

$$\lim_{h\longrightarrow 0} f(x_0+h) = f(x_0)$$

This follows from the existence of $f'(x_0)$ and the equality

$$f(x_0+h) - f(x_0) = \left[\frac{f(x_0+h)-f(x_0)}{h}\right]h,$$

where $h \neq 0$. □

PROPOSITION 2.4.3. *Let $f, g : D \to \mathbb{R}$ be differentiable functions at $x_0 \in D$, and $\lambda \in \mathbb{R}$. Then the functions*

$$f+g, \lambda f, f\cdot g : D \to \mathbb{R}$$

are also differentiable at x_0, and the following rules hold:

$$(f+g)'(x_0) = f'(x_0) + g'(x_0),$$
$$(\lambda f)'(x_0) = \lambda f'(x_0),$$
$$(f\cdot g)'(x_0) = f'(x_0)\cdot g(x_0) + f(x_0)\cdot g'(x_0).$$

If $g(x) \neq 0$, for every $x \in D$, then the function

$$\frac{f}{g} : D \to \mathbb{R}$$

is also differentiable at x_0 with

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)\cdot g(x_0) - f(x_0)\cdot g'(x_0)}{g(x_0)^2}.$$

PROOF. We use the following equalities:

$$\frac{(f+g)(x_0+h)-(f+g)(x_0)}{h} = \frac{f(x_0+h)-f(x_0)}{h} + \frac{g(x_0+h)-g(x_0)}{h},$$

$$\frac{(\lambda f)(x_0+h)-(\lambda f)(x_0)}{h} = \lambda\cdot\frac{f(x_0+h)-f(x_0)}{h},$$

$$\begin{aligned}
&\frac{(fg)(x_0+h)-(fg)(x_0)}{h}\\
&= \frac{f(x_0+h)g(x_0+h)-f(x_0)g(x_0+h)+f(x_0)g(x_0+h)-f(x_0)g(x_0)}{h}\\
&= \frac{[f(x_0+h)-f(x_0)]g(x_0+h)+f(x_0)[g(x_0+h)-g(x_0)]}{h}\\
&= \left[\frac{f(x_0+h)-f(x_0)}{h}\right]g(x_0+h)+f(x_0)\left[\frac{g(x_0+h)-g(x_0)}{h}\right].
\end{aligned}$$

If $f(x) = 1$, for every $x \in D$, then

$$\frac{\frac{1}{g(x_0+h)}-\frac{1}{g(x_0)}}{h} = \frac{1}{h}\frac{g(x_0)-g(x_0+h)}{g(x_0)g(x_0+h)}$$

$$= -\frac{1}{g(x_0)g(x_0+h)}\left[\frac{g(x_0+h)-g(x_0)}{h}\right],$$

and

$$\left(\frac{1}{g}\right)'(x_0) = \frac{-g'(x_0)}{g(x_0)^2}.$$

The general case follows from the product-rule:

$$\begin{aligned}\left(\frac{f}{g}\right)'(x_0) &= \left(f\cdot\frac{1}{g}\right)'(x_0)\\ &= f'(x_0)\cdot\frac{1}{g(x_0)} + f(x_0)\cdot\left(\frac{1}{g}\right)'(x_0)\\ &= f'(x_0)\cdot\frac{1}{g(x_0)} + f(x_0)\cdot\frac{-g'(x_0)}{g(x_0)^2}\\ &= \frac{f'(x_0)\cdot g(x_0) - f(x_0)\cdot g'(x_0)}{g(x_0)^2}.\end{aligned}$$

□

Example 10. Let $n \in \mathbb{N}^+$ and let $f_n : \mathbb{R} \to \mathbb{R}$ defined by $f_n(x) = x^n$, for every $x \in \mathbb{R}$. Then

$$f_n{}'(x_0) = nx_0^{n-1},$$

for every $x_0 \in \mathbb{R}$. If $n = 1$, then

$$f_1{}'(x_0) = \mathrm{id}_{\mathbb{R}}{}'(x_0) = 1 = 1f_1(x_0)^{1-1}.$$

For the induction-step, since

$$f_{n+1}(x) = x^{n+1} = x^n x = f_n(x)\mathrm{id}_{\mathbb{R}}(x),$$

from the product-rule we have that

$$\begin{aligned}f_{n+1}{}'(x_0) &= \big(f_n\cdot\mathrm{id}_{\mathbb{R}}\big)'(x_0)\\ &= f_n{}'(x_0)\mathrm{id}_{\mathbb{R}}(x_0) + f_n(x_0)\mathrm{id}_{\mathbb{R}}{}'(x_0)\\ &= f_n{}'(x_0)x_0 + f_n(x_0)1\\ &\overset{(I.H.)}{=} nx_0^{n-1}x_0 + x_0^n\\ &= nx_0^n + x_0^n\\ &= (n+1)x_0^n.\end{aligned}$$

Corollary 2.4.4. *Let $f, g : D \to \mathbb{R}$ be n-times differentiable functions at $x_0 \in D$, and $\lambda \in \mathbb{R}$. If $f^{(n)}(x_0)$ denotes the nth-derivative of f at x_0, where $f^{(0)} = f$, then $f+g, \lambda f, f\cdot g : D \to \mathbb{R}$ are n-times differentiable functions at x_0 and the following equalities hold:*

$$(f+g)^{(n)}(x_0) = f^{(n)}(x_0) + g^{(n)}(x_0),$$

$$(\lambda f)^{(n)}(x_0) = \lambda f^{(n)}(x_0),$$

$$(f \cdot g)^{(n)}(x_0) = \sum_{k=0}^{n} \binom{n}{k} f^{(n-k)}(x_0) \cdot g^{(k)}(x_0).$$

