Mathematics

Master in Integrated Building Systems

ETH, D-ARCH

Cornelia M. Busch

HS 2023

Contents

Chapter 0

Prerequisites

"Die Mathematiker sind eine Art Franzosen: Redet man zu ihnen, so übersetzen sie es in ihre Sprache, und dann ist es alsobald ganz etwas anderes."

Johann Wolfgang von Goethe

(Mathematicians are like Frenchmen: whatever you say to them they translate into their own language and forthwith it is something entirely different.)

0.1 Functions

0.1.1 Definitions

Whenever we open a mathematics book we see that the content is written in a language in which each word has a well-defined meaning. With the following definitions we introduce some vocabulary that mathematicians use.

Definition 0.1. A *set* is a collection of well defined distinct objects, considered as an object in its own right.

Example. Examples of sets are $\mathbb{Z}$, the set of integers or $\mathbb{R}$, the set of real numbers. The set of natural numbers $\mathbb{N}$ is a subset of $\mathbb{Z}$: we write $\mathbb{N} \subset \mathbb{Z}$. Let $a, b \in \mathbb{R}$, $a < b$, then the *closed interval*

$$[a, b] := \{x \in \mathbb{R} \mid a \leqslant x \leqslant b\}$$

is a subset of $\mathbb{R}$, $[a, b] \subset \mathbb{R}$, and so is the *open interval*

$$(a, b) := \{x \in \mathbb{R} \mid a < x < b\} \subset \mathbb{R}\,.$$

Let S be any set. We write $x \in S$ if x is an *element* of S.

Definition 0.2. A *function* f from a set A to a set B is a rule that defines for every $x \in A$ a unique $y = f(x) \in B$. We write

$$\begin{array}{rccc} f: & A & \to & B \\ & x & \mapsto & y = f(x)\,. \end{array}$$

We call

- $A = \mathrm{dom}(f)$ the *domain* of f,
- B the *codomain* or *range* of f.
- $\mathrm{im}(f) = \{y \in B \mid \exists x \in A \text{ with } f(x) = y\}$ the *image* of f
- $\mathrm{graph}(f) = \{(x, y) \in A \times B \mid y = f(x)\}$ the *graph* of f.

Example. i) Consider the function

$$\begin{array}{rccc} f_1: & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & f_1(x) := x^2\,. \end{array}$$

The function is well-defined since $\forall x \in \mathbb{R}$, $\exists!\, y \in \mathbb{R}$ with $y = x^2$. [1]

- The domain of f_1 is $\mathrm{dom}(f_1) = \mathbb{R}$.
- The range of f_1 is $\mathbb{R}$.
- The image of f_1 is

$$\mathrm{im}(f_1) = \{y \in \mathbb{R} \mid \exists x \in A \text{ with } y = f_1(x) = x^2\} = \mathbb{R}^{\geqslant 0}\,,$$

  where $\mathbb{R}^{\geqslant 0} = \{y \in \mathbb{R} \mid y \geqslant 0\}$. We see that in this example the image and the range of f_1 are different.
- The graph of f_1 is

$$\mathrm{graph}(f_1) = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid y = x^2\}\,,$$

  a parabola.

[1] This is read as follows: for all x in $\mathbb{R}$ there exists a unique $y \in \mathbb{R}$ that satisfies $y = x^2$.

ii) We now change the range in our first example and define

$$\begin{array}{rccl} f_2 : & \mathbb{R} & \to & \mathbb{R}^{\geqslant 0} \\ & x & \mapsto & f_2(x) := x^2 \, . \end{array}$$

Here the range is equal to the image of f_2. The graph of f_2 is

$$\mathrm{graph}(f_2) = \{(x, y) \in \mathbb{R} \times \mathbb{R}^{\geqslant 0} \mid y = x^2\} \, .$$

iii) What is the difference between the following two functions?

$$\begin{array}{rccl} g_1 : & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & g_1(x) := \cos(x) \end{array}$$

$$\begin{array}{rccl} g_2 : & \mathbb{R} & \to & [-1, 1] \\ & x & \mapsto & g_2(x) := \cos(x) \end{array}$$

Both functions are well-defined since the image of the cosine is the interval $[-1, 1]$.

0.1.2 Limits and continuous functions

Definition 0.3. Let $f : \mathrm{dom}(f) \to \mathbb{R}$ be a real function, and $I \subseteq \mathrm{dom}(f)$ an open interval. Let $\overline{I}$ be the union of I with its boundaries. Let $\xi \in \overline{I}$. We say that the *limit of $f(x)$ as x approaches ξ is η* and write

$$\lim_{x \to \xi} f(x) = \eta$$

if the following statement is true. For every $\varepsilon > 0$, $\varepsilon \in \mathbb{R}$, there exists a $\delta > 0$, $\delta \in \mathbb{R}$, such that for every $x \in I$

$$\text{if } 0 < |x - \xi| < \delta \quad \text{then} \quad |f(x) - \eta| < \varepsilon \, .$$

If $\xi \in I \, \mathrm{dom}(f)$, then the function f is *continuous* in ξ if and only if

$$\lim_{x \to \xi} f(x) = f(\xi) \, .$$

Example. i) A real polynomial function of degree $n \in \mathbb{N}$, $0 \leqslant n$, is defined to be

$$\begin{array}{rccl} f : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) := a_n x^n + \ldots + a_1 x + a_0 \, , \end{array}$$

where $a_i \in \mathbb{R}$, $i = 0, \ldots, n$, and $a_n \neq 0$. A real polynomial function is continuous.

ii) The cosine function,

$$\begin{aligned} \cos: \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \cos(x) \end{aligned}$$

the sine function,

$$\begin{aligned} \sin: \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \sin(x) \end{aligned}$$

the exponential function

$$\begin{aligned} \exp: \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \exp(x) = e^x \end{aligned}$$

and the logarithm

$$\begin{aligned} \ln: \mathbb{R}^{>0} &\longrightarrow \mathbb{R} \\ x &\longmapsto \ln(x) \end{aligned}$$

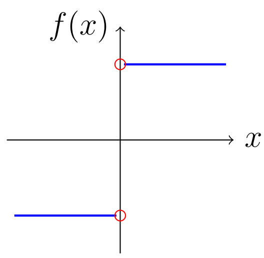
are continuous.

iii) We define

$$\begin{aligned} f: (-\infty, 0) \cup (0, \infty) &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x) \end{aligned}$$

with

$$f(x) = \begin{cases} -1 & \text{for } x < 0 \\ 1 & \text{for } x > 0\,. \end{cases}$$



Then for $x \in (-\infty, 0)$

$$\lim_{x \to 0} f(x) = -1$$

and for $x \in (0, \infty)$

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

This shows that there doesn't exist a continuous function $g: \mathbb{R} \to \mathbb{R}$ with $g(x) = f(x)$ for all $x \in \text{dom}(f)$.

0.2 Differential calculus

0.2.1 Definition

Definition 0.4. Let

$$f: \mathbb{R} \longrightarrow \mathbb{R}$$
$$x \longmapsto f(x)$$

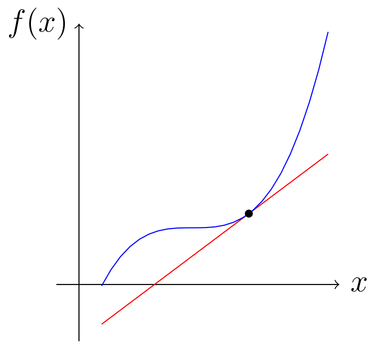
be a function from $\mathbb{R}$ to $\mathbb{R}$. The *derivative* of f is defined to be the limit

$$\lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} =: f'(x)$$

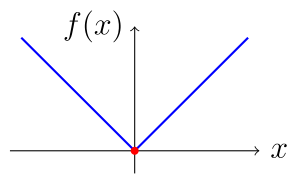
Another notation for f' is

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

The derivative of a function f in x_0 is the slope of the tangent on the graph of f in the point $(x_0, f(x_0))$.



The derivative is not defined for each function f or for every point $(x, f(x))$. Let $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x) := |x|$ with $|x| = -x$ for $x \leqslant 0$ and $|x| = x$ for $0 \leqslant x$.



Then the derivative is -1 for $x < 0$ and 1 for $0 < x$, but the derivative is not defined in $x = 0$.
The derivative measures how the value of the function changes in the neighbourhood of x.

0.2.2 Rules and examples

Some rules are well-known. Let $f : \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R} \to \mathbb{R}$, $h : \mathbb{R} \to \mathbb{R}$ be functions and $a, b \in \mathbb{R}$. We assume that the derivatives of f, g and h are defined.
The derivative of a function f in x gives the slope of the tangent to the curve $\big(x, f(x)\big)$.

Properties		
sum	$(f+g)' = f' + g'$	
constant factor	$(\lambda f)' = \lambda f'$	
product	$(fg)' = f'g + fg'$	
quotient	$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$	
chain rule	$(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$	$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}$
inverse	$g'(y) = \frac{1}{f'(g(y))}$	$\frac{dx}{dy} = \left(\frac{dy}{dx}\right)^{-1}$

Exponential functions, logarithm		
$f(x)$	$f'(x)$	Condition
c	0	c is a constant
x^n	nx^{n-1}	$n \in \mathbb{Z}$ and $x \neq 0$ if $n < 0$
x^a	ax^{a-1}	$a \in \mathbb{R}$ and $x > 0$
e^x	e^x	
a^x	$a^x \cdot \ln a$	$a > 0$
$\ln x$	$\frac{1}{x}$	$x > 0$

Trigonometric and hyperbolic functions			
$f(x)$	$f'(x)$	$f(x)$	$f'(x)$
$\sin x$	$\cos x$	$\sinh x$	$\cosh x$
$\cos x$	$-\sin x$	$\cosh x$	$\sinh x$
$\tan x$	$\frac{1}{\cos^2 x}$	$\tanh x$	$\frac{1}{\cosh^2 x}$
$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$	$\operatorname{arsinh} x$	$\frac{1}{\sqrt{1+x^2}}$
$\arccos x$	$\frac{-1}{\sqrt{1-x^2}}$	$\operatorname{arcosh} x$	$\frac{1}{\sqrt{x^2-1}}$
$\arctan x$	$\frac{1}{1+x^2}$	$\operatorname{artanh} x$	$\frac{1}{1-x^2}$

Example. i) Sum and scalar multiplication

We compute the derivative of $f(x) + \lambda g(x)$ with $f(x) = x^2$, $g(x) = \sin(x)$ and $\lambda = 3$.

$$\begin{aligned}\big(x^2 + 3\sin(x)\big)' &= \big(x^2\big)' + \big(3\sin(x)\big)' \\ &= \big(x^2\big)' + 3\big(\sin(x)\big)' \\ &= 2x + 3\cos(x)\end{aligned}$$

ii) Product rule

We compute the derivative

$$\big(f \cdot g\big)' = f'g + fg'$$

for

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

$$\begin{aligned}\big(x^2\sin(x)\big)' &= \big(x^2\big)'\sin(x) + x^2\big(\sin(x)\big)' \\ &= 2x\sin(x) + x^2\cos(x)\end{aligned}$$

- $f(x) = \cos(x)$, $g(x) = e^x$

$$\begin{aligned}\big(\cos(x)\cdot e^x\big)' &= \big(\cos(x)\big)'\cdot e^x + \cos(x)\cdot\big(e^x\big)' \\ &= -\sin(x)\cdot e^x + \cos(x)\cdot e^x \\ &= \big(\cos(x) - \sin(x)\big)e^x\end{aligned}$$

iii) Quotient rule

We compute the derivative

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

for

- $f(x) = x^2$, $g(x) = \cos(x)$

$$\begin{aligned}\left(\frac{x^2}{\cos(x)}\right)' &= \frac{(x^2)'\cos(x) - x^2(\cos(x))'}{(\cos(x))^2} \\ &= \frac{2x\cos(x) + x^2\sin(x)}{\cos^2(x)} \\ &= \frac{2x}{\cos(x)} + \frac{x^2}{\cos(x)}\tan(x)\end{aligned}$$

- $f(x) = \sin(x)$, $g(x) = x$

$$\begin{aligned}\left(\frac{\sin(x)}{\cos(x)}\right)' &= \frac{(\sin(x))'\cdot\cos(x) - \sin(x)\cdot(\cos(x))'}{(\cos(x))^2} \\ &= \frac{\cos(x)\cdot\cos(x) - \sin(x)\cdot(-\sin(x))}{\cos^2(x)} \\ &= \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} \\ &= \frac{1}{\cos^2(x)} = 1 + \tan^2(x)\end{aligned}$$

Since $\tan(x) = \frac{\sin(x)}{\cos(x)}$, this result is the derivative of the tangent function.

iv) Chain rule

We compute the derivative of the composition $(f\circ g)(x) = f\big(g(x)\big)$ of two functions

$$(f\circ g)'(x) = f'\big(g(x)\big)\cdot g'(x)$$

for

- $f(x) = \sin(x)$, $g(x) = x^2$. In this case

$$(f \circ g)(x) = f\big(g(x)\big) = \sin(x^2)$$

and

$$\begin{aligned}(f \circ g)'(x) &= \big(\sin(x^2)\big)' = \sin'(x^2) \cdot (x^2)' \\ &= \cos(x^2) \cdot (2x) \\ &= 2x\cos(x^2)\end{aligned}$$

Exchanging the order of the composition, we get a different function:

$$(g \circ f)(x) = g\big(f(x)\big) = \big(\sin(x)\big)^2 = \sin^2(x)$$

and

$$\begin{aligned}(g \circ f)'(x) &= \big(\sin^2(x)\big)' \\ &= 2\sin(x)\big(\sin(x)\big)' \\ &= 2\sin(x)\cos(x)\end{aligned}$$

- $f(x) = e^x$, $g(x) = x^2$. In this case

$$(f \circ g)(x) = f\big(g(x)\big) = e^{x^2}$$

and

$$\begin{aligned}(f \circ g)'(x) &= \big(e^{x^2}\big)' \\ &= e^{x^2}\big(x^2\big)' = e^{x^2} \cdot 2x \\ &= 2xe^{x^2}\end{aligned}$$

Exchanging the inner and the outer function, we get

$$(g \circ f)(x) = g\big(f(x)\big) = (e^x)^2 = e^{2x}$$

and

$$\begin{aligned}(g \circ f)'(x) &= \big(e^{2x}\big)' \\ &= e^{2x}\big(2x\big)' = e^{2x} \cdot 2 \\ &= 2e^{2x}\end{aligned}$$

0.2.3 Extrema

Definition 0.5. Let $U \subseteq \mathbb{R}$ be a subset of the real numbers, $f : U \longrightarrow \mathbb{R}$ a function and $x_0 \in U$. Then the following hold.

- f has a *local minimum* in $x_0 \in U$ if there is an interval $I = (a, b) \subset \mathbb{R}$ with $x_0 \in I$ and $f(x_0) \leqslant f(x)$ for all $x \in I \cap U$.
- f has a *global minimum* in $x_0 \in U$ if $f(x_0) \leqslant f(x)$ for all $x \in U$.
- f has a *local maximum* in $x_0 \in U$ if there is an interval $I = (a, b) \subset \mathbb{R}$ with $x_0 \in I$ and $f(x_0) \geqslant f(x)$ for all $x \in I \cap U$.
- f has a *global maximum* in $x_0 \in U$ if $f(x_0) \geqslant f(x)$ for all $x \in U$.

Proposition 0.6. *Let*

$$f: \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & f(x) \end{array}$$

be a function that is differentiable in $x_0 \in \mathbb{R}$, i.e. the derivative

$$\frac{d}{dx} f(x_0)$$

exists. If

$$\frac{d}{dx} f(x_0) = f'(x_0) = 0$$

and the derivative f' is differentiable in x_0 then

$$\begin{cases} \frac{d^2}{dx^2} f(x_0) = f''(x_0) > 0 & \Rightarrow f(x_0) \text{ is a local minimum of } f \\ \frac{d^2}{dx^2} f(x_0) = f''(x_0) < 0 & \Rightarrow f(x_0) \text{ is a local maximum of } f \end{cases}$$

attained in x_0. No general statement is possible for $f''(x_0) = 0$.

Example. i) The derivative of the function

$$f: \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & f(x) := x^2 \end{array}$$

is $f'(x) = 2x$ and has a zero in $x_0 = 0$, i.e. $f'(0) = 0$. The second derivative is $f''(x) = 2$, hence $f''(0) = 2 > 0$ and $f(0) = 0$ is a local minimum of f. It is attained in $x_0 = 0$. This function has no maximum.

A similar analysis shows that the function

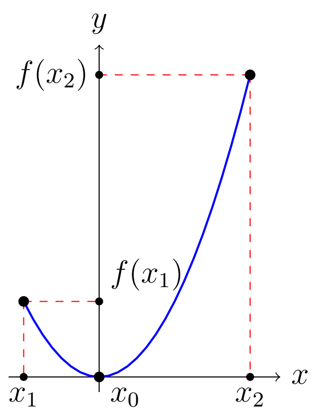
$$\begin{array}{rccl} g: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & g(x) := -x^2 \end{array}$$

has a local maximum $g(0) = 0$ that is attained in $x_0 = 0$. This function has no minimum.

ii) We now restrict the domain to closed intervals. Consider the function

$$\begin{array}{rccl} f: & [-1,2] & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) := x^2 . \end{array}$$

Since $0 \in [-1, 2]$, this function has a local minimum $f(0) = 0$ in $x_0 = 0$. We now have to evaluate f in the boundaries $x_1 = -1$ and $x_2 = 2$ of the interval $[-1, 2]$ and get $f(-1) = 1$, $f(2) = 4$. Hence $f(-1) > f(0)$, $f(0) < f(2)$ and $f(-1) < f(2)$. The function has a global minimum 0 in 0. It has a global maximum $f(2) = 4$ in $x_2 = 2$ and a local maximum $f(-1) = 1$ in $x_1 = -1$.



iii) Consider the functions

$$\begin{array}{rccl} f: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) := x^3 \end{array}$$

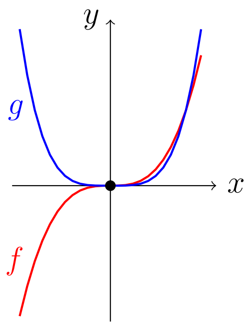
and

$$\begin{array}{rccl} g: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & g(x) := x^4 \end{array}$$

Since they are polynomial functions, f and g are continuous and so are their derivatives. The first and second derivatives of f and g are

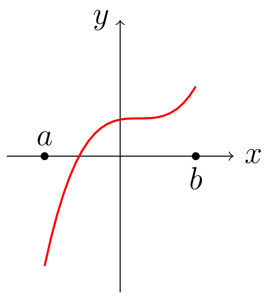
$$\begin{aligned} f'(x) &= 3x^2 , & g'(x) &= 4x^3 , \\ f''(x) &= 6x , & g''(x) &= 12x^2 . \end{aligned}$$

We consider the point $x_0 = 0$, where $f'(0) = 0$ and $g'(0) = 0$. Since $f''(0) = 0$ and $g''(0) = 0$ we have to study those functions and see that f has a saddle point in $x_0 = 0$ and g attains a local minimum 0 in $x_0 = 0$.



0.2.4 Intermediate value theorem

Theorem 0.7. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function with $f(a) < 0$ and $f(b) > 0$ (resp. $f(a) > 0$ and $f(b) < 0$). Then $p \in (a, b)$ exists with $f(p) = 0$.*



The proof uses mathematical induction. Since this is an important method, we show the proof.

Proof. We consider the case with $f(a) < 0$ and $f(b) > 0$. The other one is analogous. This theorem is proved by construction of a sequence of intervals $[a_n, b_n] \subset [a, b]$, $n \in \mathbb{N}$, with

i) $[a_n, b_n] \subset [a_{n-1}, b_{n-1}]$ for $n \geqslant 1$,

ii) $b_n - a_n = 2^{-n}(b - a)$,

iii) $f(a_n) \leqslant 0$, $f(b_n) \geqslant 0$.

We will use mathematical induction to prove that if these properties hold for n, i.e. for the interval $[a_n, b_n]$, then they also hold for $n+1$, i.e. for the interval $[a_{n+1}, b_{n+1}]$. We therefore first show that they hold for the base case $n = 0$ and then do the induction step from n to $n+1$. Then we know that the property holds for every natural n. Base case: We set $[a_0, b_0] := [a, b]$. Induction step: Let $m := (a_n + b_n)/2$ be the middle of the interval $[a_n, b_n]$ and $[a_n, b_n]$ satisfies the three properties i), ii) and iii) stated above. We then distinguish two cases.

Case 1: $f(m) \geqslant 0$. Then let $[a_{n+1}, b_{n+1}] := [a_n, m]$.

Case 2: $f(m) < 0$. Then let $[a_{n+1}, b_{n+1}] := [m, b_n]$.

