Preparatory Online Course in Mathematics

Course version: MFR-TUB(10000)
Course date: 09/2016
Local version: DE-MINT
Course variant: std
admin@ve-und-mint.de

The course was developed within the scope of the VE&MINT project.

VE&MINT is a project cooperation of MINT-Kolleg Baden-Württemberg with VEMINT-Konsortium, Leibniz Universität Hannover and Technische Universität Berlin which seeks to offer a preparatory mathematical online course that is freely accessible nationwide based on a free licence, and more generally to promote exchange of instructional material as well as software between its sites.

For further information on the online course and the project, see www.ve-und-mint.de.

The English language/bilingual version of the course was developed at the TU Berlin with the specific goal to create a course accessible and useful for refugees.

Inhaltsverzeichnis

1 Elementary Arithmetic

Module Overview

This module covers the mathematical basics of elementary arithmetic and introduces and explains the notation used throughout this online course.

1.1 Numbers, Variables, Terms

1.1.1 Introduction

Mathematics is a science in which abstract structures and the logical relations between them are investigated. Before we examine the actual subjects of this section in more detail, we shall refer briefly to the fundamental notion of a **set**.

Info1.1.1

We will often be making statements about a number of structurally similar objects. To do so in a compact manner, we can gather such objects into sets that serve as containers for the objects. Let the objects be denoted by $a, b, c, \ldots$. Then the symbol $M = \{a; b; c; \ldots\}$ denotes the set M which has the previously listed objects as its **elements**. The latter statement is written in short as $a \in M$, $b \in M$, $c \in M$ etc; thus the symbol "$\in$" reads "is an element of". (Sometimes it is more convenient to reverse the order of element and set. To do so, we can reverse the symbol $\in$, as in $M \ni a$, $M \ni b$, $M \ni c$. The meaning is the same as before, and the reversed symbol "$\ni$" reads "contains as an element" or just "contains".)

Apart from the list notation for sets, other notations exist. If, for example, the elements have to satisfy a condition B, then this is written in the form $T = \{x : x \text{ satisfies } B\}$. If x is taken (explicitly) from a more comprehensive set U, then this is also written in the form $T = \{x : x \in U \text{ and } x \text{ satisfies } B\}$ or, in short, $T = \{x \in U : x \text{ satisfies } B\}$.

Statements like "$x \in U$" or "x satisfies B" are statements in a mathematical sense, i.e. we can assign them a unique truth value "true" or "false". Let A_1 and A_2 be two such statements. In case both A_1 **and** A_2 hold, we write $A_1 \wedge A_2$. In case only one of the two statements needs to hold, i.e. A_1 **or** A_2 **or** both, we write $A_1 \vee A_2$.

For two sets M and N we write:

- $M \subseteq N$, i.e. M is a (possibly improper) **subset** of N, if every element of M is also an element of N. If at least one element of N exists that is not in M, we say that M is a proper subset of N. In this case we can also write $M \subset N$.
- $M \cup N$ for the **union** of the two sets. The union denotes the set containing all elements that are contained in at least one of the two sets.
- $M \cap N$ for the **intersection** of the two sets. The intersection denotes the set containing all elements that are contained in both of the two sets.

- $N \setminus M$ for the **complement**, i.e. the set containing all elements of N that are *not* contained in M.

Thus the union mentioned above is characterised by elements that satisfy the condition $(x \in M) \vee (x \in N)$. However, for the elements of the intersection $(x \in M) \wedge (x \in N)$ holds. In contrast, the complement above contains elements for which $(x \in N) \wedge (x \notin M)$ holds. The symbol $\notin$ denotes the negation of the element statement.

To a large extent, mathematics is concerned with the universe of numbers:

$$\ldots; 0; -3; 4; \frac{4}{5}; \sqrt{2}; e; \pi; 12.3; 10^{23}; \ldots\ .$$

However, considering different numbers in more detail reveals fundamental differences. Some numbers cannot be expressed as a closed decimal fraction, others are almost unimaginable (imaginary), still others can be counted on the fingers or can be derived as solutions of equations.

Info1.1.2

The number ranges used throughout this online course are:

$\mathbb{N} = \{1; 2; 3; \ldots\}$	the **set of all natural numbers excluding zero**,
$\mathbb{N}_0 = \{0; 1; 2; 3; \ldots\}$	the **set of all natural numbers including zero**,
$\mathbb{Z} = \{\ldots; -2; -1; 0; 1; 2; \ldots\}$	the **set of all integer numbers (integers)**,
$\mathbb{Q}$	the **set of all rational numbers (fractions, rationals)**,
$\mathbb{R}$	the **set of all real numbers (reals)**.

These number ranges are not independent of each other. Rather, they form a chain of nested number sets:

$$\mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \,.$$

One obtains these number ranges by examining the solutions of the following equations and extending the number range in such a way that a solution always exists:

Number range	Solvable equation	Unsolvable	Extension by	New range
$\mathbb{N}$	$x + 2 = 4$	$x + 2 = 1$	negative numbers	$\mathbb{Z}$
$\mathbb{Z}$	$4x = 20$	$4x = 5$	fractions	$\mathbb{Q}$
$\mathbb{Q}$	$x^2 = 4$	$x^2 = 2$	irrational numbers	$\mathbb{R}$
$\mathbb{R}$	$x^2 = 2$	$x^2 = -1$	etc.	

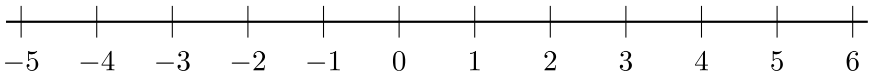
Natural numbers occur whenever numbers of objects have to be determined or things have to be labelled (using numbers). They play a great role in combinatorics: the number of possibilities for selecting 6 balls out of 49 is, for example, a natural number. Natural numbers are the bases of several number systems important either in daily life or in computer science: the binary system has base 2, the decimal system has base 10, and the hexadecimal system has base 16. Specific natural numbers, the prime numbers, are fundamental for modern encryption methods.

Arithmetic on the set of natural numbers is easy, but limits are reached if, for example, you read a temperature value of 3 °C (does it mean plus or minus degrees?) or if an equation such as $x + 5 = 1$ needs to be solved. Thus we must extend the set of natural numbers by the negative natural numbers to obtain the set of integers $\mathbb{Z}$. The **set of**

integers is denoted by

$$\mathbb{Z} := \{\ldots; -4; -3; -2; -1; 0; 1; 2; 3; 4; \ldots\}.$$

Integers are required whenever the sign (plus or minus) of a natural number matters. In $\mathbb{Z}$, numbers can be subtracted from each other without any restriction, i.e. systems of equations of the form $a + x = b$ are always solvable in $\mathbb{Z}$ ($x = b + (-a)$).



On the set of integers a comparator $<$ can be uniquely defined, so that the integers can be ordered into a chain:

$$\cdots < -3 < -2 < -1 < 0 < 1 < 2 < 3 < \ldots .$$

A rational number (rational) is the ratio of two integers:

Info1.1.3

The set of **rational numbers** is denoted by

$$\mathbb{Q} := \left\{ \frac{p}{q} \ : \ p, q \in \mathbb{Z}, q \neq 0 \right\}.$$

The elements $\frac{p}{q}$ of the set $\mathbb{Q}$ are called **fractions**, where p is the **numerator** of the fraction and q is the non-zero **denominator** of the fraction.

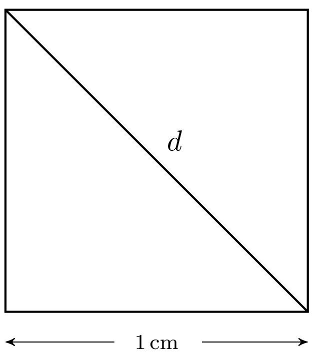
The rationals play a role whenever the numbers have to be "more precise", e.g. if temperatures have to be given in fractional amounts of °C, parts of surfaces have to be coloured, or medications have to be mixed from specific ingredients.

Note that the representation as a fraction is not unique: one number can be represented by several fractions. For example,

$$2 \ = \ \frac{4}{2} \ = \ \frac{1024}{512}$$

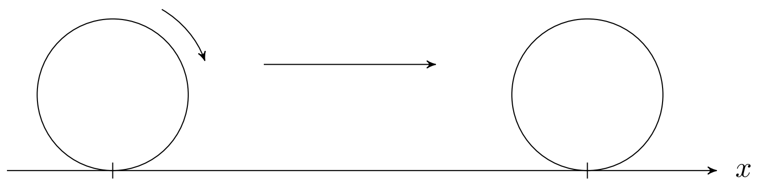
all represent the same rational number.

Also, not every number on the number line can be represented as a fraction. Considering, for example, a square with sides of length 1, the length of the diagonal d can be calculated by means of the Pythagoras' theorem:



$$d^2 = 1^2 + 1^2 = 2, \text{ or formally, } d = \sqrt{2}.$$

Another number that cannot be represented as a fraction is obtained by unrolling a wheel of diameter 1 on the number line. The result is the number π. It can be proven that these two numbers ($\sqrt{2}$ and π) cannot be represented as fractions. (In the case of $\sqrt{2}$ this proof is relatively simple.) These numbers are two examples of the so-called **irrational numbers**.



A number is irrational if it is not rational, i.e. if it cannot be represented as a fraction. The irrational numbers close the remaining gaps on the number line, where every point now corresponds to exactly one real number.

Info1.1.4

The set of **real numbers** is denoted by $\mathbb{R}$ and includes the set of rational numbers and the set of irrational numbers. It contains all numbers that can be represented on the number line.

Real numbers serve as measures for lengths, areas, temperatures, masses, etc. Throughout this course the mathematical problems are typically solved using real numbers.

A basic property of the real numbers is that they are ordered, i.e. for two reals a, b exactly one of the three relations $a < b$, $a = b$, or $a > b$ holds. Another defining property is completeness, which – roughly speaking – describes the "gaplessness" of the number line.

Info1.1.5

For two different real numbers, one sometimes considers all reals lying between these two numbers on the number line. Such a set of reals is called an **interval**. An interval is described by assigning a left interval boundary (a) and a right interval boundary (b) with $a < b$. Depending on whether each interval boundary is included, we must distinguish the following cases:

- $\{x \in \mathbb{R} : x \geq a \text{ and } x \leq b\} = [a; b]$ denotes the **closed** interval between a and

b including the interval boundaries.

- $\{x \in \mathbb{R} : x > a \text{ and } x < b\} =]a;b[$ denotes the **open** interval between a and b not including (i.e. excluding) the interval boundaries.
- $\{x \in \mathbb{R} : x \geq a \text{ and } x < b\} = [a;b[$ denotes the **left-closed and right-open** interval between a and b, including the left interval boundary but excluding the right interval boundary.
- $\{x \in \mathbb{R} : x > a \text{ and } x \leq b\} =]a;b]$ denotes the **left-open and right-closed** interval between a and b, including the right interval boundary but excluding the left interval boundary.

The last two intervals are also called **half-open** intervals.

For open interval ends, **unbounded** intervals can be considered as well. In these cases the corresponding condition in the set definition is dropped: $\{x \in \mathbb{R} : x \geq a\} = [a;\infty[$, $\{x \in \mathbb{R} : x > a\} =]a;\infty[$, $\{x \in \mathbb{R} : x \leq b\} =]-\infty;b]$, $\{x \in \mathbb{R} : x < b\} =]-\infty;b[$, $\{x \in \mathbb{R}\} = \mathbb{R} =]-\infty;\infty[$.

Moreover, the following descriptions are common: $\mathbb{R}^+ =]0;\infty[$, ${\mathbb{R}_0}^+ = [0;\infty[$, $\mathbb{R}^- =]-\infty;0[$, ${\mathbb{R}_0}^- =]-\infty;0]$.

A final remark as to the notation: In the literature you will find two different notations for the open end of an interval, either with square brackets pointing outwards or with parentheses, e.g. $[a;b[= [a;b)$, $]a;b[= (a;b)$. Both notations are correct, don't let them confuse you.

1.1.2 Variables and Terms

The use of variables, terms and equations is required to formalise expressions whose values have not been fixed.

Info1.1.6

A **variable** is a symbol (typically a letter) used as a placeholder for an indeterminate value. A **term** is a mathematical expression that can contain variables, arithmetic operations and further symbols and, after substituting variables with numbers, can be evaluated to a specific value. Terms can be combined into equations and inequalities, respectively, or they can be inserted into function descriptions, as we shall see later.

Example 1.1.7

The word problem

In a school class there are four more girls than boys and in total there are 20 children. How many girls and boys are in the class, respectively?

can be formalised, for example, by introducing the variable a for the number of girls and the variable b for the number of boys in the class and setting up the two equations $a = b + 4$ and $a + b = 20$. These equations can be solved by inserting the first equation into the second which gives $a = 12$ and $b = 8$. Out of this, the full written answer

In the school class there are 12 *girls and* 8 *boys*

can be constructed. Here, for example, $b + 4$ is a term, b itself is a variable, and $a + b = 20$ is an equation with a term on the left and a number on the right.

Variables (and sometimes also terms) are generally denoted by Latin lowercase letters x, y, z, etc. Often Greek letters are used as well, for example, to distinguish variables that represent angles from those that represent numbers.

Info1.1.8

This overview shows the (lowercase and uppercase) letters of the Greek alphabet in Greek alphabetical order:

α, A	"alpha"	β, B	"beta"	γ, Γ	"gamma"	δ, Δ	"delta"	ε, E	"epsilon"
ζ, Z	"zeta"	η, H	"eta"	ϑ, Θ	"theta"	ι, I	"iota"	$\varkappa$, K	"kappa"
λ, Λ	"lambda"	μ, M	"mu"	ν, N	"nu"	ξ, Ξ	"xi"	o, O	"omicron"
π, Π	"pi"	ϱ, P	"rho"	σ, Σ	"sigma"	τ, T	"tau"	υ, Υ	"upsilon"
φ, Φ	"phi"	χ, X	"chi"	ψ, Ψ	"psi"	ω, Ω	"omega"		

It is important that a term can be evaluated to a specific value if the variables occurring in the term are substituted with numbers:

Example 1.1.9

The following expressions are terms:

- $x \cdot (y+z) - 1$: for $x = 1$, $y = 2$, and $z = 0$ one obtains, for example, the value 1.
- $\sin(\alpha) + \cos(\alpha)$: for $\alpha = 0°$ and $\beta = 0°$ one obtains, for example, the value 1 (for the calculation of sine and cosine refer to 5).
- $1 + 2 + 3 + 4$: no variables occur, however this is a term (which always gives the value 10).
- $\frac{\alpha+\beta}{1+\gamma}$: for example, $\alpha = 1$, $\beta = 2$, and $\gamma = 3$ give the value $\frac{3}{4}$. But $\gamma = -1$ is not allowed.
- $\sin(\pi(x+1))$: this term, for example, always gives the value zero if x is substituted with an integer.
- z: a single variable is also a term.
- $1 + 2 + 3 + \cdots + (n-1) + n$ is a term, in which the variable n occurs in the term itself and defines its length as well.

Example 1.1.10

These expressions are not terms in a mathematical sense:

- $a + b = 20$ is an equation (inserting values for a and b gives no number, but the equation is simply true or false).
- $a \cdot (b + c$ is not correctly bracketed,
- "*The ratio of girls in the school class*" is not a term, but can be formalised by the term $\frac{a}{a+b}$,
- sin is not a term but a function name, in contrast $\sin(\alpha)$ is a term (which can be evaluated by inserting an angle for α).

Exercise 1.1.1
In each question, given a term and number values for the variables that occur in it, what is the evaluation of the term?

a. $\frac{\alpha+\beta}{\alpha-\beta}$ takes the value [ ] for $\alpha = 6$ and $\beta = 4$.

b. $y^2 + x^2$ takes the value [ ] for $y = 2x + 1$ and $x = -1$.

c. $1 + 2 + 3 + \cdots + (n-1) + n$ takes the value [ ] for $n = 6$.

Solution:
After inserting the given values for the variables, the term evaluates to a) $\frac{\alpha+\beta}{\alpha-\beta} = \frac{6+4}{6-4} = \frac{10}{2} = 5$, b) $y^2 + x^2 = (-1)^2 + (-1)^2 = 2$, and c) $1 + 2 + 3 + 4 + 5 + 6 = 21$.

Exercise 1.1.2
Formalise, using the variables given, the proportion of girls and the proportion of boys, the number of girls being denoted by the variable a and the number of boys by the variable b:

The proportion of girls is [ ] and the proportion of boys is [ ]
.

Solution:
The total number of children is $a + b$, hence the proportion of girls is $\frac{a}{a+b}$ and the proportion of boys is $\frac{b}{a+b}$.

Terms can be inserted into other terms as well:

Info1.1.11

When **inserting** terms, a term is substituted for a symbol in another term. If the term to be inserted contains several expressions, the replaced symbol has to be bracketed in advance.

Example 1.1.12

Substituting, for example, the right-hand side of $x = 1+2+3$ into the term $x^2 + y^2$ results in the new term $x^2 + y^2 = (1+2+3)^2 + y^2 = 36 + y^2$ and certainly not $1 + 2 + 3^2 + y^2 = 12 + y^2$.

Exercise 1.1.3
Which term is formed if the following object is inserted into the term $x^2 + y^2$?

a. The angle α both for x and y: Then $x^2 + y^2 =$ ______ .

b. The number 2 for y and the term $t + 1$ for x: Then $x^2 + y^2 =$ ______ .

c. The term $z + 1$ for x and the term $z - 1$ for y: Then $x^2 + y^2 =$ ______ .

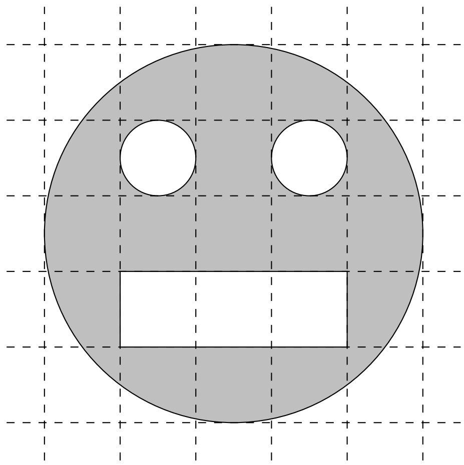
Solution:
It is safest to bracket the variables before inserting, if the new term contains several symbols:

a. $x^2 + y^2 = \alpha^2 + \alpha^2 = 2\alpha^2$.

b. $x^2 + y^2 = (x)^2 + y^2 = (t + 1)^2 + 2^2 = t^2 + 2t + 5$.

c. $x^2 + y^2 = (x)^2 + (y)^2 = (z + 1)^2 + (z - 1)^2 = z^2 + 2z + 1 + z^2 - 2z + 1 = 2z^2 + 2$.

Exercise 1.1.4
In the following figure, a square on the paper has side length x. What is the area of this figure (as a term in the variable x)?



(A single square has side length x)

A figure on squared paper.

Answer:

- The large circle has a total area of ______ ,
- each smaller circle has an area of ______ ,

- the total area of the figure is [ ] .

Hint for calculating the area:
The calculation of areas is presented in later chapters. To solve this exercise, you only need to know that a rectangle of side lengths a and b has the area $a \cdot b$ (entered as `a*b`) and a circle of radius r has the area πr^2 (entered as `pi*r^2`).

Solution:
The large circle has in total the area $\frac{25}{4}\pi x^2$ (entered as `25/4*pi*x*x`). Each smaller circle has the area $\frac{1}{4}\pi x^2$, and the whole figure has the area $(\frac{25}{4}\pi - \frac{1}{2}\pi - 3) \cdot x^2$.

1.1.3 Transformation of terms

There is always more than one way to write the same term, although some are more natural than others. For example, $x + x$ is a different arrangement of symbols than $2x$, but describes the same term, i.e. if x is substituted with a specific number, then $x + x$ and $2x$ provide the same value.

Info1.1.13

Terms are related by an equals sign if they are always evaluated to the same value.

In general, new terms are created by transformation of existing terms:

Info1.1.14

A **transformation** of a term is created by applying one or more calculation rules to the term:

- Collecting: $a + a + \cdots + a = n \cdot a$ (n is the number of summands).
- Distributive property ("expansion"): $(a+b) \cdot c = ac+bc$ and $c \cdot (a+b) = ca+cb$.
- Commutative property: $a + b = b + a$.
- Associative property ("group numbers differently in operations of the same

kind"):
$a + (b + c) = (a + b) + c = a + b + c$, also possible in multiplications.

- Calculation rules for powers and special functions.
- Calculation rules for specific types of terms (e.g. the binomial formulas).
- Calculation rules for fractions: $\frac{1}{\frac{a}{b}} = \frac{b}{a}$.

The rules will be presented in detail in the following sections. Often, the aim of this transformation is to simplify the term, to isolate individual variables, or to transform a term into a certain form:

Example 1.1.15

Examples of transformations and their uses:

- $a(a + a + a) + a^2 + a^2 + a^2 = 6a^2$: the term on the right is simpler, since it requires fewer symbols.
- $(x+3)^2 - 9 = x^2 + 6x$ (first binomial formula): both terms describe a parabola. On the left, the vertex $(-3, -9)$ of the parabola can be seen easily, on the right, the two roots ($x_1 = 0$ and $x_2 = -6$) can be seen easily.
- $1 + 3x + 3x^2 + x^3 = (1 + x)^3$: on the right, it can be seen, for example, that the function described by the term has only the root $x_1 = -1$.
- $\frac{a+1}{a} = 1 + \frac{1}{a}$: on the left, it can be seen that the term has the root $a_1 = -1$, on the right it can be seen that, for very large a, the term converges to 1 (since $\frac{1}{a}$ is very small in this case).

Exercise 1.1.5
Transform into a sum: $a \cdot (b + c) + c \cdot (a + b) =$ [ ] .
Solution:

$$a \cdot (b + c) + c \cdot (a + b) = ab + ac + ca + cb = ab + 2ac + bc$$

Exercise 1.1.6
Transform into a sum: $(x-y)(z-x)+(x-z)(y-z) \;=\; \boxed{}$.
Solution:

$$(x-y)(z-x)+(x-z)(y-z) \;=\; xz-x^2-yz+yx+xy-xz-zy+z^2 \;=\; -x^2-2yz+2xy+z^2$$

Exercise 1.1.7
Transform into a sum: $(a+b+2)(a+1) \;=\; \boxed{}$.
Solution:

$$(a+b+2)(a+1) \;=\; a^2+ba+2a+a+b+2 \;=\; a^2+ab+3a+b+2$$

1.2 Fractional Arithmetic

1.2.1 Calculating with Fractions

A fraction is a rational number written in the form $\frac{\text{numerator}}{\text{denominator}}$, where numerator and denominator are integers, and the denominator is $\neq 0$. Examples are:

$$\frac{1}{2}\,,\ \frac{5}{-10}\,,\ \frac{-17}{12}\,,\ \frac{1}{23}\,,\ \frac{4}{6}\,,\ \frac{-2}{3}\,,\ \ldots\ .$$

It can be seen very quickly that a single number can have an arbitrary number of equivalent representations. For example:

$$\frac{12}{36}=\frac{1}{3}=\frac{24}{72}=\frac{-12}{-36}=\frac{3}{9}=\frac{2}{6}=\frac{120}{360}=\ldots\ .$$

The different representations transform into each other by **reducing** and **expanding**, respectively.

Info1.2.1

Fractions are **reduced** by dividing numerator and denominator by the same non-zero integer.

Fractions are **expanded** by multiplying numerator and denominator by the same non-zero integer.

Example 1.2.2

Three friends like to share a pizza. Tom eats $\frac{1}{4}$ of the pizza, Tim eats $\frac{1}{3}$ of the pizza. How much of the pizza is left for their friend Sven, who always has the biggest appetite?
The solution is found by means of fractional arithmetic: First, we have to add two fractions, to decide how much of the pizza Tim and Tom already ate:

$$\frac{1}{4}+\frac{1}{3}=\frac{1\cdot 3}{4\cdot 3}+\frac{1\cdot 4}{3\cdot 4}=\frac{3}{12}+\frac{4}{12}=\frac{7}{12}\ .$$

Here, we can already identify the two most important steps: First we have to expand the two fractions to the so-called **least common denominator**, or, as one also says, we have to create **like** fractions. Then, if the fractions have the same denominator, we can add them by adding their numerators and maintaining the same denominator.

From the result that Tim and Tom ate $\frac{7}{12}$ of the pizza, we can calculate how much of the pizza is left for Sven by subtracting $\frac{7}{12}$ from 1:

$$1 - \frac{7}{12} = \frac{12}{12} - \frac{7}{12} = \frac{5}{12} \ .$$

Again, we must expand the fractions to the least common denominator. Then we have to subtract the two numerators. Indeed, the two friends have left much of the pizza for the always hungry Sven.soweit

The reducing of fractions can be practised in the training exercises below.

`In the online version, exercises from an exercise list will be shown here`

It becomes more difficult if indeterminate expressions (e.g. variables) occur in numerator and denominator. These can be reduced or cancelled just like numbers (but not with numbers), for example, we get

$$\frac{4x^2y^3 + 3y^2}{10y^2} = \frac{4x^2y + 3}{10}$$

by cancelling the term y^2 from numerator and denominator. The following training exercise has been extended to fractions including indeterminate expressions.

`In the online version, exercises from an exercise list will be shown here`

Info1.2.3

The **least common denominator** of two fractions is the least common multiple (lcm) of the two denominators.

The **least common multiple** (lcm) of two numbers is the smallest number that is divisible by both numbers.

The **greatest common divisor** (gcd) of two numbers is the largest number that divides both numbers without remainder.

If the determination of the lcm is too complicated, the simple product of the denominators can be used instead of the lcm in the following calculation rule:

Info1.2.4

Fractions are **added/subtracted** by finding a common denominator and then adding/substracting the numerators, i.e.

$$\frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd} \ , \ \ bd \neq 0 \ .$$

Usually fractions are expanded to the least common denominator.

For example, the least common multiple of $6 = 2 \cdot 3$ and $15 = 3 \cdot 5$ is the number $2 \cdot 3 \cdot 5 = 30$. However, the product is $6 \cdot 15 = 90$. Thus, you can calculate

$$\frac{1}{6} + \frac{1}{15} = \frac{5}{30} + \frac{2}{30} = \frac{7}{30}$$

but also

$$\frac{1}{6}+\frac{1}{15} = \frac{15}{90}+\frac{6}{90} = \frac{21}{90}$$

and finally reduce the last fraction to $\frac{7}{30}$.

Example 1.2.5

The least common multiple for the least common denominator is the smallest number that can be divided by all denominators involved. If these denominators have no factors in common, the least common multiple is simply the product of all denominators:

$$\begin{aligned}
\frac{1}{6}+\frac{1}{10} &= \frac{5}{30}+\frac{3}{30} = \frac{8}{30} = \frac{4}{15}\,, \\
\frac{1}{6}+\frac{1}{10} &= \frac{10}{60}+\frac{6}{60} = \frac{16}{60} = \frac{4}{15} \text{ (also correct)}, \\
\frac{4}{15}-\frac{1}{2} &= \frac{8}{30}-\frac{15}{30} = \frac{8-15}{30} = -\frac{7}{30}\,, \\
\frac{1}{3}+\frac{1}{9} &= \frac{3}{9}+\frac{1}{9} = \frac{4}{9}\,, \\
\frac{1}{2^2}+\frac{1}{2^4} &= \frac{2^2}{2^4}+\frac{1}{2^4} = \frac{5}{16}\,, \\
\frac{1}{2}+\frac{1}{3}+\frac{1}{7} &= \frac{21}{42}+\frac{14}{42}+\frac{6}{42} = \frac{41}{42}\,.
\end{aligned}$$

The least common denominator can also be found if the denominators include variables. Since the transformations of the fractions have to be correct for all possible values of these variables, they have to be considered as numbers without any common factors:

Example 1.2.6

Let x and y be variables, then

$$\begin{aligned}
\frac{1}{3}+\frac{1}{x} &= \frac{x}{3\cdot x}+\frac{3}{3\cdot x} = \frac{3+x}{3\cdot x}\,, \\
\frac{1}{x}+\frac{1}{y} &= \frac{y}{x\cdot y}+\frac{x}{x\cdot y} = \frac{x+y}{x\cdot y}\,, \\
\frac{1}{(x+1)^2}+\frac{1}{x+1} &= \frac{1}{(x+1)^2}+\frac{x+1}{(x+1)^2} = \frac{x+2}{(x+1)^2}\,.
\end{aligned}$$

Exercise 1.2.1
Calculate the following sums by means of the least common denominator (or the product of the denominators).

a. $\frac{1}{2} - \frac{1}{8} =$ ☐ .

b. $\frac{1}{3} + \frac{1}{5} + \frac{1}{6} =$ ☐ .

c. $\frac{1}{2x} + \frac{1}{3x} =$ ☐ .

Solution:
Finding the least common denominator and collecting/reducing gives

$$\begin{aligned} \frac{1}{2} - \frac{1}{8} &= \frac{4}{8} - \frac{1}{8} = \frac{3}{8}, \\ \frac{1}{3} + \frac{1}{5} + \frac{1}{6} &= \frac{10}{30} + \frac{6}{30} + \frac{5}{30} = \frac{21}{30} = \frac{7}{10}, \\ \frac{1}{2x} + \frac{1}{3x} &= \frac{3}{6x} + \frac{2}{6x} = \frac{5}{6x}. \end{aligned}$$

Exercise 1.2.2
In the case of like fractions, you may only add or decompose the numerators, for denominators no such rule exists. To convince yourself, calculate the following numbers by finding the least common denominator and reducing as much as possible:

a. $\frac{1}{2} + \frac{1}{3} =$ ☐ but $\frac{1}{2+3} =$ ☐ .

b. $\frac{1+2}{5+6} =$ ☐ but $\frac{1}{5} + \frac{2}{6} =$ ☐ .

Solution:
Sums of denominators may not be collected, not even in the case of like numerators. Here, we have

$$\frac{1}{2} + \frac{1}{3} = \frac{3}{6} + \frac{2}{6} = \frac{5}{6} \text{ but } \frac{1}{2+3} = \frac{1}{5}.$$

Also, the simple “splitting” of fraction parts is not allowed, here we have

$$\frac{1+2}{5+6} = \frac{3}{11} \text{ but } \frac{1}{5} + \frac{2}{6} = \frac{6}{30} + \frac{10}{30} = \frac{16}{30} = \frac{8}{15}.$$

Info1.2.7

Two fractions are **multiplied** by multiplying the numerators and multiplying the denominators, i.e.

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{a \cdot c}{b \cdot d} \;, \quad bd \neq 0 \;.$$

The division of two fractions is reduced to their multiplication:

Info1.2.8

Two fractions are **divided** by multiplying the first fraction with the reciprocal of the second fraction, i.e.

$$\frac{a}{b} : \frac{c}{d} = \frac{a}{b} \cdot \frac{d}{c} = \frac{a \cdot d}{b \cdot c} \;, \quad b, c, d \neq 0 \;.$$

The division of two fractions can be expressed as a **compound fraction** as well:

$$\frac{a}{b} : \frac{c}{d} = \frac{\frac{a}{b}}{\frac{c}{d}} \;.$$

Example 1.2.9

Taking possible reducing into account, the multiplication and the division of two fractions, respectively, takes the following form:

$$\frac{2}{3} \cdot \frac{4}{5} = \frac{2 \cdot 4}{3 \cdot 5} = \frac{8}{15} \;, \quad \frac{2}{3} : \frac{4}{5} = \frac{2}{3} \cdot \frac{5}{4} = \frac{10}{12} = \frac{5}{6} \;.$$

1.2.2 Converting Fractions

Dividing the denominator into the numerator gives a **decimal fraction** or a decimal number, respectively, for example,

$$\frac{1}{2} = 0.5\ ,\ \ \frac{1}{3} = 0.33333\ldots = 0.\bar{3}\ ,\ \ \frac{1}{7} = 0.\overline{142857}\ ,\ \ \frac{1}{8} = 0.125\ .$$

By means of these examples, it can already be seen that the division is either finite, leading to a **proper decimal fraction**, or the digits of the decimal number repeat in a certain way, leading to an **infinite repeating decimal fraction**.

Converting decimal numbers to fractions is done using the base-10 positional notation. Each decimal number has the form

...	1	2	3	4	5	.	6	7	8	9	...
	TTH	TH	H	T	O	.	t	h	th	tth	

with the abbreviations TTH ... ten thousands, TH ... thousands, H ... hundreds, T ... tens, O ... ones, t ... tenths, h ... hundredths, th ... thousandths, tth ... ten thousandths etc.
Then, the conversion is done as follows:

$$\begin{aligned} 4.375 &= 4 + \frac{3}{10} + \frac{7}{100} + \frac{5}{1000} \\ &= 4 + \frac{300+70+5}{1000} \\ &= 4 + \frac{375}{1000} \\ &= 4 + \frac{75}{200} \\ &= 4 + \frac{15}{40} = \frac{35}{8}\ . \end{aligned}$$

But what about converting an infinite repeating decimal number? It seems that we would have to add an infinite number of fractions, which is in practise of less use, of course. Therefore, in **converting infinite repeating decimal numbers to fractions** we use a trick:

Info1.2.10

Converting infinite repeating decimal numbers to fractions is done by multiplying the decimal number with a power of ten such that the repeating digits are shifted to the left of the decimal point. This leads to an equation of the form $10^k \cdot x = x + n$ for the decimal number x, that can be solved for x: $x = \dfrac{n}{10^k - 1}$ (which is a simple fraction).

Example 1.2.11

The number $0.\bar{6}$ is to be converted to a fraction. For this, you multiply the number by 10 and subtract from the result the initial number to eliminate the infinite repeating decimal:

$$\begin{array}{rrcccc} & 10 & \cdot & 0.\bar{6} & = & 6.\bar{6} \\ - & 1 & \cdot & 0.\bar{6} & = & 0.\bar{6} \\ \hline \Rightarrow & 9 & \cdot & 0.\bar{6} & = & 6.0 \end{array}$$

From the last relation, it immediately follows after division by 9: $0.\bar{6} = \frac{6}{9} = \frac{2}{3}$.

This method also works if not all digits after the decimal point repeat periodically:

Example 1.2.12

The decimal number $0.8\bar{3} = 0.83333\ldots$ is to be converted to a fraction:

$$\begin{array}{rrcccr} & 100 & \cdot & 0.8\bar{3} & = & 83.\bar{3} \\ - & 10 & \cdot & 0.8\bar{3} & = & 8.\bar{3} \\ \hline \Rightarrow & 90 & \cdot & 0.8\bar{3} & = & 75.0 \end{array}$$

Division by 90 gives the result: $0.8\bar{3} = \frac{75}{90} = \frac{5}{6}$.

Thus, the method is always the same: by multiplying by powers of ten and subsequent subtraction, the infinite repeating decimal is removed.

Exercise 1.2.3
Using the above method, find a simple and fully reduced fraction that represents the value $0.45555\ldots$.

Answer: $0.4\overline{5} =$ [ ] .

Solution:
Multiplying $x = 0.4\overline{5}$ by an appropriate power of ten gives

$$10x - x = 4.1 \;\Rightarrow\; 9x = \frac{41}{10} \;\Rightarrow\; x = \frac{41}{90}\ .$$

this fraction is already fully reduced as well.