Proof. Exercise. □

Proposition 2.4.5 (Derivative of the inverse function). *Let $D \subseteq \mathbb{R}$ be a non-trivial interval of $\mathbb{R}$ (i.e., D has more than one points), $f : D \to \mathbb{R}$ a continuous and strictly monotone function and $g = f^{-1} : f(D) \to \mathbb{R}$ its inverse function.*

$$D \xrightarrow{\ f\ } f(D) \xrightarrow{\ g\ } D.$$
$$g \circ f = \mathrm{id}_D$$

If f is differentiable at $x_0 \in D$ with $f'(x_0) \neq 0$, then g is differentiable at $y_0 = f(x_0)$ with

$$g'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(g(y_0))}.$$

Proof. Let $(\beta_n)_{n\in\mathbb{N}} \subseteq f(D) \setminus \{y_0\}$ such that $\beta_n \xrightarrow{n} y_0$. If $\alpha_n = g(\beta_n)$, for every $n \in \mathbb{N}$, then by the continuity of g at y_0 we get $\alpha_n \xrightarrow{n} x_0$. Notice that by the injectivity of g we have that $\beta_n \neq y_0 \Rightarrow \alpha_n \neq x_0$, for every $n \in \mathbb{N}$. Hence

$$\begin{aligned} g'(y_0) &= \lim_{n\longrightarrow\infty} \frac{g(\beta_n) - g(y_0)}{\beta_n - y_0} \\ &= \lim_{n\longrightarrow\infty} \frac{\alpha_n - x_0}{f(\alpha_n) - f(x_0)} \\ &= \lim_{n\longrightarrow\infty} \frac{1}{\frac{f(\alpha_n)-f(x_0)}{\alpha_n - x_0}} \\ &= \frac{1}{\lim_{n\longrightarrow\infty} \frac{f(\alpha_n)-f(x_0)}{\alpha_n - x_0}} \\ &= \frac{1}{f'(x_0)}. \end{aligned}$$

□

In the above proof we used the equality

$$f'(x_0) = \lim_{x\longrightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0},$$

for the proof of which we work as in the solution of the Exercise 2(i), Sheet 10.

Example 11. The function $\ln : \mathbb{R}^{+*} \to \mathbb{R}$, where $x \mapsto \ln(x)$, is the inverse function of the function $\exp : \mathbb{R} \to \mathbb{R}$. From the previous proposition we get

$$\ln{}'(x_0) = \frac{1}{\exp{}'(\ln(x_0))} = \frac{1}{\exp(\ln(x_0))} = \frac{1}{x_0}.$$

PROPOSITION 2.4.6 (Chain-rule). *Let $f : D \to \mathbb{R}$ and $g : E \to \mathbb{R}$ such that $f(D) \subseteq E$*

$$D \xrightarrow{\;f\;} f(D) \subseteq E \xrightarrow{\;g\;} \mathbb{R}.$$

$$g \circ f$$

If f is differentiable at $x_0 \in D$ and g is differentiable at $y_0 = f(x_0) \in E$, then the composite function $g \circ f : D \to \mathbb{R}$ is differentiable at x_0 with

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0).$$

PROOF. Let the function $h : E \to \mathbb{R}$ be defined by

$$h(e) := \begin{cases} \frac{g(e)-g(y_0)}{e-y_0} & , e \neq y_0 \\ g'(y_0) & , e = y_0. \end{cases}$$

Since g is differentiable at y_0, we get

$$\lim_{e \longrightarrow y_0} h(e) = g'(y_0) = h(y_0)$$

i.e., h is continuous at y_0. Moreover, we have that

$$\forall_{e \in E}\big(g(e) - g(y_0) = h(e)(e - y_0)\big).$$

If $e \neq y_0$, then we use the definition of $h(e)$, while if $e = y_0$, both therms of the equality are 0. Hence

$$\begin{aligned}
(g \circ f)'(x_0) &= \lim_{x \longrightarrow x_0} \frac{g(f(x)) - g(f(x_0))}{x - x_0} \\
&= \lim_{x \longrightarrow x_0} \frac{h(f(x))\big[f(x) - f(x_0)\big]}{x - x_0} \\
&= \lim_{x \longrightarrow x_0} h(f(x)) \lim_{x \longrightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \\
&= h(f(x_0))g'(x_0) \\
&= h(y_0)g'(x_0) \\
&= g'(f(x_0))g'(x_0).
\end{aligned}$$

□

Example 12. Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable in $\mathbb{R}$ and let the function $g : \mathbb{R} \to \mathbb{R}$ be defined by

$$g(x) = f(2019x + 2020),$$

for every $x \in \mathbb{R}$. Then

$$g'(x_0) = 2019 f'(2019x_0 + 2020).$$

Example 13. Let $g : \mathbb{R} \to \mathbb{R}$ be defined by

$$g(x) = \sin^2(x),$$

for every $x \in \mathbb{R}$. I.e., $h = \texttt{sq} \circ \sin$. Hence

$$g'(x_0) = 2\sin(x_0)\sin{}'(x_0) = 2\sin(x_0)\cos(x_0).$$

Example 14. Let $h : \mathbb{R} \to \mathbb{R}$ be defined by

$$h(x) = \cos^2(x),$$

for every $x \in \mathbb{R}$. I.e., $h = \texttt{sq} \circ \cos$. Hence

$$h'(x_0) = 2\cos(x_0)\cos{}'(x_0) = 2\cos(x_0)[-\sin(x_0)] = -2\sin(x_0)\cos(x_0).$$

Example 15. Let $a \in \mathbb{R}$ and $f : \mathbb{R}^{+*} \to \mathbb{R}$ be defined by

$$f(x) = x^a,$$

for every $x \in \mathbb{R}$. Then one can show (Exercise) that

$$f'(x_0) = a x_0^{a-1},$$

for every $x_0 \in \mathbb{R}^{+*}$.