It is easy to see that $[a_{n+1}, b_{n+1}]$ satisfies the properties i), ii) and iii). It follows that the sequence (a_n) is monotonically increasing and that the sequence (b_n) is monotonically decreasing. Moreover both sequences are bounded. Since bounded monotone sequences converge$^{(*)}$, we get

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n =: p\,.$$

Because of the continuity of f, we have

$$\lim_{n\to\infty} f(a_n) = \lim_{n\to\infty} f(b_n) =: f(p)\,.$$

If a sequence (a_n) satisfies $A \leqslant a_n \leqslant B$ for every $n \in \mathbb{N}$, then$^{(*)}$

$$A \leqslant \lim_{n\to\infty} a_n \leqslant B\,.$$

With this statement and property iii) we get

$$f(p) = \lim f(a_n) \leqslant 0 \text{ and } f(p) = \lim f(b_n) \geqslant 0$$

hence $f(p) = 0$. q.e.d.[2]. □

The method for the proof that the point $p \in [a, b]$ exists is also used for the construction of the zero of a continuous function.

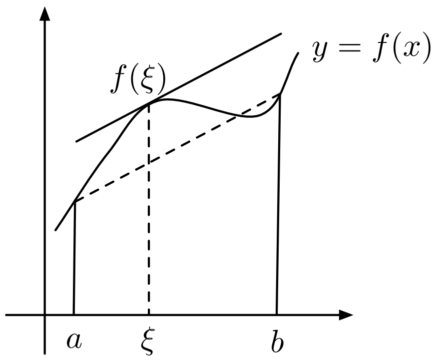
0.2.5 Mean value theorem

The mean value theorem of differential calculus is used for approximating functions.

[2]The statements in this proof that are labeled $^{(*)}$ need to be proved.

Theorem 0.8. *Let $f : [a, b] \to \mathbb{R}$ be continuous and differentiable in the inner of $[a, b]$ (i.e. in (a, b)). Then there is a point $\xi \in (a, b)$ such that*

$$f'(\xi) = \frac{f(b) - f(a)}{b - a} \quad \textit{resp.} \quad f(b) - f(a) = f'(\xi)(b - a)\,.$$



Chapter 1

Calculus

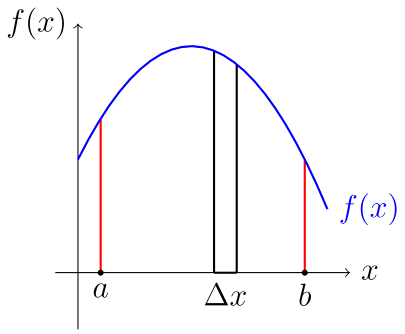
1.1 Integration

1.1.1 Definition

Given a function

$$\begin{array}{rccc} f: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) \end{array}$$

we compute the area between the graph $(x, f(x))$, $x \in [a, b]$, and the interval $[a, b]$, $a < b$.



We first cut the area in very thin stripes of width $\Delta x = \frac{b-a}{n}$ and approximate it with

$$\sum_{k=1}^{n} \Delta x \, f(\widetilde{x}_k) \, , \qquad a + (k-1)\, \Delta x \leqslant \widetilde{x}_k \leqslant a + k \, \Delta x \, .$$

If

$$f(\widetilde{x}_k) = \max\{f(x) \mid a + (k-1)\, \Delta x \leqslant x \leqslant a + k \, \Delta x\}$$

we get an upper sum and if

$$f(\widetilde{x}_k) = \min\{f(x) \mid a + (k-1)\,\Delta x \leqslant x \leqslant a + k\,\Delta x\}$$

we get a lower sum.
We define the *integral of f on the interval* $[a, b]$

$$\int_a^b f(x)\,dx := \lim_{n\to\infty} \sum_{k=1}^{n} \Delta x\, f(\widetilde{x}_k)\,.$$

If the limit exists, then the limit of the upper sum equals the limit of the lower sum.

Definition 1.1. Let $f : \mathrm{dom}(f) \to \mathbb{R}$. If $[a, b] \subseteq \mathrm{dom}(f)$, then the integral

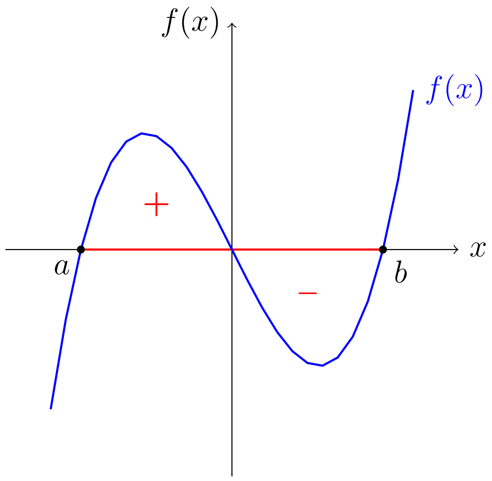
$$\int_a^b f(x)\,dx$$

is the *definite integral* of f on $[a, b]$.

The definite integral yields the area between the graph of f and the x-axis.
Let $[a, b] \in \mathbb{R}$. The integral

$$\int_a^b f(x)\,dx\,, \quad a < b\,,$$

determines the area between the x-axis and the graph of f on the interval $[a, b]$, where the area above the x-axis contributes to the integral with a positive sign and the area below the x-axis contributes with a negative sign.



Note that

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx\,.$$

1.1.2 Main theorem of integration theory

We define the area function

$$F_a : [a, \infty) \to \mathbb{R}$$

of a non-negative function $f : \mathbb{R} \to \mathbb{R}_0^+$ to be

$$F_a(x) = \int_a^x f(\tilde{x}) d\tilde{x} \,,$$

the area between the graph of f and the x-axis on the interval $[a, x]$. This function satisfies $F_a(a) = 0$ and for $b > a$

$$F_b(x) = F_a(x) - F_a(b)$$

It can be shown that

$$F_a'(x) = f(x) \,.$$

Definition 1.2. A function $F : \mathrm{dom}(F) \to \mathbb{R}$, $\mathrm{dom}(F) \subseteq \mathbb{R}$ that satisfies

$$\frac{d}{dx} F(x) = F'(x) = f(x)$$

is called the *antiderivative* of f. Two different antiderivatives of f differ by a constant. The set of all antiderivatives of f is called the *indefinite integral* of f.

$$\begin{aligned} \int f(x)\,dx &:= \{F(x) \mid \tfrac{d}{dx} F(x) = f(x)\} \\ &= \{F(x) + c \mid c \in \mathbb{R},\ F \text{ is an antiderivative of } f\} \,. \end{aligned}$$

Theorem 1.3. *Given a function $f : \mathbb{R} \to \mathbb{R}$ with antiderivative*

$$F : \mathbb{R} \to \mathbb{R} \,.$$

The definite integral of f on $[a, b]$ equals

$$\int_a^b f(x)\,dx = F(b) - F(a) \,.$$

Remark 1.4. We also write

$$f(x)\,dx = dF$$

and

$$\int_a^b dF = F(b) - F(a)\,.$$

The integral may also exist if a or b are not finite or if the function is not defined in a or in b. In those cases the integral is defined as follows.

Remark 1.5. If the limit exists, we have

$$\int_{-\infty}^b f(x)\,dx = \lim_{a\to\infty}\int_{-a}^b f(x)\,dx$$

and

$$\int_a^\infty f(x)\,dx = \lim_{b\to\infty}\int_a^b f(x)\,dx\,.$$

With the assumption that the integrals on the right hand exist, we define

$$\int_{-\infty}^\infty f(x)\,dx = \int_{-\infty}^0 f(x)\,dx + \int_0^\infty f(x)\,dx\,.$$

If $\mathrm{dom}(f) = [a,b)$, then (with the assumption that the limits exist)

$$\int_a^b f(x)\,dx = \lim_{\varepsilon\to 0}\int_a^{b-\varepsilon} f(x)\,dx\,.$$

If $\mathrm{dom}(f) = (a,b]$, then (with the assumption that the limits exist)

$$\int_a^b f(x)\,dx = \lim_{\varepsilon\to 0}\int_{a+\varepsilon}^{b} f(x)\,dx\,.$$

Example.

$$\begin{aligned}\int_0^\infty \frac{1}{1+x^2}\,dx &= \lim_{b\to\infty}\int_0^b \frac{1}{1+x^2}\,dx = \lim_{b\to\infty}\big([\arctan x]_0^b\big)\\ &= \lim_{b\to\infty}\big(\arctan b - \arctan 0\big) = \frac{\pi}{2}\,.\end{aligned}$$

$$\int_0^\infty e^{-x}\,dx = \lim_{b\to\infty}\int_0^b e^{-x}\,dx = \lim_{b\to\infty}\left(-e^{-b}+e^{-0}\right)$$
$$= 1\,.$$

For $x \in \mathbb{R}$ let

$$|x| = \begin{cases} -x & \text{for } x < 0 \\ x & \text{for } x \geqslant 0 \end{cases}$$

then

$$\begin{aligned}
\int_{-\infty}^{\infty} |x|e^{-x^2}\,dx &= \int_{-\infty}^{0} -xe^{-x^2}\,dx + \int_0^\infty xe^{-x^2}\,dx \\
&= \lim_{a\to\infty}\int_{-a}^{0} -xe^{-x^2}\,dx + \lim_{b\to\infty}\int_0^b xe^{-x^2}\,dx \\
&= \lim_{a\to\infty}\left(\left[\frac{1}{2}e^{-x^2}\right]_{-a}^{0}\right) + \lim_{b\to\infty}\left(\left[-\frac{1}{2}e^{-x^2}\right]_0^b\right) \\
&= \lim_{a\to\infty}\left(\frac{1}{2}e^{-0} - \frac{1}{2}e^{-(-a)^2}\right) + \lim_{b\to\infty}\left(-\frac{1}{2}e^{-b^2} + \frac{1}{2}e^{-0}\right) \\
&= \left(\frac{1}{2} - 0\right) + \left(0 + \frac{1}{2}\right) = 0\,.
\end{aligned}$$

Here we determine the antiderivative by substitution or we guess it.

1.1.3 Indefinite integrals

We give tables with the antiderivatives of the most common functions.

Polynomial, rational and exponential functions	
Integral	Condition
$\int x^n\,dx = \frac{x^{n+1}}{n+1} + C$	$x \in \mathbb{R},\ n \in \mathbb{Z},\ n \geqslant 0$
$\int x^n\,dx = \frac{x^{n+1}}{n+1} + C$	$x \in \mathbb{R},\ x \neq 0,\ n \in \mathbb{Z},\ n \leqslant -2$

$\int x^s\,dx = \frac{x^{s+1}}{s+1} + C$	$x \in \mathbb{R},\ 0 < x,\ s \in \mathbb{R},\ s \neq -1$
$\int \frac{1}{x}\,dx = \ln\|x\| + C$	$x \in \mathbb{R},\ x \neq 0$
$\int e^x\,dx = e^x + C$	$x \in \mathbb{R}$
$\int \ln x\,dx = x\ln x - x + C$	$x \in \mathbb{R},\ 0 < x$

Trigonometric functions	
Integral	Condition
$\int \sin x\,dx = -\cos x + C$	$x \in \mathbb{R}$
$\int \cos x\,dx = \sin x + C$	$x \in \mathbb{R}$
$\int \tan x\,dx = -\ln\|\cos x\| + C$	$x \in \bigcup_{k\in\mathbb{Z}}\left(k\pi - \frac{\pi}{2},\ k\pi + \frac{\pi}{2}\right)$
$\int \frac{1}{\cos^2 x}\,dx = \tan x + C$	$x \in \bigcup_{k\in\mathbb{Z}}\left(k\pi - \frac{\pi}{2},\ k\pi + \frac{\pi}{2}\right)$
$\int \frac{1}{\sqrt{1-x^2}}\,dx = \arcsin x + C$	$x \in (-1, 1)$
$\int \frac{-1}{\sqrt{1-x^2}}\,dx = \arccos x + C$	$x \in (-1, 1)$

$\int \frac{1}{1+x^2}\,dx = \arctan x + C$	$x \in \mathbb{R}$

Hyperbolic functions	
Integral	Condition
$\int \sinh x\,dx = \cosh x + C$	$x \in \mathbb{R}$
$\int \cosh x\,dx = \sinh x + C$	$x \in \mathbb{R}$
$\int \tanh x\,dx = \ln(\cosh x) + C$	$x \in \mathbb{R}$
$\int \frac{1}{\cosh^2 x}\,dx = \tanh x + C$	$x \in \mathbb{R}$
$\int \frac{1}{\sqrt{1+x^2}}\,dx = \operatorname{arsinh} x + C$	$x \in \mathbb{R}$
$\int \frac{1}{\sqrt{x^2-1}}\,dx = \operatorname{arcosh} x + C$	$x \in \mathbb{R},\ 1 < x$
$\int \frac{1}{1-x^2}\,dx = \operatorname{artanh} x + C$	$x \in (-1, 1)$
$\int \frac{1}{1-x^2}\,dx = \operatorname{artanh} \frac{1}{x} + C$	$x \in (-\infty, -1) \cup (1, \infty)$
$\int \frac{1}{1-x^2}\,dx = \frac{1}{2} \cdot \ln\left\lvert\frac{x+1}{x-1}\right\rvert + C$	$x \in (-\infty, -1) \cup (-1, 1) \cup (1, \infty)$

In the next sections we learn how to integrate functions that are not in this list. We will therefore use the properties of the derivative.

1.1.4 Integration by parts

If we want to integrate a product of two functions, we should check whether this can be done by integration by parts. The formula is based on the product rule

$$\big(f(x)\cdot g(x)\big)' = f'(x)\cdot g(x) + f(x)\cdot g'(x)\,.$$

If we integrate

$$f'(x)\cdot g(x) = \big(f(x)\cdot g(x)\big)' - f(x)\cdot g'(x)$$

we get

$$\int f'(x)\cdot g(x)\,dx = f(x)\cdot g(x) - \int f(x)\cdot g'(x)\,dx$$

or

$$\int f(x)\cdot g'(x)\,dx = f(x)\cdot g(x) - \int f'(x)\cdot g(x)\,dx\,.$$

This formula can be used to compute integrals of functions $f'(x)\cdot g(x)$. The difficulty of this method consist in choosing the functions $f(x)$ and $g(x)$.

Example. For the computation of the integral $\int x\,e^x\,dx$ we choose

$$f(x) = x\,,\quad g'(x) = e^x\,.$$

Hence

$$f'(x) = 1\,,\quad g(x) = e^x$$

and

$$\begin{aligned}\int x\,e^x\,dx &= x\,e^x - \int 1\cdot e^x\,dx\\ &= x\,e^x - (e^x + C) \quad = (x-1)e^x + C\,.\end{aligned}$$

We notice that in the example we introduce the constant of integration only in the last step. Indeed, in the general case it is not necessary to introduce the constant earlier. [1]

[1] In the integral

$$\int f'(x)\cdot g(x)\,dx$$

Example. i) For the integral $\int x^2\, e^x\, dx$, we choose

$$f(x) = x^2\,, \quad g'(x) = e^x$$

and we get

$$f'(x) = 2x\,, \quad g(x) = e^x\,,$$

$$\begin{aligned}\int x^2\, e^x\, dx &= x^2\, e^x - 2\int x\, e^x\, dx\\ &= x^2\, e^x - 2(x-1)e^x + C \quad = (x^2 - 2x + 2)e^x + C\,.\end{aligned}$$

Here we use the result of the previous example or we apply the integration by parts on the integral $\int x\, e^x\, dx$.

ii) The integral $\int \cos(x) \cdot e^x\, dx$ is another example where we use twice the integration by parts. The choice

$$\begin{array}{ll} f(x) = \cos(x)\,, & g'(x) = e^x\,,\\ f'(x) = -\sin(x)\,, & g(x) = e^x \end{array}$$

---

we make the computation with a constant of integration

$$\begin{array}{ll} f'(x)\,, & g(x)\,,\\ f(x) + C\,, & g'(x) \end{array}$$

and get

$$\begin{aligned}\int f'(x)\, g(x)\, dx &= (f(x) + C)\, g(x) - \int (f(x) + C)\, g'(x)\, dx\\ &= f(x)\, g(x) + C\, g(x) - \int f(x)\, g'(x) + C\, g'(x)\, dx\\ &= f(x)\, g(x) + C\, g(x) - \int f(x)\, g'(x)\, dx - \int C\, g'(x)\, dx\\ &= f(x)\, g(x) + C\, g(x) - \int f(x)\, g'(x)\, dx - C\,(g(x) + C)\\ &= f(x)\, g(x) - \int f(x)\, g'(x)\, dx - C\end{aligned}$$

where $C \in \mathbb{R}$ is a constant. If $\int f(x)\, g'(x)\, dx = h(x) + C$, we can write

$$\int f'(x) \cdot g(x)\, dx = f(x) \cdot g(x) - h(x) + C$$

where $C \in \mathbb{R}$ is a constant. The argument is analogous for

$$\int f(x) \cdot g'(x)\, dx = f(x) \cdot g(x) - \int f'(x) \cdot g(x)\, dx\,.$$

leads to

$$\begin{aligned}\int \cos(x)\cdot e^x\,dx &= \cos(x)\cdot e^x - \int \big(-\sin(x)\big)\,e^x\,dx\\ &= \cos(x)\cdot e^x + \int \sin(x)\,e^x\,dx\,.\end{aligned}$$

Our first thought is that it doesn't help, but applying the integration by parts again and making sure that we compute the derivative of the derivative and integrate the integral, i.e.,

$$\begin{aligned} h(x) &= \sin(x)\,, \quad & k'(x) &= e^x\,,\\ h'(x) &= \cos(x)\,, \quad & k(x) &= e^x\,,\end{aligned}$$

we get

$$\int \sin(x)\,e^x\,dx = \sin(x)\cdot e^x - \int \cos(x)\,e^x\,dx$$

hence

$$\begin{aligned}\int \cos(x)\,e^x\,dx &= \cos(x)\,e^x + \int \sin(x)\,e^x\,dx\\ &= \cos(x)\,e^x + \left(\sin(x)\,e^x - \int \cos(x)\,e^x\,dx\right)\\ &= \big(\cos(x)+\sin(x)\big)\,e^x - \int \cos(x)\,e^x\,dx\end{aligned}$$

We see that our integral appears on both sides of the equation with a different sign

$$\int \cos(x)\,e^x\,dx = \big(\cos(x)+\sin(x)\big)\,e^x - \int \cos(x)\,e^x\,dx\,.$$

Hence

$$2\int \cos(x)\,e^x\,dx = \big(\cos(x)+\sin(x)\big)\,e^x$$

and

$$\int \cos(x)\,e^x\,dx = \frac{1}{2}\big(\cos(x)+\sin(x)\big)\,e^x + C\,,\quad C\in\mathbb{R}\,.$$

iii) In the example $\int \ln x\,dx$ the product is not obvious. With $\int 1\cdot \ln x\,dx$ we choose

$$\begin{aligned} f'(x) &= 1\,, \quad & g(x) &= \ln(x)\,,\\ f(x) &= x\,, \quad & g'(x) &= \frac{1}{x}\,,\end{aligned}$$

hence

$$\begin{aligned}\int \ln x\,dx &= x\ln(x) - \int x\cdot\frac{1}{x}\,dx\\ &= x\ln(x) - \int 1\,dx\\ &= x\ln(x) - x + C\,,\quad C\in\mathbb{R}\,.\end{aligned}$$

1.1.5 Integration by substitution

This method is based on the chain rule of differentiation. Let $I, J \subset \mathbb{R}$ be intervals. Let $f : I \to \mathbb{R}$, $x \mapsto f(x)$, and $\varphi : J \to I$, $t \mapsto \varphi(t)$, be two functions. We know that the derivative with respect to t of their composition is

$$\frac{d}{dt}\Big(f\big(\varphi(t)\big)\Big) = f'\big(\varphi(t)\big)\varphi'(t)\,.$$

We will use this fact for the integration by substitution.

Proposition 1.6. *Let $x := \varphi(t)$. Then*

$$\int f\big(\varphi(t)\big)\,\varphi'(t)\,dt = \Big(\int f(x)\,dx\Big)_{x:=\varphi(t)}$$

and

$$\int_a^b f\big(\varphi(t)\big)\,\varphi'(t)\,dt = \int_{\varphi(a)}^{\varphi(b)} f(x)\,dx\,.$$

Hence we search for a substitution $x := \varphi(t)$ and determine the integral $\int f(x)\,dx$.