However, in **back-of-the-envelope calculations** (i.e. if you only want to roughly estimate a magnitude or the ratio of one number to the other without knowing the correct values of the decimal fractions) it is useful to multiply the numbers by the least common denominator instead of converting them to decimals:

Example 1.2.13

The fractions $\frac{2}{3}$, $\frac{32}{12}$, and $\frac{12}{15}$ are to be arranged in order of size. For this, the fractions are multiplied by the least common denominator (60, in this case). The denominators are cancelled and the fractions are converted to the integers

$$\frac{2}{3}\cdot 60 \;=\; 2\cdot 20 \;=\; 40\ ,\quad \frac{32}{12}\cdot 60 \;=\; 32\cdot 5 \;=\; 160\ ,\quad \frac{12}{15}\cdot 60 \;=\; 12\cdot 4 \;=\; 48\ .$$

Arrangement in order of size gives $40 < 48 < 160$. With this, we have $\frac{2}{3} < \frac{12}{15} < \frac{32}{12}$, since the multiplication of the fractions by the same integer 60 does not change the arrangement of the fractions (see section 3.1, which deals with inequalities and their transformation).

Exercise 1.2.4
Arrange the fractions $\frac{16}{15}$, $\frac{1}{2}$, $\frac{2}{3}$, $\frac{2}{-3}$, $\frac{60}{90}$, and $\frac{4}{3}$ in order of size:
[ ] < [ ] < [ ] = [ ] < [ ] < [ ] .

Solution:
Multiplying by the common denominator 180 gives the numbers 192, 90, 120, −120, 120

and 240, which leads to the arrangement

$$-120 \ < \ 90 \ < \ 120 \ = \ 120 \ < \ 192 \ < \ 240$$

and thus finally to

$$\frac{2}{-3} \ < \ \frac{1}{2} \ < \ \frac{60}{90} \ = \ \frac{2}{3} \ < \ \frac{16}{15} \ < \ \frac{4}{3}.$$

1.2.3 Exercises

Exercise 1.2.5
Reduce the following fractions to the lowest terms:

a. $\frac{216}{240} = \square$.
Solution:
Since $\gcd(216, 240) = 24$ we have $\frac{216}{240} = \frac{216 : 24}{240 : 24} = \frac{9}{10}$.

b. $\frac{36}{72} = \square$.
Solution:
36 divides 72, hence $\frac{36}{72} = \frac{1}{2}$.

c. $\frac{48}{144} = \square$.
Solution:
48 divides 144, hence $\frac{48}{144} = \frac{1}{3}$.

d. $\frac{-a + 2b}{-4b + 2a} = \square$ if a not equals $\square$.
Solution:
Reducing gives $\frac{-a + 2b}{-4b + 2a} = \frac{(-1) \cdot (-2b + a)}{2 \cdot (-2b + a)} = -\frac{1}{2}$. The fraction is only defined for $a \neq 2b$.

Exercise 1.2.6
Calculate and fully reduce the following expressions for appropriate numbers a, b, x, y:

a. $\frac{1}{2} - \frac{2}{7} + \frac{3}{8} + \frac{3}{4} = \square$.
Solution:
Adding up the numerators over the least common denominator gives $\frac{1}{2} - \frac{2}{7} + \frac{3}{8} + \frac{3}{4} = \frac{28}{56} - \frac{16}{56} + \frac{21}{56} + \frac{42}{56} = \frac{75}{56}$, since $\gcd(2, 7, 8, 4) = 56$.

b. $\frac{3}{13} : \frac{7}{26} = \square$.
Solution:
Dividing by a fraction is the same as multiplying by its reciprocal: $\frac{3}{13} : \frac{7}{26} = \frac{3}{13} \cdot \frac{26}{7} = \frac{3 \cdot 26}{13 \cdot 7} = \frac{3 \cdot 2}{1 \cdot 7} = \frac{6}{7}$.

c. $\left(1.\bar{4} \cdot 3 - \frac{1}{2}\right) \cdot \frac{6}{7} = \square$.

Solution:

Converting the decimal expression to a fraction gives $\left(1.\bar{4}\cdot 3-\frac{1}{2}\right)\cdot\frac{6}{7}=(1.\bar{4}\cdot 18-3)\cdot\frac{1}{7}=(26-3)\cdot\frac{1}{7}=\frac{23}{7}$.

Exercise 1.2.7

Convert the following infinite repeating decimal fractions to fractions and fully reduce them:

a. $0.\overline{4}$ = [ ] .

b. $0.\overline{23}$ = [ ] .

c. $0.12\overline{34}$ = [ ] .

d. $0.\overline{9}$ = [ ] .

Solution:

Using the trick for converting infinite repeating decimal numbers one gets the following solutions:

- $x=0.\overline{4}$, hence $10x-x=4 \;\Rightarrow\; 9x=4 \;\Rightarrow\; x=\frac{4}{9}$,
- $x=0.\overline{23}$, hence $100x-x=23 \;\Rightarrow\; 99x=23 \;\Rightarrow\; x=\frac{23}{99}$,
- $x=0.12\overline{34}$, hence $100x-x=12.22 \;\Rightarrow\; 99x=\frac{1222}{100} \;\Rightarrow\; x=\frac{1222}{9900}=\frac{611}{4950}$,
- $x=0.\overline{9}$, hence $10x-x=9 \;\Rightarrow\; 9x=9 \;\Rightarrow\; x=1$.

Note for the last part of this exercise, that $1=1.000\ldots$ and $1=0.999\ldots=0.\overline{9}$ are two decimal representations of the same number.

1.3 Transformation of terms

1.3.1 Introduction

What exactly are terms?

Info1.3.1

Terms are arithmetic expressions that are combinations of numbers, variables, brackets, and appropriate arithmetic operations.

Terms can be interpreted in two ways:

- As functional expressions: If each variable contained in the term is substituted with a specific number, the term can be evaluated to a certain value. For example, $x+x-1$ is a term; once $x = 2$ is inserted one gets the value 3. The expression $2x-1$ is a term as well, this term can be transformed into $x+x-1$, and hence it evaluates to the same value if $x = 2$ is inserted. As a symbolic expression, $x+x-1$ is different from $2x - 1$, but as functional expressions they are both the same (equivalent): No matter which value is inserted for x, both terms are always evaluated to the same value. A term can also be a value on its own, if no variables occur in it. For example, $3 \cdot (2 + 4)$ is a term with the value 18.
- As evaluation rules: A term can be interpreted as a type of instruction how to calculate a new value from given values (inserted into the variables). For example, the term $x^2 - 1$ can be read as "square the value of x and subtract one from the result". This is different from the term $(x + 1)(x - 1)$, even if both terms have the same value. The second term describes the evaluation as "add one to x and multiply the result with the value, resulting if x is subtracted by one". The two terms are mathematically equal. One writes $x^2 - 1 = (x - 1)(x + 1)$, but they represent two different ways for calculating the value. Depending on the problem setting, one of the two terms may be more convenient for solving the problem.

1.3.2 Transformation of Terms

Dealing with terms gets interesting when we investigate the equality of two terms or simplify complicated terms.

Info1.3.2

Two terms are **equal** if they can be transformed into each other by allowed transformations. Complicated terms can be simplified using calculation rules. Note in doing so:

1. Exponentiation precedes multiplication precedes addition.
2. If brackets are involved, the distributive property applies:
$$a \cdot (b \pm c) = a \cdot b \pm a \cdot c \; , \quad (a \pm b) \cdot c = a \cdot c \pm b \cdot c \; .$$
3. With $d \neq 0$ we have: $(a \pm b) : d = \frac{a}{d} \pm \frac{b}{d}$.
4. For expressions with nested brackets, first evaluate what's inside the innermost set of brackets with respect to the calculation rules and then work your way towards the outermost brackets.

Exercise 1.3.1
Remove the brackets from each of the following terms and simplify as far as possible:

a. $(1 - a) \cdot (1 - b) \; =$ [ ] .
Solution:
$$(1 - a) \cdot (1 - b) \; = \; 1 - a - b + ab \; .$$

b. $5a - (2b - (2a - 7b) + 4a) - 3b \; =$ [ ] .
Solution:
$$5a - (2b - (2a - 7b) + 4a) - 3b \; = \; 5a - 2b + 2a - 7b - 4a - 3b \; = \; 3a - 12b \; .$$

Info1.3.3

The three **binomial formulas** are:
$$(a + b)^2 = a^2 + 2ab + b^2 \; , \quad (a - b)^2 = a^2 - 2ab + b^2 \; , \quad (a + b)(a - b) = a^2 - b^2 \; .$$

Here, for a and b both numbers and whole terms can be inserted.

Example 1.3.4

A few typical applications of the binomial formulas are:

- $(1+2x)^2 = 1^2 + 2\cdot 1\cdot 2x + (2x)^2 = 1 + 4x + 4x^2$.
- $(1+2x)(1-2x) = 1^2 - (2x)^2 = 1 - 4x^2$.
- $x^4 - 1 = (x^2+1)(x^2-1) = (x^2+1)(x+1)(x-1)$, in this transformation it can be seen, that in the set of real numbers $x^4 - 1$ has only the roots $x = 1$ and $x = -1$.
- $(1+x+y)^2 = \left((1+x)+y\right)^2 = (1+x)^2 + 2(1+x)y + y^2 = 1+2x+x^2+2y+2xy+y^2$.

Exercise 1.3.2
Simplify the following term using the second binomial formula:

$(-3x+4)(4-3x) \ = \ \boxed{}$.

Solution:

$$(-3x+4)(4-3x) \ = \ (4-3x)(4-3x) \ = \ 16 - 24x + 9x^2 \ .$$

Example 1.3.5

The binomial formulas can be used to transform quadratic expressions cleverly. This is very useful if we want to calculate squares without any aid. For this, the number to be squared is split into a simple number (usually a power of ten) and the remainder:

$$\begin{aligned} 103^2 &= (100+3)^2 = 100^2 + 2\cdot 100\cdot 3 + 3^2 = 10609\ , \\ 49^2 &= (50-1)^2 = 50^2 - 2\cdot 50\cdot 1 + 1^2 = 2401\ , \\ 61^2 - 59^2 &= (61-59)(61+59) = 2\cdot 120 = 240\ . \end{aligned}$$

Exercise 1.3.3
Using the method described in Example 1.3.5, calculate $1005^2 \ = \ \boxed{}$.

Solution:

$$1005^2 \;=\; (1000+5)^2 \;=\; 1000000 + 2\cdot 1000\cdot 5 + 25 \;=\; 1010025\,.$$

In the following exercise section you can practise the transformation methods in several exercises.

1.3.3 Exercises

Exercise 1.3.4
Simplify the following terms for appropriate numbers a, b, x, y, z:

a. $\dfrac{3x - 6xy^2 + 4xyz}{-2x} = \boxed{}$.
Solution:

$$\frac{3x - 6xy^2 + 4xyz}{-2x} = -\frac{3}{2} + 3y^2 - 2yz \ .$$

b. $(3a - 2b) \cdot (4a - 6) = \boxed{}$.
Solution:

$$(3a - 2b) \cdot (4a - 6) = 12a^2 - 18a - 8ab + 12b \ .$$

c. $(2a + 3b)^2 - (3a - 2b)^2 = \boxed{}$.
Solution:

$$\begin{aligned} (2a + 3b)^2 - (3a - 2b)^2 &= (4a^2 + 12ab + 9b^2) - (9a^2 - 12ab + 4b^2) \\ &= -5a^2 + 24ab + 5b^2 \ . \end{aligned}$$

d. $\dfrac{3a}{3a + 6b} + \dfrac{2b}{a + 2b} = \boxed{}$.
Solution:
Adding up the numerators over the least common denominator gives $\dfrac{3a}{3a + 6b} + \dfrac{2b}{a + 2b} = \dfrac{3a}{3a + 6b} + \dfrac{6b}{3a + 6b} = \dfrac{3a + 6b}{3a + 6b} = 1$.

Exercise 1.3.5
For the following exercises a little more patience is required. Simplify:

a. $\dfrac{1}{2}x(4x + 3y) + \dfrac{3}{2}(5x^2 - 6xy) = \boxed{}$.
Solution:

$$\frac{1}{2}x(4x + 3y) + \frac{3}{2}(5x^2 - 6xy) = 2x^2 + \frac{3}{2}xy + \frac{15}{2}x^2 - 9xy = \frac{19}{2}x^2 - \frac{15}{2}xy \ .$$

b. $\dfrac{18x^2 - 48xy + 32y^2}{12y - 9x} \cdot \dfrac{18x + 24y}{9x^2 - 16y^2} = \boxed{}$.
Solution:

Simplifying the expression gives

$$\begin{aligned}\frac{18x^2-48xy+32y^2}{12y-9x}\cdot\frac{18x+24y}{9x^2-16y^2} &= 4\cdot\frac{9x^2-24xy+16y^2}{4y-3x}\cdot\frac{3x+4y}{9x^2-16y^2}\\ &= 4\cdot\frac{(3x-4y)^2}{4y-3x}\cdot\frac{3x+4y}{9x^2-16y^2}\\ &= -4\cdot\frac{(3x-4y)^2}{3x-4y}\cdot\frac{3x+4y}{(3x+4y)(3x-4y)} = -4\;.\end{aligned}$$

c. $(a^2+5a-2)(2a^2-3a-9)-\left(\frac{1}{2}a^2+3a-5\right)(a^2-4a+3) =$ [ ] .

Solution:

$$\begin{aligned}&(a^2+5a-2)(2a^2-3a-9)-\left(\frac{1}{2}a^2+3a-5\right)(a^2-4a+3)\\ =\;& 2a^4+10a^3-4a^2-3a^3-15a^2+6a-9a^2-45a+18\\ &-\left(\frac{1}{2}a^4+3a^3-5a^2-2a^3-12a^2+20a+\frac{3}{2}a^2+9a-15\right)\\ =\;& \frac{3}{2}a^4+6a^3-\frac{25}{2}a^2-68a+33\;.\end{aligned}$$

Exercise 1.3.6

Use a binomial formula to calculate the following squares:

a. $43^2 =$ [ ] .

Solution:

$$43^2 = (40+3)^2 = 40^2+2\cdot 40\cdot 3+3^2 = 1600+240+9 = 1849\;.$$

b. $97^2 =$ [ ] .

Solution:

$$97^2 = (90+7)^2 = 90^2+2\cdot 90\cdot 7+7^2 = 8100+1260+49 = 9409\;.$$

c. $41^2-38^2 =$ [ ] .

Solution:

$$41^2-38^2 = (41+38)(41-38) = 79\cdot 3 = 237\;.$$

Exercise 1.3.7

Apply a binomial formula to expand the product and collect like terms:

a. $(-5xy-2)^2 \;=\;$ [ ] .
Solution:

$$(-5xy-2)^2 \;=\; (-1)^2\cdot(5xy+2)^2 \;=\; 25x^2y^2+20xy+4\,.$$

b. $(-6ab+7bc)(-6ab-7bc) \;=\;$ [ ] .
Solution:

$$(-6ab+7bc)(-6ab-7bc) \;=\; (-6ab)^2-(7bc)^2 \;=\; 36a^2b^2-49b^2c^2\,.$$

c. $(-6ab+7bc)(-6ab+7bc) \;=\;$ [ ] .
Solution:

$$(-6ab+7bc)(-6ab+7bc) \;=\; (-6ab+7bc)^2 \;=\; 36a^2b^2-84ab^2c+49b^2c^2\,.$$

d. $(x^2+3)(-x^2-3) \;=\;$ [ ] .
Solution:

$$(x^2+3)(-x^2-3) \;=\; -(x^2+3)^2 \;=\; -x^4-6x^2-9\,.$$

Exercise 1.3.8
Factorise the following terms as far as possible using one of the binomial formulas:

a. $4x^2+12xy+9y^2 \;=\;$ [ ] .
Solution:

$$4x^2+12xy+9y^2 \;=\; (2x+3y)^2\,.$$

b. $64a^2-96a+36 \;=\;$ [ ] .
Solution:

$$64a^2-96a+36 \;=\; (8a-6)^2\,.$$

c. $25x^2-16y^2+15x+12y \;=\;$ [ ] .
Solution:

$$\begin{aligned}25x^2-16y^2+15x+12y &= (5x)^2-(4y)^2+3\cdot(5x+4y)\\ &= (5x+4y)(5x-4y)+3\cdot(5x+4y)\\ &= (5x+4y)(5x-4y+3)\,.\end{aligned}$$

1.3.4 Representation as a Sum and as a Product

Mathematical expressions can be written in different ways that all have their own pros and cons. We distinguish them based on which mathematical operation is to be performed last. The most common types are the representation as a sum and the representation as a product.

Info1.3.6

For a **representation as a product**, it is the multiplication that is performed last. Because of the order of operation rule, if any of the factors involves an addition or subtraction, then having the multiplication be last can only be achieved by enclosing the factors in brackets. Product is particularly useful for determining cases in which a term takes the value zero. This happens if and only if one of the factors takes the value zero.

For example, the term $(x-1)\cdot(x-2)$ is zero if $x=1$ or $x=2$. For all other values of x the term takes a non-zero value.

Info1.3.7

For a **representation as a sum** it is the addition or the subtraction that is performed last. Because of the order of operation rule, terms without brackets are automatically in this representation, provided they contain any addition or subtraction at all. From the representation as a sum the asymptotic behaviour of an expression can often be deduced. The asymptotic behaviour of a function describes how the function behaves if the variable x takes arbitrarily large values. For polynomials, for example, the asymptotic behaviour is determined by the term with the largest exponent.

To change from one representation to another, several methods exist.

Info1.3.8

Expanding means multiplying each summand of one factor by each summand of the other factor and adding up the results. In case of more than two factors, they should be multiplied out step by step (only two at a time).

Example 1.3.9

The function $f(x) = (x+3)(x-2)(x+1)$ is multiplied out as follows:

$$\begin{aligned} f(x) &= (x+3)\cdot(x-2)\cdot(x+1) \\ &= (x^2+3x-2x-6)\cdot(x+1) \\ &= (x^2+x-6)\cdot(x+1) \\ &= x^3+x^2-6x \;+\; x^2+x-6 \\ &= x^3+2x^2-5x-6\,. \end{aligned}$$

Exercise 1.3.9
Expand the following terms completely and collect like terms. Describe the asymptotic behaviour of the final expression:

a. $f(x) \;=\; (3-x)(x+1) =$ [ ] .
Solution:

$$(3-x)(x+1) \;=\; 3x+3-x^2-x \;=\; 3+2x-x^2$$

Description of the asymptotic behaviour:
As x approaches ∞ the function $f(x)$ approaches [ ] .
Solution:
If the asymptotic behaviour arises uniquely it can be described for short using the symbol "lim":

$$\lim_{x\to\infty} f(x) \;=\; -\infty\,.$$

As x approaches $-\infty$ the function $f(x)$ approaches [ ] .

Solution:
In this case we have

$$\lim_{x\to-\infty} f(x) \;=\; -\infty \;.$$

b. $(x+4)(2-x)(x+2) =$ [ ] .
Solution:

$$(x+4)(2-x)(x+2) \;=\; (x+4)(4-x^4) \;=\; 4x-x^3+16-4x^2 \;=\; 16+4x-4x^2-x^3$$

c. $(3-x)(x+1)^2 =$ [ ] .
Solution:

$$(3-x)(x+1)^2 \;=\; (3-x)(x^2+2x+1) \;=\; 3x^2+6x+3-x^3-2x^2-x \;=\; 3+5x+x^2-x^3$$

d. $t\cdot(t+1)\cdot(t^2+t+1) =$ [ ] .
Solution:

$$t\cdot(t+1)\cdot(t^2+t+1) \;=\; (t^2+t)(t^2+t+1) \;=\; t^4+t^3+t^2+t^3+t^2+t \;=\; t^4+2t^3+2t^2+t$$

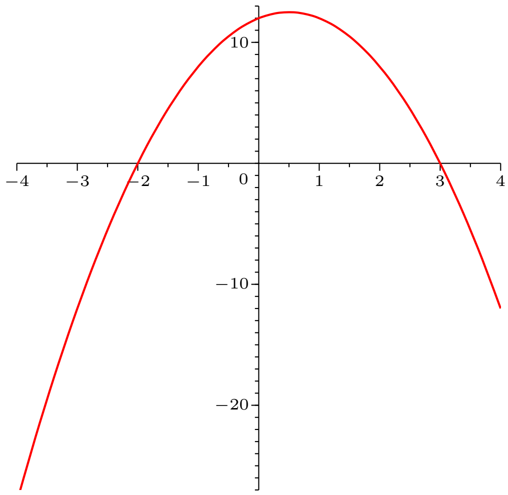
Solution:
Expanding gives $f(x) = (3-x)(x+1) = -x^2+2x+3$. The leading term is $-x^2$. Thus the graph of the function is a parabola opening downwards with the two asymptotes $-\infty$ for x approaching $\pm\infty$:

$$\lim_{x\to\infty} f(x) \;=\; -\infty \;,\quad \lim_{x\to-\infty} f(x) \;=\; -\infty \;.$$

In the other parts of this exercise we get by expanding

$$\begin{aligned}
(x+4)(2-x)(x+2) &= -x^3-4x^2+4x+16 \;, \\
(3-x)(x+1)^2 &= (3-x)(x^2+2x+1) \;=\; -x^3+x^2+5x+3 \;, \\
t\cdot(t+1)\cdot(t^2+t+1) &= t\cdot(t^3+2t^2+2t+1) \;=\; t^4+2t^3+2t^2+t \;.
\end{aligned}$$

Exercise 1.3.10
The following graph corresponds to a polynomial $g(x)$ of degree two:



Graph of the function $g(x)$.

From the graph, derive the representation of $g(x)$ as a product.

a. The graph has two zeros x_1 and x_2. Multiplied out the two factors resulting from this fact, we get the polynomial $f(x) = (x-x_1)(x-x_2) =$ [ ] .

b. The polynomial $f(x)$ does not correspond to the graph since at $x = 0$ it takes the value [ ] whereas the graph shows that the function $g(x)$ at $x = 0$ takes the value [ ] . This difference can be corrected by setting $g(x) = c \cdot f(x)$, where $c =$ [ ] .

c. This finally gives the representation of $g(x)$ as a product: $g(x) =$ [ ] .

Solution:
The graph shows the zeros $x_1 = -2$ and $x_2 = 3$. Multiplying out the two factors resulting from this fact, we get the polynomial $f(x) = (x+2)(x-3) = x^2 - x - 6$. At $x = 0$, we have $f(0) = -6$, but the graph shows $g(0) = 12$. This difference can be corrected by an additional factor of -2. In total, we have $g(x) = -2x^2 + 2x + 12$.

Exercise 1.3.11
Fully expand the expression: $(a+2b+3c)^2 =$ [ ] .

Solution:

It is simplest to multiply each summand in the first pair of brackets by each summand in the second, finally collecting like terms:

$$
\begin{aligned}
(a+2b+3c)^2 &= (a+2b+3c)\cdot(a+2b+3c) \\
&= a^2+2ab+3ac+2ab+4b^2+6bc+3ac+6bc+9c^2 \\
&= a^2+4ab+6ac+4b^2+12bc+9c^2\ .
\end{aligned}
$$

1.4 Powers and Roots

1.4.1 Exponentiation and Roots

The following section deals with expressions of the form a^s, where $a \in \mathbb{R}$. But the question is: For which numbers s can this power be reasonably defined?

Powers with positive integer exponents are defined as follows:

Info1.4.1

Let $n \in \mathbb{N}$. The n-th **power** of a number $a \in \mathbb{R}$ is the n-fold product of the number a by itself:

$$a^n = \underbrace{a \cdot a \cdot a \cdot \dots \cdot a}_{n \text{ factors}} .$$

a is called the **base** and n is called the **exponent**.

Here, some special cases exist that you should ideally know by heart:

Info1.4.2

For a zero exponent, the value of the power is one, i.e. for example, $4^0 = 1$, also $0^0 = 1$. But for a zero base, for $n > 0$, we have $0^n = 0$. For base -1, we have

$$(-1)^n \ = \ -1 \text{ if the exponent is odd} \ , \ \ (-1)^n \ = \ 1 \text{ if the exponent is even} .$$

Example 1.4.3

$$3^2 = 3 \cdot 3 = 9 \ , \ \ (-2)^3 = (-2) \cdot (-2) \cdot (-2) = -8 \ , \ \ \left(\frac{1}{2}\right)^4 = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{16} .$$

Many powers can be calculated using the calculation rule presented above — but what about 2^{-2}?

Info1.4.4

Powers with negative integer exponents are defined by the formula $a^{-n} = \frac{1}{a^n}, n \in \mathbb{N}, a \neq 0$.

Hence $2^{-2} = \frac{1}{2^2} = \frac{1}{4}$. By analogy, we have $(-2)^{-2} = \frac{1}{(-2)^2} = \frac{1}{4}$.

Exercise 1.4.1
Calculate the values of the following powers.

a. $5^3 =$ ☐ .

Solution:

$$5^3 = 5 \cdot 5 \cdot 5 = 25 \cdot 5 = 125\ .$$

b. $(-1)^{1001} =$ ☐ .

Solution:

$$(-1)^{1001} = -1 \text{ since the exponent is odd}\ .$$

c. $\left(-\frac{1}{2}\right)^{-3} =$ ☐ .

Solution:

$$(-\frac{1}{2})^{-3} = \frac{1}{(-\frac{1}{2})^3} = \frac{1}{-\frac{1}{8}} = -8\ .$$

d. $\left((-3)^2\right)^3 =$ ☐ .

Solution:

$$\left((-3)^2\right)^3 = 9^3 = 729\ .$$

However, even for a rational exponent of the form $\frac{1}{n}, n \in \mathbb{N}$, we need to extend this definition again so that we can calculate $4^{\frac{1}{2}}$, for example. This power can be expressed as a root as well, namely $4^{\frac{1}{2}} = \sqrt{4} = 2$. Generally, we have:

Info1.4.5

Let $n \in \mathbb{N}$ and $a \in \mathbb{R}$ with $a \geq 0$. The n-th root has the power representation $\sqrt[n]{a} = a^{\frac{1}{n}}$.

This leads to one inverse operation of the exponentiation: extracting a root.

Info1.4.6

The n-th **root** of a number $a \in \mathbb{R}, a \geq 0$, is the number whose n-th power is a:

$$a^{\frac{1}{n}} = \sqrt[n]{a} = b \Longrightarrow b^n = a \ .$$

a is called the **base of the root** or **radicand**, and n is called the **exponent of the root** or **order of the root**. We have $\sqrt[1]{a} = a$, and $\sqrt[2]{a} = \sqrt{a}$ is called the **square root** and $\sqrt[3]{a}$ is called the **cube root** of a.

Example 1.4.7

$$16^{\frac{1}{2}} = \sqrt[2]{16} = \sqrt{16} = 4 \ , \quad 27^{\frac{1}{3}} = \sqrt[3]{27} = 3 \ .$$

Info1.4.8

For $n \in \mathbb{N}, a, b \in \mathbb{R}$ with $a, b \geq 0$ we have the following **calculation rules**:

1. Two roots with the same exponent are multiplied by multiplying the radi-

cands and extracting the root of the product, leaving the exponent of the root unchanged:

$$\sqrt[n]{a} \cdot \sqrt[n]{b} = \sqrt[n]{a \cdot b} \,.$$

2. Two roots with the same exponent are divided by dividing the radicands and extracting the root of the quotient, leaving the exponent of the root unchanged:

$$\frac{\sqrt[n]{a}}{\sqrt[n]{b}} = \sqrt[n]{\frac{a}{b}}, b \neq 0 \,.$$

But how can we calculate the number $\left(\sqrt[10]{4}\right)^5$?

Info1.4.9

Let $m, n \in \mathbb{N}$ and $a \in \mathbb{R}, a \geq 0$.

1. The m-th power of a root is calculated by raising the radicand to the power of m, leaving the exponent of the root unchanged:

$$\left(\sqrt[n]{a}\right)^m = \sqrt[n]{a^m} = a^{\frac{m}{n}} \,.$$

2. The ***m*-th root of a root** is calculated by multiplying the exponents of the roots, leaving the radicand unchanged (**root of a root**).

$$\sqrt[m]{\sqrt[n]{a}} = \sqrt[m \cdot n]{a} \,.$$

Hence

$$\sqrt[10]{4^5} = (4^5)^{\frac{1}{10}} = 4^{\frac{5}{10}} = 4^{\frac{1}{2}} = \sqrt[2]{4} = 2 \,.$$

Example 1.4.10

A general power with rational exponent is then calculated as follows:

$$\left(\frac{1}{4}\right)^{-\frac{2}{3}} = \sqrt[3]{\left(\frac{1}{4}\right)^{-2}} = \sqrt[3]{\frac{1}{\left(\frac{1}{4}\right)^{2}}} = \sqrt[3]{\frac{1}{\frac{1}{16}}} = \sqrt[3]{16} = \sqrt[3]{2^3 \cdot 2} = 2\sqrt[3]{2}\ .$$

The calculation rules for powers with real base and rational exponent are known as exponent rules. The rules vary depending on whether powers of the same base or the same exponent are considered.

Exercise 1.4.2
Calculate the following roots (here, the result is):

a. $\left(\sqrt[5]{3}\right)^5 =$ ☐ .

Solution:

$$\left(\sqrt[5]{3}\right)^5 = (3^5)^{\frac{1}{5}} = 3^{5\cdot\frac{1}{5}} = 3\ .$$

b. $\sqrt[4]{256} =$ ☐ .

Solution:

$$\sqrt[4]{256} = \left(2^8\right)^{\frac{1}{4}} = 2^2 = 4\ .$$

1.4.2 Calculating using Powers

The following calculation rules allow one to transform and simplify expressions containing powers or roots:

Info1.4.11

For $a, b \in \mathbb{R}, a, b > 0, p, q \in \mathbb{Q}$, the following **exponent rules** hold:

$$a^p \cdot b^p = (a \cdot b)^p \ , \quad \frac{a^p}{b^p} = \left(\frac{a}{b}\right)^p \ , \quad a^p \cdot a^q = a^{p+q} \ , \quad \frac{a^p}{a^q} = a^{p-q} \ , \quad (a^p)^q = a^{p \cdot q} \ .$$

Note that, generally, $(a^p)^q \neq a^{p^q}$, i.e. multiple powers should be bracketed. For example, $\left(2^3\right)^2 = 8^2 = 64$, but $2^{\left(3^2\right)} = 2^9 = 512$.

Example 1.4.12

Without brackets one reads a^{p^q} as $a^{(p^q)}$, i.e. for example,

$$\begin{aligned} 2^{3^4} &= 2^{3 \cdot 3 \cdot 3 \cdot 3} = 2^{81} = 2417851639229258349412352 \text{ (exponent evaluated first)} \\ (2^3)^4 &= 8^4 = 4096 \text{ (bracket evaluated first)} \ . \end{aligned}$$

Alternatively, we could have used the exponent rules to calculate $(2^3)^4 = 2^{(3 \cdot 4)} = 2^{12} = 4096$.

Exercise 1.4.3
The following expressions can be simplified using the exponent rules:

a. $3^3 \cdot 3^5 \cdot 3^{-1} =$ ______ .

Solution:

$$3^3 \cdot 3^5 \cdot 3^{-1} = 3^{3+5-1} = 3^7 \ .$$

b. $4^2 \cdot 3^2 =$ ______ .

Solution:

$$4^2 \cdot 3^2 = (4 \cdot 3)^2 = 12^2 \ .$$

However, when comparing powers and roots, one should take care: not only the values, but also the signs of exponent and base control whether the value of the power is large or small:

Example 1.4.13

For a positive base and a negative exponent, the value of the power decreases when increasing the base:

$$\begin{aligned} 2^{-1} &= \frac{1}{2} = 0.5 \\ 3^{-1} &= \frac{1}{3} = 0.\bar{3} \\ 4^{-1} &= \frac{1}{4} = 0.25 \text{ etc.} \end{aligned}$$

For a negative base the sign of the power alternates when increasing the exponent:

$$\begin{aligned} (-2)^1 &= -2 \\ (-2)^2 &= 4 \\ (-2)^3 &= -8 \\ (-2)^4 &= 16 \text{ etc.} \end{aligned}$$

Extracting the root (or exponentiation with a positive number smaller than one) decreases a base > 1, but increases a base < 1:

$$\begin{aligned} \sqrt{2} &= 1.414\ldots < 2 \\ \sqrt{3} &= 1.732\ldots < 3 \\ \sqrt{0.5} &= 0.707\ldots > 0.5 \\ \sqrt{0.\bar{3}} &= 0.577\ldots > 0.\bar{3} \text{ etc.} \end{aligned}$$

Exercise 1.4.4
Arrange the following powers in order of size considering the signs of bases and exponents: 2^3, 2^{-3}, 3^2, $(-3)^2$, $(-3)^{-2}$, $3^{\frac{1}{2}}$, $2^{\frac{1}{3}}$:
[ ] < [ ] < [ ] < [ ] < [ ] <
[ ] = [ ] .

Solution:
Here, several ways to the correct solution exist. For example, we can raise all terms to the

power of 3 (analogous to the calculation in Example 1.2.13 auf Seite 27). Exponentiation to the power of 3 is allowed since it is an odd number that does not cancel the sign of the exponent. In contrast, exponentiation to the power of 2 would cancel all signs such that the resulting arrangement would be invalid. We have

$$\begin{aligned}
(2^3)^3 &= 2^{3\cdot 3} = 2^9 = 512 \\
(2^{-3})^3 &= 2^{(-3)\cdot 3} = 2^{-9} = \frac{1}{512} \text{ since the exponent is negative} \\
(3^2)^3 &= 9^3 = 81\cdot 9 = 729 \\
((-3)^2)^3 &= 9^3 = 729 \text{ since the exponent is even} \\
((-3)^{-2})^3 &= \frac{1}{((-3)^2)^3} = \frac{1}{729} \\
(3^{\frac{1}{2}})^3 &= 3^{\frac{3}{2}} > 3 \text{ (since the exponent is larger than one)} \\
(2^{\frac{1}{3}})^3 &= 2\,.
\end{aligned}$$

Comparing these values leads to the following arrangement:

$$(-3)^{-2} < 2^{-3} < 2^{\frac{1}{3}} < 3^{\frac{1}{2}} < 2^3 < 3^2 = (-3)^2\,.$$

1.4.3 Exercises

Exercise 1.4.5
Calculate the following powers:

a. $\left(-\frac{3}{5}\right)^4 = \boxed{}$.

Solution:

$$\left(-\frac{3}{5}\right)^4 = \frac{3^4}{5^4} = \frac{81}{625} \; .$$

b. $\left(2^{-2}\right)^{-3} = \boxed{}$.

Solution:

$$\left(2^{-2}\right)^{-3} = 2^{(-2)\cdot(-3)} = 2^6 = 64 \; .$$

c. $\left(-\frac{1}{2}\right)^0 = \boxed{}$.

Solution:

$$\left(-\frac{1}{2}\right)^0 = 1 \; .$$

Exercise 1.4.6
Simplify the following expressions using the exponent rules and by reducing the results. You don't need to evaluate the powers:

a. $\frac{(-2)^7}{(-2)^5} = \boxed{}$.

Solution:

$$\frac{(-2)^7}{(-2)^5} = (-2)^{7-5} = (-2)^2 \; .$$

b. $6^2 \cdot 3^{-2} = \boxed{}$.

Solution:

$$6^2 \cdot 3^{-2} = 6^2 \cdot \left(\frac{1}{3}\right)^2 = \left(6 \cdot \frac{1}{3}\right)^2 = 2^2 \; .$$

c. $\dfrac{64^3}{8^3} =$ ☐ .