Let $f : C \to \mathbb{R}$, $g : D \to \mathbb{R}$, and $h : E \to \mathbb{R}$ such that $f(C) \subseteq D$ and $F(D) \subseteq E$

$$C \xrightarrow{f} f(C) \subseteq D \xrightarrow{g} g(D) \subseteq E \xrightarrow{h} \mathbb{R}.$$

$$h \circ g \circ f$$

If f is differentiable at $x_0 \in E$, g is differentiable at $y_0 = f(x_0) \in E$, and h is differentiable at $z_0 = g(y_0)$, then one can show (Exercise) that the composite function $h \circ g \circ f : C \to \mathbb{R}$ is differentiable at x_0 with

$$(h \circ g \circ f)'(x_0) = h'(g(f(x_0))) \cdot g'(f(x_0)) \cdot f'(x_0).$$

2.5. Some geometric properties of the derivative

Definition 2.5.1. A function $f : [a, b] \to \mathbb{R}$ has a *local maximum* at $\xi \in [a, b]$, if there is $\varepsilon > 0$ such that

$$\forall_{x\in[a,b]}\big(|x - \xi| < \varepsilon \Rightarrow f(x) \leq f(\xi)\big),$$

while f has a *local minimum* at $\xi \in [a, b]$, if there is $\varepsilon > 0$ such that

$$\forall_{x\in[a,b]}\big(|x - \xi| < \varepsilon \Rightarrow f(x) \geq f(\xi)\big).$$

A function $f : [a, b] \to \mathbb{R}$ has a *local extremum* at $\xi \in [a, b]$, if f has a local maximum at ξ or f has a local minimum at ξ.

A constant function has a (local) maximum [and a (local) minimum] at every point of its domain. Clearly, a local minimum (maximum) may not be a (global) minimum (maximum).

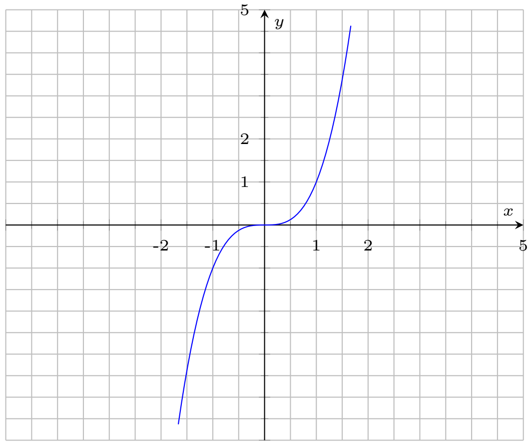
Proposition 2.5.2. *Let $f : (a, b) \to \mathbb{R}$ and $\xi \in [a, b]$ such that f has a local extremum at ξ and f is differentiable at ξ. Then $f'(\xi) = 0$.*

Proof. We suppose that f has a local maximum at ξ and for the case of a local minimum we proceed similarly. Let $\varepsilon > 0$ such that $f(x) \leq f(\xi)$, for every $x \in [a, b]$ with $|x - \xi| < \varepsilon$. We have that

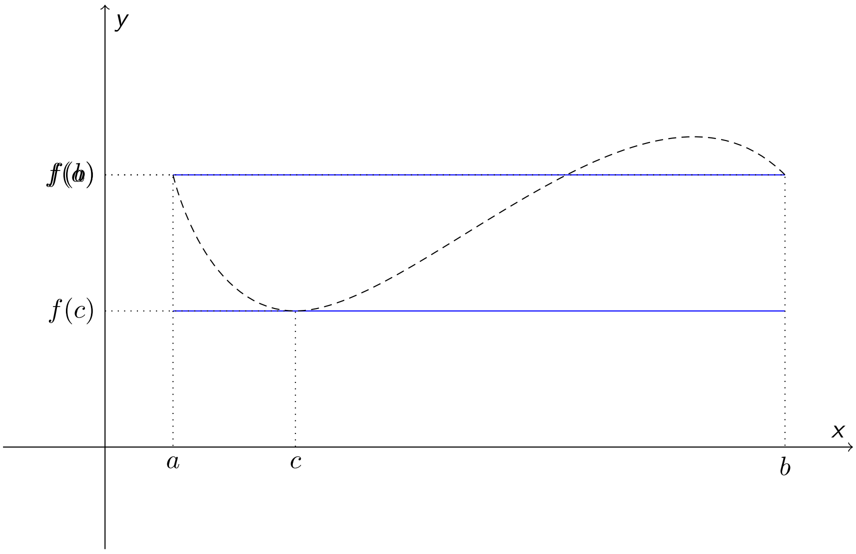
$$\begin{aligned} f'(\xi) &= \lim_{h\longrightarrow 0} \frac{f(\xi + h) - f(\xi)}{h} \\ &= \lim_{h\longrightarrow 0^+} \frac{f(\xi + h) - f(\xi)}{h} := {f'}_+(\xi) \\ &= \lim_{h\longrightarrow 0^-} \frac{f(\xi + h) - f(\xi)}{h} := {f'}_+(\xi). \end{aligned}$$

Since for appropriately small h we have that $f(\xi + h) - f(\xi) \leq 0$, if $h > 0$, then $\frac{f(\xi+h)-f(\xi)}{h} \leq 0$, hence ${f'}_+(\xi) \leq 0$, while if $h < 0$, then $\frac{f(\xi+h)-f(\xi)}{h} \geq 0$, hence ${f'}_-(\xi) \geq 0$. Consequently, $f'(\xi) = 0$. □

If a differentiable function f at ξ satisfies $f'(\xi) = 0$, this does not imply, in general, that f has a local extremum at ξ. E.g., if $f : \mathbb{R} \to \mathbb{R}$ is defined by $f(x) = x^3$, for every $x \in \mathbb{R}$, then $f'(0) = 0$, while f has not a local extremum at 0.