Example. In the integral

$$\int_a^b 2t\sin(t^2)\,dt$$

we recognise a function $\varphi(t) = t^2$ and its derivative $\varphi'(t) = 2t$. Hence with the substitution $x := \varphi(t) = t^2$ we get

$$\frac{dx}{dt} = 2t\,,\quad \Rightarrow dx = 2t\,dt = \varphi'(t)\,dt$$

and

$$\begin{aligned}\int_a^b 2t\sin(t^2)\,dt &= \int_{\varphi(a)}^{\varphi(b)} \sin(\varphi(t))\,\varphi'(t)\,dt\\ &= \int_{a^2}^{b^2} \sin(x)\,dx\\ &= \big[-\cos(x)\big]_{a^2}^{b^2}\\ &= -\cos(b^2) - \big(-\cos(a^2)\big)\\ &= \cos(a^2) - \cos(b^2)\,.\end{aligned}$$

Example. The integral

$$J := \int (\cos t + \cos^3 t)\,dt$$

can be transformed to

$$J = \int (1 + \cos^2 t)\ \cos t\,dt = \int (2 - \sin^2 t)\cos t\,dt\,.$$

The substitution

$$\sin t := x\,,\quad \cos t\,dt = dx$$

yields

$$\begin{aligned}J &= \left(\int (2 - x^2)\,dx\right)_{x:=\sin t} = \left(2x - \frac{x^3}{3}\right)_{x:=\sin t} + c\\ &= 2\sin t - \frac{1}{3}\sin^3 t + c\,.\end{aligned}$$

The definite integral

$$J_0 := \int_{\pi/6}^{\pi} (\cos t + \cos^3 t)\,dt$$

becomes with this substitution with $\sin \pi/6 = 1/2$ and $\sin \pi = 0$

$$\begin{aligned}J_0 &= \int_{1/2}^{0} (2 - x^2)\,dx = \left(2x - \frac{x^3}{3}\right)\Big|_{1/2}^{0}\\ &= -\left(1 - \frac{1}{24}\right) = -\frac{23}{24}\,.\end{aligned}$$

Example. We compute the integral

$$\int \frac{3x-1}{x^2-x+1}\,dx\,.$$

We know that $(x^2-x+1)' = 2x-1$. With the numerator $2x-1$ we would substitute $v(x) = x^2-x+1$, $dv = (2x-1)dx$ and get

$$\begin{aligned}\int \frac{2x-1}{x^2-x+1}\,dx &= \int \frac{1}{v}\,dv = \ln|v| \\ &= \ln|x^2-x+1|\,.\end{aligned}$$

We now transform the fraction in order to use this fact. Since

$$\begin{aligned}\frac{3x-1}{x^2-x+1} &= \frac{3}{2}\cdot\frac{2x-\frac{2}{3}}{x^2-x+1} = \frac{3}{2}\cdot\frac{2x-1+\frac{1}{3}}{x^2-x+1} \\ &= \frac{3}{2}\left(\frac{2x-1}{x^2-x+1} + \frac{\frac{1}{3}}{x^2-x+1}\right)\end{aligned}$$

we get

$$\begin{aligned}\int \frac{3x-1}{x^2-x+1}\,dx &= \frac{3}{2}\int\left(\frac{2x-1}{x^2-x+1} + \frac{\frac{1}{3}}{x^2-x+1}\right)dx \\ &= \frac{3}{2}\ln|x^2-x+1| + \frac{1}{2}\int \frac{1}{x^2-x+1}\,dx\end{aligned}$$

We have to compute the second integral. Here we search for a substitution that allows us to use

$$\int \frac{1}{u^2+1}\,du = \arctan u + C$$

With

$$\begin{aligned}x^2-x+1 &= \underbrace{x^2-x+\frac{1}{4}}_{(x-\frac{1}{2})^2} - \frac{1}{4} + 1 = \left(x-\frac{1}{2}\right)^2 + \frac{3}{4} \\ &= \frac{3}{4}\left(\left(\frac{x-\frac{1}{2}}{\sqrt{\frac{3}{4}}}\right)^2 + 1\right)\end{aligned}$$

we get

$$\int \frac{1}{x^2-x+1}\,dx = \frac{4}{3}\int \frac{1}{\left(\frac{x-\frac{1}{2}}{\frac{\sqrt{3}}{2}}\right)^2 + 1}\,dx$$

and substitute

$$u = \frac{2}{\sqrt{3}}\Big(x - \frac{1}{2}\Big)\,, \quad du = \frac{2}{\sqrt{3}}\,dx\,.$$

Herewith

$$\frac{4}{3}\,\frac{\sqrt{3}}{2}\int \frac{1}{u^2+1}\,du = \frac{2}{\sqrt{3}}\arctan u + C$$

and

$$\int \frac{1}{x^2-x+1}\,dx = \frac{2}{\sqrt{3}}\arctan\bigg(\frac{2}{\sqrt{3}}\Big(x-\frac{1}{2}\Big)\bigg) + C\,.$$

Therefore

$$\int \frac{3x-1}{x^2-x+1}\,dx = \frac{3}{2}\ln|x^2-x+1| + \frac{1}{\sqrt{3}}\arctan\bigg(\frac{2}{\sqrt{3}}\Big(x-\frac{1}{2}\Big)\bigg) + C\,.$$

Remark 1.7. If the function $\varphi(t)$ is invertible on the required t-interval, then

$$\int f(x)\,dx = \bigg(\int f\big(\varphi(t)\big)\,\varphi'(t)\,dt\bigg)_{t:=\varphi^{-1}(x)}$$

and

$$\int_a^b f(x)\,dx = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f\big(\varphi(t)\big)\,\varphi'(t)\,dt\,.$$

1.1.6 Partial fraction decomposition

We consider functions

$$f(x) = \frac{P_n(x)}{Q_m(x)}\,,$$

where $P_n(x)$, $Q_m(x)$ are polynomials of degree $n < m$. Let $\alpha_1, \dots, \alpha_m$ be the zeros of $Q_m(x)$ and $P_n(\alpha_i) \neq 0$ for $i = 1, \dots, m$. If $\alpha_1, \dots, \alpha_m \in \mathbb{R}$ are distinct constants, then we make the ansatz

$$\frac{P_n(x)}{Q_m(x)} = \frac{A_1}{x-\alpha_1} + \frac{A_2}{x-\alpha_2} + \dots + \frac{A_m}{x-\alpha_m}\,.$$

If the multiplicity of α_2 is $\ell > 1$, then we make the ansatz

$$\frac{P_n(x)}{Q_m(x)} = \frac{A_1}{x-\alpha_1} + \frac{B_1}{x-\alpha_2} + \frac{B_2}{(x-\alpha_2)^2} + \dots + \frac{B_\ell}{(x-\alpha_2)^\ell} + \frac{C_1}{x-\alpha_3} + \dots$$

We then determine

$$\begin{aligned}\int f(x)\,dx &= \int \frac{P_n(x)}{Q_m(x)}\,dx \\ &= \int \frac{A_1}{x-\alpha_1}\,dx + \int \frac{B_1}{x-\alpha_2}\,dx + \cdots + \int \frac{B_\ell}{(x-\alpha_2)^\ell}\,dx + \ldots\end{aligned}$$

using

$$\int \frac{A}{x-\alpha}\,dx = A \cdot \ln|x-\alpha| + c$$

and for $n \neq 1$

$$\int \frac{A}{(x-\alpha)^n}\,dx = \frac{-A}{(n-1)}(x-\alpha)^{-(n-1)} + c$$

Example. In order to compute

$$\int \frac{-1}{x^2+5x+6}\,dx$$

we first solve

$$x^2+5x+6=0 \quad \Leftrightarrow \quad x_{1,2} = \frac{-5 \pm \sqrt{25-24}}{2} = \frac{-5 \pm 1}{2}$$

hence

$$x_1 = -2\,,\; x_2 = -3$$

and

$$x^2+5x+6 = (x+2)(x+3)\,.$$

To determine constants $A, B \in \mathbb{R}$ such that

$$\frac{-1}{x^2+5x+6} = \frac{-1}{(x+2)(x+3)} = \frac{A}{x+2} + \frac{B}{x+3}$$

we compute

$$\begin{aligned}\frac{A}{x+2} + \frac{B}{x+3} &= \frac{A(x+3)}{(x+2)(x+3)} + \frac{B(x+2)}{(x+2)(x+3)} \\ &= \frac{A(x+3)+B(x+2)}{(x+2)(x+3)} \\ &= \frac{(A+B)x+3A+2B}{(x+2)(x+3)} \\ &\overset{!}{=} \frac{-1}{(x+2)(x+3)}\end{aligned}$$

Hence, by comparison of the coefficients, we get the following system of equations for A and B:

$$\begin{aligned} A + B &= 0 \\ 3A + 2B &= -1 \end{aligned}$$

The first equation yields $B = -A$ and with

$$3A + 2B = 3A - 2A = A = -1$$

we get $A = -1$, $B = 1$. Now

$$\begin{aligned} \int \frac{-1}{x^2 + 5x + 6}\, dx &= \int \frac{-1}{x+2} + \frac{1}{x+3}\, dx \\ &= \int \frac{-1}{x+2}\, dx + \int \frac{1}{x+3}\, dx \\ &= -\ln|x+2| + \ln|x+3| + c \\ &= \ln\left|\frac{x+3}{x+2}\right| + c\,. \end{aligned}$$

Example. To compute the integral

$$\int \frac{x}{(x-1)^2\,(x+1)}\, dx$$

we use the partial fraction decomposition. We determine A_1, A_2 and B such that

$$\frac{x}{(x-1)^2\,(x+1)} = \frac{A_1}{x-1} + \frac{A_2}{(x-1)^2} + \frac{B}{x+1}$$

Now

$$\begin{aligned} \frac{A_1}{x-1} + \frac{A_2}{(x-1)^2} + \frac{B}{x+1} &= \frac{A_1(x-1)(x+1) + A_2(x+1) + B(x-1)^2}{(x-1)^2\,(x+1)} \\ &= \frac{A_1(x^2-1) + A_2(x+1) + B(x^2-2x+1)}{(x-1)^2\,(x+1)} \\ &= \frac{(A_1+B)x^2 + (A_2-2B)x - A_1 + A_2 + B}{(x-1)^2\,(x+1)} \\ &\overset{!}{=} \frac{x}{(x-1)^2\,(x+1)} \end{aligned}$$

yields the system of equations

$$\begin{aligned} A_1 + B &= 0 \\ A_2 - 2B &= 1 \\ -A_1 + A_2 + B &= 0 \end{aligned}$$

With $B = -A_1$ we get

$$\begin{aligned} A_2 + 2A_1 &= 1 \\ A_2 - 2A_1 &= 0 \end{aligned}$$

hence $A_2 = 2A_1$ and $2A_2 = 1$ yields the solution

$$A_1 = \frac{1}{4}, \quad A_2 = \frac{1}{2}, \quad B = \frac{-1}{4}.$$

Now

$$\begin{aligned} \int \frac{x}{(x-1)^2\,(x+1)}\,dx &= \int \frac{(\frac{1}{4})}{x-1} + \frac{(\frac{1}{2})}{(x-1)^2} + \frac{-(\frac{1}{4})}{x+1}\,dx \\ &= \frac{1}{4}\int \frac{1}{x-1}\,dx + \frac{1}{2}\int \frac{1}{(x-1)^2}\,dx - \frac{1}{4}\int \frac{1}{x+1}\,dx \\ &= \frac{1}{4}\ln|x-1| + \frac{1}{2}\cdot\frac{-1}{(x-1)} - \frac{1}{4}\ln|x+1| + c \\ &= \frac{-1}{2(x-1)} + \frac{1}{4}\ln\left|\frac{x-1}{x+1}\right| + c \end{aligned}$$

1.2 Complex numbers

1.2.1 Solving quadratic equations

We consider the equation

$$ax^2 + bx + c = 0\,,$$

with $a, b, c \in \mathbb{R}$. If $a \neq 0$, the solutions of this quadratic equation are

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad \text{and} \quad x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}\,. \tag{1.1}$$

We get

$$x_1 = x_2 \;\Leftrightarrow\; b^2 - 4ac = 0$$

and

$$x_1 = -x_2 \;\Leftrightarrow\; b = 0\,.$$

Let $\alpha \in \mathbb{R}$, $0 \neq \alpha$, be a real number. By (1.1) the solutions of

$$x^2 - \alpha^2 = 0$$

are

$$x_1 = \frac{0 + \sqrt{0 - 4(-\alpha^2)}}{2} = \alpha \quad \text{and} \quad x_2 = \frac{0 - \sqrt{0 - 4(-\alpha^2)}}{2} = -\alpha\,.$$

Indeed, we know that

$$x^2 - \alpha^2 = (x - \alpha)(x + \alpha)\,.$$

Remark 1.8. For $\alpha = 0$, the equation $x^2 = 0$ has the solution 0 with multiplicity 2.

Example. The solutions of the equation $x^2 - 1 = 0$ are $x_1 = 1$ and $x_2 = -1$.

We consider the equation

$$x^2 + 1 = 0\,.$$

If we determine its solutions, we get by (1.1)

$$x_1 = \frac{\sqrt{-4}}{2} = \sqrt{-1} \quad \text{and} \quad x_2 = \frac{-\sqrt{-4}}{2} = -\sqrt{-1}\,.$$

Since the function $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x) = x^2$, has a global minimum in $x = 0$ with $f(0) = 0$, every $x \in \mathbb{R}$ satisfies

$$x^2 \geqslant 0\,,$$

where $x^2 = 0$ if and only if $x = 0$. Hence there exists no $x \in \mathbb{R}$ with $x^2 = -1 < 0$. This shows that $x^2 + 1 = 0$ has no real solutions. We now introduce a number i that is a solution of the equation $x^2 + 1 = 0$. It satisfies

$$i^2 = -1\,.$$

Then

$$(x - i)(x + i) = x^2 - i^2 = x^2 + 1\,,$$

hence i and $-i$ are the solutions of $x^2 + 1 = 0$. Now, we can solve for $\alpha \in \mathbb{R}$, $\alpha \neq 0$,

$$x^2 + \alpha^2 = 0\,.$$

By (1.1) it has the solutions

$$x_1 = \frac{\sqrt{-4\alpha^2}}{2} = \sqrt{-\alpha^2} \quad \text{and} \quad x_2 = \frac{-\sqrt{-4\alpha^2}}{2} = -\sqrt{-\alpha^2}\,.$$

If we use the number i again, we get

$$x_1 = \sqrt{-\alpha^2} = \sqrt{-1} \cdot \alpha = i\alpha \quad \text{and} \quad x_2 = -\sqrt{-\alpha^2} = -\sqrt{-1} \cdot \alpha = -i\alpha\,.$$

Indeed,

$$(x - i\alpha)(x + i\alpha) = x^2 - (i\alpha)^2 = x^2 + \alpha^2\,.$$

Hence $i\alpha$ and $-i\alpha$ are the solutions of $x^2 + \alpha^2 = 0$.

Definition 1.9. The number i that solves $x^2 + 1 = 0$ is called the *imaginary unit.* For any $\alpha \in \mathbb{R}$, $\alpha \neq 0$, the number $i\alpha$ is called *imaginary.*

We consider the equation $\alpha x^2 + \beta x + \gamma = 0$, with $\alpha, \beta, \gamma \in \mathbb{R}$. If $\alpha \neq 0$, then

$$\alpha x^2 + \beta x + \gamma = 0 \;\Leftrightarrow\; x^2 + \frac{\beta}{\alpha} x + \frac{\gamma}{\alpha} = 0\,.$$

Hence it suffices to consider equations

$$x^2 + bx + c = 0\,,$$

$b, c \in \mathbb{R}$. By (1.1) its solutions are

$$\begin{aligned} x_1 &= \frac{-b + \sqrt{b^2 - 4c}}{2} = -\frac{b}{2} + \frac{\sqrt{b^2 - 4c}}{2}\,,\\ x_2 &= \frac{-b - \sqrt{b^2 - 4c}}{2} = -\frac{b}{2} - \frac{\sqrt{b^2 - 4c}}{2}\,. \end{aligned}$$

These solutions are real if and only if $b^2 - 4c \geqslant 0$. Now we consider the case where $x^2 + bx + c = 0$ has no real solution, i.e. $b^2 - 4c < 0$. Then the solutions are

$$\begin{aligned} x_1 &= -\frac{b}{2} + \frac{\sqrt{-(4c - b^2)}}{2} = -\frac{b}{2} + i\,\frac{\sqrt{4c - b^2}}{2}\,,\\ x_2 &= -\frac{b}{2} - \frac{\sqrt{-(4c - b^2)}}{2} = -\frac{b}{2} - i\,\frac{\sqrt{4c - b^2}}{2}\,. \end{aligned}$$

Note that $4c - b^2 > 0$.

Remark 1.10.

i) Note that

$$-\frac{b}{2} \in \mathbb{R}$$

and, since $4c - b^2 > 0$, we also have

$$\frac{\sqrt{4c - b^2}}{2} \in \mathbb{R}\,.$$

ii) We observe that, given x_1, we can determine x_2 by replacing i with $-i$ and vice versa.

Example. By (1.1) the solutions of the quadratic equation

$$x^2 - 2x + 2 = 0$$

are

$$x = \frac{2 \pm \sqrt{(-2)^2 - 4 \cdot 2}}{2} = \frac{2 \pm \sqrt{-4}}{2} = 1 \pm \sqrt{-1}$$

hence

$$x_1 = 1 + i \quad \text{and} \quad x_2 = 1 - i\,.$$

1.2.2 Definition of complex numbers

Definition 1.11. We define the *field* $\mathbb{C}$ *of complex numbers* to be the smallest field[a] that is an extension of the field of real numbers $\mathbb{R}$ and contains the solution i of the equation $x^2 + 1 = 0$. The elements of the field $\mathbb{C}$ are called *complex numbers*.

[a] Examples of fields are the field of rational numbers $\mathbb{Q}$, the field of real numbers $\mathbb{R}$ and the field of complex numbers $\mathbb{C}$.

Since $\mathbb{C}$ is a field that contains $\mathbb{R}$ as a subfield the complex numbers can be represented as follows. For each complex number $z \in \mathbb{C}$, there are $x, y \in \mathbb{R}$ such that

$$z = x + iy\,.$$

The number $z = x + iy \in \mathbb{C}$ is a solution of the equation

$$\big(z - (x + iy)\big)\big(z - (x - iy)\big) = 0\,.$$

Note that $\overline{z} := x - iy$ is also a solution of this equation. We compute

$$\begin{aligned} 0 &= \big(z - (x + iy)\big)\big(z - (x - iy)\big) \\ &= z^2 - (x + iy)z - (x - iy)z + (x + iy)(x - iy) \\ &= z^2 - (x + iy + x - iy)z + (x^2 + i(yx - xy) - i^2y^2) \\ &= z^2 - 2x\,z + (x^2 + y^2) \end{aligned}$$

Since x, y are real, we have

$$-2x \in \mathbb{R} \quad \text{and} \quad x^2 + y^2 \in \mathbb{R}\,.$$

Hence

$$p(z) = z^2 - 2x\,z + (x^2 + y^2)$$

is a real polynomial in the variable z, i.e. $p(z) \in \mathbb{R}[z]$. With $b := -2x$ and $c := x^2 + y^2$ we get

$$p(z) = z^2 + bz + c\,.$$

By (1.1), the zeros of the polynomial are

$$\begin{aligned} z_1 &= -\frac{b}{2} + \frac{\sqrt{b^2-4c}}{2} \\ &= -\frac{(-2x)}{2} + \frac{\sqrt{(-2x)^2 - 4(x^2+y^2)}}{2} = x + \sqrt{-y^2} \\ &= x + iy \ = z\,, \end{aligned}$$

$$\begin{aligned} z_2 &= -\frac{b}{2} - \frac{\sqrt{b^2-4c}}{2} \\ &= -\frac{(-2x)}{2} - \frac{\sqrt{(-2x)^2 - 4(x^2+y^2)}}{2} = x - \sqrt{-y^2} \\ &= x - iy = \overline{z}\,. \end{aligned}$$

Remark 1.12. If $z = x \in \mathbb{R}$ ($y = 0$), then

$$p(z) = z^2 - 2x\,z + x^2 = (z-x)^2\,.$$

Complex conjugation

Definition 1.13. Let

$$z = x + iy \in \mathbb{C}\,,$$

$x, y \in \mathbb{R}$, be a complex number. We define the *complex conjugate* $\overline{z}$ of z to be

$$\overline{z} = x - iy\,.$$

We define *the complex conjugation* to be the mapping

$$\begin{aligned} \overline{\cdot}\,:\ \mathbb{C} &\longrightarrow \mathbb{C} \\ z &\longmapsto \overline{z} \end{aligned}$$

that maps z to its complex conjugate.