Solution:

$$\frac{64^3}{8^3} = \left(\frac{64}{8}\right)^3 = 8^3 \;.$$

d. $\left(\dfrac{3}{4}\right)^{\frac{1}{3}} \cdot \left(\dfrac{3}{4}\right)^{\frac{2}{3}} =$ ☐ .

Solution:

$$\left(\frac{3}{4}\right)^{\frac{1}{3}} \cdot \left(\frac{3}{4}\right)^{\frac{2}{3}} = \left(\frac{3}{4}\right)^{\frac{1}{3}+\frac{2}{3}} = \frac{3}{4} \;.$$

Exercise 1.4.7
Calculate the following roots (here, the result is always an integer):

a. $\sqrt[3]{3} \cdot \sqrt[3]{9} =$ ☐ .

Solution:

$$\sqrt[3]{3} \cdot \sqrt[3]{9} = \sqrt[3]{3 \cdot 3 \cdot 3} = 3 \;.$$

b. $\sqrt[3]{343} =$ ☐ .

Solution:

$$\sqrt[3]{343} = \sqrt[3]{7^3} = 7 \;.$$

Exercise 1.4.8
Simplify the following expressions and write each as a reduced fraction containing no powers:

a. $\dfrac{3^3 \cdot 6^3}{9 \cdot 2^3 \cdot 4^3} =$ ☐ .

Solution:

$$\frac{3^3 \cdot 6^3}{9 \cdot 2^3 \cdot 4^3} = \frac{3^4 \cdot 2^3}{2^3 \cdot 4^3} = \frac{3^4}{4^3} = \frac{81}{64} \;.$$

b. $3^2 \cdot 9^{-3} \cdot 27^6 \cdot 27^{-2} \ = \ \boxed{}$.

<u>Solution:</u>

$$3^2 \cdot 9^{-3} \cdot 27^{-2} \cdot 27^6 = 3^{2+2\cdot(-3)+3\cdot 6-6} = 3^8 \ .$$

1.5 Final Test

1.5.1 Final Test Module 1

Exercise 1.5.1
Check the box in each case to indicate whether the mathematical expressions are equations, inequalities, terms, or numbers (multiple checks are possible):

Mathematical expression	Equation	Inequality	Term	Number
$1+\frac{1}{2}-3(3-\frac{1}{2})$	☐	☐	☐	☐
$5^x - x^5$	☐	☐	☐	☐
$x^2 < \sqrt{x}$	☐	☐	☐	☐
$xyz - 1$	☐	☐	☐	☐
$b^2 = 4ac$	☐	☐	☐	☐

Exercise 1.5.2
Simplify the compound fraction $\frac{3+\frac{3}{2}}{\frac{1}{12}+\frac{1}{4}}$ to a reduced simple fraction: ☐

Exercise 1.5.3
Expand the following term completely and collect like terms:

$(x-1)\cdot(x+1)\cdot(x-2) =$ ☐ .

Exercise 1.5.4
Apply one of the binomial formulas to transform the term:

a. $(x-3)(x+3) =$ ☐ .

b. $(x-1)^2 =$ ☐ .

c. $(2x+4)^2 =$ ☐ .

Exercise 1.5.5
Rewrite the following expression containing powers and roots as a simple power with a rational exponent:

$\frac{x^3}{(\sqrt{x})^3} =$ ☐ .

2 Equations in one Variable

Module Overview

Equations arise when we require the values of two terms to be equal. If variables occur in at least one of them, an equation can be understood as the task to find out for which values of the variables the left-hand side term and the right-hand side term have the same value. Simple equations can be solved by applying transformations and solution formulas from a formulary. For a more sophisticated equation, a case analysis may be required.

2.1 Simple Equations

2.1.1 Introduction

Info2.1.1

An **equation** is an expression of the form

$$\text{left-hand side} \;=\; \text{right-hand side}$$

with mathematical expressions on both sides of the equation. These expressions generally contain variables or unknowns (e.q. x). Depending on the variable values an equation is satisfied if both sides of the equation evaluate to the same value. An equation is not satisfied if the sides of the equation evaluate to different values.

Equations describe relations between expressions or model a problem to be solved. In general, an equation itself is not true or false. Instead, some variables satisfy the equation and others do not. To test whether the equation is true or false for a single variable value this value has to be inserted into the equation. Then, both sides of the equation are evaluated to certain values. The equation is satisfied by an inserted variable value if the evaluated values coincide:

Example 2.1.2

The equation $2x - 1 = x^2$ has the right-hand side x^2 and the left-hand side $2x - 1$. Inserting $x = 1$ results in the value 1 on both sides of the equals sign, hence $x = 1$ is a solution of this equation. However, $x = 2$ is no solution since the left-hand side of the equation is evaluated to the value 4 while the right-hand side is evaluated to the value 3.

Info2.1.3

The **solution set** L of an equation is the set of all numbers satisfying the relation

$$\text{left-hand side} \;=\; \text{right-hand side}$$

if inserted into the the equation instead of the variable (e.q. x).

Typical problems concerning equations are:

- specify the solution set of an equation, i.e. find all variable values satisfying the equation,
- transform the equation, in particular, solve an equation for the variables, and
- find an equation modelling a problem described textually.

Example 2.1.4

We want to design a savings deposit in such a way that there is a fixed annual return. The bank wants to make sure that when investing over a five-year period, a saver receives exactly 600 Euros more than when investing over a two-year period.

First, the word problem is translated into an equation with the variable r denoting the annual return. The resulting equation is $5 \cdot r = 2 \cdot r + 600$. It says that five payments (left-hand side of the equation) equal two payments plus 600. (For simplicity, we omit the unit Euro during calculation.)

We can easily solve this equation can be solved for r by subtracting the term $2r$ from both sides of the equation. The resulting equation reads $3r = 600$, and dividing by 3 results in the solution $r = 200$.

Thus, the bank has to offer a return of 200 Euros per year to reach the required savings target.

Info2.1.5

Two equations are said to be **equivalent** if they have the same solution set.

An **equivalent transformation** is a special transformation that changes the equation but not its solution set. Important equivalent transformations are

- adding/subtracting terms to both sides of the equation,
- exchanging both sides of the equation,
- transformation of terms on one side of the equation, and
- substituting a term for another that is known to always have the same value.

The following transformations are equivalent transformations only if the used term is known to be non-zero (which may depend on the possible values of the variables):

- multiplying/dividing by a term (this term has to be non-zero),
- taking the reciprocal on both sides of the equation (both sides have to be non-zero).

Here, the following **notation** is used:

- equivalent equations are indicated by the symbol $\Leftrightarrow$ (which reads: if and only if, i.e. one equation is satisfied if and only if the other equation is satisfied).
- under this symbol we put a short description of the transforming operation (or, for solutions with more than one line, the transforming operation is written next to the transformation).

What matters is that the reader should be able to understand which transformation was carried out.

Example 2.1.6

This example illustrates two simple equivalent transformations written in a single line. Even though the symbol $\Leftrightarrow$ is two-sided the notation is interpreted in such a way that the transformation is applied from left to right:

$$3x - x^2 \;=\; 2x - x^2 + 1 \underset{+x^2}{\Leftrightarrow} 3x \;=\; 2x + 1 \underset{-2x}{\Leftrightarrow} x \;=\; 1\;.$$

The leftmost equation and the rightmost equation are equivalent. On the left we have the initial equation (corresponding to a certain word problem) and on the right we have an equivalent equation showing the solution immediately.

Example 2.1.7

For several complicated transformations, the transformation steps should be written one under another. In this case we use vertical bars to separate the respective transformations from the equations.

$$
\begin{aligned}
\text{Start:}\quad 12+t &= \frac{2t}{2t^2}+t && \Big\|\quad -t \\
\Leftrightarrow\quad 12 &= \frac{2t}{2t^2} && \Big\|\quad \text{sides exchanged} \\
\Leftrightarrow\quad \frac{2t}{2t^2} &= 12 && \Big\|\quad \text{left-hand side transformed} \\
\Leftrightarrow\quad \frac{1}{t} &= 12 && \Big\|\quad \text{reciprocals taken} \\
\Leftrightarrow\quad t &= \frac{1}{12}\,.
\end{aligned}
$$

Here, after the vertical bar both short symbols as, e.g. $-t$, and textual descriptions are allowed. Again, what matters is that the reader can understand which transformations were carried out and so can check that everything was done correctly.

2.1.2 Conditions in Transformations

Multiplication, division, and taking reciprocals are equivalent transformations only if the factors or terms are non-zero. In the last step of example 2.1.7 auf der vorherigen Seite, the reader can see that both sides of the equation are always non-zero. Therefore the transformation is allowed. If the variables themselves are used in the transformation, we need to make a note somewhere to remind us that the respective term must be non-zero. The solution at the end of the transformations is then only valid for variable values satisfying the transformation conditions. All other values have to be checked *separately*, typically by inserting the value into the equation:

Example 2.1.8

In this example, the necessary transformation conditions are not problematic:

$$\begin{aligned} \text{Start:} \quad 9x &= 81x^2 && \Big\| \; : x, \text{ transformation allowed if } x \neq 0 \\ \Leftrightarrow \quad 9 &= 81x && \Big\| \; : 81 \text{ and exchange sides} \\ \Leftrightarrow \quad x &= \frac{1}{9} && \text{and this value satisfies the condition } x \neq 0 \,. \end{aligned}$$

The value $x = 0$, initially rejected by the transformation condition, has to be checked separately. The equation $9x = 81x^2$ is also satisfied for $x = 0$, hence $x = 0$ is also a solution of the equation. In set notation, the solution set is $L = \{0; \frac{1}{9}\}$.

In any case, values violating a condition have to be checked separately. It may or may not turn out that they are part of the solution.

Example 2.1.9

$$\begin{aligned} \text{Start:} \quad x^2 - 2x &= 2x - 4 && \Big\| \; \text{factor out on both sides} \\ \Leftrightarrow \quad x \cdot (x-2) &= 2 \cdot (x-2) && \big| \; : (x-2), \text{ transformation only allowed if } x \neq 2 \\ \Leftrightarrow \quad x &= 2 \,. \end{aligned}$$

This value of x violates the condition $x \neq 2$. Hence, this is possibly no solution. Inserting $x = 2$ into the initial equation gives $2^2 - 2 \cdot 2 = 0$ on the left-hand side and also $2 \cdot 2 - 4 = 0$ on the right-hand side. Hence, $x = 2$ is indeed a solution, even though it violated the transformation condition.

Exercise 2.1.1
Find the solution of the equation $(x-2)(x-3) = x^2 - 9$ by transforming the right-hand side using the third binomial formula and then dividing by a common factor.

The solution is $x =$ [ ] .

Solution:
The correct transformation steps including conditions are

$$\begin{array}{lrcl|l} \text{Start:} & (x-2)(x-3) & = & x^2-9 & \text{transformation of the right-hand side} \\ \Leftrightarrow & (x-2)(x-3) & = & (x+3)(x-3) & :(x-3)\text{, transformation allowed if } x \neq 3 \\ \Leftrightarrow & x-2 & = & x+3 & -x \\ \Leftrightarrow & -2 & = & 3 \text{ is a wrong equation.} & \end{array}$$

Importantly, this equation is only wrong for $x \neq 3$. We have to check $x = 3$ separately, and indeed $x = 3$ satisfies the initial equation.

2.1.3 Proportionality and Rule of Three

A relation between two varying quantities that frequently occurs in practice is the **proportionality**. Examples include mass and volume, time and travelled distance or weight (quantity) of a product and its price. Often we are given two example values that stand in relation and need to complete another example for which only one value is given. We will illustrate this procedure with an example.

Example 2.1.10

5 kg of apples cost 3 Euro. How much do 11 kg of apples cost?

In a first step, we convert the information we have about the price of apples in the following traditional notation, which in this example we could read as “costs”:

$$5\,\text{kg} \;\hat{=}\; 3 \text{ Euro}\,.$$

In this example, the symbol in the middle can be read as “costs”, but in other examples other readings will be required. The key point is that as the value on one side of the symbol varies, the value on the other will have to vary proportionally. In the second step, we want to scale this proportional relation so as to express it in terms of a unit amount of the quantity for which another value (here: 11 kg) is given. So we multiply both sides by 1/5 and get:

$$1\,\text{kg} \;\hat{=}\; \frac{1}{5}\cdot 3 \text{ Euro} = 0.6 \text{ Euro}\,.$$

In the third and final step, we multiply both sides by the number of units specified in the problem, which in our example is 11:

$$11\,\mathrm{kg} \;\hat{=}\; 11 \cdot 0.6 \text{ Euro} = 6.6 \text{ Euro} .$$

The required price for 11 kg of apples is therefore 6.60 Euro.

We have derived the required relation by deriving a relation for one unit of a quantity from the initial relation. The procedure demonstrated here is called the **rule of three** and is taught in great detail in the schools of some countries such as Germany and France.

The same problem can also be solved by introducing a proportionality factor. Again, we consider the example above.

Example 2.1.11

The price P is proportional to the mass m. Hence, there exists a constant k with

$$P = km .$$

Since this relation also holds for the given values $m_0 = 5\,\mathrm{kg}$ and $P_0 = 3$ Euro it follows

$$\begin{aligned} P_0 = km_0 & \quad \Big\| \quad \text{multiplying by } \frac{1}{m_0} \\ \Longleftrightarrow \frac{P_0}{m_0} = k & \quad ; \end{aligned}$$

hence in this case

$$k = \frac{3}{5} = 0.6 ,$$

taken in the unit of Euro per kg. (As a scientist you would correctly write $k = 0.6$ Euro/kg, since proportionality factors generally carry a dimensional unit.) Using $m_1 = 11\,\mathrm{kg}$, one obtains finally

$$P_1 = km_1 = 0.6 \cdot 11 = 6.6 \text{ (Euro)}$$

which is the same result as for using the rule of three (see previous example).

Exercise 2.1.2
A car takes 9 minutes to travel a distance of 6 km.

a. Which distance s the car travels within 15 minutes?

   The solution is $s_{15} =$ [ ] km.

b. The proportionality factor between travelled distance s and travelling time t is the velocity v of the car.

   The velocity is $v =$ [ ] km/ h.

Solution:
¿From the given values we know that the car travels $\frac{6}{9}$ km $= \frac{2}{3}$ km within one minute and therefore $15 \cdot \frac{2}{3}$ km $= 10$ km within 15 minutes.

So, the velocity is

$$v = \frac{10\,\text{km}}{15\,\text{min}} = \frac{10\,\text{km}}{(1/4)\,\text{h}} = 40\,\frac{\text{km}}{\text{h}}\ .$$

2.1.4 Solving linear Equations

Info2.1.12

A **linear equation** is an equation in which only multiples of variables and constants occur.

For a linear equation in one variable (here the variable x) one of the following three statements holds:

- The equation has no solution.
- The equation has a single solution.
- Every value of x is a solution of the equation.

These three cases are distinguished by means of the transformation steps:

- If the transformation ends up in a statement that is wrong for all x (e.g. $1 = 0$) then the equation is unsolvable.

- If the transformation ends up in a statement that is true for all x (e.g. $1 = 1$) then the equation is solvable for all values of x.
- Otherwise, the equation can be solved, i.e. it can be transformed into the equation $x =$ value which is the solution.

Set notation2.1.13

Using set notation (with L as the conventional symbol for the solution set) these cases can be expressed as follows:

- $L = \{\}$ or $L = \emptyset$ if there is no solution,
- $L = \{\text{value}\}$ if there is a single solution,
- $L = \mathbb{R}$ if all real numbers x are a solution.

Example 2.1.14

The linear equation $3x + 2 = 2x - 1$ has one solution. This solution is obtained by equivalent transformations:

$$3x + 2 \;=\; 2x - 1 \underset{-2x}{\Leftrightarrow} x + 2 \;=\; -1 \underset{-2}{\Leftrightarrow} x \;=\; -3\,.$$

Hence, $x = -3$ is the only solution.

Example 2.1.15

The linear equation $3x + 3 = 9x + 9$ has the solution:

$$3x + 3 \;=\; 9x + 9 \underset{:(x+1)}{\Leftrightarrow} 3 \;=\; 9\,.$$

This statement is wrong. Hence, for all $x \neq -1$ (transformation condition) the equation is wrong. Inserting $x = -1$ satisfies the equation, and so the only solution is $x = -1$.

Alternatively, the equation could have been transformed as follows:

$$3x + 3 \;=\; 9x + 9 \underset{-3x-9}{\Leftrightarrow} -6 \;=\; 6x \Leftrightarrow x \;=\; -1\,.$$

Exercise 2.1.3
Transform the following linear equations and specify their solution sets:

a. The equation $x - 1 = 1 - x$ has the solution set $L =$ ☐ ,

b. The equation $4x - 2 = 2x + 2$ has the solution set $L =$ ☐ ,

c. The equation $2x - 6 = 2x - 10$ has the solution set $L =$ ☐ .

Solution:
The first equation can be transformed into $2x = 2$ or $x = 1$, respectively, so the solution set is $L = \{1\}$. The second equation can be transformed into $2x = 4$ and the solution set is $L = \{2\}$. The third equation can be transformed into $-6 = -10$ which is a false statement, hence $L = \{\}$.

Exercise 2.1.4
Find the solution of the general linear equation $ax = b$ with a and b being real numbers. Specify the values of a and b for which the following three cases occur:

- Every value of x is a solution ($L = \mathbb{R}$) if $a =$ ☐ and $b = 0$.
- There is no solution ($L = \emptyset$) if $a =$ ☐ and $b \neq$ ☐ .
- Otherwise, there is a single solution, namely $x =$ ☐ .

Solution:
Every value of x is a solution ($L = \mathbb{R}$) if $a = 0$ and $b = 0$. There is no solution ($L = \emptyset$) if $a = 0$ and $b \neq 0$. Otherwise, there is only one solution, namely $x = \frac{b}{a}$.

2.1.5 Solving quadratic Equations

Info2.1.16

A **quadratic equation** is an equation of the form $ax^2 + bx + c = 0$ with $a \neq 0$, or, in reduced form, $x^2 + px + q = 0$. This form is obtained by dividing the equation by a.

For a quadratic equation in one variable (here the variable x) one of the following three statements holds:

- The quadratic equation has no solution: $L = \{\}$.
- The quadratic equation has a single solution $L = \{x_1\}$.
- The quadratic equation has two different solutions $L = \{x_1; x_2\}$.

The solutions are obtained by applying **quadratic solution formulas**.

Info2.1.17

The *pq* **formula** for solving the equation $x^2 + px + q = 0$ reads

$$x_{1,2} \;=\; -\frac{p}{2} \pm \sqrt{\frac{1}{4}p^2 - q}\;.$$

Here, the equation has

- no (real) solution if $\frac{1}{4}p^2 - q < 0$ (taking the square root is not allowed),
- a single solution $x_1 = -\frac{p}{2}$ if $\frac{1}{4}p^2 = q$ and the square root is zero,
- two different solutions if the square root is a positive number.

The expression $D := \frac{1}{4}p^2 - q$ underneath the square root considered above is called the **discriminant**.

The solution of a quadratic equation is often described by an alternative formula:

Info2.1.18

For the equation $ax^2 + bx + c = 0$ with $a \neq 0$ the *abc* **formula** reads

$$x_{1,2} \;=\; \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}\;.$$

Here, the equation has

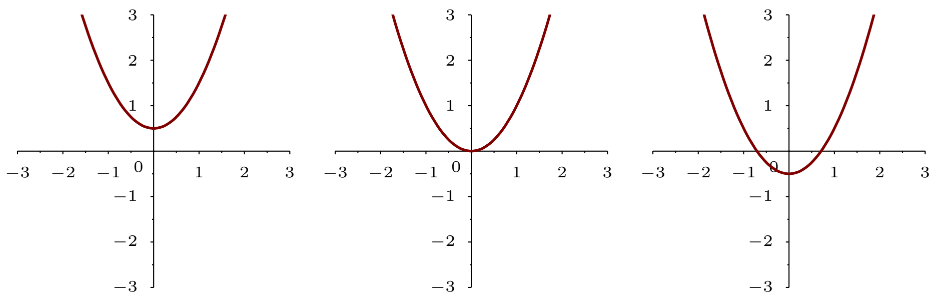
- no (real) solution if $b^2 - 4ac < 0$ (the square root of a negative number is undefined within the range of real numbers),

- a single solution $x_1 = -\frac{b}{2a}$ if $b^2 = 4ac$ and the square root is zero,
- two different solutions if the square root is a positive number.

Again, the expression $D := b^2 - 4ac$ underneath the square root considered above is called the **discriminant**.

Both formulas result in the same solutions. The pq formula is easier to learn, but is only applicable if a, the coefficient of the quadratic term, is 1. Otherwise we must divide both sides of the equation by a.

In terms of the pq formula, the three different cases correspond to three possibilies for the number of intersection points that the graph of a shifted standard parabola $f(x) = x^2 + px + q$ may have with the x axis.



Three cases: no intersection point, one intersection point, two intersection points with the x axis.

Example 2.1.19

The quadratic equation $x^2-x+1 = 0$ has no solution since the discriminant $\frac{1}{4}p^2-q = -\frac{3}{4}$ within the pq formula is negative. In contrast, the equation $x^2 - x - 1 = 0$ has two solutions

$$\begin{aligned} x_1 &= \frac{1}{2} + \sqrt{\frac{1}{4}+1} = \frac{1}{2}(1+\sqrt{5})\ , \\ x_2 &= \frac{1}{2} - \sqrt{\frac{1}{4}+1} = \frac{1}{2}(1-\sqrt{5})\ . \end{aligned}$$

Info2.1.20

The function expression of a parabola has **vertex form** if the function has the form $f(x) = a \cdot (x-s)^2 - d$ with $a \neq 0$. In this case, $(s; -d)$ is the **vertex** of the parabola. The corresponding quadratic equation for $f(x) = 0$ then reads $a \cdot (x-s)^2 = d$.

Dividing this equation by a one obtains the equivalent quadratic equation $(x-s)^2 = \frac{d}{a}$. Since the left-hand side is a square of a real number, only solutions exist if and only if the right-hand side is non-negative as well, i.e. $\frac{d}{a} \geq 0$. By taking the square root, taking the two possible signs into account, one obtains $x - s = \pm\sqrt{\frac{d}{a}}$.

So, for $\frac{d}{a} > 0$ two solutions of the equation exist:

$$x_1 \; = \; s - \sqrt{\frac{d}{a}} \; , \;\; x_2 \; = \; s + \sqrt{\frac{d}{a}} \; ;$$

they are symmetric to the x coordinate s of the vertex. For $d = 0$, only one solution exists.

The sign of a determines whether the function expression describes a parabola opening upwards or downwards.

The quadratic equation has only one single solution s if it can be transformed into the form $(x-s)^2 = 0$.

Info2.1.21

Any quadratic equation can be transformed (after collecting terms on the left-hand side and normalisation, if necessary) to vertex form by **completing the square**. For this, a constant is added to both sides of the equation such that on the left-hand side we have a term of the form $x^2 \pm 2sx + s^2$ to which the first or second binomial formula can be applied.

Example 2.1.22

Adding the constant 2 transforms the equation $x^2 - 4x + 2 = 0$ into the form $x^2 - 4x + 4 = 2$ or into the form $(x-2)^2 = 2$, respectively. From this, the two solutions $x_1 = 2 - \sqrt{2}$ and $2 + \sqrt{2}$ can be seen immediately. In contrast, the quadratic equation $x^2 + x = -2$ has no solution since completing the square results in $x^2 + x + \frac{1}{4} = -\frac{7}{4}$ or $(x + \frac{1}{2})^2 = -\frac{7}{4}$, respectively, where the right-hand side is negative for $a = 1$.

Exercise 2.1.5
Find the solutions of the following quadratic equations by completing the square after collecting terms on the left-hand side and normalisation (i.e. selecting $a = 1$):

a. $x^2 = 8x - 1$ has the vertex form ☐ = ☐ .
The solution set is $L =$ ☐ .

b. $x^2 = 2x + 2 + 2x^2$ has the vertex form ☐ = ☐ .
The solution set is $L =$ ☐ .

c. $x^2 - 6x + 18 = -x^2 + 6x$ has the vertex form ☐ = ☐
.
The solution set is $L =$ ☐ .

Solution:
The transformations are for the first equation

$$\begin{aligned} & x^2 = 8x - 1 \\ \Leftrightarrow\ & x^2 - 8x + 1 = 0 \\ \Leftrightarrow\ & x^2 - 8x + 16 = 15 \\ \Leftrightarrow\ & (x-4)^2 = 15 \\ & L = \{4 - \sqrt{15}; 4 + \sqrt{15}\} \end{aligned}$$

and for the second equation

$$\begin{aligned} & x^2 = 2x + 2 + 2x^2 \\ \Leftrightarrow\ & x^2 + 2x + 2 = 0 \\ \Leftrightarrow\ & x^2 + 2x + 1 = -1 \\ \Leftrightarrow\ & (x+1)^2 = -1 \\ & L = \{\} \end{aligned}$$

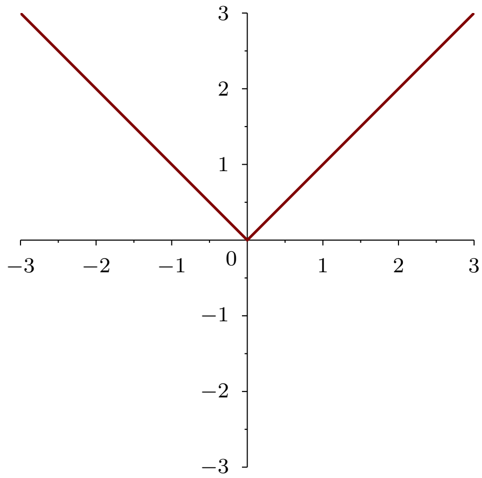
and for the third equation

$$
\begin{aligned}
& x^2 - 6x + 18 = -x^2 + 6x \\
\Leftrightarrow\ & 2x^2 - 12x + 18 = 0 \\
\Leftrightarrow\ & x^2 - 6x + 9 = 0 \\
\Leftrightarrow\ & (x-3)^2 = 0 \\
& L = \{3\}\ .
\end{aligned}
$$

2.2 Absolute Value Equations

2.2.1 Introduction

The absolute value $|x|$ assigns a variable $x \in \mathbb{R}$ its value without sign: If $x \geq 0$, then $|x| = x$, otherwise $|x| = -x$ (see figure).



The absolute value $|x|$ as a function of x.

Absolute value equations are equations in which one absolute value or several absolute values occur. Problems arise since the absolute value is calculated by distinguishing the two cases

$$|\,\text{term}\,| \quad = \quad \begin{cases} \text{term} & \text{for term} \geq 0 \\ -\text{term} & \text{for Term} < 0 \end{cases} .$$

For solving absolute value equations, these cases have to be solved step by step and analysed to find the solutions.

Example 2.2.1

Obviously, the absolute value equation $|x| = 2$ has the solution set $L = \{2; -2\}$. Just as easy, it can be seen that $|x - 1| = 3$ has the solution set $L = \{-2; 4\}$.

As soon as beside the absolute value several other terms occur a case analysis is required. In the following section we will explain in detail how this analysis is done and how it is written correctly since the case analysis will play an important role in the next modules.

2.2.2 Carry out a Case Analysis

Info2.2.2

To solve an **absolute value equation** two cases are distinguished:

- For all values of x for which the absolute value term is non-negative the absolute value can be omitted or replaced by simple brackets, respectively.
- For all values of x for which the absolute value term is negative the term is bracketed and negated.

Then, the solution sets from the case analyses will be restricted to satisfy the case conditions. Only if this procedure is finished for all cases, the solution subsets will be merged to the solution set of the initial equation.

For solving absolute value equations it is important to write down the solution steps correctly and to distinguish the cases clearly.

The following video demonstrates a detailed written solution of the absolute value equation $|2x-4| = 6$ by case analysis.

(Video cannot be played)

Video 1: Carry out a case analysis.

The case analysis presented in the video can be written briefly as follows:

$$|2x-4| = \begin{cases} 2x-4 & \text{for } x \geq 2 \\ -2x+4 & \text{for } x < 2 \end{cases} = \begin{cases} 2x-4 & \text{for } x \geq 2 \\ -2x+4 & \text{otherwise} \end{cases} .$$

Exercise 2.2.1
Describe the values of the expression $2 \cdot |x-4|$ by a case analysis:

$$2 \cdot |x-4| = \boxed{} .$$

Solution:

$$2 \cdot |x-4| \;=\; \begin{cases} 2x-8 & \text{for } x \geq 4 \\ -2x+8 & \text{for } x < 4 \end{cases}$$

Exercise 2.2.2
Reproduce the steps shown in the video 2.2.2 auf der vorherigen Seite to solve the absolute value equation $|6+3x| = 12$.

The case analysis reads shortly $|6+3x| =$ [ ] .
Solution:

$$|6+3x| \;=\; \begin{cases} 6+3x & \text{for } x \geq -2 \\ -6-3x & \text{for } x < -2 \end{cases}$$

Finding the solution for each case and checking the case conditions leads to the solution set $L =$ [ ] for the equation $|6+3x| = 12$.

Solution:

$$L \;=\; \{-6; 2\}$$

You can practise the stepwise solution of absolute value equations within the following exercise.

`In the online version, exercises from an exercise list will be shown here`

2.2.3 Mixed Equations

Info2.2.3

If an equation contains both absolute values and other expressions, the case analysis has to be done according to the absolute value terms and applied only to these.

Finally, keep in mind to cross-check the solution sets you found with the case conditions.

Example 2.2.4

Solve the equation $|x-1|+x^2=1$. Here, the case analysis is as follows:

- For $x \geq 1$, the absolute value bars can be replaced by normal brackets which results in the quadratic equation $(x-1)+x^2=1$ that is transformed into the equation $x^2+x-2=0$. Using the *pq* formula we get the two solutions

$$x_1 = -\frac{1}{2}-\sqrt{\frac{9}{4}} = -2\,,$$
$$x_2 = -\frac{1}{2}+\sqrt{\frac{9}{4}} = 1$$

  of which only x_2 satisfies the case condition.

- For $x < 1$, one obtains the quadratic equation $-(x-1)+x^2=1$ that is transformed into the equation $x^2-x=0$ or $x \cdot (x-1)=0$, respectively. The product representation indicates the two solutions $x_3=0$ and $x_4=1$. Because of the case condition only $x_3=0$ is a solution of the initial equation.

So, altogether $L=\{0;1\}$ is the solution set of the initial equation.

Exercise 2.2.3
Find the solution set of the mixed equation $|x-3| \cdot x = 9$.

a. If x is in the interval [ ] the absolute value term is non-negative.
One obtains the quadratic equation [ ] $=0$.
The solution set is [ ] .
Only the solution [ ] satisfies the case condition.

b. If x is in the interval [ ] the absolute value term is negative.
One obtains the normalised quadratic equation [ ] $=0$.
The solution set is [ ] .

So, altogether the solution set is $L=$ [ ] .

Solution:
If x is in the interval $[3;\infty[$ the absolute value term is non-negative and one obtains the quadratic equation $x^2-3x-9=0$ with the solution set $L=\{\frac{3}{2}-\sqrt{\frac{45}{4}};\frac{3}{2}+\sqrt{\frac{45}{4}}\}$. Only

the larger solution $\frac{3}{2} + \sqrt{\frac{45}{4}}$ satisfies the condition $x \geq 3$. This can also be seen without any calculator by estimating $\sqrt{\frac{45}{4}} \geq \sqrt{\frac{36}{4}} = 3$. In contrast, if x is in the interval $]-\infty; 3[$ the absolute value term is negative. One obtains the normalised quadratic equation $x^2 - 3x + 9 = 0$. Because of $\frac{1}{4}p^2 - q < 0$ in the pq formula this equation is unsolvable. Hence, the initial equation has only one solution $\frac{3}{2} + \sqrt{\frac{45}{4}}$.

Exercise 2.2.4
Find the solutions of the mixed absolute value equation $3|2x+1| = |x-5|$ by visualising the different cases on the number line and finally solving the equation by case analysis. First, visualise the case analysis for each absolute value.

The solution set is [ ] .

Solution:
Visualising the different cases for the expressions $|2x+1|$ and $|x-5|$ above each other indicates all cases to be distinguished:

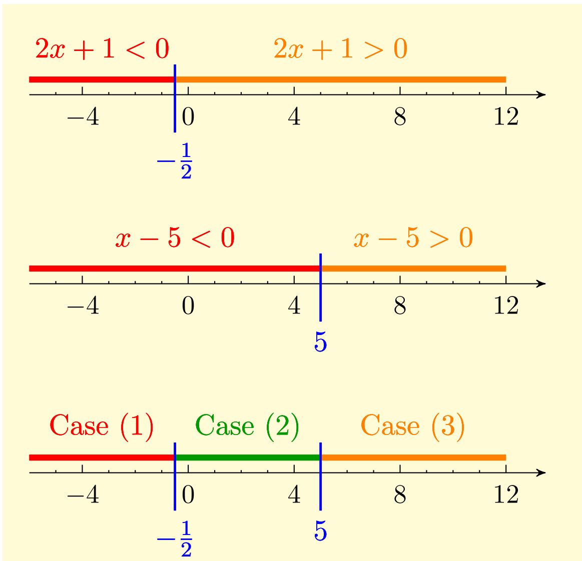


Illustration of the three cases

According to the figure above the following three cases have to be distinguished:

- Case (1): For $x < -\frac{1}{2}$ both terms in the absolute value terms are negative.
- Case (2): For $-\frac{1}{2} \leq x < 5$ the term in the second absolute value term is negative but the term in the first one is not.

- Case (3): For $5 \leq x$ both terms in the absolute value terms are non-negative.
- Obviously, there is no x for which the first term is negative and the second term is non-negative.

So, the solutions can be summarised:

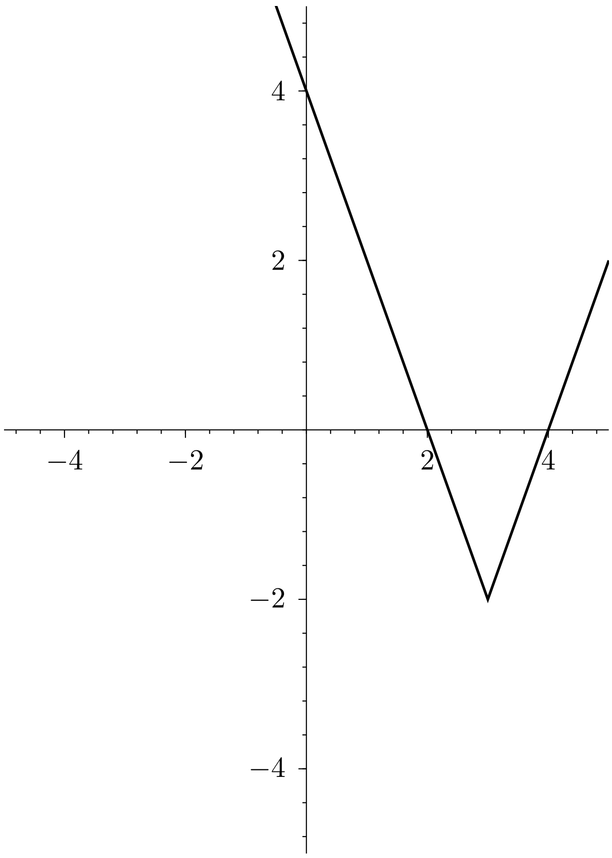
- In case (1), both absolute values reverse the sign of the terms:
  $3|2x+1| = |x-5| \Leftrightarrow 3(-(2x+1)) = -(x-5)$.
  This equation has the solution $x = -\frac{8}{5}$ satisfying the case condition.
- In case (2), only the second absolute value reverses the sign of the term:
  $3|2x+1| = |x-5| \Leftrightarrow 3(2x+1) = -(x-5)$.
  This equation has the solution $x = \frac{2}{7}$ satisfying the case condition.
- In case (3), the absolute value bars in both terms can be omitted (replaced by normal brackets):
  $3|2x+1| = |x-5| \Leftrightarrow 3(2x+1) = (x-5)$.
  This equation has the solution $x = -\frac{8}{5}$, but this solution does *not* satisfy the case condition. Thus, it will be discarded within its case analysis.