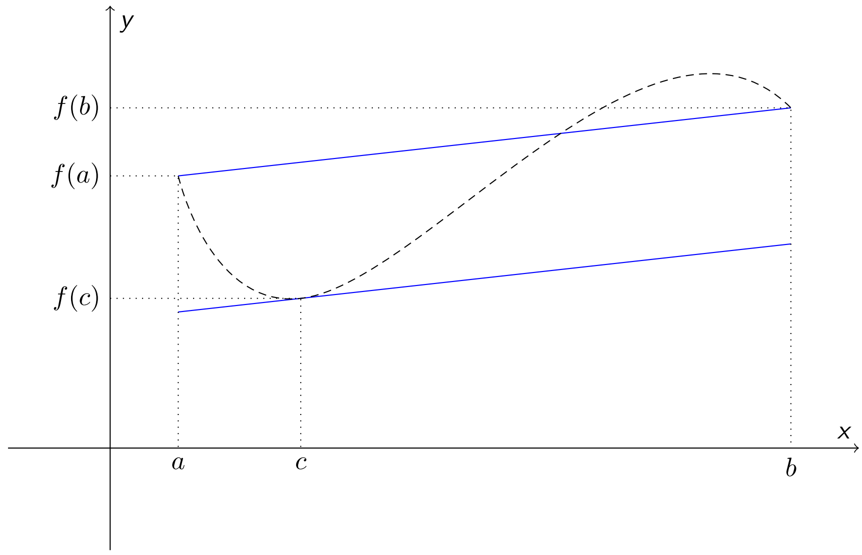
PROPOSITION 2.5.3 (Rolle's theorem). *Let* $f : [a, b] \to \mathbb{R}$, *where* $a < b$, *a continuous function with* $f(a) = f(b)$. *Let also* f *be differentiable in the open interval* (a, b). *Then there is* $c \in (a, b)$ *such that* $f'(c) = 0$.



PROOF. If f is constant, then we can take as c any element of (a, b). If f is not constant, then there is $x_0 \in (a, b)$ with $f(x_0) > f(a)$ or $f(x_0) < f(a)$. Let $f(x_0) > f(a)$ is the case. Since f is a continuous function on $[a, b]$, it has a global minimum at some $\xi \in [a, b]$. Since $x_0 \in (a, b)$, we get $\xi \in (a, b)$. By Proposition 2.5.2 we have that $f'(\xi) = 0$. We proceed similarly, if $f(x_0) < f(a)$. $\square$

Corollary 2.5.4 (Mean value theorem). *Let* $f : [a, b] \to \mathbb{R}$, *where* $a < b$, *a continuous function, which is also differentiable in the open interval* (a, b). *Then there is* $c \in (a, b)$ *such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$



Proof. Let the function $F : [a, b] \to \mathbb{R}$ defined by

$$F(x) = f(x) - \left[\frac{f(b) - f(a)}{b - a}\right](x - a).$$

for every $x \in [a, b]$. Clearly, F is continuous on $[a, b]$ and differentiable in (a, b) with

$$F'(x) = f'(x) - \frac{f(b) - f(a)}{b - a},$$

for every $x \in (a, b)$. Moreover, $F(a) = f(a) = F(b)$, hence by Rolle's theorem there is $c \in (a, b)$ such that $F'(c) = 0$. By the above formula for $F'(x)$ we get

$$F'(c) = 0 \Leftrightarrow f'(c) = \frac{f(b) - f(a)}{b - a}.$$

□

The geometric meaning of the mean value theorem is that there is a point $(c, f(c))$ in the graph of f such that the line tangent at $(c, f(c))$ is parallel to the segment from $(a, f(a))$ to $(b, f(b))$. Notice that if f satisfies the hypotheses of the mean value theorem and $f(a) = f(b)$, then Rolle's theorem follows from the mean value theorem.

COROLLARY 2.5.5. *Let $f : [a,b] \to \mathbb{R}$, where $a < b$, a continuous function, which is also differentiable in the open interval (a,b). Let $m, M \in \mathbb{R}$ such that*

$$\forall_{x \in (a,b)} \big(m \leq f'(x) \leq M\big).$$

Then for every $x_1, x_2 \in [a,b]$ with $x_1 \leq x_2$ we have that

$$m(x_2 - x_1) \leq f(x_2) - f(x_1) \leq M(x_2 - x_1).$$

PROOF. If $x_1 = x_2$, then all terms in the required inequalities are 0, hence equal to each other. Let $x_1 < x_2$. Since the restriction $f_{|[x_1,x_2]}$ of f to the subinterval $[x_1, x_2]$ of $[a,b]$ is continuous on $[x_1, x_2]$ and differentiable in (x_1, x_2), by the mean value theorem there is $c \in (x_1, x_2)$ such that

$$m \leq f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq M,$$

and what we want to show follows now immediately. □

COROLLARY 2.5.6. *Let $f : [a,b] \to \mathbb{R}$, where $a < b$, a continuous function, which is also differentiable in the open interval (a,b). If $f'(x) = 0$, for every $x \in (a,b)$, then f is constant on $[a,b]$.*