Remark 1.14. The complex conjugation has the following properties.

i) For all $z \in \mathbb{C}$,
$$\overline{\overline{z}} = z\,.$$

ii) For any $x \in \mathbb{R}$,
$$\overline{x} = x\,.$$

Hence the complex conjugation acts as the identity on $\mathbb{R}$.

iii) For any $y \in \mathbb{R}$,
$$\overline{iy} = -iy\,.$$

Addition and product

We define an *addition* $+$ in $\mathbb{C}$:

$$\begin{array}{rccc} + : & \mathbb{C} \times \mathbb{C} & \longrightarrow & \mathbb{C} \\ & (z_1, z_2) & \longmapsto & z_1 + z_2 \end{array}$$

and a *product*

$$\begin{array}{rccc} \cdot : & \mathbb{C} \times \mathbb{C} & \longrightarrow & \mathbb{C} \\ & (z_1, z_2) & \longmapsto & z_1 \cdot z_2 \end{array}$$

where for $z_1 = x_1 + iy_1 \in \mathbb{C}$, $z_2 = x_2 + iy_2 \in \mathbb{C}$, with $x_1, x_2, y_1, y_2 \in \mathbb{R}$, we have for the addition

$$z_1 + z_2 = (x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2)$$

and for the product

$$\begin{aligned} z_1 \cdot z_2 = (x_1 + iy_1) \cdot (x_2 + iy_2) &= x_1x_2 + x_1iy_2 + iy_1x_2 + iy_1iy_2 \\ &= x_1x_2 + ix_1y_2 + iy_1x_2 + i^2y_1y_2 \\ &= x_1x_2 - y_1y_2 + i(x_1y_2 + y_1x_2) \end{aligned}$$

With $z = x + iy$, $x, y \in \mathbb{R}$, we can state the distributive property

$$z \cdot (z_1 + z_2) = z \cdot z_1 + z \cdot z_2$$

and

$$(z_1 + z_2) \cdot z = z_1 \cdot z + z_2 \cdot z$$

that can be checked by the reader.

Definition 1.15. Let $z \in \mathbb{C}$ be a complex number and $\overline{z} \in \mathbb{C}$ its complex conjugate. We call

$$\operatorname{Re}(z) = \frac{1}{2}(z + \overline{z})$$

the *real part* of z and

$$\operatorname{Im}(z) = -\frac{i}{2}(z - \overline{z})$$

the *imaginary* part of z.

Remark 1.16. For $z = x + iy \in \mathbb{C}$, with $x, y \in \mathbb{R}$, we get

$$\begin{aligned}\operatorname{Re}(z) = \frac{1}{2}(z + \overline{z}) &= \frac{1}{2}\big((x + iy) + (x - iy)\big) \\ &= \frac{1}{2}\big((x + x) + i(y - y)\big) \\ &= x \in \mathbb{R}\end{aligned}$$

and

$$\begin{aligned}\operatorname{Im}(z) = -\frac{i}{2}(z - \overline{z}) &= -\frac{i}{2}\big((x + iy) - (x - iy)\big) \\ &= -\frac{i}{2}\big((x - x) + i(y + y)\big) \\ &= -\frac{i}{2}(i\,2y) = -i^2 y \\ &= y \in \mathbb{R}\,.\end{aligned}$$

Since $\overline{\overline{z}} = z$, we see that

$$\begin{aligned}\operatorname{Re}(\overline{z}) &= \frac{1}{2}(\overline{z} + \overline{\overline{z}}) = \frac{1}{2}(\overline{z} + z) = \operatorname{Re}(z)\,, \\ \operatorname{Im}(\overline{z}) &= -\frac{i}{2}(\overline{z} - \overline{\overline{z}}) = -\frac{i}{2}(\overline{z} - z) = -\operatorname{Im}(z)\,.\end{aligned}$$

Definition 1.17. Let $z \in \mathbb{C}$ be a complex number and $\overline{z} \in \mathbb{C}$ its complex conjugate. We define the *absolute value* or *modulus* of $z \in \mathbb{C}$ to be

$$|z| = \sqrt{z\,\overline{z}}\,.$$

If $z = x + iy$ for some $x, y \in \mathbb{R}$, then

$$|z| = \sqrt{x^2 + y^2}\,.$$

Remark 1.18. If $z = x \in \mathbb{R}$, then

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

i.e. the absolute value of a real number equals its absolute value as a complex number.

1.2.3 The complex plane

The cartesian plane

We know that for each complex $z \in \mathbb{C}$ there are $x, y \in \mathbb{R}$ such that

$$z = x + iy\,.$$

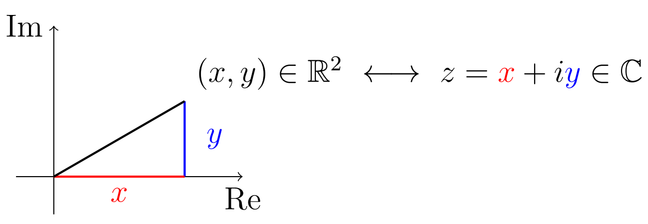
The mapping

$$\begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{C} \\ (x, y) & \longmapsto & z = x + iy \end{array}$$

is invertible

$$\begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{R}^2 \\ z & \longmapsto & (\mathrm{Re}(z), \mathrm{Im}(z))\,. \end{array}$$

It describes an isomorphism (a one-to-one correspondence) between the complex field $\mathbb{C}$ and the real plane $\mathbb{R}^2$.



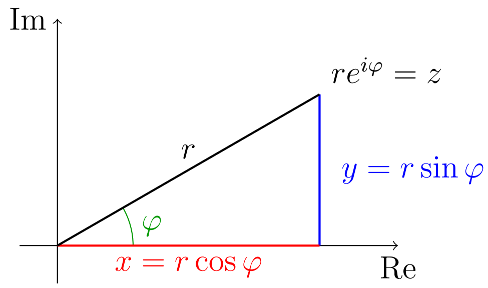
Definition 1.19. We call the x-axis the *real axis* Re and the y-axis the *imaginary axis* Im. The plane is called *complex plane*. It is sometimes known as *Argand plane* or *Gauß plane*.

The polar plane

In analogy to the polar coordinates in the plane, there is a unique polar representation for every complex number $z \in \mathbb{C}$, $z \neq 0$. It is

$$z = re^{i\varphi} = r(\cos\varphi + i\sin\varphi)$$

with $r \in \mathbb{R}_+$ and $\varphi \in]-\pi, \pi]$.



Given $z = re^{i\varphi}$, we get

$$x = r\cos\varphi \quad \text{and} \quad y = r\sin\varphi\,.$$

For a given $z = x + iy \in \mathbb{C}$ we see that the modulus (absolute value) of z

$$|z| = r = |(x, y)| = \sqrt{x^2 + y^2}$$

is the distance between (x, y) and the origin. If $z = x + iy \in \mathbb{C}$, $z \neq 0$, satisfies $x > 0$, then the argument is

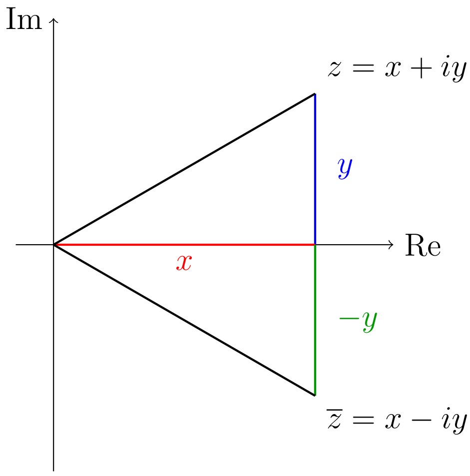
$$\varphi = \arctan\left(\frac{y}{x}\right),$$

where $\arctan : \mathbb{R} \to]-\frac{\pi}{2}, \frac{\pi}{2}[$ is the inverse of the tangent function. We get

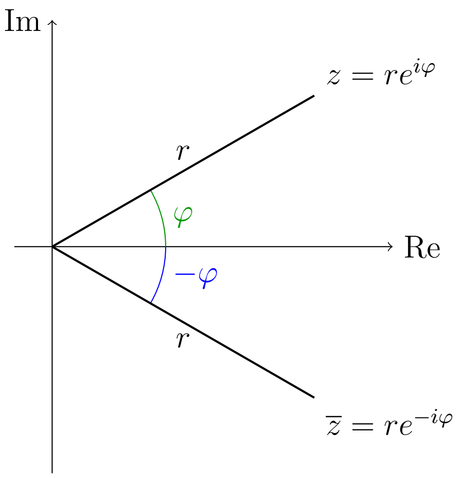
$$\varphi = \begin{cases} -\pi + \arctan(\frac{y}{x}) & \text{if } x < 0, y < 0, \\ -\frac{\pi}{2} & \text{if } x = 0, y < 0, \\ \arctan(\frac{y}{x}) & \text{if } x > 0, \\ \frac{\pi}{2} & \text{if } x = 0, y > 0, \\ \pi + \arctan(\frac{y}{x}) & \text{if } x < 0, y \geqslant 0\,. \end{cases}$$

Complex conjugation, addition and product in the plane

The complex conjugation that maps a complex number $z \in \mathbb{C}$ to its complex conjugate $\overline{z} \in \mathbb{C}$ corresponds to a reflection on the real axis. We first consider the cartesian representation.



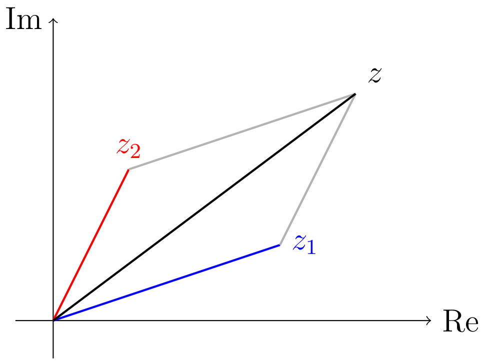
In the polar representation the complex conjugation is given as follows.



Hence

$$\overline{re^{i\varphi}} = re^{-i\varphi}\,.$$

The addition of complex numbers corresponds to an addition of vectors. We choose the cartesian form. Let $z_1, z_2 \in \mathbb{C}$, then $z = z_1 + z_2$ is determined as follows.

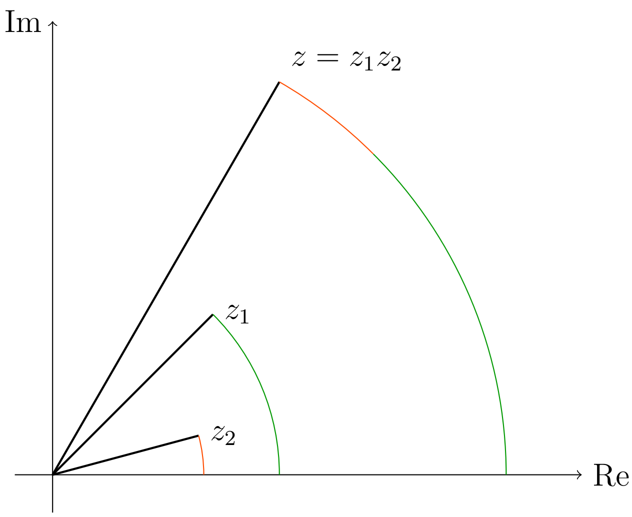


Let $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$ with $x_1, x_2, y_1, y_2 \in \mathbb{R}$. Then

$$\begin{aligned} z &= z_1 + z_2 \\ &= (x_1 + iy_1) + (x_2 + iy_2) \\ &= (x_1 + x_2) + i(y_1 + y_2) \\ &= x + iy\,. \end{aligned}$$

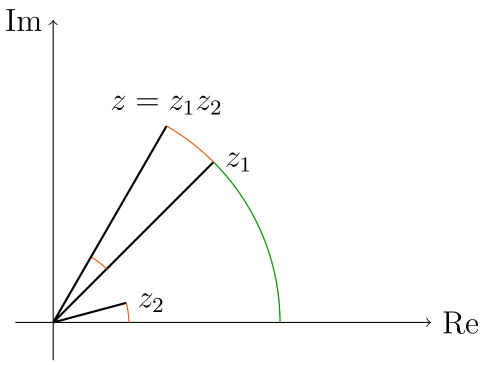
For the product we choose the polar form. Let $z_1 = r_1 e^{i\varphi_1}$ and $z_2 = r_2 e^{i\varphi_2}$, then the product is

$$\begin{aligned} z = z_1 \cdot z_2 = r_1 e^{i\varphi_1} \cdot r_2 e^{i\varphi_2} &= r_1 r_2\, e^{i\varphi_1} e^{i\varphi_2} \\ &= r_1 r_2\, e^{i(\varphi_1 + \varphi_2)} = r e^{i\varphi} \end{aligned}$$



By $z_1 \cdot z_2 = r_1 r_2\, e^{i(\varphi_1 + \varphi_2)}$ we see that the multiplication with a number of modulus $r = 1$ represents a rotation in the complex plane. Indeed the product of $z_1 = r_1 e^{i\varphi_1}$ and $z_2 = e^{i\varphi_2}$ is

$$\begin{aligned} z = z_1 \cdot z_2 &= r_1 e^{i\varphi_1} \cdot e^{i\varphi_2} \\ &= r_1 \, e^{i(\varphi_1 + \varphi_2)} \end{aligned}$$



We might also compute for some $n \in \mathbb{Z}$ the n-th power z^n of $z = re^{i\varphi}$.

$$z^n = r^n \left(e^{i\varphi}\right)^n = r^n \, e^{i\,n\varphi} \,.$$

If we set $r = 1$, then we have $z = e^{i\varphi}$ and

$$z^n = \left(e^{i\varphi}\right)^n = e^{i\,n\varphi} \,.$$

Depending on n and on $\varphi \in]-\pi, \pi]$, we might have $n\varphi \notin]-\pi, \pi]$. Example if $\varphi = \frac{\pi}{3}$, then $4\varphi = \frac{4\pi}{3} \notin]-\pi, \pi]$, but

$$e^{i\frac{4\pi}{3}} = e^{-i\frac{2\pi}{3}}$$

since $\frac{4\pi}{3} = \frac{-2\pi}{3} + 2\pi$ and

$$e^{i\varphi} = e^{i(\varphi + 2k\pi)} \quad \text{for any } k \in \mathbb{Z} \,.$$

Definition 1.20. An n-th root of 1 is a solution of the equation

$$z^n = 1 = e^{i\,k\,2\pi} \,, \quad k \in \mathbb{Z} \,.$$

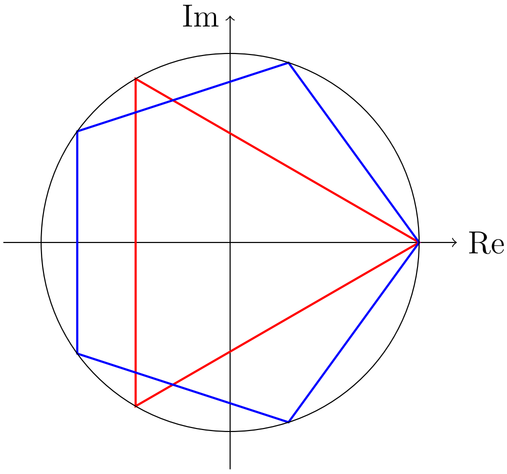
Hence every

$$z_k = e^{i\,k\frac{2\pi}{n}} \,, \quad k \in \mathbb{Z} \,,$$

is a solution of $z^n = 1 = e^{i\,k\,2\pi}$.

Remark 1.21. It is easy to see that $z_k = z_{k+n}$ and that $z_0, \ldots, z_{n-1}$ are pairwise different.

In the complex plane the solutions of the n-th root of 1 lie on a regular polygon with n vertices. Example for $n = 3$ and for $n = 5$:



Chapter 2

Multivariable Calculus

2.1 Multivariable functions

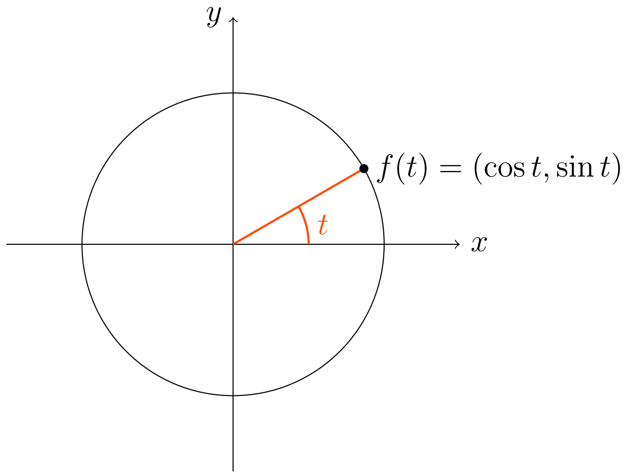
The functions studied in multivariable calculus are

Curves	$f : \mathbb{R} \to \mathbb{R}^n$	Length of curves, line integrals, curvature
Surfaces	$f : \mathbb{R}^2 \to \mathbb{R}^n$	Areas of surfaces, surface integrals, flux through surfaces, curvature
Scalar fields	$f : \mathbb{R}^n \to \mathbb{R}$	Maxima and minima, Lagrange multipliers, directional derivatives
Vector fields	$f : \mathbb{R}^m \to \mathbb{R}^n$	Any of the operations of vector calculus, gradient, divergence, curl

Example. The curve given by

$$\begin{array}{rccl} f: & \mathbb{R} & \longrightarrow & \mathbb{R}^2 \\ & t & \longmapsto & f(t) := (\cos t, \sin t) \end{array}$$

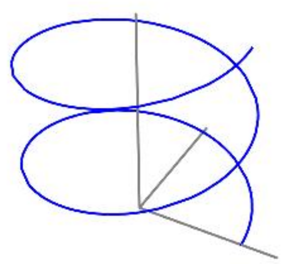
is the unit circle in the plane.



The curve parametrized with

$$\begin{array}{rccc} f: & \mathbb{R} & \longrightarrow & \mathbb{R}^3 \\ & t & \longmapsto & (\cos t, \sin t, t) \end{array}$$

is a line that ”screws upwards”.



2.2 Scalar fields

In this section we consider functions

$$\begin{array}{rccc} f: & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) \end{array}$$

that map points $x = (x_1, \dots, x_n)$ in $\mathbb{R}^n$ to scalars $f(x_1, \dots, x_n)$. If $D \subset \mathbb{R}^n$, then the graph

$$\big\{\big(x, f(x)\big) \in \mathbb{R}^n \times \mathbb{R} \mid x \in D\big\}$$

describes a surface over D.
The function f may represent the metres above sea level of a point on a map or the temperature at a point in a space.

Definition 2.1. The *level set* of the function

$$\begin{array}{rccc} f: & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) \end{array}$$

to the level $c \in \mathbb{R}$ is the set

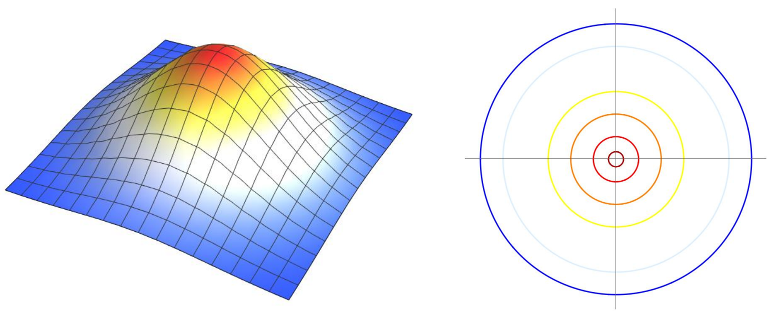
$$f^{-1}(c) := \big\{(x_1, \dots, x_n) \in \mathbb{R}^n \mid f(x_1, \dots, x_n) = c\big\}\,.$$

On the examples above it corresponds to the points at the same altitude or with the same temperature.

Example. Consider the function

$$\begin{array}{rccc} f: & \mathbb{R}^2 & \to & \mathbb{R} \\ & (x,y) & \mapsto & f(x,y) := e^{-(x^2+y^2)} \end{array}$$

What are its level sets?



The level set of f to the level $c \in \mathbb{R}$ is the set of points $(x, y) \in \mathbb{R}^2$ that satisfy $f(x, y) = c$, where c is a constant. Since

$$c = e^{-(x^2+y^2)} \quad \Leftrightarrow \quad x^2 + y^2 = C\,,$$

where $C \in \mathbb{R}$ is a constant, we get the level lines

$$x^2 + y^2 = C = r^2\,.$$

These are circles with radius r centred in $(0, 0)$.