Therefore, the solution set is $\{-\frac{8}{5}; \frac{2}{7}\}$.

2.3 Final Test

2.3.1 Final Test Module 2

Exercise 2.3.1
Find an absolute value term describing the following graph of a function as simply as possible:



Graph of the function $f(x)$.

Answer: $f(x) =$ ______ .

Exercise 2.3.2
Solve the following equations:

a. $|2x - 3| = 8$ has the solution set ______ .

b. $|x - 2| \cdot x = 0$ has the solution set ______ .

Exercise 2.3.3
A camera has a resolution of 6 megapixels, i.e. – for convenience – 6 million pixels and produces images in format 2 : 3. Which size has a quadratic pixel on a print-out of format (60 cm) $\times$ (40 cm)? Specify the side length of a pixel in millimetre.

Answer: [ ] (without the unit mm).

Exercise 2.3.4
Find the solution set of the mixed equation $|x-1| \cdot (x+1) = 3$.
Answer: $L =$ [ ] .

3 Inequalities in one Variable

Module Overview

An inequality arises when we relate two terms using one of the comparators $\leq$, $<$, $\geq$, or $>$. Simple inequalities usually have intervals as their solution sets. But solving inequalities is often more difficult than solving equations. Hence, specific types of inequalities will be explained in more detail.

3.1 Inequalities and their Solution Sets

3.1.1 Introduction

Info3.1.1

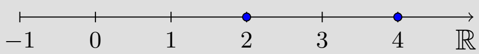
If two numbers are related by one of the **comparators** $\leq$, $<$, $\geq$, or $>$, a statement is generated that can be true or false depending on the numbers:

- $a < b$ (reads: "a is strictly less than b" or simply "a is less than b") is true if the number a is less than and not equal to b.
- $a \leq b$ (reads: "a is less than b") is true if a is less than or equal to b.
- $a > b$ (reads: "a is strictly greater than b" or simply "a is greater than b") is true if the number a is greater and not equal to b.
- $a \geq b$ (reads: "a is greater than b") is true if the number a is greater than or equal to b.

The comparators describe how the given values are related to each other on the number line: $a < b$ means that a is to the left of b on the number line.

Example 3.1.2

The statements $2 < 4$, $-12 \leq 2$, $4 > 1$, and $3 \geq 3$ are true, but the statements $2 < \sqrt{2}$ and $3 > 3$ are false.



On the number line, the number 2 is to the left of the number 4, thus $2 < 4$.

Here, $a < b$ means the same as $b > a$, likewise $a \leq b$ means the same as $b \geq a$. But it should be noted that the opposite of the statement $a < b$ is the statement $a \geq b$ and not $a > b$. If terms with a variable occur in an inequality, the problem is to find the number range of the variable such that the inequality is true.

3.1.2 Solving simple Inequalities

If the variable occurs isolated in the inequality, the solution set is an interval, see also info box 1.1.5 auf Seite 9:

Info3.1.3

The **solved inequalities** have the following **intervals** as their solution sets:

- $x < a$ has the solution set $]-\infty; a[$, i.e. all x less than a.
- $x \leq a$ has the solution set $]-\infty; a]$, i.e. all x less than or equal to a.
- $x > a$ has the solution set $]a; \infty[$, i.e. all x greater than a.
- $x \geq a$ has the solution set $[a; \infty[$, i.e. all x greater than or equal to a.

Here, x is the variable and a is a specific value. If the variable does not occur in the inequality anymore, the solution set is either $\mathbb{R} =]-\infty; \infty[$ if the inequality is satisfied, or the empty set $\{\}$ if the inequality is not satisfied.

The symbol ∞ means **infinity**. A finite interval has the form $]a; b[$ which reads "all numbers between a and b". If the interval is bounded only on one side, we can write the symbol ∞ (right-hand side) or $-\infty$ (left-hand side) as the other bound.

As in the case of equations, one tries to find a solved inequality by applying transformations that do not change the solution set. The solution set can be read off from the solved inequality.

Info3.1.4

To obtain a solved inequality from an unsolved inequality the following **equivalent transformations** are allowed:

- adding a constant to both sides of the inequality: $a < b$ is equivalent to $a + c < b + c$.
- multiplying both sides of the inequality by a positive constant: $a < b$ is equivalent to $a \cdot c < b \cdot c$ if $c > 0$.
- multiplying both sides of the inequality by a negative constant and inverting

the comparator: $a < b$ is equivalent to $a \cdot c > b \cdot c$ if $c < 0$.

Example 3.1.5

The inequality $-\frac{3}{4}x - \frac{1}{2} < 2$ is solved stepwise by the above transformations:

$$\begin{aligned} & -\frac{3}{4}x - \frac{1}{2} < 2 && \Big\| \quad +\frac{1}{2} \\ \Leftrightarrow \quad & -\frac{3}{4}x < 2 + \frac{1}{2} && \Big\| \quad \cdot\left(-\frac{4}{3}\right) \\ \Leftrightarrow \quad & x > -\frac{4}{3}\left(2+\frac{1}{2}\right) && \Big\| \quad \text{simplifying} \\ \Leftrightarrow \quad & x > -\frac{20}{6} = -\frac{10}{3} \, . \end{aligned}$$

So, the initial inequality has the solution set $\left]-\frac{10}{3};\infty\right[$. Importantly, multiplying the inequality by the negative number $-\frac{4}{3}$ inverts the comparator.

Exercise 3.1.1
Are the following inequalities true or false?

- ☐ $\frac{1}{2} > 1 - \frac{1}{3}$
- ☐ $a^2 \geq 2ab - b^2$ (where a and b are unknown numbers)
- ☐ $\frac{1}{2} < \frac{2}{3} < \frac{3}{4}$
- ☐ Let $a < b$, then also $a^2 < b^2$.

Solution:
The first inequality can be simplified to $\frac{1}{2} > \frac{2}{3}$, which, after multiplying by 6, is equivalent to $3 > 4$. This statement is false. The second inequality can be simplified by collecting all numbers on the left-hand side: $a^2 - 2ab + b^2 \geq 0$. Since $a^2 - 2ab + b^2 = (a-b)^2$, this statement is true for all a and b. Multiplying the third chain of inequalities by the least common denominator 12 results in the chain of inequalities $6 < 8 < 9$. This statement is true. In contrast, the last statement is false, since for example, for $a = -1$ and $b = 1$, the term $a^2 = 1$ is not less than $b^2 = 1$. Taking the square of terms is not an equivalent transformation.

Exercise 3.1.2

Find the solution sets of the following inequalities.

a. $2x + 1 > 3x - 1$ has the solution interval $L =$ [ ] .

b. $-3x - \frac{1}{2} \leq x + \frac{1}{2}$ has the solution interval $L =$ [ ] .

c. $x - \frac{1}{2} \leq x + \frac{1}{2}$ has the solution interval $L =$ [ ] .

Solution:
Transformation of the first inequality results in

$$\begin{aligned} & 2x + 1 > 3x - 1 && \Big\| \quad +1 \\ \Leftrightarrow \quad & 2x + 2 > 3x && \Big\| \quad -2x \\ \Leftrightarrow \quad & 2 > x \end{aligned}$$

and hence the solution interval is $L = \left]-\infty; 2\right[$. Transformation of the second inequality results in

$$\begin{aligned} & -3x - \frac{1}{2} \leq x + \frac{1}{2} && \Big\| \quad +3x - \frac{1}{2} \\ \Leftrightarrow \quad & -1 \leq 4x && \Big\| \quad \cdot \frac{1}{4} \\ \Leftrightarrow \quad & -\frac{1}{4} \leq x \end{aligned}$$

and hence $L = \left[-\frac{1}{4}; \infty\right[$. Transformation of the third inequality results in

$$\begin{aligned} & x - \frac{1}{2} \leq x + \frac{1}{2} && \Big\| \quad -x \\ \Leftrightarrow \quad & -\frac{1}{2} \leq \frac{1}{2} \,. \end{aligned}$$

This statement does not depend on $x \in \mathbb{R}$ and is always true, thus the solution set is $L = \mathbb{R} = \left]-\infty; \infty\right[$.

Info3.1.6

An inequality in one variable x is **linear** if on both sides of the inequality only multiples of x and constants occur. Each linear inequality can be transformed into a solved inequality by one of the equivalent transformations described in the info box 3.1.3 auf Seite 84.

3.1.3 Specific Transformations

The following equivalent transformations are useful if the variable occurs in the denominator of an expression. But they can only be applied under certain restrictions:

Info3.1.7

Under the restriction that none of the occurring denominators is zero (the corresponding variable values are never solutions) and the fractions on both sides have the same sign, the reciprocal can be taken on both sides of the inequality while inverting the comparator.

Example 3.1.8

For example, the inequality $\frac{1}{2x} \leq \frac{1}{3x}$ is equivalent to $2x \geq 3x$ (comparator inverted) as long as $x \neq 0$. The new inequality has the solution set $]-\infty; 0]$. However, since the value $x = 0$ was excluded (and does not belong to the domain of the initial inequality either) the solution set of $\frac{1}{2x} \leq \frac{1}{3x}$ is $L = \,]-\infty; 0[\,$.

Exercise 3.1.3
Find the solution sets of the following inequalities.

a. $\frac{1}{x} \geq \frac{1}{3}$ has the solution set $L =$ [ ] .

b. $\frac{1}{x} < \frac{1}{\sqrt{x}}$ has the solution set $L =$ [ ] .

<u>Solution:</u>
For the first inequality, the value $x = 0$ is not in the domain, hence this value is excluded. For $x > 0$, taking the reciprocal while inverting the comparator is allowed and results in $x \leq 3$. Together with the condition above the solution interval is $L = \,]0; 3]$. For $x < 0$ the reciprocal rule cannot be applied. However, it can be seen, even without any rule, that none of the values $x < 0$ can be a solution, since then $\frac{1}{x}$ is negative as well and not greater than or equal to $\frac{1}{3}$.

The domain of the second inequality is $]0; \infty[$, since only for these values of x taking the square root is defined and only for $x \neq 0$ the denominators are non-zero. On the domain, taking the reciprocal while inverting the comparator is allowed and results in $x > \sqrt{x}$. Since $\sqrt{x} > 0$, the inequality can be divided by $\sqrt{x}$ resulting in $\sqrt{x} > 1$. This inequality has the solution set $L =]1;\infty[$ which occurs also in the domain.

Please note for the last part of the exercise:

Info3.1.9

Taking the square on both sides of an inequality is not an equivalent transformation and possibly does change the solution set.

For example, $x = -2$ is no solution of $x > \sqrt{x}$, but indeed a solution of $x^2 > x$. However, this transformation can be applied if the case analysis for the transformation is carried out correctly and the domain of the initial inequality is taken into account. This method is described in more detail in the next section.

3.2 Transformation of Inequalities

3.2.1 Transformation with Case Analysis

The simple linear transformations described in the previous section are equivalent transformations. They do not change the solution set of the corresponding inequality. For nonlinear inequalities, advanced solution methods are required. Usually, these methods need a case analysis depending on the sign. This is because, in contrast to the situation for equations described in Module 2, now also the comparator can be inverted when performing transformations.

Info3.2.1

If an inequality is multiplied by a term in which the variable x occurs, a case analysis is required and for each case the transformation has to be considered separately:

- For those values of x, for which the multiplied term is positive, the comparator of the inequality is unchanged.
- For those values of x, for which the multiplied term is negative, the comparator of the inequality is inverted.
- The case that the multiplied term is zero has to be excluded during the transformation and has to be considered separately, if necessary.

The solution sets found in the individual cases have to be checked with respect to the case conditions as described for the solution of absolute value equations.

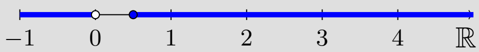
In contrast, adding terms in which the variable occurs, does not require a case analysis. Usually, transformations involving case analyses are mandatory if the variable occurs in the denominator or in a composite term.

Example 3.2.2

The inequality $\frac{1}{2x} \leq 1$ can be simplified by multiplying both sides of the inequality by the term $2x$:

- Under the condition $x > 0$ this results in the new inequality $1 \leq 2x$. It has the solution set $L_1 = \left[\frac{1}{2};\infty\right[$. The condition $x > 0$ is satisfied by all elements of the solution set.
- Under the condition $x < 0$ this results in the new inequality $1 \geq 2x$. It has the solution set $\left]-\infty;\frac{1}{2}\right]$. Because of the additional condition $x < 0$ only the elements of the set $L_2 = \left]-\infty;0\right[$ are solutions.
- The single case $x = 0$ is no solution since this value is not in the domain of the inequality. In this case multiplying the inequality by x is not allowed.

So, altogether one obtains the union set $L = L_1 \cup L_2 = \mathbb{R} \setminus \left[0;\frac{1}{2}\right[$ as solution set:



As in Module 2 the following statement holds for the solution set.

Info3.2.3

The cases have to be chosen such that all elements of the domain of the inequality are covered. For the solution set in an individual case, it has to be checked that the solution set satisfies the corresponding case condition. For any case, the resulting solution set has to be reduced to the solution subset satisfying the case condition. The union of the solution sets for the individual cases is the solution set of the initial inequality.

3.2.2 Exercises

If the inequality is multiplied by a composite term, we must investigate precisely for which values of x the case analysis must be done:

Exercise 3.2.1
Find the solution set of the inequality $\frac{1}{4-2x} < 3$. The domain of the inequality is $D = \mathbb{R} \setminus \{2\}$ since only for these values of x the denominator is non-zero. If the inequality is multiplied by the term $4 - 2x$, three cases have to be distinguished. Fill in the blanks in the following text accordingly:

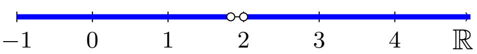
a. On the interval ______ the term is positive, the comparator remains unchanged, and the new inequality reads $1 <$ ______ . Linear transformations result in the solution set $L_1 =$ ______ . The elements of this set satisfy the case condition.

b. On the interval ______ the term is negative, the comparator is inverted. Initially, the new inequality has the solution set ______ , because of the case condition only the subset $L_2 =$ ______ is allowed.

c. The single value $x = 2$ is no solution of the initial inequality since it is not in ______ .

Sketch the solution set of the inequality and indicate the boundary points.

Solution:
On the interval $\left]-\infty; 2\right[$ the term is positive. The corresponding solution set is $\left]-\infty; \frac{11}{6}\right[$. In contrast, on the interval $\left]2; \infty\right[$ the term is negative, the comparator is inverted. Initially, the new inequality has the solution set $\left]\frac{11}{6}; \infty\right[$, because of the case condition $x > 2$ only the subset $L_2 = \left]2; \infty\right[$ is allowed. So, altogether the union set $L = L_1 \cup L_2 = \mathbb{R} \setminus \left[\frac{11}{6}; 2\right]$ is the solution set of the initial inequality excluding the boundary points:



Exercise 3.2.2
The solution set of the inequality $\frac{x-1}{x-2} \leq 1$ is $L =$ ______ .

Solution:
The domain of the inequality is $D = \mathbb{R} \setminus \{2\}$.

- For $x > 2$, multiplying the inequality by the term $x - 2$ results in $x - 1 \leq x - 2$, which is equivalent to the false statement $-1 \leq -2$. Thus, this case does not contribute a solution to the solution set.
- For $x < 2$, multiplying the inequality by the term $x - 2$ results in $x - 1 \geq x - 2$, which is equivalent to the true statement $-1 \geq -2$. Because of the case condition the solution interval for this case is only $L_2 = \,]-\infty; 2[\,$.
- The single value $x = 2$ is no solution.

So, altogether the solution set is $L = \,]-\infty; 2[$ excluding the boundary points (even though the comparator $\leq$ occurred in the initial inequality).

Exercise 3.2.3
The solution set of the inequality $\frac{1}{1-\sqrt{x}} < 1 + \sqrt{x}$ is $L =$ [ ] .

Solution:
The domain of the inequality is $D = [0; \infty[\setminus \{1\}$ since only for these values of x the square root is defined and the denominator is non-zero.

- For $0 \leq x < 1$, multiplying the inequality by the term $1 - \sqrt{x}$ results in $1 < (1 + \sqrt{x})(1 - \sqrt{x})$, which is equivalent to $1 < 1 - x$. This inequality is satisfied for $x < 0$, but these values of x violate the case condition and thus, they are not in the solution set.
- For $x > 1$, multiplying the inequality by the term $1 - \sqrt{x}$ results in $1 > 1 - x$, which is equivalent to $x > 0$. But only the values of x in the interval $]1; \infty[$ satisfy the case condition, hence $L = \,]1; \infty[$ is the only solution interval of the initial inequality.
- The single value $x = 1$ is no solution.

3.3 Absolute Value Inequalities and Quadratic Inequalities

3.3.1 Introduction

As in the approach in Module 2 and in the previous section, **absolute values** in inequalities are solved by a case analysis:

Info3.3.1

To solve an **absolute value inequality** two cases are distinguished:

- For those values of x for which the term beween absolute value bars is non-negative, the absolute value can be omitted or replaced by simple brackets, respectively.
- For those values of x for which the term between absolute value bars is negative, the term is bracketed and negated.

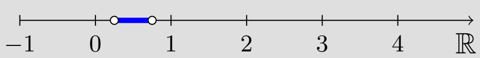
Then, the solution sets arising from the case analysis will be restricted as described in the previous module and merged to form the solution set of the initial inequality.

Example 3.3.2

To solve the absolute value inequality $|4x - 2| < 1$, two cases are distinguished:

- For $x \geq \frac{1}{2}$, the term between absolute value bars is non-negative: In this case the inequality is equivalent to $(4x - 2) < 1$ or $x < \frac{3}{4}$, respectively. Because of the case condition the solution set is only $L_1 = \left[\frac{1}{2}; \frac{3}{4}\right[$ in this case.
- For $x < \frac{1}{2}$, the term between absolute value bars is negative: In this case the inequality is equivalent to $-(4x-2) < 1$ or $x > \frac{1}{4}$, respectively. Only the subset $L_2 = \left]\frac{1}{4}; \frac{1}{2}\right[$ satisfies the case condition and contributes to the overall solution set.

The union of the two solution intervals results in the solution set $L = \left]\frac{1}{4}; \frac{3}{4}\right[$ for the initial absolute value inequality:



Exercise 3.3.1
To solve the absolute value inequality $|x-1| < 2|x-1| + x$ two cases are distinguished:

a. On the interval [ ], both terms in the absolute value terms are non-negative. The solution set of the inequality is in this case L_1 = [ ].

b. On the interval [ ], both terms in the absolute value terms are negative. The solution set of the inequality is in this case L_2 = [ ].

The union of the two intervals results in the solution interval L = [ ].

Solution:
For $x \in [1; \infty[$, both terms in the absolute value terms are non-negative, one obtains the inequality $x-1 < 2(x-1)+x$, which is equivalent to $x > \frac{1}{2}$. Because of the case condition one obtains $L_1 = [1; \infty[$ as solution set. For $x \in \,]-\infty; 1[$, both terms in the absolute value terms are negative and one obtains $-(x-1) < -2(x-1) + x$. This inequality is equivalent to the inequality $x - 1 < x$ which is always true. Thus, the solution set for the second case is $L_2 = \,]-\infty; 1[$.

Since $L = L_1 \cup L_2 = \mathbb{R} = \,]-\infty; \infty[$ the inequality is always satisfied.

3.3.2 Quadratic Absolute Value Inequalities

Info3.3.3

An inequality is called **quadratic** in x if it can be transformed into $x^2 + px + q < 0$. (Other comparators are allowed as well.)

Hence, quadratic inequalities can be solved in two ways: by investigating the roots and the orientation behaviour of the polynomial (i.e., whether the parabola opens upwards or downwards) and by completing the square. Often completing the square is simpler:

Info3.3.4

To solve an inequality by **completing the square** one tries to transform it into the form $(x+a)^2 < b$. Taking the square root then results in the absolute value inequality $|x+a| < \sqrt{b}$ with the solution set $\left]-a-\sqrt{b};-a+\sqrt{b}\right[$ if $b \geq 0$. If $b < 0$, the inequality is unsolvable.

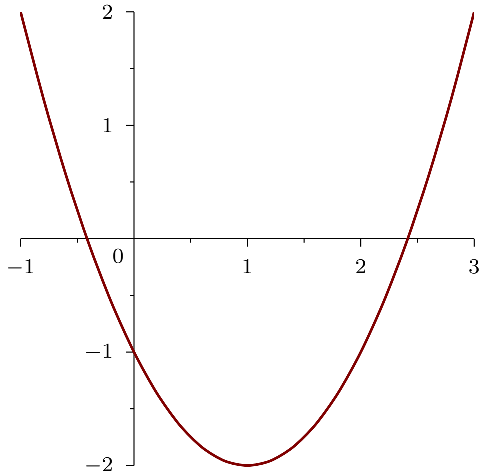
The inverted inequality $|x+a| > \sqrt{b}$ has the solution set $\left]-\infty;-a-\sqrt{b}\right[\cup \left]-a+\sqrt{b};\infty\right[$. For $\leq$ and $\geq$ the corresponding boundary points have to be included.

Always note the calculation rule $\sqrt{x^2} = |x|$ described in Module 1.

Example 3.3.5

Find the solution of the inequality $2x^2 \geq 4x+2$. Collecting the terms on the left-hand side and dividing the inequality by 2 results in $x^2-2x-1 \geq 0$. Completing the square on the left-hand side to the second binomial formula results in the equivalent inequality $x^2-2x+1 \geq 2$ or $(x-1)^2 \geq 2$, respectively. Taking the square root results in the absolute value inequality $|x-1| \geq \sqrt{2}$ with the solution set $L = \left]-\infty;1-\sqrt{2}\right] \cup \left[1+\sqrt{2};\infty\right[$.

On the other hand, the inequality $x^2-2x-1 \geq 0$ can be investigated as follows: The left-hand side describes a parabola opened upwards. The roots $x_{1,2} = 1 \pm \sqrt{2}$ can be found using the pq formula:



Since the parabola opens upwards, the inequality $x^2 - 2x - 1 \geq 0$ is satisfied by the values of x in the parabola branches left and right to the roots, i.e. by the set $L = \left]-\infty; 1 - \sqrt{2}\right] \cup \left[1 + \sqrt{2}; \infty\right[$.

Info3.3.6

Depending on the roots of $x^2 + px + q$, the orientation of the parabola and the comparator, the quadratic inequality $x^2 + px + q < 0$ (including other comparators) has one of the following solution sets:

- the set of real numbers $\mathbb{R}$,
- two branches $]-\infty; x_1[\,\cup\,]x_2; \infty[$ (including the boundary points for $\leq$ and $\geq$),
- an interval $]x_1; x_2[$ (including the boundary points for $\leq$ and $\geq$ if applicable),
- a single point x_1,
- the pointed set $\mathbb{R} \setminus \{x_1\}$,
- the empty set $\{\}$.

Fill in the blanks in the following text describing the solution of a quadratic inequality by investigating the behaviour of the parabola:

Exercise 3.3.2
Find the solution set of the inequality $x^2+6x < -5$. Transformation results in the inequality [ ] < 0. Using the pq formula one obtains the set of roots [ ] . The left-hand side describes a parabola opening [ ] . It belongs to an inequality involving the comparator $<$, hence the solution set is L = [ ] .

Solution:
Transformation results in $x^2 + 6x + 5 < 0$. Using the pq formula one obtains the roots $x_{1,2} = -3 \pm \sqrt{9-5}$, i.e. $x_1 = -1$ and $x_2 = -5$. The left-hand side describes a parabola opening upwards. It satisfies the inequality involving $<$ only on the interval $]-5;-1[$ excluding the boundary points.

3.3.3 Further Types of Inequalities

Many other types of inequalities can be transformed into quadratic inequalities. Sometimes, case analyses have to be done or excluded values in the domain have to be taken into account:

Info3.3.7

An inequality containing **fractions**, where the variable x occurs in the denominator of composite terms, can be transformed into a form without fractions by multiplying the inequality by the least common denominator. However, in doing so, the roots of the denominators have to be excluded from the domain of the new inequality.

Additionally, if the inequality is multiplied by a term, different cases have to be distinguished depending on the sign of the term.

Example 3.3.8

The inequality $2 - \frac{1}{x} \leq x$ can be transformed by multiplying the inequality by x. Here, three cases have to be distinguished:

- For $x > 0$, the comparator in the inequality is unchanged. The new inequality reads $2x - 1 \leq x^2$ and is equivalent to $x^2 - 2x + 1 \geq 0$ or $(x-1)^2 \geq 0$, respectively. This inequality is always satisfied. Because of the case condition one obtains the solution set $L_1 =]0; \infty[$.

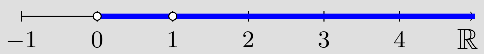
- For $x < 0$, the comparator in the inequality is inverted. The new inequality reads $2x-1 \geq x^2$ and is equivalent to $x^2-2x+1 \leq 0$ or $(x-1)^2 \leq 0$, respectively. This inequality is only satisfied for $x = 1$. But this value is excluded by the case condition, i.e. $L_2 = \{\}$.
- The single value $x = 0$ is not in the domain of the initial inequality and hence it is no solution.

So, altogether one obtains the union set $L =]0; \infty[$ as solution set of the initial inequality.

Inequalities involving composite fraction and root terms often do not have solution sets of the types described in info box 3.3.6 auf Seite 96:

Example 3.3.9

Find the solution set of the inequality $\sqrt{x} + \frac{1}{\sqrt{x}} > 2$. The domain of the inequality is $]0; \infty[$. Multiplying by $\sqrt{x}$ results in the inequality $x + 1 > 2\sqrt{x}$. Here, no case analysis is required since $\sqrt{x} > 0$ is in the domain. Transformation results in $x - 2\sqrt{x} + 1 > 0$ or $(\sqrt{x} - 1)^2 > 0$, respectively, which is satisfied for all $x \neq 1$ in the domain. Hence, the solution set of the initial inequality is $L =]0; \infty[\setminus \{1\}$:



3.4 Final Test

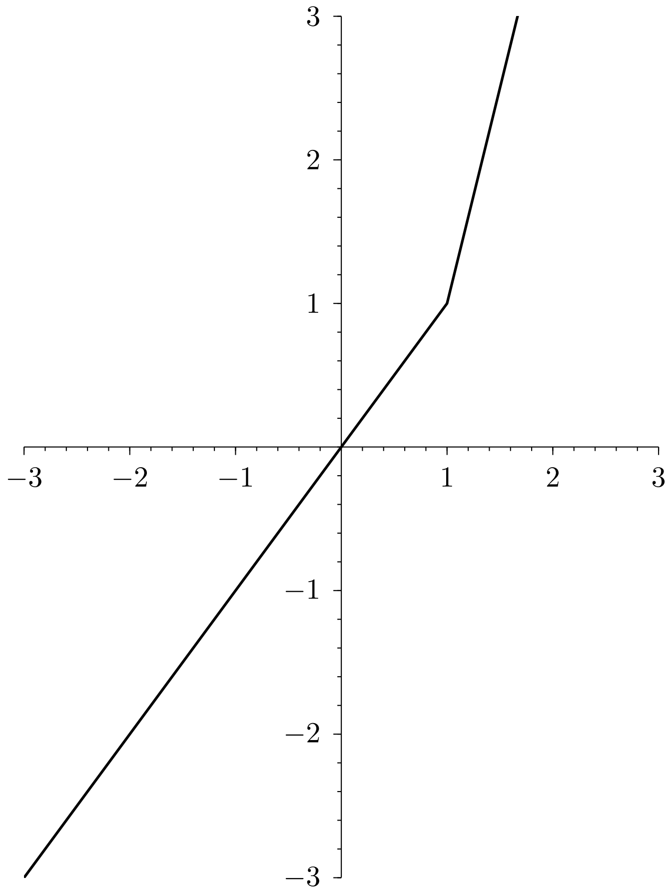
3.4.1 Final Test Module 3

Exercise 3.4.1
Find the value of the parameter α such that the inequality $2x^2 \leq x - \alpha$ has exactly one solution:

a. The parameter value is α = ______ .

b. In this case x = ______ is the only solution of the inequality.

Exercise 3.4.2
Find an absolute value function $g(x)$ describing the following graph as easy as possible.



Graph of the function $g(x)$.

Try to find a representation of the form $g(x) = |x + a| + bx + c$. The kink in the graph indicates how the absolute value term looks like.

a. Find the solution set of the inequality $g(x) \leq x$ by means of the graph.
The solution set is L = [ ] .

b. $g(x)$ = [ ] .

Exercise 3.4.3
Which positive real numbers x satisfy the following inequalities?

a. $|3x-6| \leq x+2$ has the solution set L = [ ] (written as an interval).

b. $\frac{x+1}{x-1} \geq 2$ has the solution set L = [ ] (written as an interval).

4 System of Linear Equations

Module Overview

4.1 What are Systems of Linear Equations?

4.1.1 Introduction

A problem in which several variables occur at the same time!? And on top of that, a whole slew of equations is involved!? Problems of this kind do not only occur in science and engineering but also in other academic fields and in every day life! And they all have to be solved!

But calm down now: it won't get difficult! However, it is true that in very diverse fields you will often encounter situations and problems which can be mathematically modelled by several equations in several variables. Here, we consider a simple first example.

Example 4.1.1

A young group of stuntmen want to pimp up their breakneck cycling stunt by purchasing new rims which create garish light effects for their unicycles and bicycles. For a total of 10 unicycles and bicycles 13 rims are required. How many unicycles and bicycles do the group have?

The first step is to translate the information given in the description of the problem into mathematical equations. Let x denote the desired number of unicycles and y denote the number of bicycles. Then the first information given in the problem reads

$$\text{equation (1)} \quad : \quad x + y = 10$$

since the group have 10 cycles in total. Moreover, a unicycle has one rim and a bicycle has two rims. Since 13 rims are to be purchased in total, it is also known that

$$\text{equation (2)} \quad : \quad x + 2y = 13.$$

Thus, from the problem description two equations arise relating the two variables x (number of unicycles) and y (number of bicycles).

Of course, sooner or later you want to know how many unicycles and bicycles the group of stuntmen really has. In the given example you can guess the values of x and y by a little trial and error. Here we are actually interested in the **methods** for solving problems like the example above **systematically**.

4.1.2 Contents

Before we can really start, let us clarify the terminology.

Info4.1.2

Several equations relating a specific number of variables **at the same time** form a so-called **system of equations**. If the variables in every equation occur only linearly, i.e. at most to the power of 1, and are only multiplied by (constant) numbers, the system is called a **system of linear equations**, or **LS** (linear system).

The two equations in the first example 4.1.1 auf der vorherigen Seite form a system of linear equations in the variables x and y. In contrast, the three equations

$$x + y + z = 3, \quad x + y - z = 1, \text{ and } x \cdot y + z = 2$$

do form a system of equations in the variables x, y, and z, but the system is **not** linear since in the third equation the term $x \cdot y$ occurs, which is **bilinear** in the variables x and y and hence violates the condition of **linearity**.

By the way, in a system of equations the number of equations need not be equal to the number of variables; we will return to this later on.

Info4.1.3

If the number of equations in a system equals the number of variables, the system of equations is called a **square**.

Exercise 4.1.1
Which of the following systems is a system of linear equations?

- [ ] $x + y - 3z = 0$, $2x - 3 = y$, and $1.5x - z = 22 + y$,
- [ ] $\sin(x) + \cos(y) = 1$ and $x - y = 0$,
- [ ] $2z - 3y + 4x = 5$ and $z + y - x^2 = 25$.

Systems of linear equations are distinguished from general systems of equations by their relative simplicity. Nevertheless, they play an important role in fields as diverse as medical science (e.g. for CT scans), engineering (e.g. for describing sound propagation in complex designed spaces), or physics (e.g. concerning the question of which wave lengths excited atoms can emit). It is, without doubt, worth dealing with systems of linear equations intensely.

For systems of equations generally, the question focuses on which values the variables must take such that all equations of the system are simultaneously satisfied. Such a set of values for the variables is called a **solution of a system of equations**.

Before we solve systems of equations a detail should be noted: depending on the problem, it may not be useful to accept all variable values. In the first example 4.1.1 auf Seite 103 the variables x and y are the numbers of unicycles and bicycles the group of stuntmen owns. Such numbers can only be non-negative integers, i.e. elements of $\mathbb{N}_0$. Hence, in this case the number range for the solutions has to be restricted to $\mathbb{N}_0$ in advance (namely, for both x and y).

Info4.1.4

The possible number range for the solutions of a system of equations is called the **base set** of the system. The **domain** is the subset of the base set for which all the terms of the equations of the system are **defined**. For systems of linear equations, base set and domain coincide. Finally, the **solution set** is the subset of the domain which merges the **solutions** of the system. The solution set is denoted by L.

If the base set is not explicitly specified – and cannot be derived from the problem description – we will assume implicitly that the base set is $\mathbb{R}$, the number range of the real numbers.

4.2 LS in two Variables

4.2.1 Introduction

At first we will restrict ourselves to systems of linear equations in **two** variables.

Info4.2.1

Generally, a system of linear equations (LS) consisting of two equations in the variables x and y has the following structure:

$$\begin{aligned} a_{11} \cdot x + a_{12} \cdot y &= b_1 \ , \\ a_{21} \cdot x + a_{22} \cdot y &= b_2 \ . \end{aligned}$$

Here, a_{11}, a_{12}, a_{21}, and a_{22} are the so-called coefficients of the system of linear equations, which are, as b_1 and b_2 on the right-hand sides of the equations, often real numbers given by the problem description. If the right-hand sides b_1 and b_2 both are equal to 0 ($b_1 = 0 = b_2$), the system of linear equations is called **homogeneous**, and **inhomogeneous** otherwise.

Because of their linearity, each of the two equations of the system in info box 4.2.1 can be interpreted as the equation of a line in the x-y-plane. If, for example, the first equation is solved for y,

$$y = -\frac{a_{11}}{a_{12}} x + \frac{b_1}{a_{12}} \ ,$$

one immediately sees from this explicit form that it describes a line with the slope $m = -a_{11}/a_{12}$ and the y-intercept $y_0 = b_1/a_{12}$.

Note that solving the equation for y is only allowed if $a_{12} \neq 0$. For $a_{12} = 0$, the first equation reads $a_{11} \cdot x = b_1$; for $a_{11} \neq 0$ this is equivalent to $x = (b_1/a_{11})$, i.e. x is a constant. This equation also describes a line, namely a line parallel to the y-axis with y-intercept (b_1/a_{11}).

And what about the case in which both $a_{12} = 0$ and $a_{11} = 0$? Then, we also have $b_1 = 0$, since otherwise the first equation would result in a contradiction. But for $a_{11} = a_{12} = b_1 = 0$, the first equation is (for all values of x and y) always identically satisfied ($0 = 0$), and hence useless.