PROOF. By our hypothesis we have that

$$\forall_{x \in (a,b)} \big(0 \leq f'(x) \leq 0\big).$$

Let $x_1, x_2 \in [a,b]$. By the previous corollary we get

$$0 = 0(x_2 - x_1) \leq f(x_2) - f(x_1) \leq 0(x_2 - x_1) = 0,$$

hence $f(x_1) = f(x_2)$. □

2.6. The Riemann integral

DEFINITION 2.6.1. A function $\phi : [a,b] \to \mathbb{R}$, where $a < b$, is called a *step-function*, if there is a *partition* (*Unterteilung*)

$$a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$$

of the interval $[a,b]$, such that ϕ is constant in every sub-interval (x_{i-1}, x_i), where $i \in \{1, \ldots, n\}$. Let $\phi(x) := c_i$, for every $x \in (x_{i-1}, x_i)$. The values of ϕ at the points $x_0, x_1, \ldots, x_n$ of the partition are arbitrary real numbers. Let $\mathcal{T}[a,b]$ the set of all step-function $\phi : [a,b] \to \mathbb{R}$. The *integral* $\int_a^b \phi(x)dx$ of a step-function $\phi \in \mathcal{T}[a,b]$ is define by

$$\int_a^b \phi(x)dx = \sum_{i=1}^{n} c_i(x_i - x_{i-1}).$$

PROPOSITION 2.6.2. *The integral $\int_a^b \phi(x)dx$ of a step-function $\phi \in \mathcal{T}[a,b]$ is independent from the partition of $[a,b]$.*

Proof. Let the following partitions of $[a,b]$:

$$(P): \qquad a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b,$$

$$(Q): \qquad a = y_0 < y_1 < \ldots < y_{m-1} < y_m = b,$$

and let

$$\phi(x) = c_i, \quad x \in (x_{i-1}, x_i), \quad i \in \{1,\ldots,n\},$$

$$\phi(y) = d_j, \quad y \in (y_{j-1}, y_j), \quad j \in \{1,\ldots,m\}.$$

We show that

$$\sum_{i=1}^{n} c_i(x_i - x_{i-1}) := \int_P \phi(x)dx = \int_Q \phi(y)dy := \sum_{j=1}^{m} d_j(y_j - y_{j-1}).$$

We suppose first that

$$P \leq Q :\Leftrightarrow \forall_{i\in\{1,\ldots,n\}} \exists_{k:\{1,\ldots,n\}\to\{1,\ldots,m\}} \big(x_i = y_{k_i}\big).$$

In this case we have that

$$x_{i-1} = y_{k_{i-1}} < y_{k_{i-1}+1} < \ldots < y_{k_i} = x_i,$$

and

$$d_j = c_i, \quad \text{for every } j \text{ with } k_{i-1} < j < k_i.$$

Then we get

$$\begin{aligned}
\int_Q \phi(y)dy &= \sum_{j=1}^{m} d_j(y_j - y_{j-1}) \\
&= \sum_{i=1}^{n} \sum_{j=k_{i-1}+1}^{k_i} c_i(y_j - y_{j-1}) \\
&= \sum_{i=1}^{n} c_i(x_i - x_{i-1}) \\
&= \int_P \phi(x)dx.
\end{aligned}$$

Suppose next that P, Q are arbitrary partitions of $[a,b]$. Then $P \cup Q$ is a new partition of $[a,b]$ such that

$$P \leq P \cup Q \ \ \& \ \ Q \leq P \cup Q.$$

By the previous case we get

$$\int_{P\cup Q} \phi(z)dz = \int_P \phi(x)dx \ \ \& \ \ \int_{P\cup Q} \phi(z)dz = \int_Q \phi(x)dx,$$

hence $\int_P \phi(x)dx = \int_Q \phi(y)dy$. $\square$

From the geometric point of view, the integral $\int_a^b \phi(x)dx$ of a step-function ϕ on $[a,b]$ is the algebraic sum of the areas between the x-axis and the graph of ϕ.

Let the partition

$$a = x_0 < x_1 = b$$

of the interval $[a,b]$. The constant function $\phi_c : [a,b] \to \mathbb{R}$, where $\phi_c(x) = c$, for every $x \in [a,b]$, is a step-function with

$$\int_a^b \phi_c(x)dx = \sum_{i=1}^{1} c(x_1 - x_0) = c(b-a).$$

If $\phi, \psi \in \mathcal{T}[a,b]$ and $\lambda \in \mathbb{R}$, it is easy to show that

(i) $\phi + \psi \in \mathcal{T}[a,b]$ and

$$\int_a^b \big(\phi + \psi)x)dx = \int_a^b \phi(x)dx + \int_a^b \psi(x)dx,$$

(ii) $\lambda\phi \in \mathcal{T}[a,b]$ and

$$\int_a^b \big(\lambda\phi\big)(x)dx = \lambda \int_a^b \phi(x)dx,$$

(iii)

$$\phi \leq \psi \Rightarrow \int_a^b \phi(x)dx \leq \int_a^b \phi(x)dx,$$

where

$$\phi \leq \psi :\Leftrightarrow \forall_{x \in [a,b]} \big(\phi(x) \leq \psi(x)\big).$$

Let $f : [a,b] \to \mathbb{R}$ an arbitrary bounded function i.e., there are $m, M \in \mathbb{R}$ such that

$$m \leq f(x) \leq M, \qquad x \in [a,b].$$

If $\phi_m \in \mathcal{T}[a,b]$ is the constant function with value m on $[a,b]$ and if $\phi_M \in \mathcal{T}[a,b]$ is the constant function with value M on $[a,b]$, then

$$\phi_m \leq f \leq \phi_M.$$

Let the sets

$$A(f) = \left\{ \int_a^b \phi(x)dx \mid \phi \in \mathcal{T}[a,b] \ \& \ \phi \geq f \right\},$$

$$B(f) = \left\{ \int_a^b \phi(x)dx \mid \phi \in \mathcal{T}[a,b] \ \& \ \phi \leq f \right\}.$$