2.2.1 Partial derivatives

For functions of one variable we know that the derivative $f'(x^0)$ of a function f in a given point x^0 helps to understand how the value $f(x)$ of f changes when we move to a point x that is in a neighbourhood of x^0. We now want to get a similar information for functions in more than one variable.
Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function in n variables. We fix a point

$$x^0 = (x_1^0, \dots, x_n^0) \in \mathbb{R}^n$$

and consider the line

$$L_i := \{(x_1^0, \dots, x_{i-1}^0, x_i, x_{i+1}^0, \dots, x_n^0) \in \mathbb{R}^n \mid x_i \in \mathbb{R}\}\,.$$

This line is parallel to the x_i-axis and goes through x^0. Then the set

$$\mathcal{C}_i(x^0) := \{(x_1^0, \dots, x_{i-1}^0, x_i, x_{i+1}^0, \dots, x_n^0, f(\dots x_{i-1}^0, x_i, x_{i+1}^0, \dots)) \mid x_i \in \mathbb{R}\}$$

is a curve over the line L_i. It is the graph of the function

$$\begin{array}{rccl} \varphi_i : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x_i & \longmapsto & f(x_1^0, \dots, x_{i-1}^0, x_i, x_{i+1}^0, \dots, x_n^0)\,. \end{array}$$

We consider the derivative of φ_i with respect to the variable x_i in the point x^0. This is called the *partial derivative* of f in x^0 with respect to x_i and written

$$f_i(x^0) \qquad \text{or} \qquad \frac{\partial f}{\partial x_i}(x^0)\,.$$

It is defined to be the limit

$$f_i(x^0) := \lim_{\Delta x \to 0} \frac{f(x_1^0, \dots,\ x_i^0 + \Delta x\,, \dots, x_n^0) - f(x_1^0, \dots, x_i^0, \dots, x_n^0)}{\Delta x}\,.$$

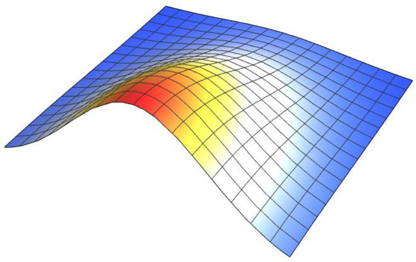
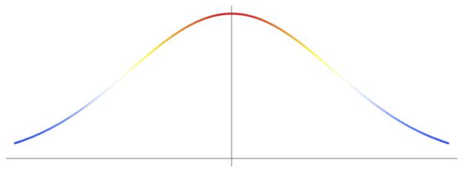
The tangent at $\mathcal{C}_i(x^0)$ in $p = \big(x^0, f(x^0)\big)$ is given by

$$\left\{ \big(x^0, f(x^0)\big) + \frac{\partial}{\partial x_i} f(x^0)(x_i - x_i^0) \,\middle|\, x_i \in \mathbb{R} \right\}\,.$$

The tangents $\mathcal{C}_i(x^0)$, $i = 1, \dots, n$, span the tangent vector space

$$T_pS$$

at the surface S in p.
We explain this theory on the following example.

Example. We cut the graph of the function

$$\begin{array}{rccl} f: & \mathbb{R}^2 & \to & \mathbb{R} \\ & (x,y) & \mapsto & f(x,y) := e^{-(x^2+y^2)}\,. \end{array}$$

along the plane $y = 0$. Determine the partial derivatives. We choose the point $p_0 = (x_0, y_0)$. The line

$$L_1 := \{(x, y_0) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$$

is parallel to the x-axis and passes through p_0. Then the set

$$\mathcal{C}_1(p_0) := \{(x, y_0, f(x, y_0)) \mid x \in \mathbb{R}\}$$

is a curve over the line L_1. It is the graph of the function

$$\begin{array}{rccl} \varphi_1: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x,y_0) = e^{-(x^2+y_0^2)}\,. \end{array}$$

The partial derivative of f in $p_0 = (x_0, y_0)$ with respect to x is

$$f_x(p_0) = \frac{\partial}{\partial x} f(p_0) = -2x\, e^{-(x^2+y^2)}\Big|_{(x_0,y_0)} = -2x_0\, e^{-(x_0^2+y_0^2)}\,.$$

With an analogous argument we see that the partial derivative of f in $p_0 = (x_0, y_0)$ with respect to y is

$$f_y(p_0) = \frac{\partial}{\partial y} f(p_0) = -2y\, e^{-(x^2+y^2)}\Big|_{(x_0,y_0)} = -2y_0\, e^{-(x_0^2+y_0^2)}\,.$$

2.2.2 The gradient

Definition 2.2. We assume that all partial derivatives $\frac{\partial}{\partial x_i} f$, $i = 1, \ldots, n$ of the function

$$\begin{array}{rccc} f: & D & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) \end{array}$$

exist and that they are continuous. Then the vector

$$\nabla f(x^0) := \left(\frac{\partial}{\partial x_1} f(x^0), \ldots, \frac{\partial}{\partial x_n} f(x^0) \right)$$

is defined and called the *gradient* of f in x^0.

Example. The gradient of the function

$$\begin{array}{rccc} f: & \mathbb{R}^2 & \to & \mathbb{R} \\ & (x, y) & \mapsto & f(x, y) := e^{-(x^2+y^2)}\,. \end{array}$$

in (x, y) is

$$\begin{aligned} \nabla f(x, y) &= \left(\frac{\partial}{\partial x} f(x, y), \frac{\partial}{\partial y} f(x, y) \right) = \big(f_x(x, y), f_y(x, y) \big) \\ &= \left(-2x\, e^{-(x^2+y^2)}, -2y\, e^{-(x^2+y^2)} \right) \\ &= 2\, e^{-(x^2+y^2)} \big(-x, -y \big) \\ &= \frac{2}{e^{(x^2+y^2)}} \big(-x, -y \big) \end{aligned}$$

In any point $(x, y) \in \mathbb{R}^2$ the gradient points to the origin $(0, 0)$ and its length depends on the norm $\sqrt{x^2 + y^2}$ of the vector (x, y), hence on the distance of (x, y) to the origin.

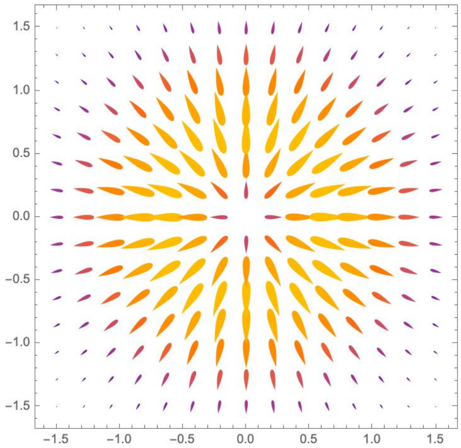
Theorem 2.3. *The gradient $\nabla f(x^0)$ is perpendicular to the level set*

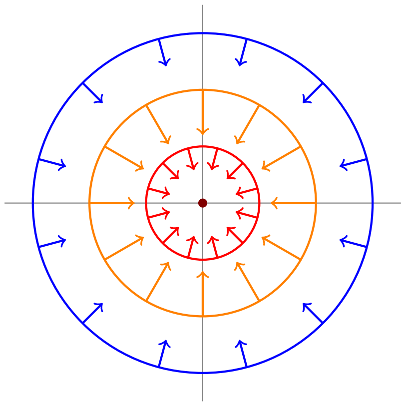
$$f^{-1}\big(f(x^0)\big) := \big\{ (x_1, \ldots, x_n) \in \mathbb{R}^n \mid f(x_1, \ldots, x_n) = f(x^0) \big\}\,.$$

Example. The level sets of the function

$$\begin{array}{rccc} f: & \mathbb{R}^2 & \to & \mathbb{R} \\ & (x, y) & \mapsto & f(x, y) := e^{-(x^2+y^2)} \end{array}$$

are circles and the gradient of f in the point (x, y) is parallel to $(-x, -y)$ and points to the origin. As we can see it on the figure, the gradient is perpendicular to the circles.





2.2.3 The directional derivative

By definition, the partial derivative with respect to x_i of a function

$$\begin{array}{rccc} f: & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) \end{array}$$

in x^0 indicates how the value of the function changes when only the x_i-coordinate in x^0 changes, hence when x^0 is moved along a line that is parallel to the x_i-axis. But how does the value of the function change when we move x^0 along any $v \in \mathbb{R}^n$? The answer is given by the directional derivative.

Definition 2.4. The *directional derivative* of a scalar function

$$\begin{array}{rccc} f: & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) \end{array}$$

along a vector $v \in \mathbb{R}^n$ is the function $\nabla_v f : \mathbb{R}^n \to \mathbb{R}$ defined by the limit

$$\nabla_v f(x) = \lim_{h \to 0} \frac{f(x + hv) - f(x)}{h\,|v|}\,.$$

The division by $|v|$ ensures that the result only depends on the direction of v.

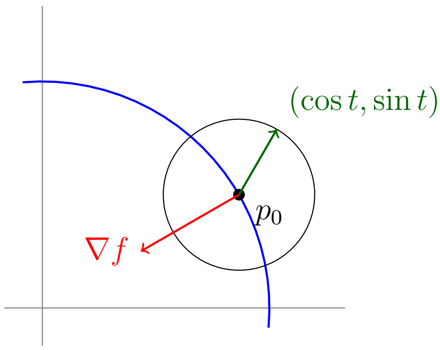
Example. The directional derivative of the function

$$\begin{array}{rccl} f: & \mathbb{R}^2 & \to & \mathbb{R} \\ & (x,y) & \mapsto & f(x,y) := e^{-(x^2+y^2)} \end{array}$$

along $v = (\cos t, \sin t)$ equals

$$\begin{aligned} \nabla_v f(x,y) &= \nabla f(x,y) \cdot v \\ &= \Big(-2x\,e^{-(x^2+y^2)}, -2y\,e^{-(x^2+y^2)}\Big) \cdot (\cos t, \sin t) \\ &= \frac{2}{e^{(x^2+y^2)}}\big(-x, -y\big) \cdot (\cos t, \sin t) \\ &= \frac{-2(x\cos t + y \sin t)}{e^{(x^2+y^2)}} \end{aligned}$$

since $|v| = \sqrt{\cos^2 t + \sin^2 t} = 1$.



Remark 2.5. If the function f is differentiable at x^0, then the directional derivative exists along any vector v and

$$\nabla_v f(x^0) = \nabla f(x^0) \cdot \frac{v}{|v|}$$

where ∇f denotes the gradient of f and "$\cdot$" is the scalar product (dot product) of vectors. The division by $|v|$ ensures that the result does not depend on the magnitude of v.

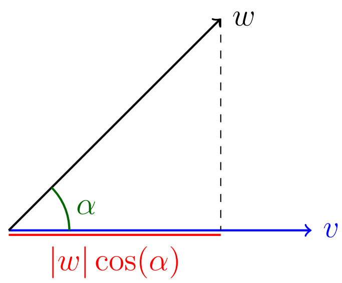
We consider the vector space $\mathbb{R}^n$. The scalar product $v \cdot w$ of two vectors $v, w \in \mathbb{R}^n$, $v = (v_1, \ldots, v_n)$, $w = (w_1, \ldots, w_n)$ is defined as follows.

$$v \cdot w = v_1 w_1 + \ldots + v_n w_n$$

This is the algebraic definition of the scalar product. The geometric definition is the following.

$$v \cdot w = |v|\,|w| \cos \alpha$$

where α is the angle between v and w. The length $|w| \cos \alpha$ is the scalar projection of w on v.



Remark 2.6. The gradient ∇f is also denoted $\text{grad}(f)$. It indicates the direction of maximal slope.

Indeed, given x^0 and f, the directional derivative

$$\begin{aligned}\nabla_v f(x^0) &= \nabla f(x^0) \cdot \frac{v}{|v|} = \left|\nabla f(x^0)\right| \left|\frac{v}{|v|}\right| \cos \alpha \\ &= \left|\nabla f(x^0)\right| \cos \alpha\end{aligned}$$

is maximal if the angle α between $\nabla f(x^0)$ and v is 0, i.e., if v has the same direction as the gradient.

2.2.4 The total differential

Definition 2.7. The *total differential* of a function $f : \mathbb{R}^3 \to \mathbb{R}$ in 3 variables in (x_0, y_0, z_0) is defined to be

$$df = f_x(x_0, y_0, z_0)\, dx + f_y(x_0, y_0, z_0)\, dy + f_z(x_0, y_0, z_0)\, dz\,.$$

The total differential of a function in 2 variables in (x_0, y_0) is

$$df = f_x(x_0, y_0)\,dx + f_y(x_0, y_0)\,dy\,.$$

2.2.5 A chain rule for partial derivatives

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a differentiable function. Consider the composition

$$f\big(x(s,t), y(s,t)\big)$$

for differentiable functions $x, y : \mathbb{R}^2 \to \mathbb{R}$. Then

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial s} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial s}$$
$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial t} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial t}$$

We show the first equation. We consider the limit

$$\begin{aligned}
&\lim_{\Delta s \to 0} \frac{1}{\Delta s}\Big(f\big(x(s+\Delta s, t), y(s+\Delta s, t)\big) - f\big(x(s,t), y(s,t)\big)\Big) \\
&\quad = \lim_{\Delta s \to 0} \frac{1}{\Delta s}\Big(f\big(x(s+\Delta s, t), y(s+\Delta s, t)\big) - f\big(x(s,t), y(s+\Delta s, t)\big)\Big) \\
&\qquad + \lim_{\Delta s \to 0} \frac{1}{\Delta s}\Big(f\big(x(s,t), y(s+\Delta s, t)\big) - f\big(x(s,t), y(s,t)\big)\Big) \\
&\quad \approx \lim_{dx \to 0} \frac{\frac{\partial x}{\partial s}}{dx}\Big(f\big(x(s,t)+dx, y(s+\Delta s, t)\big) - f\big(x(s,t), y(s+\Delta s, t)\big)\Big) \\
&\qquad + \lim_{dy \to 0} \frac{\frac{\partial y}{\partial s}}{dy}\Big(f\big(x(s,t), y(s,t)+dy\big) - f\big(x(s,t), y(s,t)\big)\Big) \\
&\quad = \frac{\partial f}{\partial x}\big(x(s,t), y(s,t)\big) \cdot \frac{\partial x}{\partial s}(s,t) + \frac{\partial f}{\partial y}\big(x(s,t), y(s,t)\big) \cdot \frac{\partial y}{\partial s}(s,t)
\end{aligned}$$

The proof of the rule for the partial derivative with respect to t is analogous. For the special case $x(t)$, $y(t)$ we have

$$\frac{df}{dt}\big(x(t), y(t)\big) = \frac{\partial f}{\partial x}\big(x(t), y(t)\big) \cdot x'(t) + \frac{\partial f}{\partial y}\big(x(t), y(t)\big) \cdot y'(t)\,.$$

Chapter 3

Differential equations

3.1 Definition

Definition 3.1. A *differential equation* is an equation

$$F\big(x, y(x), y'(x), \ldots, y^{(n)}(x)\big) = 0$$

relating the independent variable x, the unknown function

$$y : \mathbb{R} \longrightarrow \mathbb{R}\,.$$

and its derivatives.
The ith derivative of y is denoted by $y^{(i)}$.
The *order* of the differential equation is the order of the highest derivative that appears in the relation F. (Here it is $y^{(n)}$ and therefore the order is n.)
The differential equation is called *ordinary* if y is a function of a single variable. If $y : \mathbb{R}^n \to \mathbb{R}$ is a function in more than one variable, then the differential equation is called a *partial* differential equation.

3.2 First order differential equations

Definition 3.2. A differential equation of order 1 has the following general form:

$$y'(x) = f\big(x, y(x)\big)\,,$$

where $f : \mathbb{R}^2 \to \mathbb{R}$ with $\mathrm{dom}(f) = D \subseteq \mathbb{R}^2$.

The solution y of a differential equation of order one is a curve that has a given slope in every point.

Theorem 3.3. *The following statements are true under the assumption of some technical conditions for the function f.*

i) The solutions of a differential equation $y'(x) = f(x, y(x))$ form a one-parameter family of curves $y_c : x \to y_c(x)$.

ii) The initial value problem (IVP)

$$\begin{cases} y' = f(x, y) \\ y(x_0) = y_0 \end{cases}$$

has a unique solution $x \mapsto y(x)$.

3.2.1 Separation of variables

Some differential equations are of the form

$$y'(x) = f(y) \cdot g(x)\,.$$

These have the following properties:

- The extrema of $y(x)$ are in the zeros of $g(x)$.
- The zeros of $f(y)$ are solutions since for y_0 with $f(y_0) = 0$ holds:

$$\text{Let } y(x) = y_0, \text{ then } y'(x) = 0 = f(y_0) \cdot g(x).$$

W.l.o.g. we assume $f(y) \neq 0$ and define

$$h(y) = \frac{1}{f(y)}\,.$$

We separate

$$\begin{aligned} \frac{dy}{dx} &= f(y) \cdot g(x) \\ h(y)\,dy &= g(x)\,dx\,. \end{aligned}$$

Integration yields

$$\int h(y)\,dy = \int g(x)\,dx\,.$$

Let $H(y)$ be an antiderivative of $h(y)$ and $G(x)$ an antiderivative of $g(x)$. Let $y = y(x)$. Then for all x on an interval (a, b) the equation

$$H(y(x)) = G(x)$$

has to be satisfied. We compute the derivative with respect to x:

$$\frac{dH}{dy} \cdot \frac{dy}{dx} = \frac{dG}{dx},$$

i.e.

$$h(y) \cdot y' = g(x).$$

This is equivalent to our differential equation.
The solution of the initial value problem

$$\begin{cases} y' &= g(x)f(y) \\ y(x_0) &= y_0 \end{cases}$$

is given by

$$\int_{y_0}^{y(x)} \frac{ds}{f(s)} = \int_{x_0}^{x} g(t)\,dt.$$

Example. The differential equation

$$y' = \frac{-x}{y}$$

is separable:

$$\int y\,dy = -\int x\,dx.$$

The solution is a curve in $\mathbb{R}^2$ with slope

$$f(x, y) = \frac{-x}{y}$$

in any point $(x, y) \in \mathbb{R}^2$ in which f is defined. The solution satisfies

$$y^2 = -x^2 + c^2$$

i.e.

$$y = \sqrt{c^2 - x^2}$$

The constant c is defined by the initial value (x_0, y_0).

Type $\mathbf{y' = f(ax + by + c)}$

We consider the differential equation

$$y' = f(ax + by + c)$$

and introduce the unknown function

$$u(x) = ax + by(x) + c\,.$$

Then

$$u' = a + by'\,,$$

i.e.

$$y' = \frac{u' - a}{b}\,,$$

We get a differential equation for u

$$\frac{u' - a}{b} = f(u) \Longleftrightarrow u' = a + b \cdot f(u)\,.$$

This equation is separable:

$$\int \frac{du}{a + b \cdot f(u)} = x\,.$$

If it is possible to compute this integral ($= I(u)$), the equation

$$I(ax + by + c) = x$$

has to be solved with respect to y (i.e. $y = \ldots$).

Example. We solve the initial value problem

$$\begin{cases} y' + y &= 1 + x,\ x > 0 \\ y(0) &= 2\,. \end{cases}$$

The differential equation is

$$y' = x - y + 1\,.$$

We define

$$u = x - y + 1$$

and have

$$u' = 1 - y'$$

i.e.

$$y' = 1 - u'\,.$$

The differential equation is

$$1 - u' = f(u) \Longleftrightarrow u' = 1 - f(u)$$

with $f(u) = u$. The equation

$$u' = 1 - u$$

is separable:

$$\begin{aligned}
\int \frac{1}{1-u}\,du &= \int dx \\
-\ln(1-u) &= x + c_1 \\
\ln(-x+y) &= -x - c_1 \\
-x + y &= e^{-x-c_1} \\
y &= x + Ce^{-x}\,.
\end{aligned}$$

The constant C is defined by the initial value $y(0) = 2$. The solution of the problem is

$$y = x + 2e^{-x}\,.$$

Type $\mathbf{y' = f\left(\frac{y}{x}\right)}$

The differential equation

$$y' = f\left(\frac{y}{x}\right)$$

is also solved by substitution and separation. We choose the unknown function

$$u = \frac{y(x)}{x}\,.$$

Then

$$y' = (u \cdot x)' = u' \cdot x + u\,,$$

and the differential equation for u is

$$u' \cdot x + u = f(u) \Longleftrightarrow u' = \frac{f(u) - u}{x}\,.$$

Herewith

$$\underbrace{\int \frac{du}{f(u) - u}}_{I(u)} = \int \frac{dx}{x} = \ln|x| + \text{const} = \ln(c \cdot |x|)\,.$$

Sometimes it is possible to solve

$$I\left(\frac{y}{x}\right) = \ln(c \cdot |x|)$$

with respect to y.

3.3 Linear differential equations

Definition 3.4. A *linear differential equation* is defined by a linear polynomial in the unknown function $y(x)$ and its derivatives $y^{(i)}(x)$. This is an equation of the form

$$p_n(x)y^{(n)}(x) + \ldots + p_2(x)y''(x) + p_1(x)y'(x) + p_0(x)y(x) = q(x)$$

where $p_0(x), \ldots, p_n(x)$ and $q(x)$ are arbitrary differentiable functions. They do not need to be linear.
The differential equation is called *inhomogeneous* if $q(x) \neq 0$ and it is called *homogeneous* if $q(x) = 0$, i.e. if

$$p_n(x)y^{(n)}(x) + \ldots + p_2(x)y''(x) + p_1(x)y'(x) + p_0(x)y(x) = 0$$

Given an inhomogeneous linear differential equation, we get the corresponding homogeneous linear differential equation by replacing the function $q(x)$ with the zero function.

The following property holds for linear differential equations.

Proposition 3.5. *The following is true for inhomogeneous linear differential equations.*

the general solution of an inhomogeneous linear differential equation
$=$
the general solution of the corresponding homogeneous linear differential equation
$+$
the particular solution of the inhomogeneous linear differential equation.