The case of the second equation in info box 4.2.1 is similar:

$$y = -\frac{a_{21}}{a_{22}} x + \frac{b_2}{a_{22}} \ .$$

Altogether, one obtains two lines representing the two linear equations. The question for the solvability and for the solution of the system of linear equations, namely **the question for the simultaneous validity of the two equations** reads as **the question for the existence and the position of the intersection point of the two lines**. To this, let us investigate a specific example.

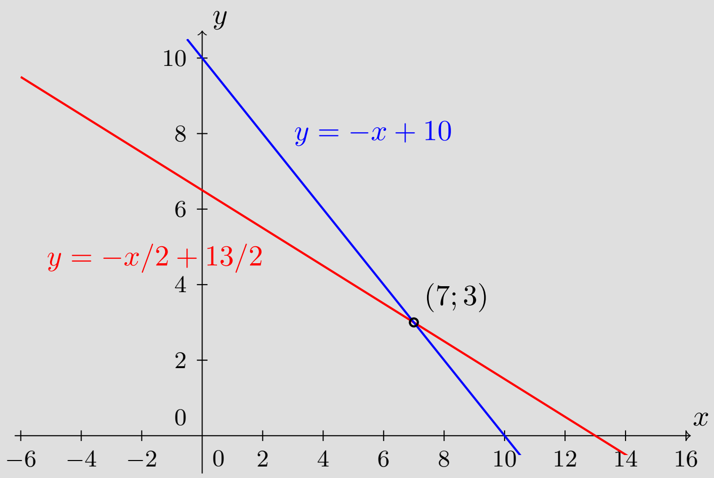
Example 4.2.2

The system of linear equations in the first example 4.1.1 auf Seite 103 reads:

$$\left.\begin{array}{rcl} x+y &=& 10 \\ x+2y &=& 13 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{rcl} y &=& -x+10 \\ y &=& -\frac{1}{2}x+\frac{13}{2} \end{array}\right. .$$

(Here, the general coefficients and right-hand sides of the system 4.2.1 auf der vorherigen Seite have the specific values $a_{11} = 1, a_{12} = 1, a_{21} = 1, a_{22} = 2, b_1 = 10$, and $b_2 = 13$.)

The equations describe two lines with the slopes $m_1 = -1$ and $m_2 = -\frac{1}{2}$, respectively, and the y-intercepts $y_{0,1} = 10$ and $y_{0,2} = \frac{13}{2}$, respectively.



The figure shows that the two lines do indeed intersect, and one reads off the coordinates of the intersection point as $(x = 7; y = 3)$. Accordingly, the system of linear equations considered here has a unique solution. The solution set consists of exactly one pair of numbers, namely $L = \{(x = 7; y = 3)\}$.

This intuitive approach is very well suited to discussing all cases which can generally occur: two lines in the x-y-plane can intersect each other – and then the intersection point is necessarily **unique** –, or the two lines are parallel and thus do not have any intersection point, or the two lines coincide – so to speak – they intersect in an infinite number of points. There are no other cases.

Accordingly, the corresponding system of linear equations has one of the following solution sets:

Info4.2.3

An **inhomogeneous system of linear equations** has either one unique solution, no solution, or an infinite number of solutions.

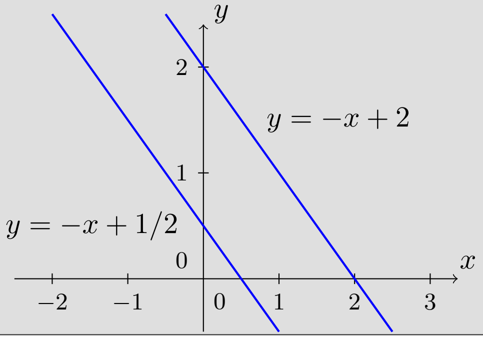
An **homogeneous system of linear equations has always at last one solution**, namely the so-called **trivial solution** $x = 0$, $y = 0$. Moreover, such a homogeneous system **can** also have **an infinite number of solutions**.

This will be illustrated by two further examples starting directly with the systems of linear equations:

Example 4.2.4

In both cases the **base set** is the set of the real numbers $\mathbb{R}$.

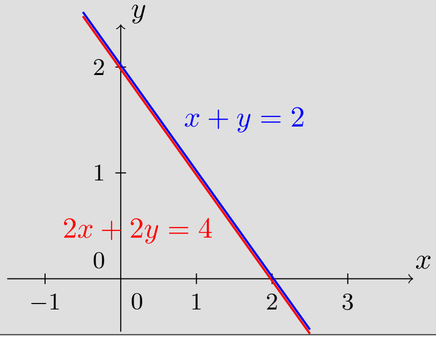
$$\left.\begin{array}{rcl} x+y &=& 2 \\ 2x+2y &=& 1 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{rcl} y &=& -x+2 \\ y &=& -x+\frac{1}{2} \end{array}\right. .$$



The two lines have the same slope $m = -1$ but the y-intercepts differ ($y_0 = 2$ or $y_0 = 1/2$, respectively). The two lines are parallel. Thus the system of linear equations has **no** solution:

$$L = \emptyset \ .$$

$$\left.\begin{array}{rcl} x+y &=& 2 \\ 2x+2y &=& 4 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{rcl} y &=& -x+2 \\ y &=& -x+2 \end{array}\right. .$$



The two lines have both the same slope $m = -1$ and the same y-intercept $y_0 = 2$. The two lines coincide. The system of linear equations has an infinite number of solutions which can be written, for example, as follows:

$$L = \{(t; -t+2) \,:\, t \in \mathbb{R}\} \ .$$

For the example on the right, other parametrisations of the solution set are possible and allowed. It basically just depends on how to describe the points of the (congruent) lines appropriately. The above description of the solution set L simply used the equation of the line itself and the control variable was denoted by t instead of x.

And what about the above mentioned restrictions due to the base set? Let us look at the following example.

Example 4.2.5

At a local fair, an exceptionally clever stallholder promises almost dreamlike prizes for an absurdly low initial wager if one, yes, only one of the passers-by is able to unravel the following little mystery:

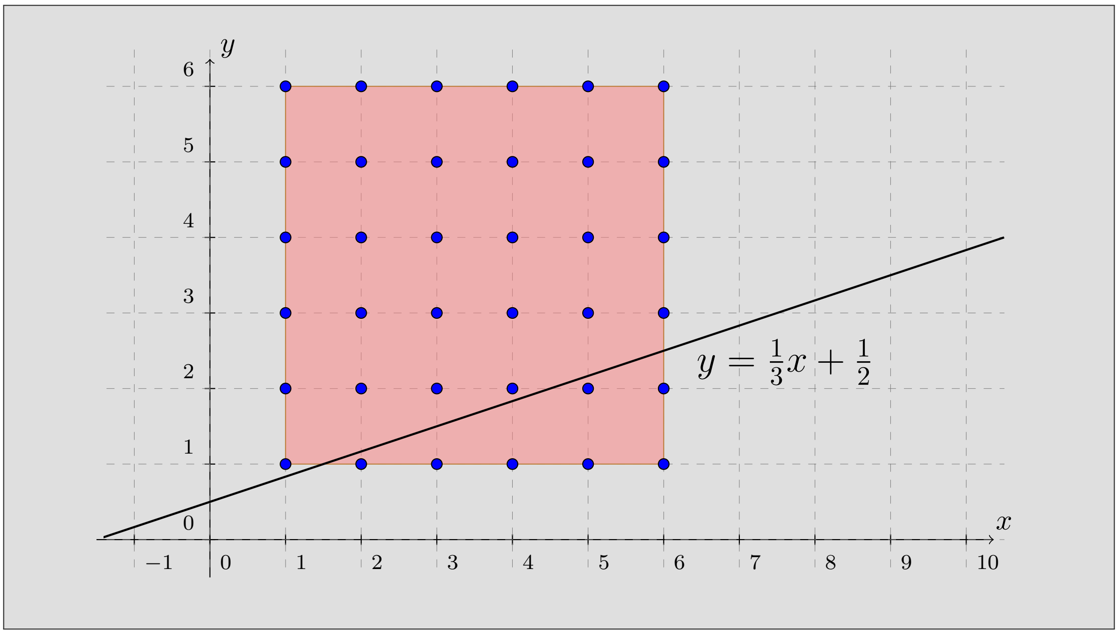
I rolled a die twice. If I subtract twice the first number I rolled from six times the second number, I get the number 3*. If I add* 6 *to four times the first number, I get the twelve times the second number. Which two numbers did I roll?*

Let the first number he rolled be denoted by x and the second by y. Then the statements of the stallholder can be translated very quickly into a set of equations:

$$\left.\begin{array}{rcl} 6y-2x & = & 3 \\ 4x+6 & = & 12y \end{array}\right\} \Leftrightarrow \left\{\begin{array}{rcl} y & = & \frac{1}{3}x+\frac{1}{2} \\ y & = & \frac{1}{3}x+\frac{1}{2} \end{array}\right. .$$

One realises that the corresponding system of linear equations – interpreted geometrically – results in two congruent lines. At first glance, it thus seems to have an infinite number of solutions.

But now the base set has to be taken into account: since both x and y represent numbers on a die, the two variables can only take values from the set $\{1;2;3;4;5;6\}$. If one looks at the line $y=\frac{1}{3}x+\frac{1}{2}$ in the x-y-plane, one sees that no possible pair of numbers thrown on a die fall onto this line. Hence, the solution set in this case is indeed the empty set $L=\emptyset$.



4.2.2 Substitution Method and Comparison Method

Until now we studied the **solvability** and the **graphical solution** of systems of linear equations of the form 4.2.1 auf Seite 106. Next, we need to study these systems algebraically. To this end, we investigate a further example.

Example 4.2.6

For the renovation of their house, the Müller family had to take out two mortgages with a total amount of 50,000 Euro. The interest they have to pay annually is in total 3,700 Euro. The interest rate for one mortgage agreement is 5% annually, and 8% annually for the other. What are the amounts of the individual mortgages?

Let the amounts of the individual mortgages be denoted by x and y. As we know from the problem description, the sum of the two amounts is 50,000 Euro. Hence, the first equation reads:

$$\text{equation (1)} \quad : x + y = 50{,}000 \quad (\text{Euro})\ .$$

The interest burden from the mortgage agreement at a rate of 5% is $0.05 \cdot x$, and the interest burden from the other one at a rate of 8% is $0.08 \cdot y$. From the problem description we know that the two values add up to 3,700. This results in a second equation:

$$\text{equation (2)} \quad : 0.05x + 0.08y = 3{,}700 \quad (\text{Euro}) .$$

Again, one ends up with a system of linear equations as in example 4.2.1 auf Seite 106.

To solve the system algebraically, the first equation is solved for y. This results in an equation $(1')$, which is equivalent to equation (1):

$$\text{equation } (1') \quad : y = 50{,}000 - x .$$

This equation for y can be substituted into equation (2). In the resulting equation only the variable x occurs, so it can be solved for x:

$$\begin{aligned} & 0.05x + 0.08(50{,}000 - x) = 3{,}700 \\ \Leftrightarrow \quad & 0.05x + 4{,}000 - 0.08x = 3{,}700 \\ \Leftrightarrow \quad & 0.03x = 300 \\ \Leftrightarrow \quad & x = 10{,}000 . \end{aligned}$$

Substituting the solution for x into equation $(1')$ results in

$$\begin{aligned} & y = 50{,}000 - 10{,}000 \\ \Leftrightarrow \quad & y = 40{,}000 . \end{aligned}$$

Hence, the mortgage amounts are 10,000 Euro (mortgage with interest of 5% annually) and 40,000 Euro (mortgage with interest of 8% annually).

The previous example illustrates the characteristics of the so-called **substitution method**:

Info4.2.7

In the **substitution method**, as a first step one of the two linear equations is solved for one of the variables – or a multiple of one of the variables. As a second step the solution is **substituted** into the other linear equation. Only three cases can occur:

(i) In the resulting equation (after collecting like terms) the other variable still occurs. Solving the resulting equation for this other variable results in the first part of the solution. The second part is obtained, for example, by substituting the solution of the first part into the equation from the first step. The solution is unique. (If this solution does not belong to the base set, it has to be excluded.)

(ii) In the resulting equation (after collecting like terms) the other variable does not occur any more and the equation is a contradiction. Then the system of linear equations has no solution.

(iii) In the resulting equation (after collecting like terms) the other variable does not occur any more and the equation is always true. Then the system of linear equations has an infinite number of solutions (unless the definition of the base set results in some restrictions).

This approach involves certain details. It is not defined which of the linear equations is to be solved for which of the variables – or multiples of the variable. As long as the applied transformations are equivalent any of the possible ways will result in the same solution. Preferring a specific way is partly a matter of taste and partly a matter of skills: a clever choice can simplify some intermediate calculations.

Concerning the **substitution method**, cases (ii) and (iii) mentioned above shall be illustrated again by means of the systems of linear equations in example 4.2.4 auf Seite 108:

Example 4.2.8

For both systems of linear equations the base set is $\mathbb{R}$.

$$\begin{aligned} \text{equation (1)}: \quad x+y &= 2 \\ \text{equation (2)}: \quad 2x+2y &= 1 \end{aligned} .$$

Solving equation (1) for x results in $x=2-y$. Substituting this equation into equation (2) results in:

$$\begin{aligned} & 2(2-y)+2y=1 \\ \Leftrightarrow\quad & 4-2y+2y=1 \\ \Leftrightarrow\quad & 4=1\ . \end{aligned}$$

This is a contradiction. The LS has **no** solution.

$$\begin{aligned} \text{equation (1)}: \quad x+y &= 2 \\ \text{equation (2)}: \quad 2x+2y &= 4 \end{aligned} .$$

Solving equation (1) for y results in $y=2-x$. Substituting this equation into equation (2) results in:

$$\begin{aligned} & 2x+2(2-x)=4 \\ \Leftrightarrow\quad & 2x+4-2x=4 \\ \Leftrightarrow\quad & 4=4\ . \end{aligned}$$

This is always true. The LS has **an infinite number** of solutions.

The substitution method is not the only approach for solving systems of linear equations. In the following section another method is discussed, which is closely related to the graphical solution of a LS.

Info4.2.9

In the **comparison method**, as a first step **both** linear equations are solved for one of the variables – or a multiple of one of the variables. As a second step the two resulting equations will be **equated**. Then the three cases discussed for the substitution method can occur.

This approach involves certain details as well. For example, it is not defined for which variable the linear equations are to be solved.

For illustration, the first example is solved again, this time by means of the comparison method:

Example 4.2.10

The system of linear equations in the first example reads:

$$\begin{aligned} x + y &= 10 \ , \\ x + 2y &= 13 \ . \end{aligned}$$

Both equations are solved for x

$$\begin{aligned} x &= 10 - y \ , \\ x &= 13 - 2y \ , \end{aligned}$$

and the right-hand-sides of the two equations are equated

$$10 - y = 13 - 2y$$

which results in $y = 3$. The solution for y can be substituted into one of the equations solved for x which results in $x = 7$.

Exercise 4.2.1
Find the solution set of the following system of linear equations

$$\begin{aligned} 7x + 2y &= 14 \ , \\ 3x - 5y &= 6 \end{aligned}$$

using the comparison method.

<u>Solution:</u>
For example, both equations are solved for x: For this, the first equation is multiplied by (1/7) and solved for x:

$$x = \frac{14}{7} - \frac{2}{7}y \Leftrightarrow x = 2 - \frac{2}{7}y \ : \ \text{equation } (1') \ .$$

In contrast, the second equation is multiplied by $\frac{1}{3}$ before solving it for x:

$$x = \frac{6}{3} + \frac{5}{3}y \Leftrightarrow x = 2 + \frac{5}{3}y \ : \ \text{equation } (2') \ .$$

Equating the two right-hand sides of equation $(1')$ and $(2')$ results in

$$2 - \frac{2}{7}y = 2 + \frac{5}{3}y \Leftrightarrow 0 = (\frac{2}{7} + \frac{5}{3})y \Leftrightarrow y = 0 \ .$$

With this solution for y, for example, the first equation results in

$$7x + 2 \cdot 0 = 14 \Leftrightarrow 7x = 14 \Leftrightarrow x = 2 \ .$$

Hence, the given system of linear equations is solvable uniquely and the solution set is $L = \{(x = 2; y = 0)\}$.

Alternatively, before equating them, the two equations could have been solved for y (or a multiple of x or a multiple of y). The solution is always the same.

4.2.3 Addition Method

We will now discuss another, third method for solving systems of linear equations algebraically. But this method will develop its full potential only for larger systems, i.e. systems of many equations in many variables since it can be systematised very well. Here, we will discuss the general approach. First, let us see an example.

Example 4.2.11

Find the solution of the system of linear equations

$$\begin{aligned} \text{equation (1)}: \quad 2x + y &= 9 \ , \\ \text{equation (2)}: \quad 3x - 11y &= 1 \ , \end{aligned}$$

where the base set is the range of the real numbers $\mathbb{R}$.

This time, the approach is as follows: The first equation is multiplied by the factor 11 and this results in the equation (1') that is equivalent to equation (1):

$$\begin{array}{rrcll} & (2x+y)\cdot 11 & = & 9\cdot 11 & \\ \Leftrightarrow & 22x+11y & = & 99 & : \text{ equation } (1')\ . \end{array}$$

Subsequently, the new equation $(1')$ is added to equation (2), i.e. the **sum** of the left-hand sides of $(1')$ and (2) is equated to the **sum** of the right-hand sides of $(1')$ and (2). In doing so, the variable y is cancelled out. This was the reason for selecting the factor 11 in the previous step.

$$3x - 11y + 22x + 11y = 1 + 99 \Leftrightarrow 25x = 100 \Leftrightarrow x = 4\ .$$

The get the solution for y the just obtained solution for x can be substituted, e.g. into equation (1):

$$2\cdot 4 + y = 9 \Leftrightarrow 8 + y = 9 \Leftrightarrow y = 1\ .$$

Thus, this system of linear equations has a unique solution $L = \{(x = 4; y = 1)\}$.

As for the other methods,the approach here is not uniquely defined: for example, equation (1) could have been multiplied by 3 and equation (2) by (-2)

$$\begin{array}{rclcrcll} (2x+y)\cdot 3 & = & 9\cdot 3 & \Leftrightarrow & 6x+3y & = & 27 & : \text{ equation } (1'')\ , \\ (3x-11y)\cdot(-2) & = & 1\cdot(-2) & \Leftrightarrow & -6x+22y & = & -2 & : \text{ equation } (2'')\ . \end{array}$$

In the subsequent **addition** of equation $(1'')$ and equation $(2'')$ the variable x could have been eliminated:

$$6x + 3y - 6x + 22y = 27 - 2 \Leftrightarrow 25y = 25 \Leftrightarrow y = 1\ .$$

To get the solution for x, the solution for y then could have been substituted, e.g. into equation (2)

$$3x - 11\cdot 1 = 1 \Leftrightarrow 3x = 12 \Leftrightarrow x = 4\ .$$

Info4.2.12

In the addition method, one of the linear equations is transformed by multiplying it by an arbitrary factor such that in the subsequent **addition** of the other equation (at least) one variable is eliminated. (Sometimes it is easier to multiply **both** equations by arbitrary factors before **adding** them.) As for the substitution method in info box 4.2.7 auf Seite 111 (or the comparison method in info box 4.2.9 auf Seite 113), three cases can occur, resulting in a solution set L containing exactly one element, no element, or an infinite number of elements.

4.2.4 Exercises

Exercise 4.2.2
Solve the following systems of linear equations using the substitution method:

a. $3x + y = 4$ and $-x + 2y = 1$,

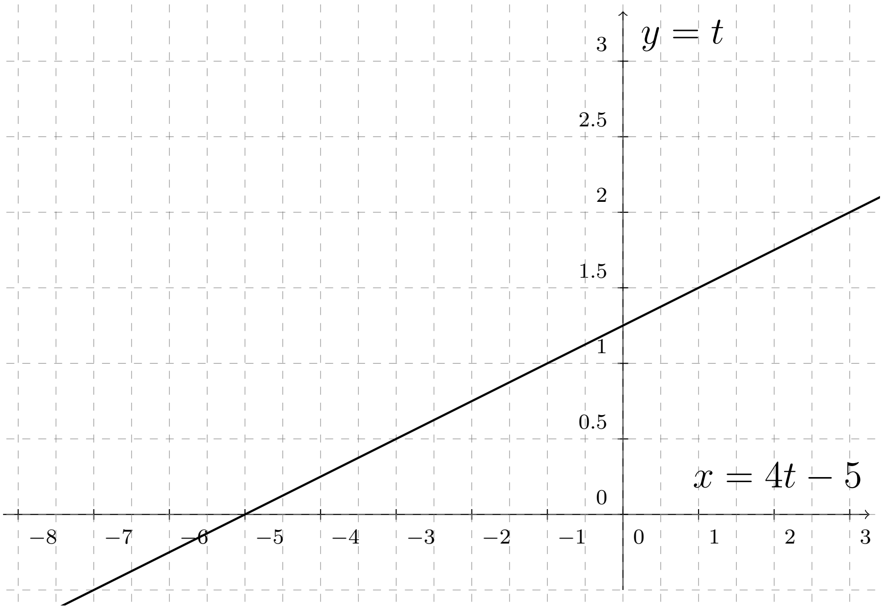
b. $-x + 4y = 5$ and $2x - 8y = -10$.

Solution:

a. For example, solving the first equation ($3x + y = 4$) for y results in $y = 4 - 3x$. Then this equation can be substituted into the second equation ($-x + 2y = 1$): $-x + 2(4 - 3x) = 1 \Leftrightarrow -x + 8 - 6x = 1 \Leftrightarrow -7x = -7 \Leftrightarrow x = 1$. With this value of x, for example, the second equation results in $3 \cdot 1 + y = 4 \Leftrightarrow y = 1$. So, the solution set is $L = \{(1; 1)\}$.

   Of course, one could start differently: For example, one could solve the first equation for x and substitute the solution for x into the second equation, to get a solution for y; or one could start generally with the second equation and solve this in the first step for x or for y. Hence, there are some details in the approach.

b. Solving, for example, the first equation ($-x + 4y = 5$) for x results in $x = 4y - 5$. Then, this equation can be substituted into the second equation ($2x - 8y = -10$): $2(4y - 5) - 8y = -10 \Leftrightarrow 8y - 10 - 8y = -10 \Leftrightarrow 0 = 0$. So, this is not a new statement, in other words: the second equation does not contain new information. Thus, in this case the solution set L contains an infinite number of solution pairs $(x; y)$ which can be parametrised by a real number t. If one chooses, for example, $y = t$, the solution set reads $L = \{(4t - 5; t) \,:\, t \in \mathbb{R}\}$. This solution set can be visualised as a line in two-dimensional space:



Accordingly, other parametrisations of the solution set are possible, for example, by choosing $x \in \mathbb{R}$ as free parameter and characterising the above line by its slope and its y-intercept, i.e. $L = \{(x; \frac{1}{4}x + \frac{5}{4}) \ : \ x \in \mathbb{R}\}$.

Exercise 4.2.3
Solve the following systems of linear equations using the addition method:

a. $2x + 4y = 1$ and $x + 2y = 3$,

b. $-7x + 11y = 40$ and $2x + 5y = 13$.

Solution:

a. Multiplying the second equation ($x + 2y = 3$) by (-2) results in equation $(2')$: $-2x - 4y = -6$. Then the last equation is added to the first equation ($2x + 4y = 1$): $2x + 4y - 2x - 4y = 1 - 6 \Leftrightarrow 0 = -5$. This is a contradiction! Hence, the solution set for this system of equations is the empty set $L = \emptyset$.

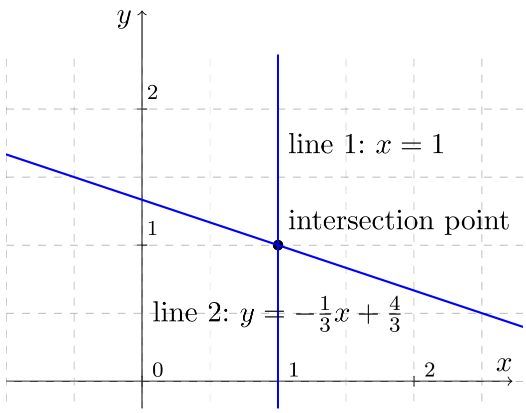
b. Multiplying the first equation ($-7x + 11y = 40$) by 2 results in the equation $(1')$: $-14x + 22y = 80$. Multiplying the second equation by 7 results in the equation $(2')$: $14x + 35y = 91$. Subsequently, adding equation $(1')$ and equation $(2')$ results in $-14x + 22y + 14x + 35y = 80 + 91 \Leftrightarrow 57y = 171 \Leftrightarrow y = 3$. Substituting this value of y, for example, in the second equation results in $2x + 5 \cdot 3 = 13 \Leftrightarrow 2x = 13 - 15 \Leftrightarrow 2x = -2 \Leftrightarrow x = -1$. Hence, the solution set L reads $L = \{(-1; 3)\}$.

Exercise 4.2.4

Solve the following system of linear equations graphically: $2x = 2$ and $x + 3y = 4$.

<u>Solution:</u>

The first equation ($2x = 2$) is equivalent to $x = 1$: This equation describes a line parallel to the y-axis through the point $(1; 0)$ on the x-axis. The second equation ($x + 3y = 4$) can be transformed into $y = -\frac{1}{3}x + \frac{4}{3}$. This equation also describes a line, this time its slope is $-\frac{1}{3}$ and its y-intercept is $\frac{4}{3}$. The graphs are as in the following figure.



¿From this figure one reads off the coordinates of the intersection point as $(x = 1; y = 1)$. Hence, the solution set is $L = \{(1; 1)\}$.

4.3 LS in three Variables

4.3.1 Introduction

In the following section we will slightly increase the level of difficulty and discuss slightly more complex systems.

Example 4.3.1

While playing, three children find a wallet with 30 Euro in it. The first child says: “If I keep the money for myself, I will have twice as much money as you both!” whereupon the second child proudly boasts: “And if I simply pocket the found money, I will have three times as much money as you both!” The third child can only smile smugly: “And if I take the money, I will be five times as rich as you two!” How much money did the children own before they found the wallet?

Let the Euro amounts which the three children owned before the find be denoted by x, y, and z, respectively. The statement of the first child can be translated into an algebraic equation as follows:

$$x + 30 = 2(y + z) \Leftrightarrow x - 2y - 2z = -30 \ : \ \text{equation (1)} \ .$$

Likewise, the statement of the second child can be translated into

$$y + 30 = 3(x + z) \Leftrightarrow -3x + y - 3z = -30 \ : \ \text{equation (2)} \ .$$

And finally, the statement of the third child is translated into

$$z + 30 = 5(x + y) \Leftrightarrow -5x - 5y + z = -30 \ : \ \text{equation (3)} \ .$$

So there arises a system of three linear equations in tree variables denoted here by x, y, and z.

The reader who is interested in the solution of this little puzzle will find it below worked out in detail using both the **substitution method** (see example 4.3.5 auf Seite 127) and the addition method (see example 4.3.7 auf Seite 128).

Info4.3.2

A system of three linear equations in the three variables x, y, and z has the following form:

$$\begin{aligned} a_{11} \cdot x + a_{12} \cdot y + a_{13} \cdot z &= b_1 \ , \\ a_{21} \cdot x + a_{22} \cdot y + a_{23} \cdot z &= b_2 \ , \\ a_{31} \cdot x + a_{32} \cdot y + a_{33} \cdot z &= b_3 \ . \end{aligned}$$

Here, a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, and a_{33} are the **coefficients** and b_1, b_2, and b_3 the right-hand sides of the **system of linear equations**.

Again, the **system of linear equations** is called **homogeneous** if the right-hand sides b_1, b_2, and b_3 are zero ($b_1 = 0$, $b_2 = 0$, $b_3 = 0$). Otherwise, the system is called **inhomogeneous**.

4.3.2 Solvability and Comparison Method, graphical Interpretation

As described in section 4.2 auf Seite 106 for systems of two linear equations in two variables, the question for solvability and the solution of the system can be traced back very clearly to the question for existence and position of the intersection point of two lines. And of course, one should think about whether for systems of three linear equations a similar graphical interpretation can be found.

If the previous two-dimensional space (for the variables x and y) is supplemented by another dimension or variable, namely z, then using a linear equation in this three variables

$$a_{11} \cdot x + a_{12} \cdot y + a_{13} \cdot z = b_1$$

a **plane** is represented **in general form**. This equation is similar to the equation of a line we already investigated. For $a_{13} \neq 0$, this equation can be solved for z:

$$z = \tfrac{b_1}{a_{13}} - \tfrac{a_{11}}{a_{13}} \cdot x - \tfrac{a_{12}}{a_{13}} \cdot y \ ,$$

which is the **explicit form** of the equation of the very same plane. The last equation assigns every pair $(x; y)$, i.e. every point in the x-y-plane, a value z, i.e. a height in three-dimensional space. Thus, a surface above the x-y-plane is created, which is a plane due to the linearity of the equation.

Now, not only the first equation of the system 4.3.2 has to be satisfied but **simultaneously** also the second and third equation which can be graphically interpreted as planes as well. If we are now interested in the solution of a system of three linear equations, we have to investigate – in the graphical interpretation – the **intersection behaviour** of three planes. On this point, let us first investigate an example.

Example 4.3.3

Find the solution set of the following system of linear equations

$$\begin{aligned} \text{equation (1)}: \quad x+y-z &= 0\,, \\ \text{equation (2)}: \quad x+y+z &= 6\,, \\ \text{equation (3)}: \quad 2x-y+z &= 4\,. \end{aligned}$$

The base set is the set of real numbers $\mathbb{R}$.

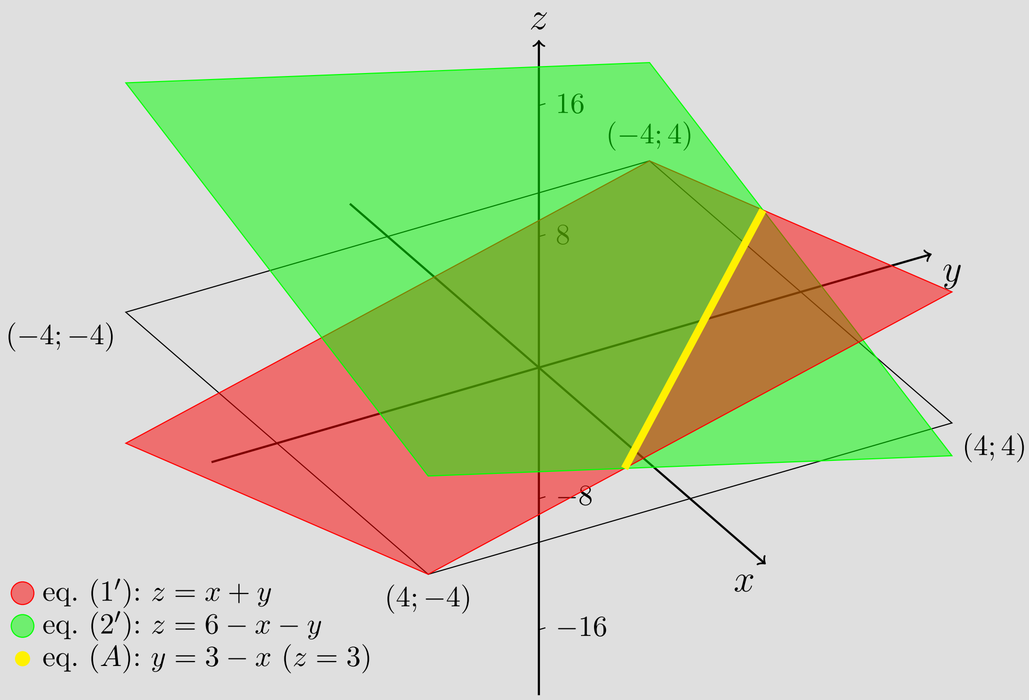
Each of the three equations can be solved for z easily:

$$\begin{aligned} \text{equation } (1'): \quad z &= x+y\,, \\ \text{equation } (2'): \quad z &= 6-x-y\,, \\ \text{equation } (3'): \quad z &= 4-2x+y\,. \end{aligned}$$

Equating the right-hand sides of equation $(1')$ and equation $(2')$ corresponds graphically to the determination of the **intersection line** of the two planes described by these equations:

$$x+y=6-x-y \Leftrightarrow 2x+2y=6 \Leftrightarrow y=3-x \;:\; \text{equation } (A)\,.$$

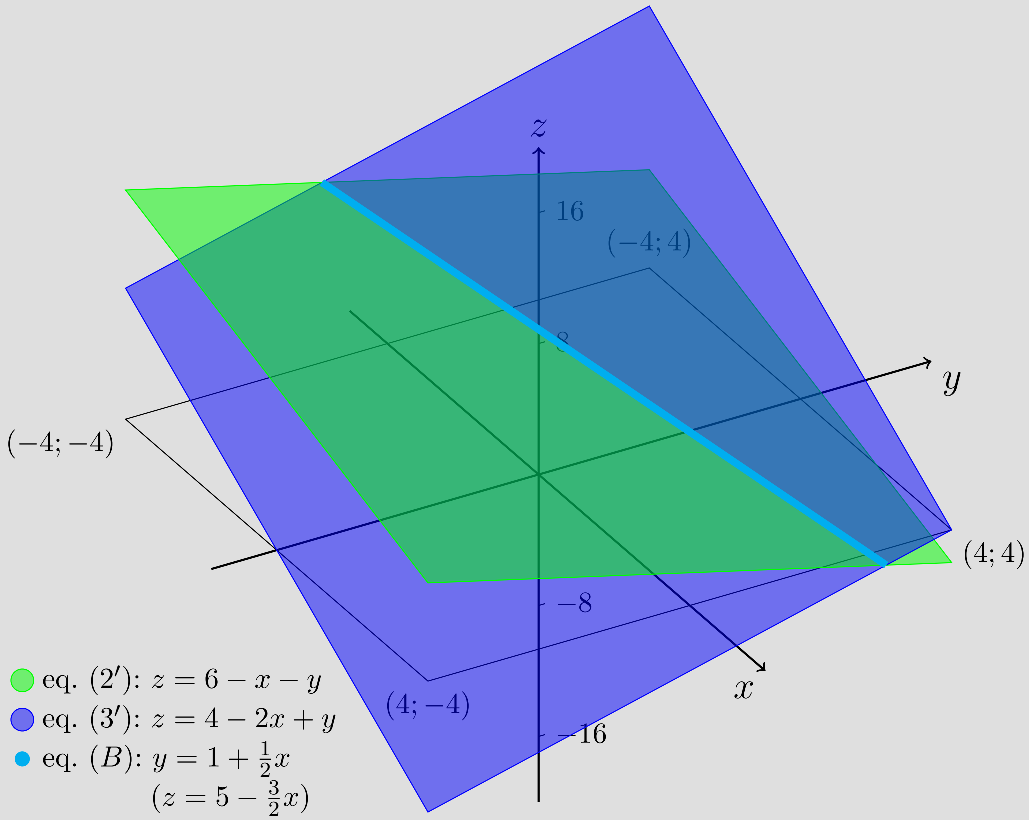
Substituting this relation into equation $(1')$ or equation $(2')$ results in an equation for the z-coordinate of the intersection line; in this case we have $z=3$. The following figure shows this intersection line described by equation (A) as the intersection of the two non-parallel planes described by equation $(1')$ and equation $(2')$.