$A(f)$ is a non-empty subset of $\mathbb{R}$, because $\phi_M \in \mathcal{T}[a,b]$ with $\phi_M \geq f$ and

$$M(b-a) = \int_a^b \phi_M(x)dx \in A(f).$$

$A(f)$ is a *bounded below* (nach unten beschränkte) subset of $\mathbb{R}$, because for every $\int_a^b \phi(x)dx \in A(f)$ we have that

$$\phi \geq f \geq \phi_m \Rightarrow \int_a^b \phi(x)dx \geq \int_a^b \phi_m(x)dx = m(b-a).$$

Similarly, $B(f)$ is a non-empty subset of $\mathbb{R}$, because $\phi_m \in \mathcal{T}[a,b]$ with $\phi_m \leq f$ and

$$m(b-a) = \int_a^b \phi_m(x)dx \in B(f).$$

$B(f)$ is also a *bounded above* (nach oben beschränkte) subset of $\mathbb{R}$, because for every $\int_a^b \phi(x)dx \in B(f)$ we have that

$$\phi \leq f \leq \phi_M \Rightarrow \int_a^b \phi(x)dx \leq \int_a^b \phi_M(x)dx = M(b-a).$$

DEFINITION 2.6.3 (Supremum, Infimum). Let $A \subseteq \mathbb{R}$. A number $s \in \mathbb{R}$ is called *supremum* (*infimum*) of A, if s is the least upper bound (gratest lower bound) of A. The real number s is the *least upper bound* of A, if the following conditions are satisfied:

(i) s is an upper bound of A $(a \in A \Rightarrow a \leq s)$.

(ii) If s' is another upper bound of A, then $s \leq s'$.

Similarly, the real number t is the *gratest lower bound* of A, if the following conditions are satisfied:

(i) t is a lower bound of A $(a \in A \Rightarrow a \geq t)$.

(ii) If t' is another lower bound of A, then $t' \leq t$.

Clearly, the least upper bound (greatest lower bound) of a are uniquely determined. For them we use the notation

$$\sup(A) \quad [\text{bzw. } \inf(A)].$$

For example, we have that

$$\sup(0,1) = 1 \quad \& \quad \inf(0,1) = 0.$$

THEOREM 2.6.4. *A non-empty and bounded above (below) subset $A \subseteq \mathbb{R}$ has a supremum (infimum).*

PROOF. With the use of the Completeness Axiom (see [**1**], pp. 89-90). □

DEFINITION 2.6.5 (Upper-integral, Lower-integral). Let $f : [a,b] \to \mathbb{R}$ be a bounded function. We define

$$\overline{\int_a^b} f(x)dx = \inf A(f) = \inf \left\{ \int_a^b \phi(x)dx \mid \phi \in \mathcal{T}[a,b] \ \& \ \phi \geq f \right\},$$

$$\underline{\int_a^b} f(x)dx = \sup B(f) = \sup \left\{ \int_a^b \phi(x)dx \mid \phi \in \mathcal{T}[a,b] \ \& \ \phi \leq f \right\}.$$

For every step-function $\phi \in \mathcal{T}[a,b]$ we have that (Exercise)

$$\overline{\int_a^b} \phi(x)dx = \underline{\int_a^b} \phi(x)dx = \int_a^b \phi(x)dx.$$

Let the Dirichlet-Function $\texttt{Dir} : [0,1] \to \mathbb{R}$ on $[0,1]$, defined by

$$\texttt{Dir}(x) := \begin{cases} 1 & , x \in \mathbb{Q} \cap [0,1] \\ 0 & , x \in \mathbb{I} \cap [0,1], \end{cases}$$

One can show (Exercise) that

$$\overline{\int_0^1} \texttt{Dir}(x)dx = 1 \quad \& \quad \underline{\int_0^1} \texttt{Dir}(x)dx = 0,$$

hence

$$\overline{\int_0^1} \texttt{Dir}(x)dx \neq \underline{\int_0^1} \texttt{Dir}(x)dx.$$

Definition 2.6.6. A bounded function $f : [a,b] \to \mathbb{R}$ is called *Riemann-integrable*, or simply *integrable*, if

$$\overline{\int_a^b} f(x)dx = \underline{\int_a^b} f(x)dx.$$

In this case we write

$$\int_a^b f(x)dx = \overline{\int_a^b} f(x)dx.$$

A step-function is Riemann-integrable, while the Dirichlet-Function on $[0,1]$ is not.

Proposition 2.6.7. *(i) A continuous function $f : [a,b] \to \mathbb{R}$ is Riemann-integrable.*

(ii) A monotone function $f : [a,b] \to \mathbb{R}$ is Riemann-integrable.

Proof. See [**1**], pp. 198-199. □

Proposition 2.6.8. *Let $f, g : [a,b] \to \mathbb{R}$ be integrable functions and $\lambda \in \mathbb{R}$.*

(i) The function $f + g : [a,b] \to \mathbb{R}$ is integrable and

$$\int_a^b \big(f+g\big)x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx.$$

(ii) The function $\lambda f : [a,b] \to \mathbb{R}$ is integrable and

$$\int_a^b \big(\lambda f\big)x)dx = \lambda \int_a^b f(x)dx.$$

(iii)

$$f \leq g \Rightarrow \int_a^b f(x)dx \leq \int_a^b g(x)dx,$$

where

$$f \leq g :\Leftrightarrow \forall_{x \in [a,b]} \big(f(x) \leq g(x)\big).$$

(iv) The $|f| : [a, b] \to \mathbb{R}$ *is integrable and*

$$\left| \int_a^b f(x)dx \right| \leq \int_a^b |f(x)|dx.$$

PROOF. We use the corresponding properties of the step-functions. For the details see [**1**], pp. 199-201. □