Proof. We take any two solutions $y_1(x)$ and $y_2(x)$ of the linear differential equation

$$p_n(x)y^{(n)}(x) + \ldots + p_2(x)y''(x) + p_1(x)y'(x) + p_0(x)y(x) = q(x)$$

and consider their difference $(y_1 - y_2)(x) = y_1(x) - y_2(x)$. We put it in the left-hand side of the equation and get

$$\begin{aligned}
p_n(x)(y_1 - y_2)^{(n)}(x) &+ \ldots + p_1(x)(y_1 - y_2)'(x) + p_0(x)(y_1 - y_2)(x) \\
&= p_n(x)(y_1^{(n)} - y_2^{(n)})(x) + \ldots + p_1(x)(y_1' - y_2')(x) + p_0(x)(y_1 - y_2)(x) \\
&= \underbrace{p_n(x)y_1^{(n)}(x) + \ldots + p_1(x)y_1'(x) + p_0(x)y_1(x)}_{=q(x)} \\
&\quad - \big(\underbrace{p_n(x)y_2^{(n)}(x) + \ldots + p_1(x)y_2'(x) + p_0(x)y_2(x)}_{=q(x)}\big) \\
&= q(x) - q(x) = 0
\end{aligned}$$

Hence $y_1 - y_2$ is a solution of the corresponding homogeneous differential equation

$$p_n(x)y^{(n)}(x) + \ldots + p_2(x)y''(x) + p_1(x)y'(x) + p_0(x)y(x) = 0$$

This shows that given a particular solution y_p of the inhomogeneous linear differential equation we get any other solution y of this equation by adding a solution y_h of the corresponding homogeneous equation to y_p. Hence the claim follows. □

With a similar argument we see that the following is true.

Remark 3.6. If $y_1(x)$ and $y_2(x)$ are two solutions of a homogeneous differential equation, then so is $\lambda\, y_1(x) + \mu\, y_2(x)$, $\lambda, \mu \in \mathbb{R}$.

3.3.1 Variation of constants

We consider the following linear inhomogeneous differential equation of first order:

$$y'(x) + p(x)y(x) = q(x)\,.$$

First we consider the corresponding homogeneous differential equation:

$$y'(x) + p(x)y(x) = 0\,.$$

This equation is separable. We get

$$\begin{aligned}
\frac{dy}{dx} &= -p(x)y(x) \\
\int \frac{1}{y}\,dy &= \int -p(x)\,dx
\end{aligned}$$

and herewith

$$y(x) = C \cdot e^{-P(x)}\,,$$

where

$$P(x) = \int_{x_0}^{x} p(x)dx$$

is an antiderivative of $p(x)$. We call $y(x) = C \cdot e^{-P(x)}$ the general solution of the homogeneous equation.

To solve the inhomogeneous problem we make the following ansatz:

$$y(x) = C(x) \cdot e^{-P(x)}\,,$$

where the function $C(x)$ has to be defined. This is the variation of constants. With the differential equation

$$y'(x) + p(x)y(x) = q(x)$$

we get

$$C'(x) = q(x)\, e^{P(x)}$$

and herewith

$$C(x) = \int q(x)\, e^{P(x)}\, dx\,.$$

The particular solution of the inhomogeneous equation is

$$y(x) = e^{-P(x)} \int_{x_0}^{x} q(s)e^{P(s)}\, ds\,.$$

Proposition 3.5 yields in our case

$$y(x) = Ce^{-P(x)} + e^{-P(x)} \int_{x_0}^{x} q(s)e^{P(s)}\, ds\,.$$

The constant C is determined by the initial value $y(x_0) = y_0$.

Example. Find the general solution of the following differential equation.

$$y' - 2xy = x$$

The general solution of the corresponding homogeneous differential equation

$$y' - 2xy = 0$$

is

$$y_h(x) = Ce^{x^2}\,.$$

Hence, by the variation of constants, we guess a special solution

$$y_p(x) = C(x)e^{x^2}\,.$$

of the inhomogeneous differential equation. This function has to satisfy the differential equation $y' - 2xy = x$. We compute

$$\begin{aligned} y_p' - 2xy_p &= C'(x)e^{x^2} + 2xC(x)e^{x^2} - 2xC(x)e^{x^2} \\ &= C'(x)e^{x^2} \\ &\overset{!}{=} x \end{aligned}$$

Now we use the separation of variables to solve

$$C'(x)e^{x^2} = x$$

and get

$$\begin{aligned} C'(x) &= xe^{-x^2} \\ C(x) &= \int_{x_0}^{x} se^{-s^2}\,ds = -\frac{1}{2}e^{-x^2} + C_1\,. \end{aligned}$$

Hence

$$y_p(x) = C(x)e^{x^2} = -\frac{1}{2} + C_1e^{x^2}$$

and the general solution is

$$y(x) = -\frac{1}{2} + Ce^{x^2}\,.$$

Example. Find the maximal solution of the initial value problem

$$y' - y \cdot \tan x = \cos^2 x \quad \text{with} \quad y(0) = 1\,.$$

The solution of the homogeneous differential equation

$$y' - y \cdot \tan x = 0$$

is

$$y_h(x) = C \cdot e^{-P(x)}\,,$$

where $P(x)$ is an antiderivative of $p(x) = -\tan x$. We know that

$$-P(x) = \int_0^x \tan s\,ds = -\ln|\cos x| \text{ for } x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \quad (x_0 = 0)$$

hence

$$y_h(x) = C \cdot e^{-\ln|\cos x|} = C \cdot \frac{1}{|\cos x|} \text{ for } x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

The particular solution of the inhomogeneous equation is

$$\begin{aligned} y_p(x) &= e^{-P(x)} \int_{x_0}^{x} q(s) e^{P(s)}\, ds \\ &= \frac{1}{|\cos x|} \int_{x_0}^{x} \cos^2(s)\, |\cos s|\, ds\,, \end{aligned}$$

$x_0 = 0$. The general solution is

$$y(x) = C \cdot \frac{1}{|\cos x|} + \frac{1}{|\cos x|} \int_0^x \cos^2(s)\, |\cos s|\, ds$$

with $C = 1$ since $y_h(0) = 1$. Moreover $\cos x > 0$ for $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ and

$$y(x) = \frac{1}{\cos x} + \frac{1}{\cos x} \int_0^x \cos^3(s)\, ds\,.$$

We compute

$$\begin{aligned} \int_0^x \cos^3(s)\, ds &= \int_0^x \cos(s)(1 - \sin^2 s)\, ds \\ &= \int_0^x \cos(s)\, ds - \int_0^x \cos(s) \sin^2 s\, ds \\ &= \sin x - \frac{1}{3} \sin^3 x \end{aligned}$$

and get

$$y(x) = \frac{1 + \sin x - \frac{1}{3}\sin^3 x}{\cos x} \quad \text{with } x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

3.3.2 Linear differential equations of order n with constant coefficients

A linear differential equation of order n is

$$x^{(n)}(t) + a_{n-1}(t) x^{(n-1)}(t) + \ldots + a_1(t)\dot{x}(t) + a_0(t) x(t) = r(t)\,.$$

Remark 3.7.

- Linear means that there are no terms of the form $x^{(k)}(t) \cdot x^{(l)}(t)$.
- If $r(t) = 0$, then the differential equation is homogeneous. If $r(t) \neq 0$, then the differential equation is inhomogeneous.

- If $a_{n-1}(t), \ldots, a_0(t)$ are constant (don't depend on t), then the differential equation has constant coefficients.

We consider the homogeneous differential equation of order n with constant coefficients.

$$x^{(n)} + a_{n-1}x^{(n-1)} + \ldots + a_1\dot{x} + a_0x = 0$$

The equation $\dot{x} = \lambda x$ has the general solution $x(t) = Ce^{\lambda t}$. We make the ansatz

$$x(t) = e^{\lambda t}, \quad \lambda \in \mathbb{R}$$

and get

$$\lambda^n e^{\lambda t} + a_{n-1}\lambda^{n-1}e^{\lambda t} + \ldots a_1\lambda e^{\lambda t} + a_0 e^{\lambda t} = 0$$

and

$$e^{\lambda t}(\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0) = 0\,.$$

Since $e^{\lambda t} \neq 0$, this equation is equivalent to

$$\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0 = 0\,.$$

Definition 3.8. We call

$$Q(\lambda) := \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0$$

the *characteristic polynomial* of the differential equation.

$$x^{(n)} + a_{n-1}x^{(n-1)} + \ldots + a_1\dot{x} + a_0x = 0\,.$$

There are different cases:

i) $Q(\lambda)$ has n different real zeros.

ii) $Q(\lambda)$ has zeros with multiplicity > 1.

iii) $Q(\lambda)$ has complex zeros.

Case i): If $Q(\lambda) = 0$ has n different real zeros, the differential equation has n linear independent solutions:

$$e^{\lambda_1 t}, \ldots, e^{\lambda_n t}$$

The general solution is

$$x(t) = \sum_{i=1}^{n} c_i e^{\lambda_i t}$$

with n constants $c_1, \ldots, c_n$.
Case ii): If λ_k is a zero with multiplicity p, then

$$e^{\lambda_k t}, te^{\lambda_k t}, t^2 e^{\lambda_k t}, \ldots, t^{p-1} e^{\lambda_k t}$$

are the p solutions corresponding to λ_k. Their contribution to the general equation is

$$C_0 e^{\lambda_k t} + C_1 t e^{\lambda_k t} + C_2 t^2 e^{\lambda_k t} + \cdots + C_{p-1} t^{p-1} e^{\lambda_k t}$$

Case iii): The simple complex zeros

$$\lambda = \alpha \pm i\beta$$

correspond to the real solutions

$$e^{\alpha t} \cos \beta t \quad \text{and} \quad e^{\alpha t} \sin \beta t \, .$$

These are linearly independent. If the multiplicity of the complex zeros is $p > 1$ we have the solutions

$$\begin{array}{llll} e^{\alpha t} \cos \beta t \, , & t e^{\alpha t} \cos \beta t \, , & \ldots , & t^{p-1} e^{\alpha t} \cos \beta t \, , \\ e^{\alpha t} \sin \beta t \, , & t e^{\alpha t} \sin \beta t \, , & \ldots , & t^{p-1} e^{\alpha t} \sin \beta t \, . \end{array}$$

Depending on the inhomogeneous term $r(t)$ there are different ansatz to make for the particular solution of the inhomogeneous equation. Here we give some of the most important examples

a) $r(t) = e^{at}$, $a \in \mathbb{R}$ or $a \in \mathbb{C}$. Then the ansatz is

$$x = c \cdot e^{at}$$

where c is a constant that has to be defined.

b) $r(t) = e^{\alpha t} \left\{ \begin{array}{l} \sin \omega t \\ \cos \omega t \end{array} \right.$. This is similar to the previous case since

$$e^{\alpha t} \cdot \sin \omega t = \operatorname{Im} e^{(\alpha + i\omega)t}, \quad e^{\alpha t} \cdot \cos \omega t = \operatorname{Re} e^{(\alpha + i\omega)t} \, .$$

c) $r(t)$ is a polynomial in t of degree m. If $a_0 = \cdots = a_q = 0$, $a_{q+1} \neq 0$, then the ansatz is

$$x(t) = b_{m+1} t^{m+q+1} + \cdots + b_1 t^{1+q} + b_0 t^q \, .$$

Example. Determine all the solutions of the differential equation

$$y'' - y' - 6y = e^{-x}\,.$$

that are bounded on the interval $[0, \infty[$ and satisfy $y(0) = 0$.
We first consider the homogeneous equation

$$y'' - y' - 6y = 0\,.$$

The zeros of the characteristic polynomial

$$\lambda^2 - \lambda - 6 = 0$$

are

$$\lambda_{1,2} = \frac{1 \pm \sqrt{1+24}}{2} = \begin{cases} 3 \\ -2 \end{cases}\,.$$

The general solution of the homogeneous equation is

$$y_h(x) = c_1 e^{3x} + c_2 e^{-2x}\,.$$

The ansatz for the special solution of the inhomogeneous equation is

$$y(x) = ce^{-x}\,.$$

Then $y' = -ce^{-x}$, $y'' = ce^{-x}$ and

$$\begin{aligned} y'' - y' - 6y &= ce^{-x} + ce^{-x} - 6ce^{-x} \\ &= -4ce^{-x} \stackrel{!}{=} e^{-x}\,. \end{aligned}$$

The particular solution is

$$y_p(x) = -\frac{1}{4}e^{-x}$$

and the general solution is

$$y(x) = y_h(x) + y_p(x) = c_1 e^{3x} + c_2 e^{-2x} - \frac{1}{4}e^{-x}\,.$$

We study the solution on the interval $[0, \infty[$:

$$\begin{array}{ll} e^{\alpha x} > 0 & \text{for } \alpha \in \mathbb{R},\ x \in \mathbb{R} \\ e^0 = 1 & \\ \lim_{x\to\infty} e^{\alpha x} = \infty & \text{for } \alpha \in \mathbb{R},\ \alpha > 0 \\ \lim_{x\to\infty} e^{\alpha x} = 0 & \text{for } \alpha \in \mathbb{R},\ \alpha < 0 \end{array}$$

therefore $y(x)$ is bounded if and only if $c_1 = 0$. The condition $y(0) = 0$ yields the solution

$$y(x) = \frac{1}{4}e^{-2x} - \frac{1}{4}e^{-x}\,.$$

Chapter 4

Linear algebra

4.1 Linear functions

Linear algebra is the theory of linear functions.

Definition 4.1. Let

$$\begin{array}{rccc} f: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) \end{array}$$

be a function mapping the real numbers on real numbers. The function f is called *linear* if and only if for all $x, y \in \mathbb{R}$ and $\lambda \in \mathbb{R}$

$$f(\lambda x + y) = \lambda f(x) + f(y)\,.$$

The definition for functions $\mathbb{Q} \to \mathbb{Q}$ on the rational numbers or functions $\mathbb{C} \to \mathbb{C}$ on the complex numbers is analogous.

Example. Let $a \in \mathbb{R}$ (or $a \in \mathbb{Q}$, $a \in \mathbb{C}$) be a constant. Then the function given by

$$f(x) = ax, \qquad x \in \mathbb{R}\,(\mathbb{Q},\, \mathbb{C})$$

is a linear function.

Example. The function

$$\begin{array}{rccl} ||\cdot||: & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & ||x|| := \sqrt{x_1^2 + \cdots + x_n^2}\,, \end{array}$$

is the norm of the vector x. The norm is not a linear function.

A function associates to every element of the domain exactly one element of the codomain.

Definition 4.2. A function $f : A \to B$ is said to be *injective* (or *one-to-one*) if for all $x_1, x_2 \in A$, whenever $f(x_1) = f(x_2)$, then $x_1 = x_2$. In symbols

$$\forall x_1, x_2 \in A, \quad f(x_1) = f(x_2) \Rightarrow x_1 = x_2 .$$

Example.

- The function $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x) = x + 5$ is injective. Indeed

$$f(x_1) = f(x_2) \Rightarrow x_1 + 5 = x_2 + 5 \Rightarrow x_1 = x_2$$

- The function $g : \mathbb{R} \to \mathbb{R}$, $x \mapsto g(x) = x^2$ is not injective, since for x and $-x \in \mathbb{R}$ we have $g(x) = x^2$ and $g(-x) = (-x)^2 = x^2$. Hence $g(x) = g(-x)$, but $x \neq -x$.

Definition 4.3. A function $f : A \to B$ is said to be *surjective* (or *onto*) if for every $y \in B$, there exists an $x \in A$, such that $f(x) = y$. In symbols

$$\forall y \in B, \ \exists x \in A, \quad f(x) = y .$$

Example.

- The function $g : \mathbb{R} \to \mathbb{R}$, $x \mapsto g(x) = x^2$ is not surjective. Indeed the negative number -4 is an element of the codomain $\mathbb{R}$, but there is no x in the domain $\mathbb{R}$, such that $x^2 = -4$ since every square of a real number is 0 or positive.

- Let $\mathbb{R}^+ := \{x \in \mathbb{R} \mid x \geqslant 0\}$ denote the set of non-negative real numbers. The function $h : \mathbb{R} \to \mathbb{R}^+$, $x \mapsto h(x) = x^2$ is surjective. Indeed for any $y \in \mathbb{R}^+$ there is a $x \in \mathbb{R}$ with $h(x) = x^2 = y$.

Definition 4.4. A function $f : A \to B$ is said to be *bijective* if it is injective and surjective.

Example.

- The function $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x) = x + 5$ is surjective.

- The function $g : \mathbb{R} \to \mathbb{R}$, $x \mapsto g(x) = x^2$ is not bijective since it is not surjective.

- The function $h : \mathbb{R} \to \mathbb{R}^+$, $x \mapsto h(x) = x^2$ is not bijective since it is not injective.

Remark 4.5. A bijection from the set A to the set B has an inverse function from B to A.

Definition 4.6. Let

$$\begin{array}{rccl} g: & \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ & x & \longmapsto & g(x) \end{array}$$

be a function mapping a vector with n components

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

on vectors with m components (coordinates)

$$g(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_m(x) \end{pmatrix} = \begin{pmatrix} g_1(x_1, \ldots, x_n) \\ \vdots \\ g_m(x_1, \ldots, x_n) \end{pmatrix},$$

where

$$\begin{array}{rccl} g_i: & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & g_i(x), \end{array}$$

$i = 1, \ldots, m$, is a function in m variables. The sets $\mathbb{R}^n$ and $\mathbb{R}^m$ are *vector spaces*. The natural number n is called the *dimension* of the vector space $\mathbb{R}^n$. The function g is called *linear* if and only if for all $x, y \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$

$$g(\lambda x + y) = \lambda g(x) + g(y).$$

The function g is linear if and only if the funtions g_i, $i = 1, \ldots, m$, are linear.

Example. We consider the function

$$\begin{array}{rccl} f: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & x & \longmapsto & f(x) \end{array}$$

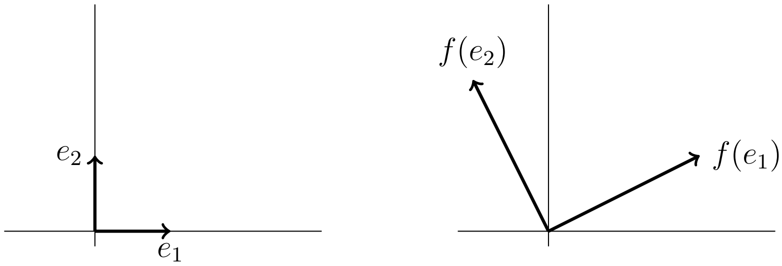
defined by

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longmapsto \begin{pmatrix} 2x_1 - x_2 \\ x_1 + 2x_2 \end{pmatrix}$$

In particular

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \longmapsto \begin{pmatrix} 2 \\ 1 \end{pmatrix} = f(e_1) \quad \text{and} \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \longmapsto \begin{pmatrix} -1 \\ 2 \end{pmatrix} = f(e_2)$$

We represent the vectors e_1, e_2 and their images in different pictures.



4.2 Linear equations

4.2.1 Introduction

Given $a, b \in \mathbb{R}$, ($\mathbb{Q}$, or $\mathbb{C}$). When does the linear equation in x

$$a \cdot x = b$$

have a solution? Is there more than one solution?
Let a and b be rational, real or complex numbers. If it is possible to divide by a, i.e. $a \neq 0$ or a is invertible, then the equation has a unique solution

$$x = a^{-1} \cdot b = \frac{b}{a} \,.$$

If $a = 0$ and $b = 0$ we have the equation

$$0 \cdot x = 0$$

and this is true for any value of x. Therefore the equation $0 \cdot x = 0$ has infinitely many solutions. If $a = 0$ and $b \neq 0$, then the equation

$$0 \cdot x = b \neq 0$$

has no solutions.
Linear algebra is a theory for solving linear equations.

4.2.2 Systems of equations

Now we consider a system of linear equations in two variables x_1, x_2.

$$a_{11}x_1 + a_{12}x_2 = b_1 \tag{4.1}$$
$$a_{21}x_1 + a_{22}x_2 = b_2 \,. \tag{4.2}$$

We will see later that this corresponds to

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = a_{22}b_1 - a_{12}b_2 \tag{4.3}$$
$$(a_{11}a_{22} - a_{12}a_{21})x_2 = -a_{21}b_1 + a_{11}b_2 \tag{4.4}$$

Now we have two equations, one for x_1 and one for x_2. We know that the system has infinitely many solutions if and only if the following three equations hold

$$\begin{aligned} a_{11}a_{22} - a_{12}a_{21} &= 0 \\ a_{22}b_1 - a_{12}b_2 &= 0 \\ -a_{21}b_1 + a_{11}b_2 &= 0 \end{aligned}$$

The system has no solution if and only if

$$a_{11}a_{22} - a_{12}a_{21} = 0$$

and

$$(a_{22}b_1 - a_{12}b_2 \neq 0 \qquad \text{or} \qquad -a_{21}b_1 + a_{11}b_2 \neq 0)$$

hold.
The system has a unique solution if and only if

$$a_{11}a_{22} - a_{12}a_{21} \neq 0\,.$$

Of course we can continue that way for 3 equations in 3 variables and n equations in n variables.