The totally analogous statement holds if the right-hand sides of equation $(2')$ and equation $(3')$ are **equated**. Then, one obtains for the **intersection line** of the planes (2) and (3):

$$6 - x - y = 4 - 2x + y \Leftrightarrow x - 2y = -2 \Leftrightarrow y = 1 + \tfrac{1}{2}x \ : \ \text{equation } (B)\ .$$

Substituting this relation into equation $(2')$ or equation $(3')$ results in an equation for the z-coordinate of the intersection line; in this case we have $z = 5 - \frac{3}{2}x$. The following figure shows the intersection line described by equation (B) as the intersection of the two non-parallel planes described by equation $(2')$ and equation $(3')$.



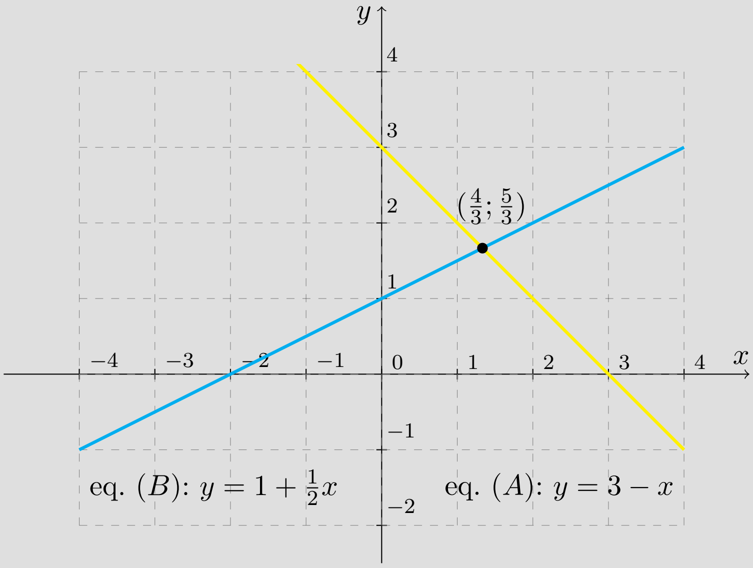
Since all equations in the initial linear system have to hold simultaneously, the two equations for the two intersection lines of the three planes also have to hold simultaneously. Graphically, this is the case at the intersection point of the intersection lines. This intersection point is found by **equating** the right-hand sides of equation (A) and equation (B):

$$3 - x = 1 + \tfrac{1}{2}x \Leftrightarrow \tfrac{3}{2}x = 2 \Leftrightarrow x = \tfrac{4}{3}\ .$$

The value of y can be calculated by inserting the value of x, for example, in equation (A):

$$y = 3 - \tfrac{4}{3} \Leftrightarrow y = \tfrac{5}{3} \ .$$

The following figure shows the intersection lines described by equation (A) and equation (B) in the x-y-plane (view from above) and their intersection point.



Finally, the value of z results by inserting the values of x and y, for example, in equation $(1')$:

$$z = \tfrac{4}{3} + \tfrac{5}{3} \Leftrightarrow z = \tfrac{9}{3} \Leftrightarrow z = 3 \ .$$

Thus, the given system of linear equations has a unique solution. The solution set is $L = \{(x = \tfrac{4}{3}; y = \tfrac{5}{3}; z = 3)\}$.

The graphical interpretation is now used to describe the solution sets of systems of three linear equations that can occur:

- If (at least) **two** of the three **planes are parallel** to each other (without being congruent), the system has **no solution**: Planes that are parallel (without being congruent) do not intersect and hence the equations describing the planes cannot hold simultaneously.
- If **two** of the three **planes are congruent**, the intersection set with the third (non-parallel) plane is an **intersection line**. All points on this intersection line

are solutions of the system. Hence, the **solution set** is infinite.

- If the **three planes are congruent**, all points of the (congruent) planes are solutions of the system. Again, the **solution set** is infinite.
- A unique solution can only exist in this last case: The three (non-parallel and non-congruent) planes have **three intersection lines** (intersection of plane (1) and plane (2), intersection of plane (2) and plane (3), and intersection of plane (1) and plane (3)):
  - If **two of the three intersection lines are parallel**, the system has **no solution**.
  - If **two of the three intersection lines are congruent**, the system has **an infinite number of solutions**.
  - If **the three intersection lines intersect in one point**, the **solution is unique** and the **solution set** consists of **exactly one element**.

Despite this illustrative graphical interpretation, the case analysis for the solution set is rather complex. Therefore, algebraic methods for investigating the solvability of systems of linear equations and for finding their solution sets will be all the more important, in particular, if the systems will get larger and the graphical interpretation gets more complex or even impossible. The **addition method**, which will be discussed again below, is one of these suitable methods.

By the way, in the above example 4.3.3 auf Seite 122 it is not necessary to find the third intersection line and to check whether this third intersection line intersects the two other lines in their intersection point: This is automatically true since equating the right-hand sides of equation $(1')$ and equation $(2')$ (first intersection line/equation (A)) and equating the right-hand sides of equation $(2')$ and equation $(3')$ (second intersection line/equation (B)) guaranties the validity of the equation of the third intersection line (right-hand side of $(1')$ = right-hand side of $(3')$):

$$x + y = 4 - 2x + y \Leftrightarrow 3x = 4 \ : \ \text{equation } (C) \ .$$

In the example, the comparison method was used since it relates very closely to the geometric interpretation. **Equating** explicit equations of planes or lines, respectively, corresponds precisely to the determination of intersection lines and points.

Info4.3.4

In the comparison method, as a first step the three linear equations are solved for one of the variables – or for a multiple of one of the variables. Then, the resulting new equations are **equated** in pairs. It is sufficient to equate two pairs. Altogether, a **system of two linear equations** in only two variables results, which can be subsequently be investigated using methods described in section 4.2 auf Seite 106.

Exercise 4.3.1
Find the solution set of the following system of linear equations

$$\begin{aligned} -x+z &= 2\,, \\ -x+y+2z &= 1\,, \\ y+z &= 11\,. \end{aligned}$$

Use the comparison method and be clever!

<u>Solution:</u>
The first equation does not depend an the variable y. Hence, it would be clever to eliminate y from the second and third equations as well. For this, the second and third equations are solved for y:

$$y = 1 + x - 2z \quad \text{and} \quad y = 11 - z\,.$$

Then the right-hand sides are equated:

$$1 + x - 2z = 11 - z \Leftrightarrow -x + z = -10\,.$$

Next, this equation has to be used together with the first equation of the initial system. Obviously, the left-hand sides of these two equations (the combination $-x+z$ of the variables x and z) are identical whereas the right-hand sides (the values 2 and -10) differ. This is a contradiction and hence the system of linear equations has no solution: $L = \emptyset$.

Other clever approaches exist.

4.3.3 Substitution Method

The **substitution method** was already discussed in section 4.2.2 auf Seite 110 for systems of two linear equations. For systems of three linear equations like the system in example 4.3.2 auf Seite 121, the approach is basically the same.

Example 4.3.5

Let us return to the first example 4.3.1 auf Seite 120 of this section. The system of linear equations for the puzzle of the three tempted children reads:

$$\begin{aligned} \text{equation (1)}: \quad x - 2y - 2z &= -30 \,, \\ \text{equation (2)}: \quad -3x + y - 3z &= -30 \,, \\ \text{equation (3)}: \quad -5x - 5y + z &= -30 \,. \end{aligned}$$

For example, one can start by solving equation (1) for x:

$$x = 2y + 2z - 30 \;:\; \text{equation } (1') \,.$$

Substituting this equation in equation (2) and equation (3) eliminates the variable x from these equations:

$$-3(2y + 2z - 30) + y - 3z = -30 \Leftrightarrow -5y - 9z = -120 \;:\; \text{equation } (2') \,,$$

$$-5(2y + 2z - 30) - 5y + z = -30 \Leftrightarrow -15y - 9z = -180 \;:\; \text{equation } (3') \,.$$

This step reduces the initial system to a system of **two** linear equations in the **two** variables y and z which can be solved using the methods from the previous section 4.2 auf Seite 106. For example, equation $(2')$ can be solved for y:

$$y = 24 - \tfrac{9}{5}z \,.$$

This relation can be **substituted** into equation $(3')$:

$$-15(24 - \tfrac{9}{5}z) - 9z = -180 \Leftrightarrow 360 - 27z + 9z = 180 \Leftrightarrow 18z = 180 \Leftrightarrow z = 10 \,.$$

So, the value of y is

$$y = 24 - \tfrac{9}{5} \cdot 10 = 24 - 9 \cdot 2 = 6$$

and finally the value of x is (using, for example, equation $(1')$):

$$x = 2 \cdot 6 + 2 \cdot 10 - 30 = 12 + 20 - 30 = 2 \,.$$

Hence, the system of linear equations has a unique solution. The solution set consists of exactly one element, namely $L = \{(x = 2; y = 6; z = 10)\}$.

Since the right-hand sides of the three equations are all equal to -30, the one and other reader may ask himself whether it wouldn't be more practical to equate all the right-hand sides in pairs and to continue with the resulting equations.

But this approach is not helpful and – if you are not careful – possibly even wrong. In any case, the number of variables would not decrease by this approach. But this is exactly what the substitution method and the comparison method are for: In both approaches (and in the addition method as well), as a first and second step one variable is eliminated such that the initial system reduces to a system of two linear equations (and hence to a simpler problem).

By the way, the approach for the solution of this reduced problem does not depend on the initial approach. In other words, it is allowed and can be even clever to start with one method, e.g. the substitution method, to reduce the system of three linear equations to a system of two linear equations and to solve this simpler system using another method, e.g. the substitution method. In this sense, the methods can be mixed.

Info4.3.6

In the substitution method, as a first step one of the three linear equations is solved for one of the variables – or for a multiple of one of the variables. As a second step the resulting relation is **substituted** into one of the two other linear equations. It results a **system of only two linear equations** in the (remaining) **two** variables. This system can be solved using one of the methods described in section 4.2 auf Seite 106.

4.3.4 Addition Method

The idea of the addition method, which we already discussed a little before (see section 4.2.3 auf Seite 114), is to **add** equations of the system such that the number of variables occurring in the system is reduced. For this, one of the equations often has to be multiplied by a cleverly chosen factor before **adding** these equations.

The addition method for a system of three linear equations in three variables shall be presented in a form that can be easily applied to larger systems. To illustrate the approach, we discuss again the system in the first example 4.3.1 auf Seite 120, i.e. the example of the three little crooks.

Example 4.3.7

The system of linear equations to be solved reads then:

$$\begin{array}{lrcl} \text{equation (1)}: & x-2y-2z &=& -30\,, \\ \text{equation (2)}: & -3x+y-3z &=& -30\,, \\ \text{equation (3)}: & -5x-5y+z &=& -30\,. \end{array}$$

Equation (1) is left unchanged in the following. But equation (2) is to be replaced by a new equation resulting from the **addition** of equation (2) and equation (1) multiplied by a factor of 3 – shortly noted as $(2)+3\cdot(1)$:

$$(-3x+y-3z)+3\cdot(x-2y-2z) = -30+3\cdot(-30) \Leftrightarrow -5y-9z = -120 \;:\; \text{equation } (2')\,.$$

Likewise, equation (3) will be replaced by $(3)+5\cdot(1)$, i.e. by the **sum** of equation (3) and equation (1) multiplied by a factor of 5:

$$(-5x-5y+z)+5\cdot(x-2y-2z) = -30+5\cdot(-30) \Leftrightarrow -15y-9z = -180 \;:\; \text{equation } (3')\,.$$

The system now reads as follows:

$$\begin{array}{lrcl} \text{equation (1)}: & x-2y-2z &=& -30\,, \\ \text{equation } (2'): & -5y-9z &=& -120\,, \\ \text{equation } (3'): & -15y-9z &=& -180\,. \end{array}$$

Equation $(2')$ and equation $(3')$ do not depend on the variable x anymore – that was the intention and the reason for choosing the factors 3 and 5 above, respectively.

The subsystem that consists of the two equations $(2')$ and $(3')$ in the two variables y and z could now be solved using one of the other methods, e.g. the substitution method. But here, it should be solved completely using the addition method. For this, equation $(2')$ and equation (1) will be left unchanged in the following. In contrast, equation $(3')$ has to be replaced, namely by the sum $(3')+(-3)\cdot(2')$:

$$(-15y-9z)+(-3)\cdot(-5y-9z) = -180+(-3)\cdot(-120) \Leftrightarrow 18z = 180 \;:\; \text{equation } (3'')\,.$$

Thus, the system has changed again,

$$\begin{array}{lrcl} \text{equation (1)}: & x-2y-2z &=& -30\,, \\ \text{equation } (2'): & -5y-9z &=& -120\,, \\ \text{equation } (3''): & 18z &=& 180\,, \end{array}$$

and now has – at least concerning the left-hand side – a kind of **triangular form**.

Solving for the variables is now very simple: The last equation (equation $(3'')$) only depends on a single variable, namely z, and hence can be solved for z immediately: $z = 10$.

This value of z is then inserted in the equation in the line above (equation $(2')$) that immediately provides the value of y: $-5y-9\cdot 10 = -120 \Leftrightarrow -5y = -30 \Leftrightarrow y = 6$.

Finally, inserting the values of y and z in the first equation (equation (1)) immediately provides the solution for the remaining variable, in the example this is the variable x: $x - 2 \cdot 6 - 2 \cdot 10 = -30 \Leftrightarrow x = 2$.

An attentive reader may ask themselves whether – and if so, why – one is allowed to replace an equation in a system of equations by another equation. In the example above this occurs three times, e.g. if equation (2) is replaced by a combination of equation (2) and three times equation (1), i.e. equation $(2')$.

Finding the solution of a system of linear equations requires that all equations of the system **hold simultaneously**, i.e. in example 4.3.7 auf Seite 128 it is required that equation (1) **and** equation (2) hold which clearly implies that also

$$\text{equation } (2) + 3 \cdot \text{ equation } (1) \Leftrightarrow \text{ equation } (2')\ .$$

holds. If now equation (1) **and** equation $(2')$ hold simultaneously, then immediately follows that equation (1) and

$$\text{equation } (2') + (-1) \cdot \text{ equation } (1) \Leftrightarrow 3 \cdot \text{ equation } (2) \Leftrightarrow \text{ equation } (2)$$

hold simultaneously as well. Hence, one is allowed to replace equation (2) by equation $(2')$ in the systems of equations.

Importantly, one can see here: if the **two** equations – equation (1) and equation (2) – were both replaced by equation (2'), information would be lost and a mistake would be made. (The requirement of **only** $(2')$ instead of (1) **and** (2) is much weaker.) This is the reason why in the "new" systems some equations are left unchanged: equation (1) **and** equation (2) are in the corresponding systems equivalent to equation (1) and equation $(2')$. The same is true for the other replacement in the example above – and generally for such transformations of systems of linear equations by means of the addition method.

Info4.3.8

In the **addition method**, pairs of linear equations of the system are added while multiplying (at least) one of the equations by a clever chosen factor (or clever chosen factors) such that in the resulting equations (at least) one variable is eliminated. It has to be ensured that in the solution process no information is lost, i.e. the number of (information relevant) equations is fixed. For this, it is most clever to bring the system into **triangular from**. Then, the solution can be found very easily.

4.3.5 Exercises

Exercise 4.3.2
Find the solution set of the following system of linear equations

$$\begin{aligned} 2x - y + 5z &= 1 \,, \\ 11x + 8z &= 2 \,, \\ -4x + y - 3z &= -1 \end{aligned}$$

using

a. the substitution method,

b. the addition method.

Solution:

a. For example, the first equation ($2x - y + 5z = 1$) is solved for y resulting in $y = 2x + 5z - 1$, which is then substituted into the third equation ($-4x + y - 3z = -1$):

$$-4x + (2x + 5z - 1) - 3z = -1 \Leftrightarrow -2x + 2z = 0 \Leftrightarrow z = x \,.$$

The last result, which is already solved for the variable z, is substituted into the second equation ($11x + 8z = 2$) that does not depend on y:

$$11x + 8x = 2 \Leftrightarrow 19x = 2 \Leftrightarrow x = \tfrac{2}{19} \,.$$

Then also

$$z = \tfrac{2}{19} \,,$$

and for y one obtains from the first equation:

$$y = 2 \cdot \tfrac{2}{19} + 5 \cdot \tfrac{2}{19} - 1 = \tfrac{4+10-19}{19} = \tfrac{-5}{19} \,.$$

Hence, the system of linear equations has a unique solution and the solution set is $L = \{(x = \frac{2}{19}; y = -\frac{5}{19}; z = \frac{2}{19})\}$.
However, other approaches are equally possible.

b. By adding the first and the third equation the variable y is eliminated:

$$(2x - y + 5z) + (-4x + y - 3z) = 1 + (-1) \Leftrightarrow -2x + 2z = 0 \,.$$

Multiplying the last equation by (-4) results in

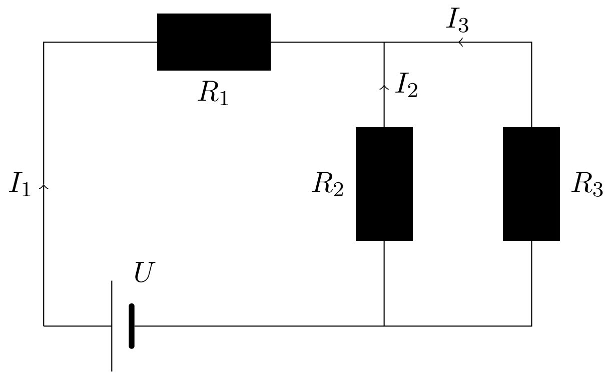
$$8x - 8z = 0 \,,$$

and adding the second equation ($11x + 8z = 2$) to this result the variable z is eliminated:

$$(8x - 8z) + (11x + 8z) = 0 + 2 \Leftrightarrow 19x = 2 \Leftrightarrow x = \tfrac{2}{19} \ .$$

As in the first part of this exercise, the variables z and y can be determined subsequently; of course, the solution set is again $L = \{(x = \frac{2}{19}; y = -\frac{5}{19}; z = \frac{2}{19})\}$. However, other approaches are equally possible.

Exercise 4.3.3
Consider the following circuit:



It consists of a source providing a voltage of $U = 5.5\,\mathrm{V}$ and three resistors $R_1 = 1\,\Omega$, $R_2 = 2\,\Omega$, and $R_3 = 3\,\Omega$. Find the currents I_1, I_2, and I_3 in the loops.

Hints: The relations between voltages, resistances, and currents in such circuits are described by the so-called **Kirchhoff's rules** which in this example result in the following three equations:

$$\begin{aligned} I_1 - I_2 - I_3 &= 0 &&: \text{ equation (1)} , \\ R_1 I_1 + R_2 I_2 &= U &&: \text{ equation (2)} , \\ R_2 I_2 - R_3 I_3 &= 0 &&: \text{ equation (3)} . \end{aligned}$$

Additionally, the relation between the physical units Volt (V) (voltage), Ampère (A) (current) und Ohm (Ω) (resistance) is used: $1\,\Omega = (1\,\mathrm{V})/(1\,\mathrm{A})$.

Solution:
For example, equation (1) is solved for I_1

$$I_1 = I_2 + I_3 \ ,$$

which is substituted into equation (2):

$$R_1(I_2 + I_3) + R_2 I_2 = U \Leftrightarrow (R_1 + R_2)I_2 + R_1 I_3 = U \ : \ \text{equation } (2') \ .$$

This last equation (2′) and equation (3) only depend on the variables I_2 and I_3. They form a system of two linear equations in two variables, which is now solved: e.g. equation (3) can be solved for I_3:

$$I_3 = \frac{R_2}{R_3} I_2 \ : \ \text{equation } (3') \ ,$$

which is then substituted into equation (2′):

$$\begin{aligned} & (R_1 + R_2) I_2 + R_1 \cdot \frac{R_2}{R_3} I_2 = U \\ \Leftrightarrow \quad & \left(R_1 + R_2 + R_1 \cdot \frac{R_2}{R_3} \right) I_2 = U \\ \Leftrightarrow \quad & \left(\frac{R_1 R_3}{R_3} + \frac{R_2 R_3}{R_3} + \frac{R_1 R_2}{R_3} \right) I_2 = U \\ \Leftrightarrow \quad & \frac{R_1 R_2 + R_1 R_3 + R_2 R_3}{R_3} I_2 = U \\ \Leftrightarrow \quad & I_2 = \frac{R_3}{R_1 R_2 + R_1 R_3 + R_2 R_3} U \ . \end{aligned}$$

Hence, for I_3 it follows, using equation (3′),

$$I_3 = \frac{R_2}{R_3} \cdot \frac{R_3}{R_1 R_2 + R_1 R_3 + R_2 R_3} U = \frac{R_2}{R_1 R_2 + R_1 R_3 + R_2 R_3} U$$

and for I_1

$$\begin{aligned} I_1 = I_2 + I_3 \quad &= \quad \frac{R_3}{R_1 R_2 + R_1 R_3 + R_2 R_3} \cdot U + \frac{R_2}{R_1 R_2 + R_1 R_3 + R_2 R_3} \cdot U \\ &= \quad \frac{R_2 + R_3}{R_1 R_2 + R_1 R_3 + R_2 R_3} \cdot U \ . \end{aligned}$$

Now, the values of the resistances of the resistors ($R_1 = 1\,\Omega$, $R_2 = 2\,\Omega$, and $R_3 = 3\,\Omega$) and the voltage ($U = 5.5\,\mathrm{V}$) can be inserted. Using $1\,\mathrm{V} = 1\,\Omega \cdot 1\,\mathrm{A}$ this results for I_1 in:

$$I_1 = \frac{(2+3)\,\Omega}{(1 \cdot 2 + 1 \cdot 3 + 2 \cdot 3)\,\Omega^2} \cdot 5.5 \cdot \Omega \cdot \mathrm{A} = 2.5\,\mathrm{A} \ .$$

Likewise, for I_2 and I_3 this results in: $I_2 = 1.5\,\mathrm{A}$ and $I_3 = 1\,\mathrm{A}$.
However, other approaches are equally possible.

Exercise 4.3.4
Find the solution set of the following system of equations using the addition method:

$$\begin{aligned} x + 2z &= 5 \ , \\ 3x + y - 2z &= -1 \ , \\ -x - 2y + 4z &= 7 \ . \end{aligned}$$

Solution:
Since the first equation ($x + 2z = 5$) does not depend on the variable y at all, it is reasonable to eliminate y from the second and third equation as well. For this, the second equation ($3x + y - 2z = -1$) is multiplied by 2 and added to the third equation ($-x - 2y + 4z = 7$):

$$\begin{aligned} & (-x - 2y + 4z) + 2 \cdot (3x + y - 2z) = 7 + 2 \cdot (-1) \\ \Leftrightarrow\quad & -x - 2y + 4z + 6x + 2y - 4z = 7 - 2 \\ \Leftrightarrow\quad & 5x = 5 \\ \Leftrightarrow\quad & x = 1 \; . \end{aligned}$$

At the same time the variable z is also eliminated. Inserting the value for x in the first equations results in

$$1 + 2z = 5 \Leftrightarrow 2z = 4 \Leftrightarrow z = 2 \; .$$

Using the second equation the value of y is

$$3 \cdot 1 + y - 2 \cdot 2 = -1 \Leftrightarrow 3 + y - 4 = -1 \Leftrightarrow y = 0 \; .$$

Hence, the solution set is $L = \{(x = 1; y = 0; z = 2)\}$.

4.4 More general Systems

4.4.1 Introduction

To conclude this module, we will briefly discuss two further details concerning systems of linear equations.

Firstly, in systems of linear equations free parameters can occur. These are variable quantities – so-called tuning parameters – which may strongly affect the behaviour of the systems, and especially the solution sets. Sometimes it is advantageous not to fix all coefficients and right-hand sides of the equations, but to keep them variable to investigate what happens for different values. Sometimes not all coefficients and right-hand sides are known from the problem description.

Secondly, in a system of linear equations either the number of linear equations or the number of variables need not be the same.

4.4.2 Systems with a Free Parameter

We start with an example that is very easy but demonstrates the essential point concerning free parameters in systems of linear equations.

Example 4.4.1

Find the solution set of the following system of linear equations

$$\begin{array}{lrcl} \text{equation (1)}: & x-2y & = & 3\,, \\ \text{equation (2)}: & -2x+4y & = & \alpha \end{array}$$

depending on the parameter α.

Multiplying equation (1) by the factor 2 and adding equation (2) results in

$$2\cdot(x-2y)+(-2x+4y)=2\cdot 3+\alpha \Leftrightarrow 2x-4y-2x+4y=6+\alpha \Leftrightarrow 0=6+\alpha\,.$$

Now, two cases have to be distinguished:

Case A ($\alpha \neq -6$): If the given free parameter α is not equal to -6, the last equation is a contradiction. In this case the system of linear equations has **no solution**, i.e. $L=\emptyset$.

Case B ($\alpha = -6$): If the given free parameter α is equal to -6, the last equation is always satisfied ($0=0$). In fact, the two initial equations in this case are multiples of each other such that only one of them indeed carries information. Accordingly, **the solution set is infinite**: $L = \{(x = 3 + 2t; y = t) \,:\, t \in \mathbb{R}\}$.

The example shows that the solution set may strongly depend on the value of the free parameter.

Such a free parameter can occur not only on one of the right-hand sides of the system of linear equations, but also on the left-hand sides, multiple times or in a function both on the left-hand sides and on the right-hand sides. Also several parameters can occur in a system at the same time.

Let us now consider a slightly more complex example.

Example 4.4.2

Find the solution set of the following system of linear equations

$$\begin{aligned} x + y + \alpha z &= 1 \,, \\ x + \alpha y + z &= 1 \,, \\ \alpha x + y + z &= 1 \end{aligned}$$

depending on the value of the parameter α.

For this, e.g. the first equation is solved for the variable x

$$x = 1 - y - \alpha z \ : \ \text{equation } (1') \,,$$

and the result is **substituted** into the second and third equation:

$$\begin{aligned} (1 - y - \alpha z) + \alpha y + z &= 1 \ \Leftrightarrow \ -(1-\alpha)y + (1-\alpha)z = 0 &&: \text{equation } (2') \,, \\ \alpha(1 - y - \alpha z) + y + z &= 1 \ \Leftrightarrow \ (1-\alpha)y + (1-\alpha^2)z = 1 - \alpha &&: \text{equation } (3') \,. \end{aligned}$$

This results in a system of **two** linear equations in **two** variables y and z. It can immediately be seen that for the value $\alpha = 1$ something happens. Hence, a case analysis is required.

Case 1 ($\alpha = 1$): In this case the two equations $(2')$ and $(3')$ are satisfied identically $(0 = 0)$ and provide no further information. The only relation between the variables x, y, and z is equation $(1')$ or equation (1), respectively, that reads for $\alpha = 1$ as follows:

$$x + y + z = 1 \ : \ \text{equation } (\hat{1}) \,.$$

Hence, the solution set has an infinite number of elements. The set can be described using **two free parameters**, e.g.

$$L = \{(s;t;1-s-t) \,:\, s,t \in \mathbb{R}\} \,.$$

Geometrically, the solution set is exactly the plane described by equation $(\hat{1})$.

Case 2: $(\alpha \neq 1)$: In this case both equation $(2')$ and equation $(3')$ can be divided by $(1-\alpha)$. Using the third binomial formula $((1-\alpha^2) = (1-\alpha)(1+\alpha))$ results in

$$\begin{aligned} -y + z &= 0 \quad &: \text{ equation } (2'') \,, \\ y + (1+\alpha)z &= 1 \quad &: \text{ equation } (3'') \,. \end{aligned}$$

According to equation $(2'')$ one has $y = z$. This is substituted into equation $(3'')$:

$$z + (1+\alpha)z = 1 \Leftrightarrow (2+\alpha)z = 1 \,:\, \text{ equation } (\star) \,.$$

Again, one has to take care and an additional case analysis is required since $\alpha = -2$ and $\alpha \neq -2$ have different consequences:

Case 2a $(\alpha = -2)$: In this (sub)case equation $(\star)$ reads $0 = 1$. This is a contradiction and the initial system of equations has no solution, i.e. $L = \emptyset$.

Case 2b: $\alpha \neq -2$: In this (sub)case the last equation can be solved easily for z:

$$z = \frac{1}{2+\alpha} \,.$$

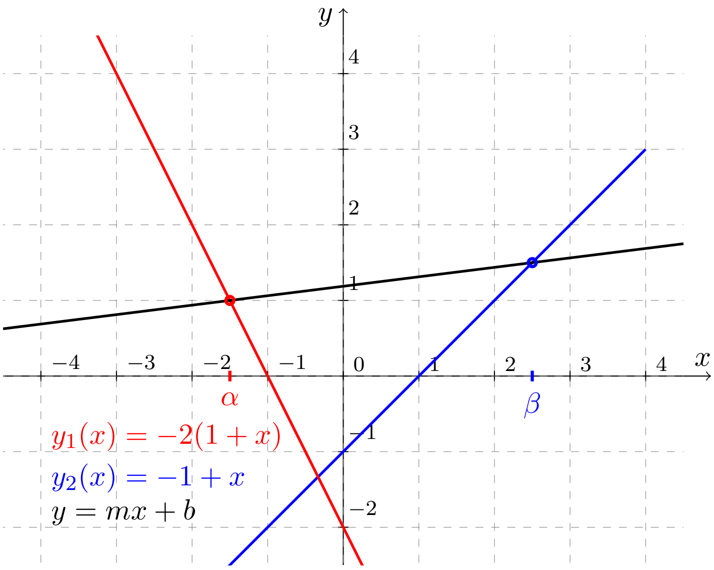
So, y $(y = z)$ and x $(x = 1 - y - \alpha z)$ are determined. The initial system of linear equations has a unique solution, namely $L = \{(x = \frac{1}{2+\alpha}; y = \frac{1}{2+\alpha}; z = \frac{1}{2+\alpha})\}$.

The previous example indicates the relevance of a clear and precise case analysis. Depending on the value of α $(\alpha = 1$, $\alpha = -2$, or $\alpha \in \mathbb{R} \setminus \{-2;1\})$ the solution set is completely different! In the first case it is an infinite set, in the second the empty set, and in the third the solution set consists of exactly one element!

By the way, the exceptionality of the case $\alpha = 1$ could have seen directly from the initial system of linear equations: For $\alpha = 1$, the same equation occurs three times, namely $x + y + z = 1$, i.e. two of the three equations in the initial system do not contribute any information and thus, they are unnecessary. For $\alpha = 1$, only the equation $x + y + z = 1$ relates the three variables.

4.4.3 Exercises

Exercise 4.4.1
Find the y-intercept b and the slope m of a line described by the equation $y = mx + b$ which is defined by two points. The first point at $x_1 = \alpha$ lies on the line described by the equation $y_1(x) = -2(1 + x)$. The second point at $x_2 = \beta$ lies on the line described by the equation $y_2(x) = x - 1$. The following figure illustrates the situation.



a. Find the system of equations for the parameters b and m.

   The first equation reads $m\alpha + b =$ [ ] ;
   the second equation reads $m\beta + b =$ [ ] .

b. Solve this system of equations for b and m. For which values of α and β does the system have a unique solution, no solution, or an infinite number of solutions?

   For $\alpha = -2$ and $\beta = 2$ one obtains, for example, the solution $m =$ [ ] and $b =$ [ ] , the case $\alpha = 2$ and $\beta = -2$ results in the solution $m =$ [ ] and $b =$ [ ] .

   The LS has an infinite number of solutions if $\alpha =$ [ ] and $\beta =$ [ ] .
   The corresponding solutions can be parameterised by $m = r$ and $b =$ [ ] , $r \in \mathbb{R}$.

c. What is the graphical interpretation of the last two cases, i.e. no solution and an infinite number of solutions?

Solution:

a. ¿From the condition

$$y(x_1 = \alpha) = y_1(x_1 = \alpha) \text{ and } y(x_2 = \beta) = y_2(x_2 = \beta)$$

the LS for m and b results in:

$$\begin{aligned} m\alpha + b &= -2(1+\alpha) \quad \text{equation } (1)\ , \\ m\beta + b &= -1 + \beta \quad \text{equation } (2)\ . \end{aligned}$$

b. If equation (2) is replaced by the difference of equation (2) and equation (1), one obtains

$$\begin{aligned} m\alpha + b &= -2(1+\alpha) \quad \text{equation } (1)\ , \\ m(\beta - \alpha) &= 1 + 2\alpha + \beta \quad \text{equation } (2')\ . \end{aligned}$$

With this, the LS has already triangular form.

A. For $\beta - \alpha \neq 0 \ \Leftrightarrow\ \beta \neq \alpha$ equation $(2')$ can be divided by $(\beta - \alpha)$, which for the slope m results in

$$m = \frac{1 + 2\alpha + \beta}{\beta - \alpha}\ .$$

For example, this result can be substituted into equation (1) resolved for b:

$$b = -2(1+\alpha) - m\alpha = -2(1+\alpha) - \frac{1 + 2\alpha + \beta}{\beta - \alpha}\alpha\ .$$

The obtained values of m and b represent a unique solution of the considered LS.

For $\alpha = -2$, $\beta = 2$, the solution is $m = -1/4, b = 5/2$, and for $\alpha = 2$, $\beta = -2$ the solution is $m = -3/4, b = -9/2$.

B. In this case $\beta - \alpha = 0 \ \Leftrightarrow\ \beta = \alpha$. So, the left-hand side of equation $(2')$ is zero. Thus, the following additional case analysis is required:

Ba. If the right-hand side of equation is non-zero, i.e. $(2')$ $1 + 2\alpha + \beta \neq 0$, the LS has no solution, i.e. $L = \emptyset$. For $\beta = \alpha$ one obtains $\beta = \alpha \neq -\frac{1}{3}$.

Bb. If the right-hand side of equation $(2')$ is zero, the LS only consists of equation (1) with $\beta = \alpha = -\frac{1}{3}$. In this case $m = \lambda$, $\lambda \in \mathbb{R}$ can be chosen arbitrary, and b results directly from equation (1): $b = \frac{1}{3}\lambda - \frac{4}{3}$. Thus, the solution set of the LS is

$$L = \left\{ \left(m = \lambda; b = \frac{1}{3}\lambda - \frac{4}{3} \right) \ :\ \lambda \in \mathbb{R} \right\}\ .$$

c. The right-hand sides of equation (1) and equation (2) contain the y-values of the initial lines at $x_1 = \alpha$ and $x_2 = \beta$, i.e. $y_1(\alpha)$ and $y_2(\beta)$, such that the right-hand side of equation (2′) is the difference $y_2(\beta) - y_1(\alpha)$. In case B, i.e. for $\beta = \alpha$, one then obtains $y_2(\alpha) - y_1(\alpha)$. So, the LS has no solution (case Ba) if $y_2(\alpha) - y_1(\alpha) \neq 0$ $\Leftrightarrow y_2(\alpha) \neq y_1(\alpha)$, i.e. the two lines have different y-values. On the other hand an infinite number of solutions exist (case Bb) if the two initial equations intersect at $x = \alpha$.

Exercise 4.4.2

Find the solution set of the following LS depending on the parameter $t \in \mathbb{R}$.

$$\begin{aligned} x - y + tz &= t && \text{equation (1) ,} \\ tx + (1-t)y + (1+t^2)z &= -1 + t && \text{equation (2) ,} \\ (1-t)x + (-2+t)y + (-1+t-t^2)z &= t^2 && \text{equation (3) .} \end{aligned}$$

The LS only has solutions for the following values of the parameter: $t \in$ [ ] .