PROPOSITION 2.6.9. *Let* $a < c < b$ *and* $f : [a, c] \to \mathbb{R}$. *Then* f *ist integrable if and only if its restrictions* $f_{|[a,b]}$ *on* $[a, b]$ *and* $f_{|[b,c]}$ *on* $[b, c]$ *are integrable, and*

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx.$$

PROOF. We use the corresponding propertiy of the step-functions. For the details see [**1**], p. 207. □

DEFINITION 2.6.10. Let $a < b$ and $f : [a, b] \to \mathbb{R}$ a bounded function. We define

$$\int_a^a f(x)dx = 0,$$

and

$$\int_b^a f(x)dx = -\int_a^b f(x)dx.$$

2.7. Integration and Differentiation

PROPOSITION 2.7.1. *Let* $f : [a, b] \to \mathbb{R}$ *be a continuous function and* $c \in [a, b]$. *If* $x \in [a, b]$, *let*

$$F(x) = \int_c^x f(t)dt.$$

The function $F : [a, b] \to \mathbb{R}$ *is differentiable and* $F' = f$. *We call* F *the indefinite integral of* f.

PROOF. See [**1**], p. 209. □

DEFINITION 2.7.2. A differentiable function $F : [a, b] \to \mathbb{R}$ is a *primitive* function of $f : [a, b] \to \mathbb{R}$, if $F' = f$.

The indefinite integral of f is a primitive function of f.

PROPOSITION 2.7.3. *Let* $F : [a, b] \to \mathbb{R}$ *be a primitive function of* $f : [a, b] \to \mathbb{R}$. *A function* $G : [a, b] \to \mathbb{R}$ *is a primitive function of* f *if and only if* $F - G$ *is a constant.*

PROOF. (i) Let $F - G = c$, where $c \in \mathbb{R}$. Then $G' = (F - c)' = F' = f$.
(ii) If G is a primitive function of f, then $G' = f = F'$. Hence, $(F - G)' = 0$. By Corollary 2.5.6 we get $F - G$ is constant. □

Theorem 2.7.4 (Fundamental theorem of Differential and Integral Calculus (FTDIC)). *Let $f : [a, b] \to \mathbb{R}$ be a continuous function and F a primitive function of f. Then*

$$\int_a^b f(x)dx = F(b) - F(a).$$

Proof. For every $x \in [a, b]$ let

$$G(x) = \int_a^x f(t)dt.$$

Since F is a primitive function of f, by Proposition 2.7.3 there is $c \in \mathbb{R}$ such that

$$F - G = c.$$

Hence we have that

$$\begin{aligned} F(b) - F(a) &= (G(b) + c) - (G(a) + c) \\ &= G(b) - G(a) \\ &= \int_a^b f(t)dt - \int_a^a f(t)dt \\ &= \int_a^b f(t)dt - 0 \\ &= \int_a^b f(t)dt. \end{aligned}$$

□

We use the notation:

$$F(x)\Big|_a^b := F(b) - F(a).$$

Hence the equality of Theorem 2.7.4 is written also as

$$\int_a^b f(x)dx = F(x)\Big|_a^b.$$

Examples of using (FTDIC):

$$\int_0^1 1dx = \int_0^1 (\mathrm{id}_{\mathbb{R}})'dx = \mathrm{id}_{\mathbb{R}}(x)\Big|_0^1 = \mathrm{id}_{\mathbb{R}}(0) - \mathrm{id}_{\mathbb{R}}(1) = 1 - 0 = 1.$$

$$\int_0^1 xdx = \int_0^1 \left(\frac{1}{2}x^2\right)'dx = \left(\frac{1}{2}x^2\right)\Big|_0^1 = \frac{1}{2}1^2 - \frac{1}{2}0^2 = \frac{1}{2}.$$

$$\int_0^1 x^2dx = \int_0^1 \left(\frac{1}{3}x^3\right)'dx = \left(\frac{1}{3}x^3\right)\Big|_0^1 = \frac{1}{3}1^3 - \frac{1}{3}0^3 = \frac{1}{3}.$$

$$\int_0^1 x^n dx = \int_0^1 \left(\frac{1}{n+1}x^{n+1}\right)' dx = \left(\frac{1}{n+1}x^{n+1}\right)\Big|_0^1 =$$
$$= \frac{1}{n+1}1^{n+1} - \frac{1}{n+1}0^{n+1} = \frac{1}{n+1}.$$
$$\int_1^2 \frac{1}{x} dx = \int_1^2 [\ln(x)]' dx = \ln(x)\Big|_1^2 := \ln(2) - \ln(1) = \ln(2) - 0 = \ln(2).$$

If $x < 0$, then by the chain rule we get

$$[\ln(-x)]' = \frac{1}{-x}(-1) = \frac{1}{x}.$$

Hence

$$\int_{-2}^{-1} \frac{1}{x} dx = \ln(-x)\Big|_{-2}^{-1} := \ln(1) - \ln(2) = 0 - \ln(2) = -\ln(2).$$

We write the two previous cases in one, as follows:

$$\int \frac{dx}{x} = \ln(|x|), \qquad 0 \text{ is not in the interval of the integration.}$$

As an application of (FTDIC) and the chain rule we have that

$$\int_a^b \frac{g'(t)}{g(t)} dt = \ln\big(|g(t)|\big)\Big|_a^b = \ln\big(|g(b)|\big) - \ln\big(|g(a)|\big),$$

where $g : [a, b] \to \mathbb{R}$ is a *continuously differentiable* function i.e., g' is a continuous function (hence the function $\frac{g'(t)}{g(t)}$ is integrable), such that $g(t) \neq 0$, for every $t \in [a, b]$.