Example. We want to solve

$$\begin{aligned} 3x_1 + x_2 &= 2 \\ 5x_1 + 2x_2 &= 3 \end{aligned}$$

We first eliminate x_1 in the second row:

$$\begin{aligned} 3x_1 + x_2 &= 2 \\ 3x_1 + \frac{6}{5}x_2 &= \frac{9}{5} \end{aligned}$$

$$\begin{aligned} 3x_1 + x_2 &= 2 \\ 0x_1 + \frac{1}{5}x_2 &= -\frac{1}{5} \end{aligned}$$

$$\begin{aligned} 3x_1 + x_2 &= 2 \\ x_2 &= -1 \end{aligned}$$

and then x_2 in the first row:

$$\begin{aligned} 3x_1 \qquad &= 3 \\ x_2 &= -1 \end{aligned}$$

and get

$$\begin{aligned} x_1 &= 1 \\ x_2 &= -1\,. \end{aligned}$$

In the example we consider a function

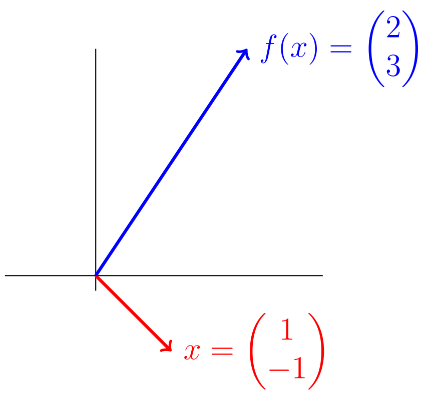
$$\begin{array}{rccc} f: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & x & \longmapsto & f(x) \end{array}$$

defined by

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longmapsto \begin{pmatrix} 3x_1 + x_2 \\ 5x_1 + 2x_2 \end{pmatrix}$$

and search for a vector $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ whose image is

$$f(x) = \begin{pmatrix} 3x_1 + x_2 \\ 5x_1 + 2x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}\,.$$



This will be done more systematically in the next section.

4.3 Matrices

4.3.1 Vector spaces

What is the meaning of what we have done?
We consider a linear function

$$\begin{array}{rccl} f: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) := ax\,, \end{array}$$

$a \in \mathbb{R}$. Then the equation $ax = b$ has a solution if and only if b belongs to the image of f, i.e. if there is $x \in \mathbb{R}$ with $f(x) = b$.
For the system of two equations in two variables we consider a linear function

$$\begin{array}{rccl} f: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & x & \longmapsto & f(x) := Ax\,. \end{array}$$

with

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$$

a vector and

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

a 2×2-matrix. With this notation the system of equations can be written as

$$Ax = b\,,$$

where

$$b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{R}^2\,.$$

We write

$$Ax = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}$$

and

$$Ax = b \quad \Longleftrightarrow \quad \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}\,.$$

This is equivalent to

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

This can be written as

$$\sum_{j=1}^{2} a_{ij}x_j = b_i\,, \quad i = 1,2\,.$$

The system

$$\begin{aligned} 3x_1 + x_2 &= 2 \\ 5x_1 + 2x_2 &= 3 \end{aligned}$$

can be written

$$Ax = b$$

with

$$A := \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}\,, \quad b := \begin{pmatrix} 2 \\ 3 \end{pmatrix}\,, \quad x := \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\,.$$

The solution is

$$x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}\,.$$

A system of m equations in n variables

$$\sum_{j=1}^{n} a_{ij}x_j = b_i\,, \quad i = 1, \ldots m\,.$$

can be written as

$$Ax = b\,,$$

with

$$A := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$$

a $m \times n$-matrix,

$$x := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

a vector in the vector space $\mathbb{R}^n$ and

$$b := \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m$$

a vector in the vector space $\mathbb{R}^m$.
The product Ax equals

$$Ax = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix}$$

The system of equations can be written

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} .$$

This is equivalent to

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{2m}x_n &= b_m . \end{aligned}$$

Here we have a linear function

$$\begin{aligned} f: \quad \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ x &\longmapsto f(x) := Ax . \end{aligned}$$

The equation $Ax = b$ has a solution if and only if $b \in \mathbb{R}^m$ belongs to the image of the function f, i.e. if there is a $x \in \mathbb{R}^n$ with $f(x) = b$.

Definition 4.7. A $n \times n$ matrix

$$A := (a_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}$$

is called an *upper triangular* matrix if and only if

$$a_{ij} = 0 \text{ for } i > j$$

and *lower triangular* matrix if and only if

$$a_{ij} = 0 \text{ for } i < j\,.$$

The matrix

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}$$

is upper triangular and the matrix

$$\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{n1} & \dots & a_{n\,(n-1)} & a_{nn} \end{pmatrix}$$

is a lower triangular matrix.

4.3.2 Linear dependency

Definition 4.8. We consider a set of n vectors $\alpha_1, \dots \alpha_n \in \mathbb{R}^m$

$$\alpha_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \quad \dots, \quad \alpha_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

This set is called *linearly dependent* if there exist $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ such that

$$\lambda_1\alpha_1 + \lambda_2\alpha_2 + \dots + \lambda_n\alpha_n = 0$$

and there is $i \in \{1, \dots n\}$ with $\lambda_i \neq 0$ (at least one of the λ_i is not 0). The set of vectors $\alpha_1, \dots \alpha_n \in \mathbb{R}^m$ is called *linearly independent* if and only if

$$\lambda_1\alpha_1 + \lambda_2\alpha_2 + \dots + \lambda_n\alpha_n = 0$$

implies

$$\lambda_1 = \lambda_2 = \dots = \lambda_n = 0\,.$$

Example. The set

$$\alpha_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} , \quad \alpha_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} , \quad \alpha_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$$

is linearly independent. Indeed

$$\begin{aligned} \lambda_1\alpha_1 + \lambda_2\alpha_2 + \lambda_3\alpha_3 &= \lambda_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 + \lambda_3 \\ \lambda_2 + \lambda_3 \\ \lambda_1 - \lambda_3 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} . \end{aligned}$$

This implies $\lambda_1 = \lambda_2 = \lambda_3 = 0$.

Definition 4.9. The set

$$< \alpha_1, \ldots, \alpha_n >:= \left\{ \lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_n\alpha_n \,\middle|\, \lambda_1, \ldots, \lambda_n \in \mathbb{R} \right\}$$

is called the *span* of $\alpha_1, \ldots, \alpha_n$. It is the set of all linear combinations of $\alpha_1, \ldots, \alpha_n$. The dimension of $V :=< \alpha_1, \ldots, \alpha_n >$ is the maximal number m such that there are sets of m linearly independent vectors in V. Such a set is called a *basis* of V.

Remark 4.10. If $\{\alpha_1, \ldots, \alpha_n\}$ is a basis of the n-dimensional vector space V, then for every vector $x \in V$ there exist $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that

$$x = \sum_{i=1}^{n} \lambda_i\alpha_i = \lambda_1\alpha_1 + \ldots + \lambda_n\alpha_n .$$

The dimension of $\mathbb{R}^n$ is n and the unit-vectors

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix} , \; e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} , \ldots , \; e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \tag{4.5}$$

form a basis of $\mathbb{R}^n$. It is called the *standard basis* of $\mathbb{R}^n$. The vector

$$\alpha = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

is written in the standard basis of $\mathbb{R}^n$ and

$$\alpha = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ \vdots \\ a_n \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots + a_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \sum_{i=1}^{n} a_i e_i \,.$$

Let A be the matrix whose columns are the vectors $\alpha_1, \ldots, \alpha_n$, i.e.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} \alpha_1, & \alpha_2, & \ldots, & \alpha_n \end{pmatrix}$$

Then

$$\alpha_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix}$$

is the image of e_i under the mapping

$$\begin{array}{rccl} f: & \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ & x & \longmapsto & f(x) := Ax \end{array}$$

and the image of $\mathbb{R}^n$ under f is

$$\begin{aligned} \mathrm{im}(f) &:= \left\{ y \in \mathbb{R}^m \,\middle|\, \exists x \in \mathbb{R}^n \,,\; f(x) = Ax = y \right\} \\ &= < \alpha_1, \ldots, \alpha_n > \;. \end{aligned}$$

Therefore the dimension of the image is the dimension of $< \alpha_1, \ldots, \alpha_n >$. Moreover $\dim(\mathrm{im}(f)) \leqslant n$.

4.3.3 Addition and product of matrices

Given two $m \times n$-matrices $A = (a_{ij})$ and $B = (b_{ij})$ we define the sum $A + B$ to be

$$A + B = (a_{ij} + b_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} \,.$$

Let $\alpha \in \mathbb{R}$ be a scalar then

$$\alpha A = (\alpha\, a_{ij})_{\substack{i=1,\ldots,m\\ j=1,\ldots,n}}\,.$$

Definition 4.11. The product of the $m \times l$-matrix

$$A = (a_{ik})_{\substack{i=1,\ldots,m\\ k=1,\ldots,l}} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1l} \\ a_{21} & a_{22} & \cdots & a_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{ml} \end{pmatrix}$$

and the $l \times n$-matrix

$$B = (b_{kj})_{\substack{k=1,\ldots,l\\ j=1,\ldots,n}} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{l1} & b_{l2} & \cdots & b_{ln} \end{pmatrix}$$

is the $m \times n$-matrix $AB = C$

$$(c_{ij})_{\substack{i=1,\ldots,m\\ j=1,\ldots,n}} = C = AB = \Big(\sum_{k=1}^{l} a_{ik} b_{kj}\Big)_{\substack{i=1,\ldots,m\\ j=1,\ldots,n}}\,.$$

The component c_{ij} in AB is given by the scalar product of the i-th row of A with the j-th column of B.

Remark 4.12. In general

$$A\,B \neq B\,A\,.$$

The $n \times n$-identity is the matrix $\mathrm{Id}_n = (c_{ij})_{i,j=1,\ldots,n}$ with

$$\mathrm{Id}_n\, A = A\,\mathrm{Id}_n = A\,.$$

We have

$$c_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j\,, \end{cases}$$

$$\mathrm{Id}_n := \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Given a $n \times n$-matrix A. If there is a $n \times n$-matrix B with $BA = \mathrm{Id}_n$, then

$$AB = BA = \mathrm{Id}_n$$

and $B = A^{-1}$ is called the *inverse* of A.
A method to compute the inverse of a matrix is the following. It is much easier to show on an example than to explain. Given a matrix

$$\begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}$$

that we want to invert. We write

$$\begin{array}{c|c}
\begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
\begin{pmatrix} 3 & 1 \\ 3 & \frac{6}{5} \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & \frac{3}{5} \end{pmatrix} \\
\begin{pmatrix} 3 & 1 \\ 0 & \frac{1}{5} \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ -1 & \frac{3}{5} \end{pmatrix} \\
\begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ -5 & 3 \end{pmatrix} \\
\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 6 & -3 \\ -5 & 3 \end{pmatrix} \\
\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix}
\end{array}$$

Now the inverse of

$$\begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}$$

is

$$\begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix}.$$

4.3.4 The system of equations

We consider the equation $Ax = b$ with

$$A := \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \qquad x := \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad b := \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Then we multiply

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

from the left with $\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$ and get

$$\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

$$\begin{pmatrix} a_{11}a_{22} - a_{12}a_{21} & 0 \\ 0 & a_{11}a_{22} - a_{12}a_{21} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{22}b_1 - a_{12}b_2 \\ -a_{21}b_1 + a_{11}b_2 \end{pmatrix}.$$

This is equivalent to

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = a_{22}b_1 - a_{12}b_2$$
$$(a_{11}a_{22} - a_{12}a_{21})x_2 = -a_{21}b_1 + a_{11}b_2\,.$$

Herewith we get the results of section 4.2.2.

4.3.5 The transpose of a matrix

Definition 4.13. We define the *transpose* A^{t} of the matrix

$$A := (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

to be the matrix

$$A^{\mathrm{t}} := (a_{ji})_{\substack{j=1,\dots,n \\ i=1,\dots,m}}.$$

It follows immediately that

$$(A^{\mathrm{t}})^{\mathrm{t}} = A\,.$$

The transpose of A is obtained in exchanging the rows with the columns. The i-th row of A is the i-th colum of A^{t} and the j-th column of A is the j-th row of A^{t}.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{\mathrm{t}} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

Given the column vector

$$a = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_i \\ \vdots \\ \alpha_n \end{pmatrix} ,$$

the transpose a^{t} of a is the row vector

$$a^{\mathrm{t}} = (\alpha_1 \,, \, \ldots \,, \, \alpha_i \,, \, \ldots \,, \, \alpha_n) \,.$$

We define the scalar product of a and b,

$$b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_n \end{pmatrix}$$

to be

$$\begin{aligned} \langle a, b \rangle := a^{\mathrm{t}}\, b = (\alpha_1 \,, \, \ldots \,, \, \alpha_i \,, \, \ldots \,, \, \alpha_n) & \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_n \end{pmatrix} \\ &= \alpha_1\beta_1 + \cdots + \alpha_i\beta_i + \cdots + \alpha_n\beta_n \\ &= \sum_{i=1}^{n} \alpha_i\beta_i \,. \end{aligned}$$

4.3.6 Determinant

Let A be a $n \times n$-matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} \alpha_1, & \alpha_2, & \ldots, & \alpha_n \end{pmatrix}$$

The determinant $\det(A)$ of a $n \times n$-matrix A is a function that maps the matrix, the set of its n column-vectors to an element of the underlying field ($\mathbb{R}$). This function satisfies the following properties for $1 \leqslant i, j \leqslant n$

i) For $a \in \mathbb{R}$

$$\det(\alpha_1, \ldots, a\,\alpha_i, \ldots, \alpha_n) = a \cdot \det(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_n)$$

ii) With $\alpha_i = \alpha'_i + \alpha''_i$

$$\det(\alpha_1, \ldots, \alpha'_i + \alpha''_i, \ldots, \alpha_n) \\ = \det(\alpha_1, \ldots, \alpha'_i, \ldots, \alpha_n) + \det(\alpha_1, \ldots, \alpha''_i, \ldots, \alpha_n)\,.$$

iii) Let $\alpha_i = \alpha_j$ for $i \neq j$. Then

$$\det(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_j, \ldots, \alpha_n) = 0\,.$$

iv) Let $e_1 \ldots, e_n$ be the unit-vectors (4.5), then

$$\det(e_1, \ldots, e_n) = 1\,.$$

The determinant has the following important properties.

Theorem 4.14. *The determinant of a $n \times n$-matrix*

$$(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_n)$$

is 0 if and only if the vectors $\alpha_1, \ldots, \alpha_n$ are linearly dependent. For the transpose of A we have

$$\det A = \det A^{\mathrm{t}}\,.$$

For the product of the $n \times n$-matrices A and B

$$\det(AB) = \det A \cdot \det B\,.$$

Let $\alpha \in \mathbb{R}$ be a scalar, then

$$\det(\alpha\, A) = \alpha^n \det A\,.$$

Example.

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc\,.$$

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}\,.$$

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{array}{l} a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} \\ -a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{23}a_{32}a_{11}\,. \end{array}$$

For $n \times n$-matrices there is an analogous formula for developing by columns or rows. For a $n\times n$-matrix $A = (a_{ij})$ we define A_{ij} to be the $(n-1)\times(n-1)$-matrix that we get from A if we take off the ith row and the jth column.

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix}$$

$$A_{ij} = \begin{pmatrix} a_{11} & \cdots & a_{1(j-1)} & a_{1(j+1)} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{(i-1)1} & \cdots & a_{(i-1)(j-1)} & a_{(i-1)(j+1)} & \cdots & a_{(i-1)n} \\ a_{(i+1)1} & \cdots & a_{(i+1)(j-1)} & a_{(i+1)(j+1)} & \cdots & a_{(i+1)n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n(j-1)} & a_{n(j+1)} & \cdots & a_{nn} \end{pmatrix}$$

Then developing by the jth column

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$$

or by the ith row

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}\,.$$

Now, we use this formula to compute the determinant of a 3×3-matrix. We develop by the first column

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= (-1)^{1+1} a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} \\ &\quad + (-1)^{2+1} a_{21} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} \\ &\quad + (-1)^{3+1} a_{31} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix}\,. \end{aligned}$$

In analogy one can develop by any other column or a row. With the second row we get

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = (-1)^{2+1} a_{21} \det\begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} + (-1)^{2+2} a_{22} \det\begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{pmatrix} + (-1)^{2+3} a_{23} \det\begin{pmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{pmatrix}.$$

The result will be the same in both cases.

Example. We compute the determinant

$$\det\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$$

Since the second column contains two zeros, we choose

$$\det\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix} = 0 + (-1)^{2+2} \cdot 1 \cdot \det\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} + 0 = -1 - 1 = -2.$$

Example. We compute the determinant

$$\det\begin{pmatrix} -1 & 2 & 1 \\ 1 & 5 & 3 \\ 2 & 0 & 0 \end{pmatrix}$$

Since the third row contains two zeros, we choose

$$\det\begin{pmatrix} -1 & 2 & 1 \\ 1 & 5 & 3 \\ 2 & 0 & 0 \end{pmatrix} = (-1)^{3+1} \cdot 2 \cdot \det\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} + 0 + 0 = 2(6-5) = 2.$$

4.4 Eigenvalues and eigenvectors

4.4.1 Change of basis

We defined the standard-basis of $\mathbb{R}^n$

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Definition 4.15. A maximal set of linearly independent vectors $\tau_1, \ldots, \tau_n$ in a vector space is called *basis*.

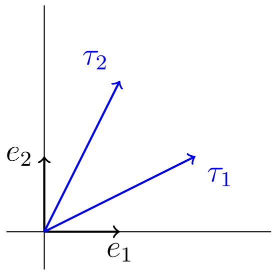
We illustrate the following explanations with an example.
If $\tau_1, \ldots, \tau_n$ are a basis of the vector space $\mathbb{R}^n$, then for any vector $\alpha \in \mathbb{R}^n$, we can find scalars $a_1, \ldots, a_n$ such that

$$\alpha = a_1\tau_1 + \ldots + a_n\tau_n \,.$$

The vectors

$$\tau_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \text{and} \quad \tau_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

are linearly independent since they are not parallel. They form a basis of $\mathbb{R}^2$.



Indeed we check that

$$e_1 = \frac{2}{3}\tau_1 - \frac{1}{3}\tau_2$$
$$e_2 = -\frac{1}{3}\tau_1 + \frac{2}{3}\tau_2$$

and this also shows that τ_1, τ_2 form a basis since e_1 and e_2 form a basis.