For the smallest value of the parameter t the solution is
$x =$ [ ] , $y =$ [ ] , $z = r$, $r \in \mathbb{R}$.
For the greatest value of the parameter t the solution is
$x =$ [ ] , $y =$ [ ] , $z = r$, $r \in \mathbb{R}$.

Solution:
As a first step, equation (3) is replaced by the sum of equation (2) and equation (3):

$$\begin{aligned} x - y + tz &= t && \text{equation (1) ,} \\ tx + (1-t)y + (1+t^2)z &= -1 + t && \text{equation (2) ,} \\ x - y + tz &= -1 + t + t^2 && \text{equation (3}') \text{ .} \end{aligned}$$

Since the left-hand side of equation (1) and equation (3′) coincide, it is useful to replace equation (3′) by the sum of the negative of equation (1) and equation (3′):

$$\begin{aligned} x - y + tz &= t && \text{equation (1) ,} \\ tx + (1-t)y + (1+t^2)z &= -1 + t && \text{equation (2) ,} \\ 0 &= -1 + t^2 && \text{equation (3}'') \text{ .} \end{aligned}$$

subsequently, replacing equation (2) by the sum of $(-t)$-times equation (1) and equation (2) results in a LS in triangular form:

$$\begin{aligned} x - y + tz &= t && \text{equation (1) ,} \\ y + z &= -1 + t - t^2 && \text{equation (2}') \text{ ,} \\ 0 &= -1 + t^2 && \text{equation (3}'') \text{ .} \end{aligned}$$

Finally, adding equation (2′) to equation (1) results in a further simplification,

$$\begin{aligned} x+(1+t)z &= -1+2t-t^2 && \text{equation } (1')\ , \\ y+z &= -1+t-t^2 && \text{equation } (2')\ , \\ 0 &= -1+t^2 && \text{equation } (3'')\ . \end{aligned}$$

This equivalent LS only has a solution if equation (3′′) is satisfied as well, i.e. $0=-1+t^2$ $\Leftrightarrow t^2=1$. In all other cases the solution set is the empty set, i.e.

$$L=\emptyset \text{ für } t\in\mathbb{R}\setminus\{-1;1\}\ .$$

If equation (3′′) is now satisfied, i.e. $t=1$ or $t=-1$, then z can be chosen arbitrarily, i.e. $z=\lambda$, $\lambda\in\mathbb{R}$. Solving equation (1′) and equation (2′) for x and y, respectively, for $t=1$ results in the expressions $x=-2z$ and $y=-1-z$, so the solution set is

$$L=\{(x=-2\lambda; y=-1-\lambda; z=\lambda)\ :\ \lambda\in\mathbb{R}\} \text{ for } t=1\ .$$

Accordingly, for $t=-1$ one obtains $x=-4$ and $y=-3-z$, so the solution set is

$$L=\{(x=-4; y=-3-\lambda; z=\lambda)\ :\ \lambda\in\mathbb{R}\} \text{ for } t=-1\ .$$

4.5 Final Test

4.5.1 Final Test Module 4

Exercise 4.5.1
Find the solution set of the following system of linear equations:

$$\begin{aligned} -x+2y &= -5\,, \\ 3x+y &= 1\,. \end{aligned}$$

The solution set ☐ is empty,
☐ contains exactly one element: $x =$ ☐ , $y =$ ☐ ,
☐ contains an infinite number of solution pairs $(x;y)$.

Exercise 4.5.2
Find the two-digit number such that its digit sum is 6 and exchanging the tens and the units digit results in a number which is 18 less. Answer: ☐ .

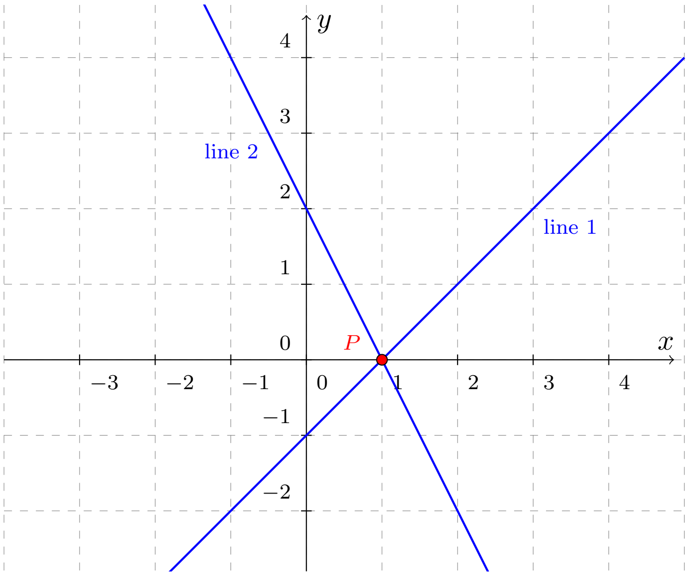
Exercise 4.5.3
Find the value of the real parameter α for which the following system of linear equations

$$\begin{aligned} 2x+y &= 3\,, \\ 4x+2y &= \alpha \end{aligned}$$

has an infinite number of solutions.

Answer: $\alpha =$ ☐ .

Exercise 4.5.4
The following figure shows two lines in two-dimensional space.



Find the two equations describing the lines.

Line 1: $y =$ ☐ ,

Line 2: $y =$ ☐ .

What is the number of solutions of the corresponding system of linear equations?

The system of linear equations has
☐ no solution,
☐ exactly one solution, or
☐ an infinite number of solutions.

Exercise 4.5.5

Find the solution set of the following system of linear equations consisting of three equations in three variables:

$$\begin{aligned} x + 2z &= 3\,, \\ -x + y + z &= 1\,, \\ 2y + 3z &= 5\,. \end{aligned}$$

The solution set
☐ is empty,
☐ contains exactly one solution: $x =$ ☐ , $y =$ ☐ , $z =$ ☐ ,
☐ contains an infinite number of solutions $(x; y; z)$.

5 Geometry

Module Overview

The first sections of this chapter will introduce you to elementary geometry, while referring to the previous chapters. As a main topic, we first deal with the properties of triangles before calculating areas of polygons and volumes of simple geometric solids. Advanced problems are solved by means of trigonometric functions. These will give us a first taste of the later modules on calculus and analytic geometry.

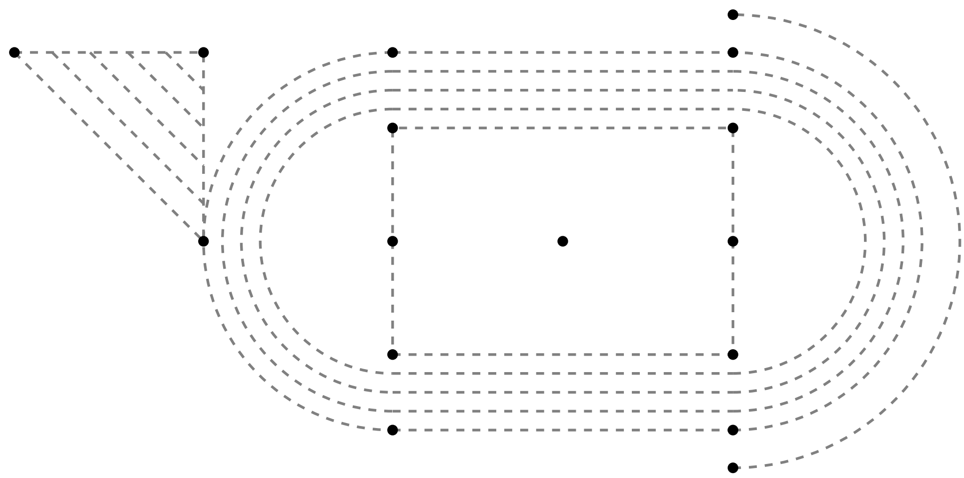
5.1 Elements of Plane Geometry

5.1.1 Introduction

Looking at the stars on a clear, moonless night conveys a vivid impression of the elementary objects of geometry, namely the points. Since time immemorial, people mentally connected the points of light in the night sky by lines which they interpreted as the contours of highly diverse characters. Every building, with its vertex corners, edges, and faces, provides evidence of the practical use of this “heavenly” geometry, ...

On the other hand, the invention of pencils, wax tablets, papyrus or paper enabled people to capture their thoughts and observations “on paper” and to show them to others. For example, the desire to realise a drawing as a physical building resulted in the concept of a plan. A plan is a drawing of an idealised image showing, for example, how a stadium shall look from above.

For the construction of a stadium, significant points are staked out in the terrain. The current status of the project is shown in the following drawing containing the contours and significant points from a plan.



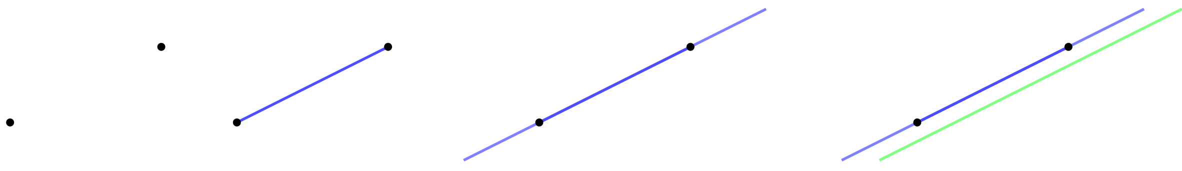
(Measurement) points and lines from a construction plan of a stadium

The drawing can be considered as an idealised image of reality. Along these lines, we will first recapitulate some basic concepts of geometry. Then, applying these concepts, we will construct more complicated figures and geometric solids.

5.1.2 Points and Lines

In geometry, a place or a position in a plane is idealised to the most basic object, namely a point. A single point itself cannot be characterised any further.

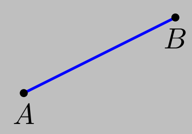
For several points, relations between these points can be considered in different ways — and points can be used to define new objects such as line segments and lines (see figure below). Mathematically, these objects are sets of points.



First, we consider a line segment and the distance between points. To do this, we need a comparison tool for measuring distance. In mathematics, this tool is a comparative length called the unit length. For applications, appropriate length units such as metres or centimetres are chosen, depending on the task in hand.

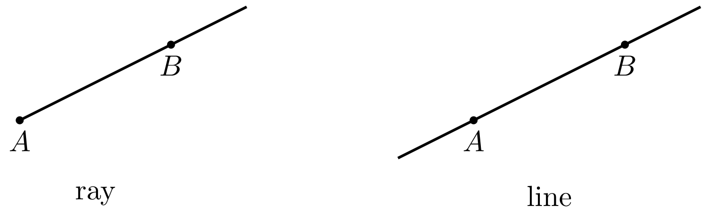
Line Segments and Distances5.1.1

Given two points A and B, the **line segment** $\overline{AB}$ between A and B is the shortest path between the two points A and B.



The length of the line segment $\overline{AB}$ is denoted by $[\overline{AB}]$. The **line length** equals the distance between the two points A and B.

A ray of light emitted by a distant star or by the sun is an appropriate notion of a **ray** starting at the initial point A and proceeding through a second point B indefinitely. A ray is also called a **half-line**.

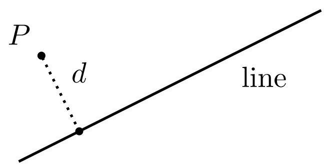


Continuing the path of a line segment $\overline{AB}$ on both ends indefinitely results in a line.

Line5.1.2

Let A and B be two points (i.e. point A is different from point B). Then, A and B define exactly one **line** AB.

Considering, beside A and B, an additional point P, we can ask for the distance d of the point P from the line AB, which is defined as the shortest path between P and one of the points of the line AB.



Given three points P, Q, and S in the plane, the lines SP and SQ can be defined.

The two lines have the point S in common. If the point Q is also on the line SP, then SQ and SP denote one and the same line. If the point Q does not belong to the line SP, the line SQ is different from the line SP. Then, the two lines have only the point S in common. The point S is called an **intersection point**.

If any two lines g and h do not have any points in common, the smallest distance between points on g and h, respectively, is called the distance between the lines g and h. Hence, g and h do not have any point in common if they have a distance larger than 0. Two lines are called **parallel** if every point on one of the two lines has the same distance from the other line.

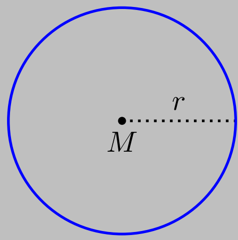
A single line can be described by the distance of two points M and M' as well: The set of all points with the same distance from two points M and M' is a line.

In geometry, it is a typical approach to define new objects by means of certain properties such as the distance. In this way, a circle can also be described very easily.

Circle5.1.3

Let a point M and a positive real number r be given.

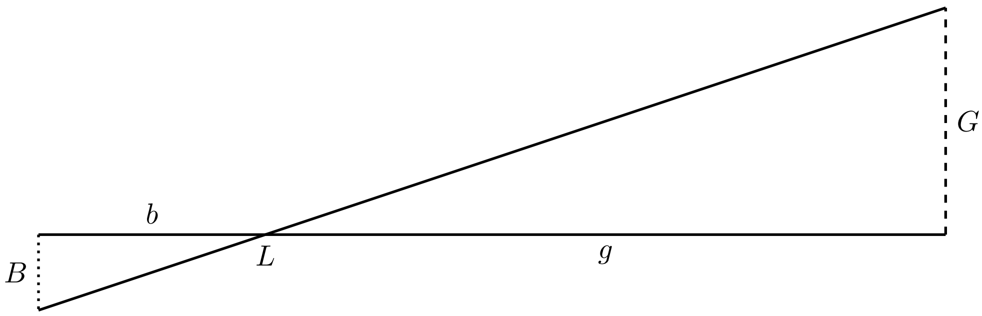
Then, the set of all points at distance r from point M is a **circle** around M with **radius** r.



5.1.3 Intercept Theorems

A pinhole camera provides a small image of the outside space. The ratio of the size of the image B to the size of the object G equals the ratio of the distance b from the pinhole L to the distance g from L:

$$\frac{B}{G} = \frac{b}{g} \ .$$



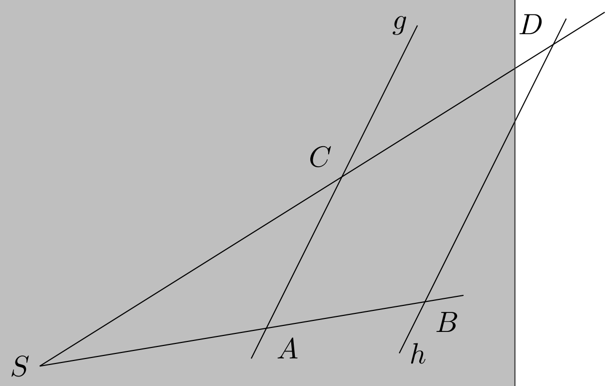
Properties of images arising from uniform scaling can also be described by means of the intercept theorems (see also Figure 5.3.4 auf Seite 174).

What all examples applying the intercept theorems have in common is that rays (or lines) with an intersection point are intersected by parallel lines.

Intercept Theorems5.1.4

Let S be the common initial point of the two rays s_1 and s_2 proceeding through the points A and C, respectively. The point B is on the ray s_1 and the point D is on the ray s_2. First, we consider the line segments between the points on the two rays and then the line segments between the rays.

For two points P and Q, $\overline{PQ}$ is the line segment from P to Q and $[\overline{PQ}]$ denotes the length of this line segment.



If the lines g and h are parallel, the following statements hold:

- The ratio of the line segments on one of the two rays equals the corresponding ratio of the line segments on the other:

$$\frac{[\overline{SA}]}{[\overline{SC}]} = \frac{[\overline{AB}]}{[\overline{CD}]} = \frac{[\overline{SB}]}{[\overline{SD}]} \ .$$

  This can also be expressed in the form:

$$\frac{[\overline{SA}]}{[\overline{AB}]} = \frac{[\overline{SC}]}{[\overline{CD}]} \quad \text{and} \quad \frac{[\overline{SA}]}{[\overline{SB}]} = \frac{[\overline{SC}]}{[\overline{SD}]} \ .$$

- The ratio of the line segments on the parallel lines equals the ratio of the corresponding line segments starting from S on a single ray

$$\frac{[\overline{SA}]}{[\overline{SB}]} = \frac{[\overline{AC}]}{[\overline{BD}]} = \frac{[\overline{SC}]}{[\overline{SD}]} \ .$$

  This can also be expressed in the form:

$$\frac{[\overline{SA}]}{[\overline{AC}]} = \frac{[\overline{SB}]}{[\overline{BD}]} \quad \text{and} \quad \frac{[\overline{SC}]}{[\overline{CA}]} = \frac{[\overline{SD}]}{[\overline{DB}]} \ ,$$

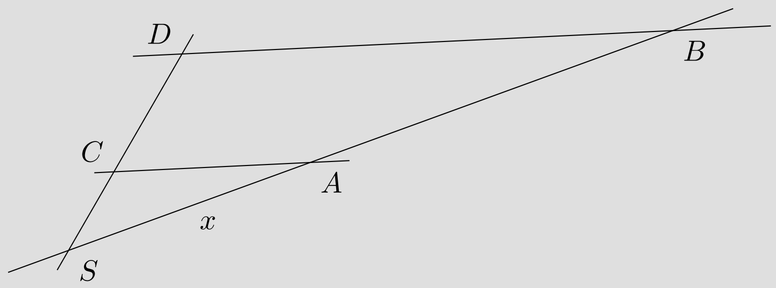
  where $[\overline{AC}] = [\overline{CA}]$ and $[\overline{BD}] = [\overline{DB}]$.

The statements of the intercept theorems also hold if two lines intersecting in a point S are considered instead of the two rays. An application example of this case is the pinhole camera mentioned above.

In this way, distances between points can be calculated without measuring the length of the line segments directly.

Example 5.1.5

Let four points A, B, C, and D be given. These points define the two lines AB and CD intersecting at the point S. Furthermore, it is known that the lines AC and BD are parallel. Between the points the following distances were measured: $[\overline{AB}] = 51$, $[\overline{SC}] = 12$, and $[\overline{CD}] = 18$.



From this, the distance between A and S can be calculated. Let x denote the required distance. Then, according to the intercept theorems, we have

$$\frac{x}{[\overline{AB}]} = \frac{[\overline{SC}]}{[\overline{CD}]} \ ,$$

from which

$$x = \frac{[\overline{SC}]}{[\overline{CD}]} \cdot [\overline{AB}] = \frac{12}{18} \cdot 51 = \frac{2}{3} \cdot 51 = 34$$

follows.

5.1.4 Exercises

Exercise 5.1.1
The son of the house is looking at the tree on the neighbouring property. He observes that the tree is completely covered by the hedge separating the two properties only if he stands close enough to the hedge. Now he is looking for the point at which he *just* cannot see the tree any more.

The boy is 1.40 metres tall. If the boy stands 2.50 metres away from the hedge, which is 2.40 metres high, 1 metre wide and clipped into a pointed shape at the top, the tree disappears from his sight.

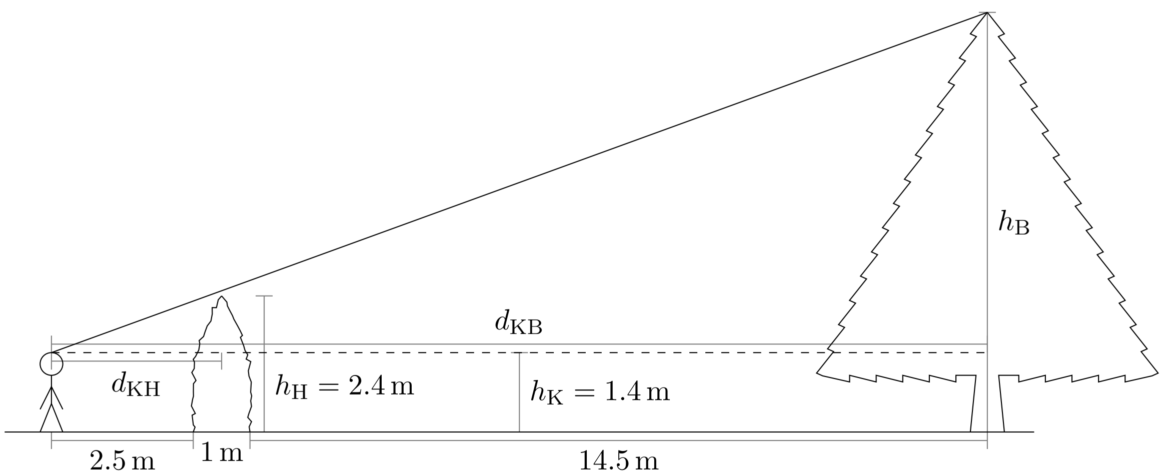
What is the height of the tree if the middle of the trunk is 14.5 metres away from the hedge?

Please carry out the calculation using variables and insert the values only at the end!

Result: ☐ m.

Hint:
Take the width of the hedge into account!

Solution:
The line segments are denoted as shown in the figure below.



Applying the second intercept theorem results in „$\frac{\text{full}}{\text{at front}} = \frac{\text{long}}{\text{short}}$“ :

$$\frac{d_{KB}}{d_{KH}} = \frac{h_B - h_K}{h_H - h_K} \qquad \text{or} \qquad h_B = (h_H - h_K) \cdot \frac{d_{KB}}{d_{KH}} + h_K\ .$$

The values are $d_{\mathrm{KH}} = 2.5\,\mathrm{m} + \frac{1\,\mathrm{m}}{2} = 3\,\mathrm{m}$ and $d_{\mathrm{KB}} = 2.5\,\mathrm{m} + 1\,\mathrm{m} + 14.5\,\mathrm{m} = 18\,\mathrm{m}$. Hence, it follows that

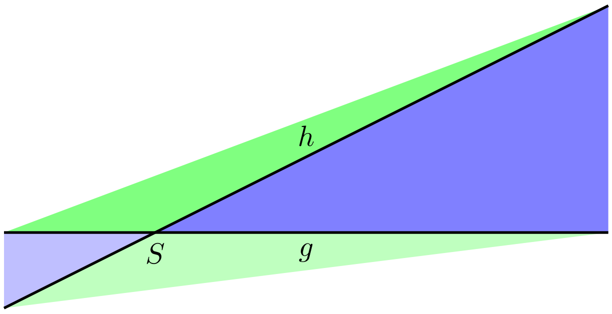
$$h_{\mathrm{B}} = (2.4\,\mathrm{m} - 1.4\,\mathrm{m}) \cdot \frac{18\,\mathrm{m}}{3\,\mathrm{m}} + 1.4\,\mathrm{m} = 1\,\mathrm{m} \cdot 6 + 1.4\,\mathrm{m} = 7.4\,\mathrm{m}\,.$$

5.2 Angles and Angle Measurement

5.2.1 Introduction

Lines intersecting at a point S divide the plane in a characteristic way. To describe how, the concept of an angle is introduced. The question how to measure angles can be answered in different ways, which in the end are all based on the subdivision of circles.

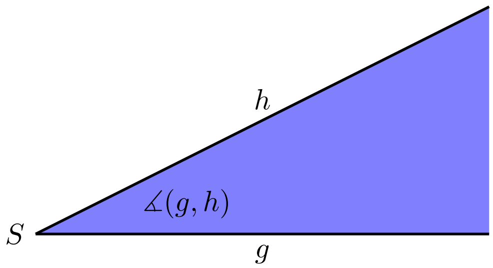
In this module, degree measure and radian measure are described.



Every coloured region represents one of the angles defined by the lines g and h.

5.2.2 Angles

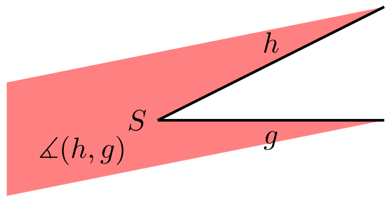
Two rays (half-lines) g and h in the plane starting from the same initial point S enclose an **angle** $\measuredangle\,(g, h)$.



Angle enclosed by the rays g and h.

For the notation of the angle $\measuredangle\,(g, h)$, the order of g and h is relevant. $\measuredangle\,(g, h)$ denotes the angle shown in the figure above. It is defined by rotating the half-line g counter-clockwise to the half-line h.

In contrast, $\measuredangle\,(h, g)$ denotes the angle from h to g as illustrated by the figure below.



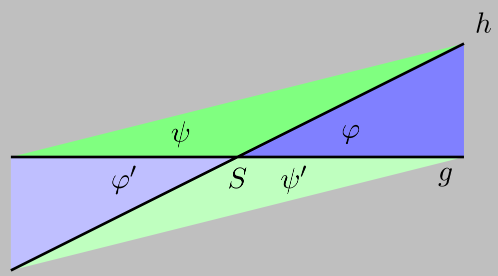
Angle enclosed by the rays h and g

The point S is called a **vertex** of the angle, and the two half-lines enclosing the angle are called the **arms** of the angle. If A is a point on the line g and B is a point on the line h, then the angle $\measuredangle(g, h)$ can also be denoted by $\measuredangle(ASB)$. In this way, angles between line segments $\overline{SA}$ and $\overline{SB}$ are described.

Angles are often denoted by lower-case Greek letters to distinguish them from variables, which are generally denoted by lower-case Latin letters (see Table 1.1.8 auf Seite 12 in module 1). Further angles can be found by considering angles formed by intersecting lines.

Vertical Angles and Supplementary Angles5.2.1

Let g and h be two lines intersecting in a point S.



- The angles φ and φ' are called **vertical angles**.
- The angles φ and ψ are called **supplementary angles** with respect to g.

The figure above contains further vertical and supplementary angles.

Exercise 5.2.1
Find all vertical and supplementary angles occurring in the figure above.

Solution:
In addition to φ and φ', ψ and ψ' are also vertical angles. Beside the angles φ and ψ, the angles φ' and ψ' are also supplementary angles of g. Moreover, ψ and φ' as well as ψ' and φ are supplementary angles.

Some special angles have their own dedicated name. For example, the angle bisector w is the half-line whose points have the same distance from the two given half-lines g and h. Then, it can be said that w bisects the angle between g and h.

Names of Special Angles5.2.2

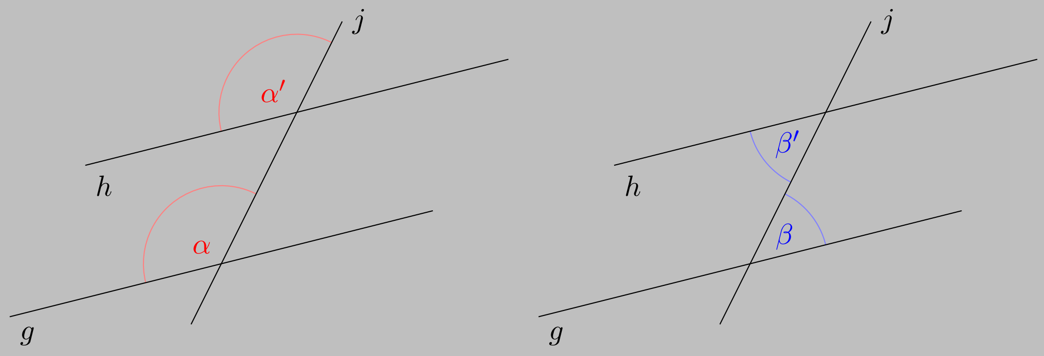
Let g and h be half-lines with the intersection point S.

- The angle covering the entire plane is called the **complete angle**.
- If the rays g and h form a line, the angle between g and h is called a straight angle.
- The angle between two half-lines bisecting a straight angle is called the **right angle**. One also says that g and h **are perpendicular (or orthogonal) to each other**.

Next, three lines are considered. Two of the three lines are parallel, while the third line is not parallel to the others. It is called a transversal. These lines form eight cutting angles. Four of the eight angles are equal.

Angles at Parallel Lines5.2.3

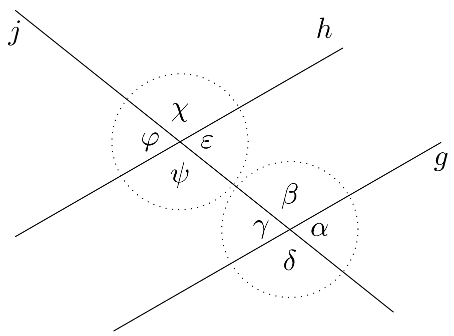
Let two parallel lines g and h be given cut by another transversal line j.



- Then the angle α' is called a **corresponding angle** of α and
- the angle β' is called an **alternate angle** of β.

Since the lines g and h are parallel, the angles α and α' are equal. Likewise, the angles β and β' are equal.

Exercise 5.2.2
The figure shows two parallel lines g and h cut by another line j. Explain which angles are equal and which angles are corresponding angles or alternate angles to each other, respectively.



Solution:

- The angles α, γ, ε, and φ are equal as well as the angles β, δ, χ, and ψ.
- The angles β and ψ as well as the angles γ and ε are alternate angles.
- The angles α and ε are corresponding angles, likewise the angles β and χ, δ and

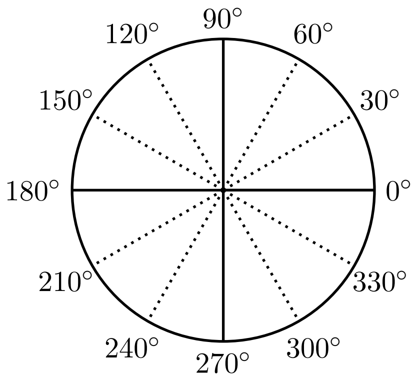
ψ as well as γ and φ.

5.2.3 Angle Measurement

We already explained the notation $\measuredangle(g,h)$ for the angle defined by rotating g counter-clockwise to h. This explanation provides an idea of how to measure angles, i.e. how to compare angles quantitatively.

Think of the face of an analogue watch with its twelve evenly spaced hour marks. Likewise, the circumference of a circle can be evenly subdivided. In this way, a certain scale for angles is obtained. Depending on the applied scaling, the magnitude of an angle can be specified in different units.

Degree Measure. A disk is subdivided into 360 equal segments. A rotation by one segment defines an angle of 1 degree. This is written as $1°$. The figure below shows angles of multiples of $30°$.



Radian Measure. In ancient Babylonia, Egypt, and Greece people had already observed that the ratio of the circumference U of a circle to its diameter D is always the same, and hence circumference and diameter of a circle are proportional to each other. This ratio is called π.

The Number π5.2.4

Let a circle with circumference U and diameter D be given. Then, the ratio of the circumference U of a circle to its diameter D is

$$\pi = \frac{U}{D} = \frac{U}{2r},$$

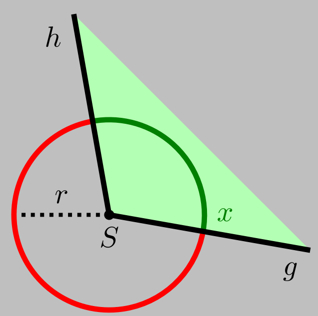
where $r = \frac{1}{2}D$ is the radius of the circle.

The number π is not a rational number. It cannot be expressed as a finite or periodic decimal fraction. From numerical calculations we know that the value of π is approximately $\pi \approx 3.141592653589793$.

If the circle has a radius of exactly 1, the circumference is 2π. Now, for the **radian measure** the circumference of a circle with radius 1 is subdivided. For the radian measure of an angle $\measuredangle(g, h)$ the length of an **arc** "cut" by this angle is used. As a result, the radian measure assigns to every angle a number between 0 and 2π. In scientific applications, the symbol rad is used to express explicitly that the angle is measured in radian measure.

Radian Measure5.2.5

Let g and h be two half-lines starting in the same initial point S and enclosing the angle $\measuredangle(g, h)$. If a circle with radius $r = 1$ is drawn around S, the two half-lines cut the circle into two pieces. Now, the angle is described by the one arc x that transforms g into h by a counter-clockwise rotation (indicated by a green line in the figure below). In other words, vertex S is always on the left if one moves on the arc x from g towards h.



The length of the arc x is the **radian measure** of the angle $\measuredangle(g, h)$.

By means of an angle measure (such as the radian and degree measures introduced previously), angles can simply be classified into different types and named accordingly.

For repetition and completeness, all names, including ones previously discussed, are listed below.

Names of Different Types of Angles5.2.6

For angles whose radian measure is in a certain range, the following names are introduced:

- An angle with a radian measure greater than 0 and less than $\frac{\pi}{2}$ is called an **acute angle**.
- An angle with a radian measure of exactly $\frac{\pi}{2}$ is called a **right angle**.
- An angle with a radian measure greater than $\frac{\pi}{2}$ and less than π is called an **obtuse angle**.
- An angle with a radian measure greater than π and less than 2π is called a **reflex angle**.

Two half-lines are said to be **perpendicular to each other** if they form a right angle.

Two half-lines form a line if they enclose an angle of radian measure π.

From the radian measure of the angle $\measuredangle(g, h)$, the radian measure of the angle $\measuredangle(h, g)$ can also be determined. From definition 5.2.5 auf der vorherigen Seite it is known that

$$\measuredangle(h, g) = 2\pi - \measuredangle(g, h) \ .$$

In the figure of definition 5.2.5 auf der vorherigen Seite the radian measure of the angle $\measuredangle(h, g)$ is the length of the red arc of the circle with radius $r = 1$.

The wording in the last sentences might seem awkward. The reason for that lies probably in the fact that we do distinguish precisely between an angle and its measure, e.g. the radian measure in this case.

When it comes to calculating a required value for line segments, the same notation is often used for a segment and its length. Mostly this is clear, and it helps to describe or to illustrate a problem efficiently. Importantly, the unit of the angle has to be known or explicitly specified. Often, such an agreement – a so-called convention – is also used if it is known from the context that a certain angle has to be calculated using a certain angle measure.

Convention5.2.7

If a calculation does not depend on a certain measure or the unit of the angles is specified in advance, the term angle is used for short denoting both the angle itself and its value in the specified measure.

Hence, for example, we can write $\measuredangle(g, h) = 90°$ and speak about the right angle $\measuredangle(g, h)$ enclosed by the lines g and h at the same time. The same idea applies for the radian measure.

The value of an angle can be converted from radian measure to degree measure (and vice versa) by considering the ratios of its value to the value of the complete angle in the respective angle measure.

The conversion from radian measure to degree measure is described below.

Relation between Radian Measure and Degree Measure5.2.8

Let g and h be two half-lines enclosing the angle $\measuredangle(g, h)$. The radian measure of the angle is denoted by x and the degree measure of the angle is denoted by α.

Then, the ratio of x to 2π equals the ratio of α to $360°$, and thus:

$$\frac{x}{2\pi} = \frac{\alpha}{360°} \;.$$

Hence,

$$x = \frac{\pi}{180°} \cdot \alpha \quad \text{and} \quad \alpha = \frac{180°}{\pi} \cdot x \;.$$

Therefore, the values in radian measure are proportional to the ones in degree measure. Thus, the conversion using the respective proportionality factors $\frac{\pi}{180°}$ and $\frac{180°}{\pi}$ is very simple.

Exercise 5.2.3

The angle $\measuredangle(g,h)$ equals 60° in degree measure. Calculate the angle in radian measure:

$\measuredangle(g,h) =$ ______ .

Solution:
From

$$\frac{\measuredangle(g,h)}{2\pi} = \frac{60^\circ}{360^\circ}$$

we have

$$\measuredangle(g,h) = \frac{60^\circ}{360^\circ} \cdot 2\pi = \frac{1}{6} \cdot 2\pi = \frac{\pi}{3} \; .$$

Exercise 5.2.4
The angle β equals $\pi/4$ in radian measure. Find its value in degree measure.