PROPOSITION 2.7.5 (Substitution rule). *Let $f : [a', b'] \to \mathbb{R}$ be a continuous function and $g : [a, b] \to [a', b']$ a continuously differentiable function. Then*

$$\int_a^b f(g(t))g'(t)dt = \int_{g(a)}^{g(b)} f(x)dx.$$

PROOF. Let $F : [a', b'] \to \mathbb{R}$ be a primitive function of f. For the composite function $F \circ g : [a, b] \to \mathbb{R}$ the chain rule gives

$$(F \circ g)'(t) = F'(g(t))g'(t) = f(g(t))g'(t).$$

By Theorem 2.7.4 we have that

$$\begin{aligned}\int_a^b f(g(t))g'(t)dt &= (F \circ g)(t)\Big|_a^b \\ &= (F \circ g)(b) - (F \circ g)(a) \\ &= F(g(b)) - F(g(a)) \\ &= \int_{g(a)}^{g(b)} f(x)dx.\end{aligned}$$

□

Example 1: If $f(x) = \frac{1}{x}$, where $x \neq 0$, and $g : [a, b] \to \mathbb{R}$ is as in the last example before Proposition 2.7.5, then we have that

$$\begin{aligned}\int_a^b \frac{g'(t)}{g(t)} dt &= \int_a^b f(g(t))g'(t)dt \\ &= \int_{g(a)}^{g(b)} f(x)dx \\ &= \int_{g(a)}^{g(b)} \frac{1}{x} dx \\ &= \ln\big(|x|\big)\Big|_{g(a)}^{g(b)} \\ &= \ln\big(|g(b)|\big) - \ln\big(|g(a)|\big).\end{aligned}$$

Example 2: If $c \in \mathbb{R}$, then

$$\int_a^b f(t+c)dt = \int_{a+c}^{b+c} f(x)dx.$$

If $g(t) = t + c$, for every $t \in \mathbb{R}$, then $g'(t) = 1$ and

$$\begin{aligned}\int_a^b f(t+c)dt &= \int_a^b f(g(t))g'(t)dt \\ &= \int_{g(a)}^{g(b)} f(x)dx \\ &= \int_{a+c}^{b+c} f(x)dx.\end{aligned}$$

Example 3: If $c \neq 0$, then

$$\int_a^b f(ct)dt = \frac{1}{c}\int_{ac}^{bc} f(x)dx.$$

If $g(t) = ct$, for every $t \in \mathbb{R}$, then $g'(t) = c$ and

$$\begin{aligned}\int_a^b f(ct)dt &= \frac{1}{c}\int_a^b f(g(t))g'(t)dt \\ &= \frac{1}{c}\int_{g(a)}^{g(b)} f(x)dx \\ &= \frac{1}{c}\int_{ac}^{bc} f(x)dx.\end{aligned}$$

PROPOSITION 2.7.6. *Let $f, g : [a, b] \to \mathbb{R}$ be continuously differentiable functions. Then*

$$\int_a^b f(x)g'(x)dx = f(x)g(x)\Big|_a^b - \int_a^b g(x)f'(x)dx.$$

PROOF. If $F = f \cdot g$, then

$$F'(x) = f'(x)g(x) + f(x)g'(x) \Leftrightarrow f(x)g'(x) = F'(x) - f'(x)g(x),$$

for every $x \in [a, b]$. Hence

$$\begin{aligned}\int_a^b f(x))g'(x)dx &= \int_a^b [F'(x) - f'(x)g(x)]dx \\ &= \int_a^b F'(x)dx - \int_a^b f'(x)g(x)dx \\ &= F(x)\Big|_a^b - \int_a^b g(x)f'(x)dx \\ &= f(x)g(x)\Big|_a^b - \int_a^b g(x)f'(x)dx.\end{aligned}$$

□

Example 1. If $a, b > 0$, then

$$\begin{aligned}\int_a^b \ln(x)dx &= \int_a^b \ln(x)x'dx \\ &= x\ln(x)\Big|_a^b - \int_a^b \ln{}'(x)xdx \\ &= x\ln(x)\Big|_a^b - \int_a^b \frac{1}{x}xdx \\ &= x\ln(x)\Big|_a^b - \int_a^b dx \\ &= x\ln(x)\Big|_a^b - x\Big|_a^b \\ &= [x\ln(x) - x]\Big|_a^b \\ &= [x(\ln(x) - 1)]\Big|_a^b.\end{aligned}$$

Example 2. Let the integral

$$I = \int e^x \cos(x)dx.$$

We have that

$$\begin{aligned} I &= \int (e^x)' \cos(x) dx \\ &= e^x \cos(x) - \int e^x [-\sin(x)] dx \\ &= e^x \cos(x) + \int e^x \sin(x) dx \\ &= e^x \cos(x) + J. \end{aligned}$$

Moreover, we have that

$$\begin{aligned} J &= \int e^x \sin(x) dx \\ &= \int (e^x)' \sin(x) dx \\ &= e^x \sin(x) - \int e^x \cos(x) dx \\ &= e^x \sin(x) - I. \end{aligned}$$

Hence

$$I = e^x \cos(x) + e^x \sin(x) - I \Leftrightarrow I = \frac{e^x}{2}[\cos(x) + \sin(x)].$$

Bibliography

[1] O. Forster: *Analysis I*, Springer, 2016.
[2] S. Lang: *Undergraduate Analysis*, Springer, 1983.
[3] S. Lang: *A First Course in Calculus*, Springer, 1986.
[4] H. Schwichtenberg: *Constructive analysis with witnesses*, Lecture notes, LMU, 2019.