We consider the linear mapping

$$\begin{aligned} f: \ \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ x &\longmapsto f(x) := Ax \,, \end{aligned}$$

where A is a $n \times n$-matrix. The matrix A corresponding to f depends on the choice of the basis of $\mathbb{R}^n$. We know that the jth column α_j of A is the image $f(e_j)$ of the jth basis vector e_j, i.e.,

$$\alpha_j = f(e_j) = \sum_{i=1}^{n} a_{ij} e_i \,.$$

We define the mapping

$$\begin{array}{rccc} f: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & x & \longmapsto & f(x) \end{array}$$

by

$$f(e_1) := 2e_1 + 7e_2\,, \quad f(e_2) := e_1 + 4e_2\,.$$

This defines the mapping because f is linear, i.e., $f(x_1e_1 + x_2e_2) = x_1f(e_1) + x_2f(e_2)$.
In the basis e_1, e_2 we can write f as $x \mapsto f(x) := Ax$ with

$$A = \begin{pmatrix} 2 & 1 \\ 7 & 4 \end{pmatrix}$$

since

$$e_1 := \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 2 \\ 7 \end{pmatrix} = 2e_1 + 7e_2$$

$$e_2 := \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 4 \end{pmatrix} = e_1 + 4e_2$$

Let $\tau_1, \ldots, \tau_n$ be another basis of $\mathbb{R}^n$. It is possible to express the mapping f in the basis $\tau_1, \ldots, \tau_n$ by a matrix B. The jth column β_j of B is the image $f(\tau_j)$ of τ_j written as a linear combination of the $\tau_1, \ldots, \tau_n$:

$$\beta_j = \sum_{k=1}^{n} b_{kj}\tau_k\,.$$

Let τ_k be written in the basis $e_1, \ldots, e_n$, i.e.,

$$\tau_k = \sum_{i=1}^{n} t_{ik}e_i = \begin{pmatrix} t_{1k} \\ \vdots \\ t_{nk} \end{pmatrix}\,.$$

Then

$$\begin{aligned} \beta_j &= \sum_{k=1}^{n} b_{kj}\tau_k = \sum_{k=1}^{n} b_{kj}\left(\sum_{i=1}^{n} t_{ik}e_i\right) \\ &= \sum_{i=1}^{n}\sum_{k=1}^{n} t_{ik}b_{kj}e_i \end{aligned}$$

is the representation of β_j in the standard basis and

$$\begin{aligned} f(\tau_j) &= f\left(\sum_{k=1}^{n} t_{kj} e_k\right) = \sum_{k=1}^{n} t_{kj} f(e_k) \\ &= \sum_{k=1}^{n} t_{kj} \left(\sum_{i=1}^{n} a_{ik} e_i\right) \\ &= \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} t_{kj} e_i \,. \end{aligned}$$

We now want to express the mapping f in the basis $\tau_1 := 2e_1 + e_2$ and $\tau_2 := e_1 + 2e_2$. Using their representation in the standard basis, we determine the image of τ_1 and τ_2.

$$\begin{aligned} f(\tau_1) = A\tau_1 = \begin{pmatrix} 2 & 1 \\ 7 & 4 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 18 \end{pmatrix} = 5e_1 + 18e_2 \\ f(\tau_2) = A\tau_2 = \begin{pmatrix} 2 & 1 \\ 7 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 15 \end{pmatrix} = 4e_1 + 15e_2 \end{aligned}$$

The rows in the matrix B that describes f in the basis τ_1, τ_2, are the images $f(\tau_1)$, $f(\tau_2)$ of these vectors written in the basis τ_1, τ_2. Hence we determine $b_{11}, b_{21}, b_{12}, b_{22} \in \mathbb{R}$ such that

$$\begin{aligned} f(\tau_1) &= b_{11}\tau_1 + b_{21}\tau_2 \\ f(\tau_2) &= b_{12}\tau_1 + b_{22}\tau_2 \end{aligned}$$

We therefore solve

$$\begin{aligned} \begin{pmatrix} 5 \\ 18 \end{pmatrix} &= b_{11} \begin{pmatrix} 2 \\ 1 \end{pmatrix} + b_{21} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ \begin{pmatrix} 4 \\ 15 \end{pmatrix} &= b_{12} \begin{pmatrix} 2 \\ 1 \end{pmatrix} + b_{22} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \end{aligned}$$

and get

$$f(\tau_1) = -\frac{8}{3}\tau_1 + \frac{31}{3}\tau_2$$
$$f(\tau_2) = -\frac{7}{3}\tau_1 + \frac{26}{3}\tau_2$$

Hence the matrix B is

$$B = \begin{pmatrix} -\frac{8}{3} & -\frac{7}{3} \\ \frac{31}{3} & \frac{26}{3} \end{pmatrix}$$

Let T be the matrix with column-vectors $\tau_1, \ldots, \tau_n$, i.e.,

$$T = (\tau_1\,,\,\ldots\,,\,\tau_n) = \begin{pmatrix} t_{11} & \cdots & t_{1j} & \cdots & t_{1n} \\ \vdots & & \vdots & & \vdots \\ t_{i1} & \cdots & t_{ij} & \cdots & t_{in} \\ \vdots & & \vdots & & \vdots \\ t_{n1} & \cdots & t_{nj} & \cdots & t_{nn} \end{pmatrix}.$$

Since $\tau_1, \ldots, \tau_n$ are a basis and herewith linearly independent, we know that T is invertible and $\det T \neq 0$. According to the definition of the product of matrices, the entry in the ith row and the jth column of the product TB is

$$(TB)_{ij} = \sum_{k=1}^{n} t_{ik} b_{kj}$$

and at the same place in the product AT we find

$$(AT)_{ij} = \sum_{k=1}^{n} a_{ik} t_{kj}\,.$$

Hence

$$\beta_j = \sum_{i=1}^{n} (TB)_{ij} e_i$$
$$f(\tau_j) = \sum_{i=1}^{n} (AT)_{ij} e_i\,.$$

Since $\beta_j = f(\tau_j)$ we have $(TB)_{ij} = (AT)_{ij}$, $i,j = 1, \ldots, n$. We get the relation

$$TB = AT\,.$$

Then the matrix B that describes the mapping f in the basis $\tau_1, \dots, \tau_n$ is given by

$$B = T^{-1}\,A\,T$$

and herewith

$$T\,B\,T^{-1} = A\,.$$

Here we use that in general the product of $n \times n$-matrices is not commutative, i.e., $MN \neq NM$.

In the example the matrix T is

$$T = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

and its inverse is

$$T^{-1} = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

Hence

$$AT = \begin{pmatrix} 2 & 1 \\ 7 & 4 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 4 \\ 18 & 15 \end{pmatrix}$$

and

$$TB = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} -\frac{8}{3} & -\frac{7}{3} \\ \frac{31}{3} & \frac{26}{3} \end{pmatrix} = \begin{pmatrix} 5 & 4 \\ 18 & 15 \end{pmatrix}$$

You might check that $B = T^{-1}AT$.

Theorem 4.16. *The determinant is invariant under the change of basis.*

Proof. Since $\det(AB) = \det A \det B$ we have

$$\det A' = \det(T^{-1}\,A\,T) = \det T^{-1}\,\det A\,\det T = \det A$$

where the last equation follows from

$$1 = \det \mathrm{Id}_n = \det T^{-1}\,\det T$$

which is equivalent to

$$\det T^{-1} = \frac{1}{\det T}\,.$$

□

Example. In the standard basis e_1, e_2, e_3 the mapping

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$

is given by

$$x := \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto Ax := \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

We choose a new basis

$$\tau_1 := e_1 - e_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \tau_2 := e_2 - e_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \tau_3 := e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Then the transformation from the standard basis to the basis $\{\tau_1, \tau_2, \tau_3\}$ is given by the matrix

$$T := \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}.$$

We have $\det T = 1$ and in the basis $\{\tau_1, \tau_2, \tau_3\}$ the mapping f is given by

$$x := \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto T^{-1}AT\,x\,.$$

Here

$$T^{-1} := \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

and

$$\begin{aligned} T^{-1}AT &= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & 2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 0 & -2 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}. \end{aligned}$$

4.4.2 Scalings

We know that a linear mapping maps the vector space $\mathbb{R}^n$ to a subspace of $\mathbb{R}^m$. In this section we study some examples of linear mappings.
The first example is the scaling. The scaling with a factor λ along the direction $< v >$ maps v to λv. Given a mapping

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

If there exist n linearly independent vectors $v_1, \ldots, v_n$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that

$$f(v_i) = \lambda_i v_i \,, \qquad i = 1, \ldots n,$$

then f acts on $< v_i >$ as a scaling by the factor λ_i.
In the basis $v_1, \ldots, v_n$ the matrix representation of f is

$$A' = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} .$$

Let

$$T = (\, v_1 \,, \, \ldots \,, \, v_n \,)$$

be the matrix with columns $v_1, \ldots, v_n$. Then T gives the change of basis from the standard basis to $v_1, \ldots, v_n$. In the standard basis the mapping f is given by the matrix

$$A = T \, A' \, T^{-1} \,.$$

In the next section we see how to find for a given matrix A the vectors $v_1, \ldots, v_n$ and the factors $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.

4.4.3 Eigenvalues and eigenvectors

Let $f : \mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto f(x) := Ax$. For the $n \times n$-matrix A we consider the equation

$$Av = \lambda v \,,$$

i.e. we search for factors $\lambda \in \mathbb{R}$ and vectors $v \in \mathbb{R}^n$ such that the equation $Av = \lambda v$ holds. Then the mapping f acts as a scaling on v and on every vector on the line

$$\mathcal{L}_v = \{ av \mid a \in \mathbb{R} \}$$

defined by v i.e. for every $w \in \mathcal{L}_v$ we have

$$f(w) = Aw = \lambda w \,.$$

We now describe a method to find such factors λ and the corresponding vectors $v \in \mathbb{R}^n$. We have

$$A\,v = \lambda\,v \quad \Longleftrightarrow \quad (A - \lambda\,\mathrm{Id}_n)\,v = 0\,.$$

Of course this is always true for $v = 0$. There exists $v \neq 0$ with $(A - \lambda\,\mathrm{Id}_n)\,v = 0$ if and only if $\det(A - \lambda \mathrm{Id}_n) = 0$. Hence we do the following steps.

I. Compute the determinant $\det(A - \lambda \mathrm{Id}_n)$ that is a polynomial $p_A(\lambda)$ in λ.

II. Determine the zeros $\lambda_1, \ldots, \lambda_n$ of $p_A(\lambda)$.

III. For every λ_i, $i = 1, \ldots, n$, solve the equation

$$A v_i = \lambda_i v_i\,.$$

Definition 4.17. The *eigenvalues* of a matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ are the scalars λ that satisfy the equation

$$\det(A - \lambda\,\mathrm{Id}_n) = 0\,.$$

A vector that satisfies the equation

$$(A - \lambda\,\mathrm{Id}_n)\,v = 0$$

for a given matrix A and an eigenvalue λ of A, is called *eigenvector* of A to the eigenvalue λ.

The eigenvalues of the $n \times n$-matrix A are the solutions of the equation

$$p_A(\lambda) = \det(A - \lambda\,\mathrm{Id}_n) = 0\,.$$

This equation is a polynomial in λ of degree n.

Definition 4.18. Given a matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$, the polynomial

$$p_A(\lambda) = \det(A - \lambda\,\mathrm{Id}_n) = 0\,.$$

is called the *characteristic polynomial* of A.

Example. We compute the eigenvalues and eigenvectors of the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}$$

I. The characteristic polynomial $p_A(\lambda)$ is

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda \, \mathrm{Id}) \\ &= \det \begin{pmatrix} 1-\lambda & 1 \\ 1 & 3-\lambda \end{pmatrix} = (1-\lambda)(3-\lambda) - 1 \\ &= \lambda^2 - 4\lambda + 2 \end{aligned}$$

II. The zeros of the characteristic polynomial $p_A(\lambda) = \lambda^2 - 4\lambda + 2$ are

$$\lambda_{1,2} = \frac{4 \pm \sqrt{16-8}}{2} = 2 \pm \sqrt{2}$$

$$\lambda_1 = 2 + \sqrt{2}\,, \quad \lambda_2 = 2 - \sqrt{2}\,.$$

III. We first determine the eigenvector v_1 to the eigenvalue $\lambda_1 = 2+\sqrt{2}$. We solve $Av_1 = \lambda_1 v_1$ that is equivalent to $(A - \lambda_1 \, \mathrm{Id})v_1 = 0$. We compute

$$\begin{aligned} A - \lambda_1 \, \mathrm{Id} &= \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} - \begin{pmatrix} 2+\sqrt{2} & 0 \\ 0 & 2+\sqrt{2} \end{pmatrix} \\ &= \begin{pmatrix} -1-\sqrt{2} & 1 \\ 1 & 1-\sqrt{2} \end{pmatrix} \end{aligned}$$

For $v_1 = \begin{pmatrix} x \\ y \end{pmatrix}$,

$$\begin{aligned} (A - \lambda_1 \, \mathrm{Id})v_1 &= \begin{pmatrix} -1-\sqrt{2} & 1 \\ 1 & 1-\sqrt{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \begin{pmatrix} (-1-\sqrt{2})x + y \\ x + (1-\sqrt{2})y \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned}$$

hence

$$\begin{aligned} (-1-\sqrt{2})x + y &= 0 \\ x + (1-\sqrt{2})y &= 0 \end{aligned}$$

and, since $(-1-\sqrt{2})(1-\sqrt{2}) = 1$, this system is equivalent to

$$y = (1+\sqrt{2})x$$

and we choose

$$v_1 = \begin{pmatrix} 1 \\ 1+\sqrt{2} \end{pmatrix}\,.$$

We now determine the eigenvector v_2 to the eigenvalue $\lambda_2 = 2 - \sqrt{2}$. We solve $(A - \lambda_2 \,\mathrm{Id})v_2 = 0$. We compute

$$\begin{aligned} A - \lambda_2 \,\mathrm{Id} &= \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} - \begin{pmatrix} 2-\sqrt{2} & 0 \\ 0 & 2-\sqrt{2} \end{pmatrix} \\ &= \begin{pmatrix} -1+\sqrt{2} & 1 \\ 1 & 1+\sqrt{2} \end{pmatrix} \end{aligned}$$

For $v_2 = \begin{pmatrix} x \\ y \end{pmatrix}$,

$$\begin{aligned} (A - \lambda_2 \,\mathrm{Id})v_2 &= \begin{pmatrix} -1+\sqrt{2} & 1 \\ 1 & 1+\sqrt{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \begin{pmatrix} (-1+\sqrt{2})x + y \\ x + (1+\sqrt{2})y \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned}$$

and, since $(-1+\sqrt{2})(1+\sqrt{2}) = 1$, this system is equivalent to

$$y = (1 - \sqrt{2})x$$

and we choose

$$v_2 = \begin{pmatrix} 1+\sqrt{2} \\ -1 \end{pmatrix}$$

Let

$$\begin{aligned} f:\ \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ x &\longrightarrow Ax \end{aligned}$$

be the mapping that is defined by A in the standard basis. Then, in the basis of eigenvectors v_1, v_2, the mapping f is given by

$$B = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} 2+\sqrt{2} & 0 \\ 0 & 2-\sqrt{2} \end{pmatrix}.$$

The transformation matrix is

$$T = \begin{pmatrix} 1 & 1+\sqrt{2} \\ 1+\sqrt{2} & -1 \end{pmatrix}$$

and its inverse is

$$T^{-1} = \frac{1}{-4-2\sqrt{2}} \begin{pmatrix} -1 & -1-\sqrt{2} \\ -1-\sqrt{2} & 1 \end{pmatrix}.$$

It is now easy to check that

$$B = T^{-1}AT.$$

Definition 4.19. The *algebraic multiplicity* of the eigenvalue λ_i of A is the k that is maximal with

$$(\lambda - \lambda_i)^k \text{ is a factor of } p_A(\lambda) = \det(A - \lambda \operatorname{Id}_n)\,.$$

The *geometric multiplicity* is the dimension of the subspace spanned by the eigenvectors of A to the eigenvalue λ.

Remark 4.20. The geometric multiplicity of an eigenvalue is less or equal to the algebraic multiplicity of the eigenvalue.

4.5 Applications of linear algebra

4.5.1 Systems of linear differential equations

Definition 4.21. A system of first order differential equations for the functions $t \to x_1(t)$, $t \to x_2(t)$ consists of two differential equations of the form

$$\begin{aligned} x_1' &= f_1(t, x_1, x_2) \\ x_2' &= f_2(t, x_1, x_2) \end{aligned}$$

where f_1 and f_2 are two functions in the three variables t, x_1, x_2. A solution of the system is a pair of functions

$$t \to x_1(t), \quad t \to x_2(t)$$

that satisfies both equations simultaneously.

Example. Consider the differential equation of order 2

$$x'' = F(t, x, x')\,.$$

If we introduce the new functions $t \to x_1(t) = x(t)$ and $t \to x_2(t) = x'(t)$, we get a system of first order differential equations

$$\begin{aligned} x_1' &= x_2 \\ x_2' &= F(t, x_1, x_2) \end{aligned}$$

Hence it is possible to consider systems of first order differential equations instead of differential equations of higher order.

Linear autonomous system of differential equations with constant coefficients

We search for the functions $x_1(t)$ and $x_2(t)$ that satisfy

$$\begin{aligned}\dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + b_1\\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + b_2\end{aligned}$$

In this system the functions $f_1(x_1(t), x_2(t))$ and $f_2(x_1(t), x_2(t))$ do not depend explicitly on the variable t and they are linear. We put

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}, \quad \dot{x}(t) = \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Hence the system is

$$\dot{x}(t) = Ax(t) + b\,.$$

We choose to solve an example with $b = 0$.

Example. Consider the system

$$\begin{aligned}\dot{x}_1 &= x_1 + 2x_2\\ \dot{x}_2 &= 2x_1 + x_2\,.\end{aligned}$$

It corresponds to $\dot{x} = Ax$ with

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

We transform this matrix in a triangular or diagonal matrix. The eigenvalues of A are the zeros of

$$\begin{aligned}p_A(\lambda) &= \det(A - \lambda \mathrm{Id}) = (1-\lambda)^2 - 4 = \lambda^2 - 2\lambda - 3\\ &= (\lambda+1)(\lambda-3)\,.\end{aligned}$$

Hence the eigenvalues are $\lambda_1 = -1$ and $\lambda_2 = 3$. The eigenvectors v_i, $i = 1, 2$, satisfy $Av_i = \lambda_i v_i$. Solving this equation we get

$$v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

With the transformation

$$T = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

we get

$$B = T^{-1}AT = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix} .$$

The solution of the system

$$\begin{aligned} \dot{y}_1 &= -y_1 \\ \dot{y}_2 &= 3y_2 \end{aligned}$$

is $y_1 = C_1 e^{-t}$ and $y_2 = C_2 e^{3t}$. Since $x_1 = y_1 + y_2$ and $x_2 = -y_1 + y_2$ we get

$$\begin{aligned} x_1 &= y_1 + y_2 = C_1 e^{-t} + C_2 e^{3t} \\ x_2 &= -y_1 + y_2 = -C_1 e^{-t} + C_2 e^{3t} . \end{aligned}$$

These functions are a solution of the system of differential equations.

4.5.2 Local extrema

Definition 4.22. Let $U \subseteq \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$ a function. A point $x \in U$ is called *local maximum* of f if an environment $V \subset U$ of x exists with

$$f(x) \geqslant f(y) \quad \text{for all } y \in V .$$

A point $x \in U$ is called *local minimum* of f if an environment $V \subset U$ of x exists with

$$f(x) \leqslant f(y) \quad \text{for all } y \in V .$$

Definition 4.23. Let $U \subseteq \mathbb{R}^n$ be an open set and let

$$\begin{aligned} f : \quad U &\to \mathbb{R} \\ (x_1, \ldots, x_n) &\mapsto f(x_1, \ldots, x_n) \end{aligned}$$

be a function whose first and second partial derivatives exist and are continuous. The *Hessian matrix* of f in $x \in U$ is the $n \times n$-matrix

$$(\operatorname{Hess} f)(x) := \left(\frac{\partial^2}{\partial x_i \partial x_j} f(x) \right)_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}} .$$

This matrix is symmetric since for $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant n$

$$\frac{\partial^2}{\partial x_i \partial x_j} f(x) = \frac{\partial^2}{\partial x_j \partial x_i} f(x)\,.$$

Theorem 4.24. *Let $U \subset \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$ a partial differentiable function. If f has a local extremum in the point x (i.e. a local maximum or a local minimum), then*

$$\nabla f(x) = 0\,.$$

Theorem 4.25. *Let $U \subset \mathbb{R}^n$ be open, $f : U \to \mathbb{R}$ a function whose first and second partial derivatives exist and are continuous. Let $x \in U$ with*

$$\nabla f(x) = 0\,.$$

- *If $(\operatorname{Hess} f)(x)$ is positive definite, then f has a local minimum in x.*
- *If $(\operatorname{Hess} f)(x)$ is negative definite, then f has a local maximum in x.*
- *If $(\operatorname{Hess} f)(x)$ is indefinite, then f doesn't have a local extremum in x.*

In the other cases there may not be a local extremum.

Definition 4.26. A symmetric matrix A is

- *positive definite* if all its eigenvalues are positive,
- *negative definite* if all its eigenvalues are negative,
- *indefinite* if there is at least one positive and one negative eigenvalue.

Example. i) The function $f(x, y) := c + x^2 + y^2$ has a local minimum in $(0, 0)$ since $\nabla f(0, 0) = (0, 0)$ and the Hessian

$$(\operatorname{Hess} f)(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

is positive definite.

ii) The function $g(x, y) := c - x^2 - y^2$ has a local maximum in $(0, 0)$ since $\nabla g(0, 0) = (0, 0)$ and the Hessian

$$(\operatorname{Hess} g)(0,0) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$

is negative definite.

iii) The function $h(x, y) := c + x^2 - y^2$ satisfies $\nabla h(0, 0) = (0, 0)$ and the Hessian

$$(\operatorname{Hess} h)(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

is indefinite. The function has a saddle point in $(0, 0)$.

Bibliography

[1] Tom M. Apostol, *Calculus, Volume 1, One-Variable Calculus with an Introduction to Linear Algebra*, 2nd edition, Wiley.

[2] Tom M. Apostol, *Multi-Variable Calculus and Linear Algebra with Applications*, 2nd edition, Wiley.

[3] Serge Lang, *A First Course in Calculus*, 5th edition, Springer New York, 1986 (Online resource)

[4] Serge Lang, *Introduction to Linear Algebra* 2nd edition, Springer New York, 1986 (Online resource)

[5] William L. Briggs, Lyle Cochran, Bernard Gillett, *Calculus, Early transcendentals*, 2nd edition, 2014, Pearson Education Limited.

[6] Ulrich L. Rohde, *Introduction to differential calculus : Systematic studies with engineering applications for beginners*, 2012, Wiley (Online resource)

[7] Ulrich L. Rohde, *Introduction to integral calculus : Systematic studies with engineering applications for beginners*, 2012, Wiley (Online resource)

[8] George B. Thomas, Maurice D. Weir, Joel Hass *Thomas' Calculus : Early Transcendentals : Single Variable*, 13th ed., 2014, Pearson.

[9] George B. Thomas, Maurice D. Weir, Joel Hass *Thomas' Calculus : Multivariable*, 13th ed., 2014, Pearson.