$\beta =$ ______ $^\circ$.

Solution:
From

$$\frac{\pi/4}{2\pi} = \frac{\beta}{360^\circ}$$

we obtain

$$\beta = \frac{\pi/4}{2\pi} \cdot 360^\circ = \frac{1}{8} \cdot 360^\circ = 45^\circ \; .$$

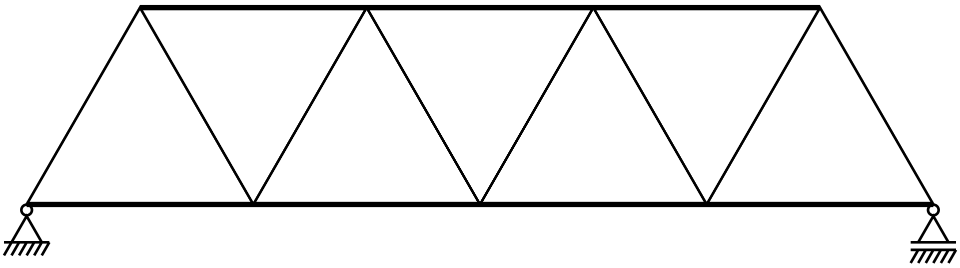
Exercise 5.2.5
The values of the six angles $\alpha_1, \ldots, \alpha_6$ are specified either in degree measure or in radian measure. Convert their values to the other measure.

	α_1	α_2	α_3	α_4	α_5	α
Radian measure	π		$\frac{2\pi}{3}$		$\frac{11\pi}{12}$	
Degree measure		324		270		3

5.3 All about Triangles

5.3.1 Introduction

Technical structures such as trusses and some bridges are built from triangles as design elements (see figure below).



Conversely, the question arises how an arbitrary surface can be subdivided into triangles. For many geometrical calculations this question is useful. Some examples are given in Section 5.4 auf Seite 177.

Furthermore, the problem of partitioning arbitrary surfaces into simple "basic elements" results in constructive answers in applications that are relevant far beyond simple geometric considerations. A first impression of such relevance gives us the integral calculus described in chapter 8 auf Seite 337 together with its application to the calculation of surface areas. There, the first approximation to the integral is a partition of the area into rectangles (each consisting of two triangles, to stay on topic). For the three-dimensional computer aided modelling of surfaces, for example in the manufacturing of car bodies, partitions into triangles (triangulations) are the basis of many calculations and deceptively realistic looking virtual animations.

5.3.2 Triangles

Many statements on geometric figures and solids arise from the properties of triangles. A triangle is the "simplest closed figure" which can be determined by three non-collinear points (i.e. the points do not lie on a single straight line).

First, we will present the important terms. Then we will determine under which conditions a triangle is uniquely defined and how individual angles and sides can be calculated. Here the intercept theorems are an important tool, since they can also be considered as statements on relations between different triangles.

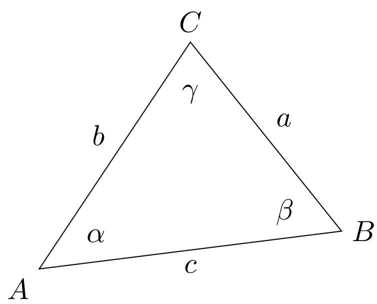
In Section 5.6 auf Seite 201 we will then investigate functional relations between side lengths and angles enabling us to answer advanced questions relevant to applications.

Triangle5.3.1

A **triangle** is constructed by joining three non-collinear points A, B, and C. The triangle is denoted by ABC.

- The three points are called the **vertices** of the triangle, and the three lines are called the **sides** of the triangle.
- Any two sides of the triangle form two angles. The smaller angle is called the **interior angle** (or simply angle for short) and the greater angle is called the **exterior angle**.
- The sum of the three interior angles is always 180° or π.

The vertices and sides of a triangle are often denoted as follows: vertices are denoted by upper-case Latin letters in mathematical positive direction (counter-clockwise). The side opposite a vertex is denoted by its lower-case Latin letter, and the interior angle of the vertex is denoted by the corresponding lower-case Greek letter.
Since exterior angles are far less important than interior angles, the **interior angles** are simply called the **angles** of the triangle.



The sum of all (interior) angles is always 180° or π. Hence, at most one angle can be equal to or greater than 90° or $\frac{\pi}{2}$. Consequently, triangles are classified according to their greatest interior angle into three types:

Names of Triangles5.3.2

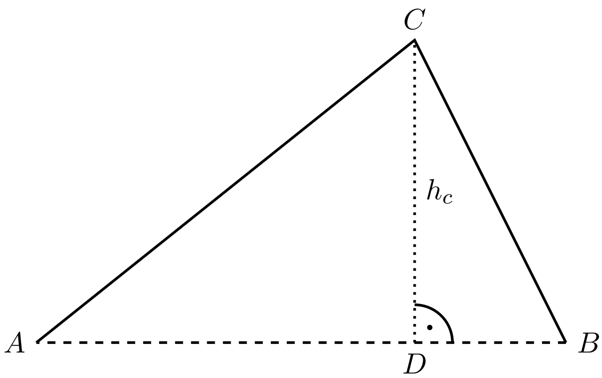
Triangles are named according to their angles as follows:

- A triangle that only has angles less than $\frac{\pi}{2}$ is called an **acute**.
- A triangle that has a right angle is called a **right-angled** triangle or simply a right triangle.

  In a right triangle the two sides enclosing the right angle are called the **catheti** or **legs**, and the side opposite to the right angle is called the **hypotenuse**.

- A triangle that has an angle greater than $\frac{\pi}{2}$ is called an **obtuse**.

As an example, let us consider the simple structure of a car jack with the shape of a triangle (see figure below): It consists of two rods connected by a joint. The two other endpoints of the rods can be pulled together. The greater the angle of a rod with respect to the street is, the higher the joint is above the ground.



Thus, in a triangle ABC the shortest line segment between vertex C and the line defined by the side c opposite to C is called the **altitude (or height) of the triangle** h_c on the (base) side c. The second endpoint D of the line segment h_c is called the **perpendicular foot**. The altitudes h_a and h_b are defined accordingly.

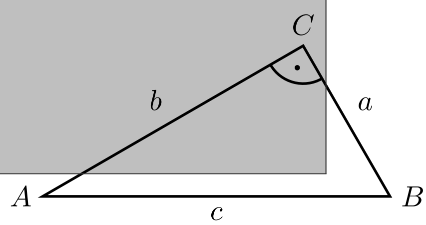
One can also say that altitudes are those line segments that are perpendicular to the line of a side and have the vertex opposite to the relevant side as an endpoint.

5.3.3 Pythagoras' Theorem

One statement relating the lengths of the sides in a right triangle is provided by **Pythagoras' theorem**. A commonly-used formulation of the theorem is given here.

Pythagoras' Theorem5.3.3

Consider a right triangle with the right angle at vertex C.



The sum of the areas of the squares on the legs a and b equals the area of the square on the hypotenuse c. This statement can be written as an equation (see also the triangle in the figure):

$$a^2 + b^2 = c^2 \ .$$

If the sides of the triangle are denoted in another way, the equation has to be adapted accordingly!

Example 5.3.4

Suppose we have a right triangle with legs (short sides) of length $a = 6$ and $b = 8$.

The length of the hypotenuse can be calculated by means of Pythagoras' theorem:

$$c = \sqrt{c^2} = \sqrt{a^2 + b^2} = \sqrt{36 + 64} = \sqrt{100} = 10 \ .$$

Exercise 5.3.1
Consider a right triangle ABC with the right angle at vertex C, hypotenuse $c = \frac{25}{3}$, and altitude (height) $h_c = 4$. The line segment $\overline{DB}$ has the length $q = [\overline{DB}] = 3$. Here, D is the perpendicular foot of the altitude h_c. Calculate the length of the two legs a and b.

Solution:
We apply Pythagoras' theorem to the triangle DBC that has a right angle at the vertex D. Then, we have

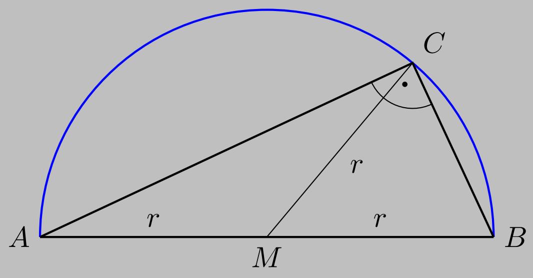
$$a = \sqrt{h_c^2 + q^2} = \sqrt{4^2 + 3^2} = \sqrt{25} = 5 \ .$$

Now, we apply Pythagoras' theorem to the given right triangle ABC:

$$b = \sqrt{c^2 - a^2} = \sqrt{\left(\frac{25}{3}\right)^2 - 5^2} = \sqrt{\frac{400}{9}} = \frac{20}{3}.$$

Thales' theorem is another important theorem that makes a statement on right triangles.

Thales' Theorem5.3.5



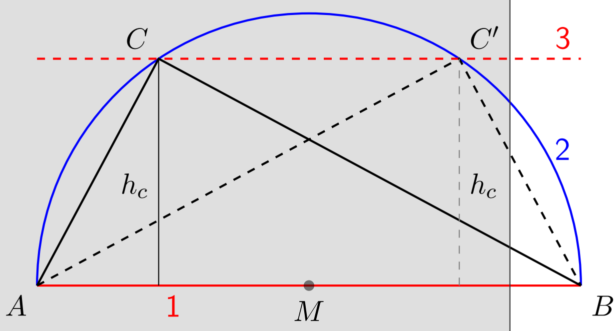
If the triangle ABC has a right angle at the vertex C, then vertex C lies on a circle with radius r whose diameter $2r$ is the hypotenuse $\overline{AB}$.

The converse statement is also true. Construct a half-circle above a line segment $\overline{AB}$. If the points A and B are joint to an arbitrary point C on the half-circle, then the resulting triangle ABC is always right-angled.

Example 5.3.6

Construct a right triangle with a given hypotenuse $c = 6\,\text{cm}$ and altitude $h_c = 2.5\,\text{cm}$.

1. First, draw the hypotenuse
$$c = \overline{AB}\ .$$
2. Let the middle of the hypotenuse be the centre of a circle with radius $r = c/2$.
3. Then draw a parallel to the hypotenuse at distance h_c. This parallel intersects Thales' circle in two points C and C'.



Together with the points A and B, each of these intersections points forms a triangle possessing the required properties, i.e. two solutions exist. Two further solutions are obtained if the construction is repeated drawing a second parallel below the hypotenuse. The constructed triangles are different in position but concerning shape and size these triangles are "congruent" (see also Section 5.3.10 auf Seite 171).

Exercise 5.3.2
Find the maximum altitude (height) h_c of a right triangle with hypotenuse c.

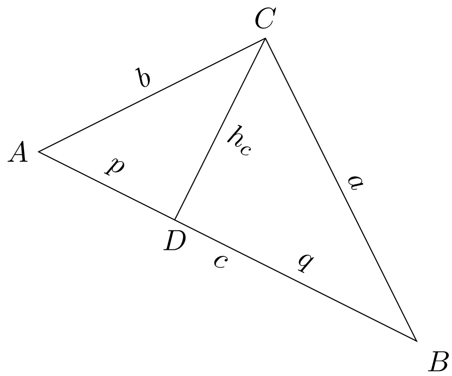
Solution:
The maximum altitude h_c is the radius of the Thales circle on the hypotenuse. Hence, $h_c \leq \frac{c}{2}$.

Further material.:

In a right triangle, some statements beyond Pythagoras' theorem hold.

To study them, we will use the notation illustrated below:

Consider a right triangle with the right angle at the vertex C. The altitude h_c intersects the hypotenuse of the triangle ABC in the point D, called the perpendicular foot. Furthermore, let $p = [\overline{AD}]$ and $q = [\overline{BD}]$.



Right Triangle Altitude Theorem5.3.7

The area of the square on the altitude equals the area of the rectangle created by the two hypotenuse segments:
$$h^2 = p \cdot q \ .$$

Cathetus Theorem5.3.8

The area of the square on a leg (cathetus) equals the area of the rectangle created by the hypotenuse and the hypotenuse segment adjacent to the leg:
$$a^2 = c \cdot q \ , \ \ b^2 = c \cdot p \ .$$

Example 5.3.9

Let a right triangle with the legs $a = 3$ and $b = 4$ be given.

The length of the hypotenuse can be calculated by means of Pythagoras' theorem:

$$c = \sqrt{a^2 + b^2} = \sqrt{9 + 16} = \sqrt{25} = 5 \ .$$

According to the cathetus theorem the hypotenuse segments p and q are:

$$q = \frac{a^2}{c} = \frac{9}{5} = 1.8 \quad \text{and} \quad p = \frac{b^2}{c} = \frac{16}{5} = 3.2 \ .$$

According to the altitude theorem the altitude h_c is:

$$h_c = \sqrt{p \cdot q} = \sqrt{\frac{9}{5} \cdot \frac{16}{5}} = \sqrt{\frac{144}{25}} = \frac{12}{5} = 2.4 \ .$$

Exercise 5.3.3
Find the length of the two legs of a given right triangle with hypotenuse $c = 10.5$, altitude $h_c = 5.04$, and hypotenuse segment $q = 3.78$.

Solution:

$$\text{Cathetus theorem:} \quad a = \sqrt{c \cdot q} = \sqrt{10.5 \cdot 3.78} = 6.3 \ ;$$

$$\text{Pythagoras' theorem:} \quad b = \sqrt{c^2 - a^2} = \sqrt{10.5^2 - 6.3^2} = 8.4 \ .$$

5.3.4 Congruence and Similar Triangles

Each triangle includes three sides and three angles. The exterior angles are already defined by the interior angles such that the "shape" of a triangle is determined by six characteristics. If two triangles coincide in all these characteristics, they are said to be **congruent**. For that, the position of the triangles is not relevant, i.e. congruent triangles can be transformed into each other by rotation, reflection, and translation.

If four of the six characteristics are known, the triangle is uniquely determined up to rotation or reflection, i.e. its position in the plane. Then, all triangles with these characteristics are congruent. In some cases, only three characteristics are sufficient to determine the triangle uniquely. These cases are described by the following **theorems for con-**

gruent triangles.

Theorems for Congruent Triangles5.3.10

Up to its position in the plane, a triangle is uniquely defined if one of the following situations is at hand:

- At least four of the six characteristics (three angles and three sides) are known.
- The lengths of all three sides are known.
(This theorem is usually called "sss" for "side, side, side".)
- Two angles and the length of the included side are known.
(This theorem is usually called "asa" for "angle, side, angle".)
- The lengths of two sides and the included angle are known.
(This theorem is usually called "sas" for "side, angle, side".)
- The lengths of two sides and a non-included angle are known such that only one side is a leg of the given angle and the second side is greater than the given leg.

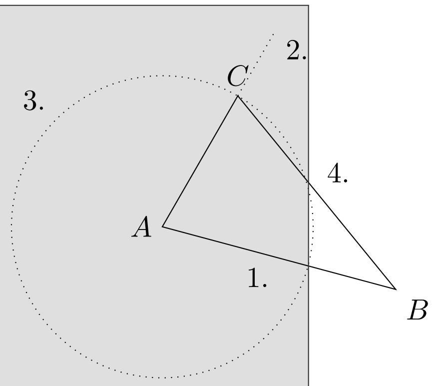
  (This theorem is called "Ssa", where the upper-case "S" indicates that the side opposite to the given angle is the greater one.)

If only two characteristics of a triangle are known, or three characteristics are known that do not correspond to one of the cases described above, than a number of different triangles with these characteristics exist which are not congruent.

The next example will illustrate how a triangle can be constructed applying the theorems for congruent triangles. Then another example will be considered, where only three angles are known and hence none of the theorems described above apply.

Example 5.3.11

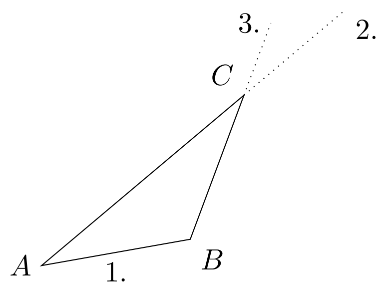
Let the sides b, c, and the angle α be given. According to the "sas" theorem the triangle is constructed as follows: 1. Draw a line, in this example side c. 2. Attach the angle α to the corresponding vertex (A). 3. Draw a circle around vertex A with a radius corresponding to the length of the second side (in this case, side b). 4. The intersection point of this circle with the second leg of the angle α is the third vertex (C) of the triangle. (The first leg of α is the side c.)



Exercise 5.3.4
Construct a triangle with side $c = 5$ and the two angles $\alpha = 30°$ and $\beta = 120°$ using the notation introduced above.

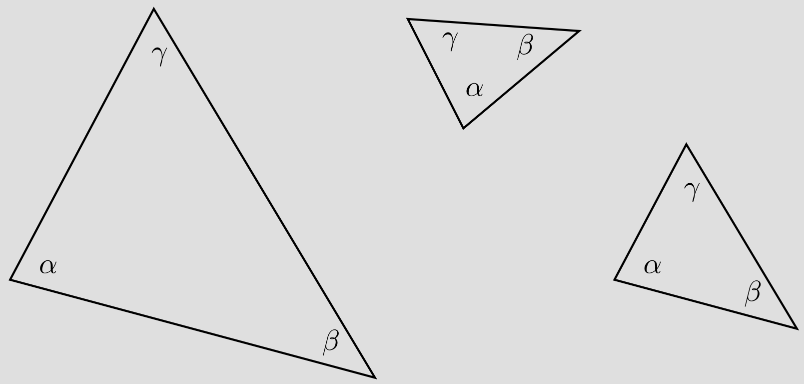
Solution:

1. Draw the given line segment c. 2. Attach the corresponding angles α and β to either side of the segment . 3. The intersection point of the two new legs is the third vertex C of the triangle.



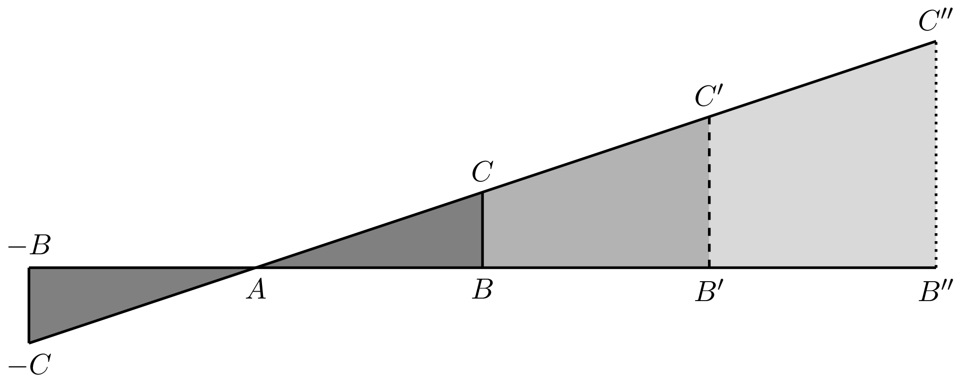
Example 5.3.12

Let three angles $\alpha = 77°$, $\beta = 44°$, and $\gamma = 59°$ be given summing up to 180°. This case does not correspond to one of the cases in the theorems for congruent triangles 5.3.10 auf der vorherigen Seite. A few examples for triangles with the given angles are shown below.



Actually, an infinite number of triangles with the given angles do exist. They are not congruent to each other, i.e. they cannot be transformed into each other by rotation or reflection.

However, the triangles look similar in a way. Such **similar** triangles are also obtained if, for example, all the side ratios are known. This fact results from the intercept theorems as illustrated by the figure below.



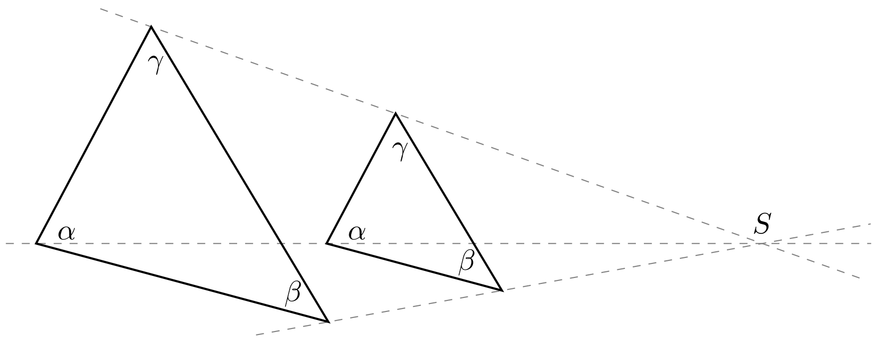
Similarity Theorems for Triangles5.3.13

Two triangles are called **similar** to each other if they

- have two (and because of the triangle postulate also three) congruent angles, or

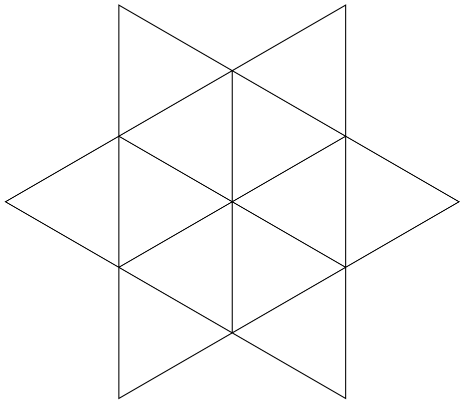
- have three sides whose lengths have the same **ratio**, or
- have one congruent angle and two adjacent sides whose lengths have the same **ratio**, or
- have two sides whose lengths have the same **ratio** and the angles opposite to the greater side are congruent.

The right and the left triangle in Example 5.3.12 auf Seite 172 have a special relationship. The left triangle is transformed into the other by uniform scaling with the centre of enlargement S and the scaling factor k.

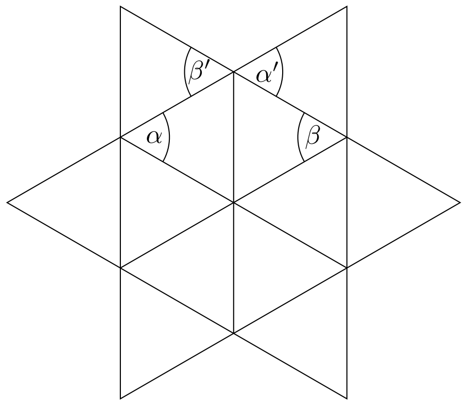


5.3.5 Exercises

Exercise 5.3.5
Find corresponding angles and alternate angles in the figure below.



Solution:

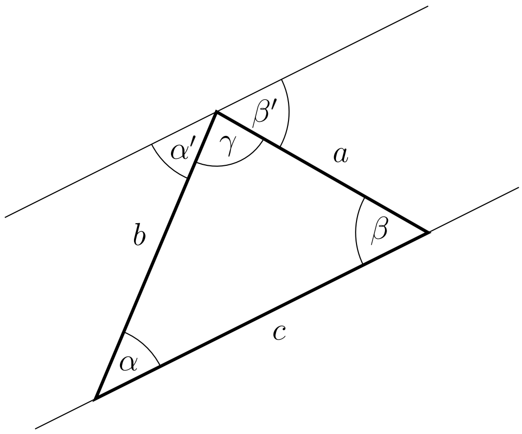


For example, the angles α and α' are corresponding angles. Likewise, angles β and β'.
For example, the angles α' and β are alternate angles. Likewise, angles α and β'.

Exercise 5.3.6
Prove that the sum of interior angles in a triangle is always π or 180°. Use the concept of alternate angles.

Hint:
Draw a parallel to one of the sides of the triangle passing through the third vertex and consider the angles at this vertex.

Solution:



Drawing a parallel to the side c passing through the vertex C one obtains an alternate angle α' to α and an alternate angle β' to β. The angles α', γ, and β' form a straight angle. Therefore,

$$\alpha' + \gamma + \beta' = \pi \ .$$

Furthermore, it is known that $\alpha' = \alpha$ and $\beta' = \beta$. Hence, $\alpha + \gamma + \beta = \pi$.

5.4 Polygons, Area and Circumference

5.4.1 Introduction

In nature, various figures in different shapes can be found. There, rounded shapes are particularly evident. When it comes to partitioning a surface completely, some boundaries can be found that can be approximated as line segments. The honeycomb structures created by insects are a famous example. Technical applications are often based on figures bounded by straight line segments.

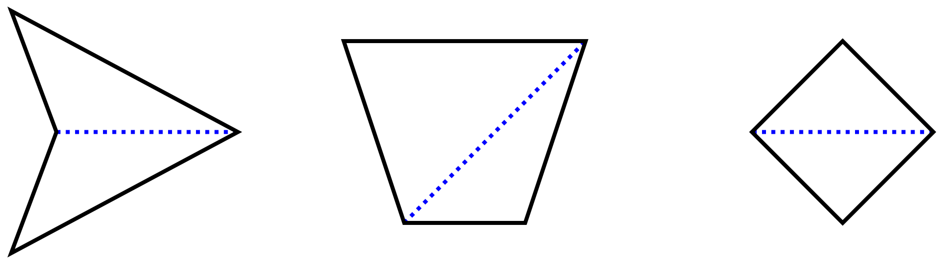
In this section we consider some special cases of polygons which can be used to describe surfaces bounded by straight line segments. To do this, we will first specify some characteristic features. Then, we will address the question of calculating the area of a polygon.

5.4.2 Quadrilaterals

In the previous section 5.3 auf Seite 164, triangles were considered. They were defined by three non-collinear points. Connecting any two of the three points by a line segment always results in a single closed path in which every point connects exactly two line segments. The line segments in the path have only their endpoints in common. Furthermore, the line segments do not intersect.

For more than three points this is not always true. Even only four points can be connected in such a way that line segments intersect or more than one closed path exists.

In the figure below all given points are to be connected by a single closed path without any intersections.



Obviously, a quadrilateral can be divided into two triangles. Generally, one obtains two triangles if the vertex with the greatest angle is connected to the opposite vertex by a line segment. Such an additional line segment between two vertices of the quadrilateral which are not connected to each other is called a **diagonal** of the quadrilateral. From the fact that the sum of (interior) angles in a triangle equals π or $180°$ then results that the sum of (interior) angles in a quadrilateral is twice this sum, i.e. 2π or $360°$.

Quadrilaterals5.4.1

Consider **quadrilaterals** constructed by connecting the four given points by line segments forming a single, closed and non-intersecting path trough these four points. There, any three of the four points connected by two line segments must be non-collinear.

As for triangles, the interior angles of quadrilaterals are simply denoted as angles if not otherwise specified in context.

Like triangles, quadrilaterals are used in technical structures in many ways. Therefore, additional terms are commonly used to specify different types of quadrilaterals.

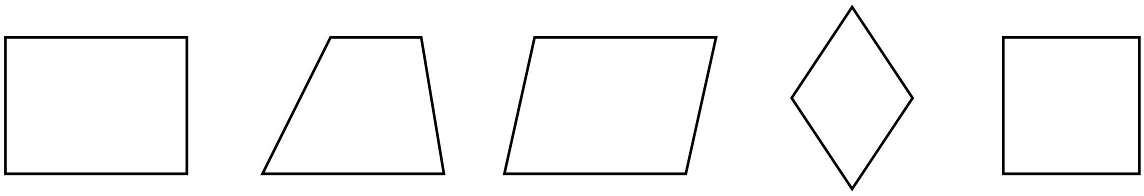
Also as for triangles, quadrilaterals are classified by the lengths of their sides or by the magnitudes of angles. There are differences between the classifications of triangles and quadrilaterals. For example, quadrilaterals can have parallel sides, or have more than one vertex with a right angle.

Special Types of Quadrilaterals5.4.2

Quadrilaterals with the following properties have their own terms: A quadrilateral is called

- a **trapezoid** if at least one pair of opposite sides is parallel;
- a **parallelogram** if two pairs of opposite sides are parallel;
- a **rhombus** or an **equilateral quadrilateral** or a **diamond** if all four sides are of equal length;
- a **rectangle** if all four (interior) angles are right angles;
- a **square** if it is a rectangle with four sides of equal length;
- a **unit square** if it is a square with sides of length 1.

Thus, for the unit square also a measure has to be fixed.



rectangle trapezoid parallelogram rhombus square

There are several relations between the quadrilaterals introduced above:

Relations between rectangles5.4.3

Between different quadrilaterals the following relations exist:

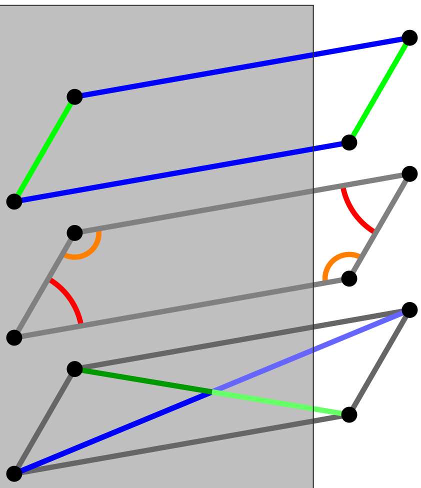
- Every square is a rectangle.
- Every square is a rhombus.
- Every rhombus is a parallelogram.
- Every rectangle is a parallelogram.
- Every parallelogram is a trapezoid.

These quadrilaterals can be characterised by means of the properties of their sides, angles, or diagonals in many ways.

Parallelogram5.4.4

A quadrilateral is a parallelogram if and only if

- opposite sides are parallel;
- opposite sides are of equal length;
- opposite (interior) angles are equal;
- two adjacent (interior) angles sum up to π or $180°$, respectively;
- diagonals bisect each other.

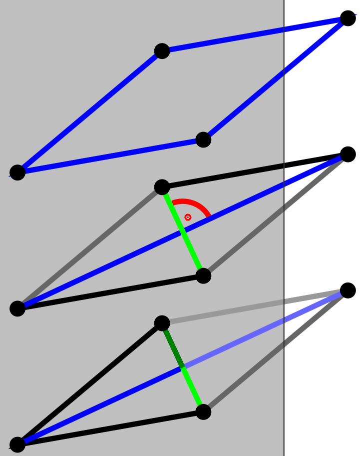


Rhombuses can be described as a special type of parallelograms.

Rhombus5.4.5

A quadrilateral is a rhombus if and only if

- all sides are of equal length;
- it is a parallelogram in which the diagonals are perpendicular;
- at least two adjacent sides are of equal length and the diagonals bisect each other.

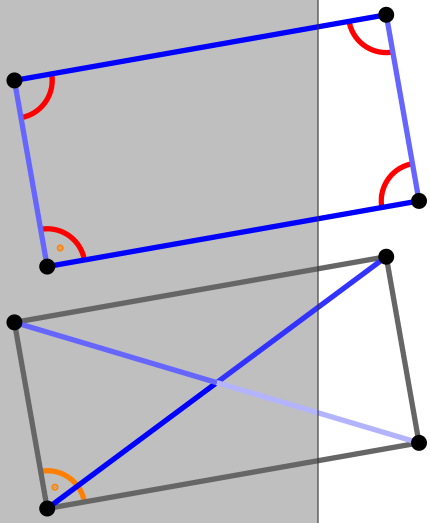


In the case of rectangles one often thinks of right angles since the term rectangle comes from the Latin word rectangulus, which is a combination of rectus (right) and angulus (angle). Apart from that, rectangles can simply be described by means of the properties of their diagonals.

Rectangle5.4.6

A quadrilateral is a rectangle if and only if

- all (interior) angles are equal;
- it is a parallelogram containing at least one right angle;
- it is a parallelogram whose diagonals are of equal length;
- the diagonals are of equal length and bisect each other;
- the diagonals bisect each other and at least one (interior) angle is a right angle.

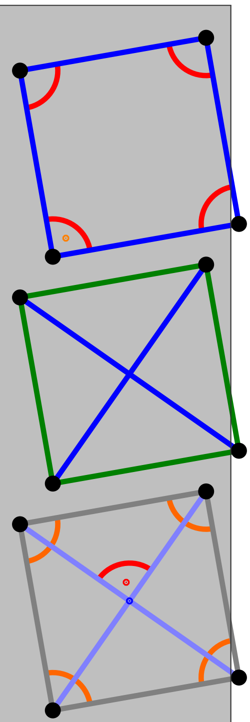


Squares are both special types of rectangles and special types of rhombuses.

Square5.4.7

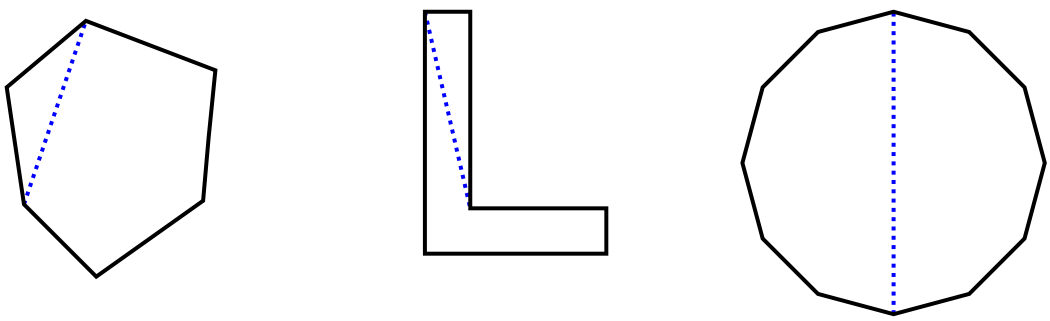
A quadrilateral is a square if and only if

- all sides are of equal length and
  - all (interior) angles are equal or
  - at least one (interior) angle is a right angle;
- the diagonals are of equal length and in addition all sides are of equal length;
- the diagonals are perpendicular and
  - bisect each other and are of equal length or
  - all (interior) angles are equal;
- it is a rhombus whose diagonals are of equal length;
- it is both a rhombus and a rectangle.



5.4.3 Polygons

For triangles, even one vertex or side contributes to the essential properties of the whole triangle, for example, one vertex with a right angle. For quadrilaterals, a single vertex no longer has such strong specifying properties. Instead, there is a greater variety of shapes. If "many" points are connected to a closed figure by line segments, there are many possibilities to create various figures and even to approximate round shapes.



In this general case, a detailed classification as for triangles or quadrilaterals is barely possible. The new possibilities such as the approximation of round shapes do also lead to new interesting questions. One does not consider single polygons but construction principles for a series of many polygons. On the other hand, every polygon can be divided into triangles if required, as we have seen already for quadrilaterals. Thus, a property of a single vertex is often considered in terms of what this means for the polygon in the whole.

For classification, the question is convenient whether a certain condition is satisfied by **all** vertices or not, and what this means for the polygons. For example, polygons are classified according to the magnitudes of their angles, e.g. whether all angles of the vertices are less than π or 180°. If so, all diagonals pass through the inside of the polygon. Otherwise, at least one diagonal exists in the outside.

The figure above shows examples of polygons exhibiting different properties. In the polygon to the left all (interior) angles are less than π or 180°. In this case the polygon is said to be convex. In contrast, the polygon in the middle contains a vertex with an angle greater than π or 180°. In the polygon to the right all angles are equal, leading to a very evenly shaped polygon.

Polygons5.4.8

Let n points in the plane be given, where n is a natural number with $n \geq 3$. Here, we consider **polygons** constructed by connecting points by line segments such that a closed, non-self-intersecting (simple) path is formed, and every point is adjacent to exactly two segments, where every three points connected by successive segments are to be non-collinear.

A polygon is also called an n-**gon**.

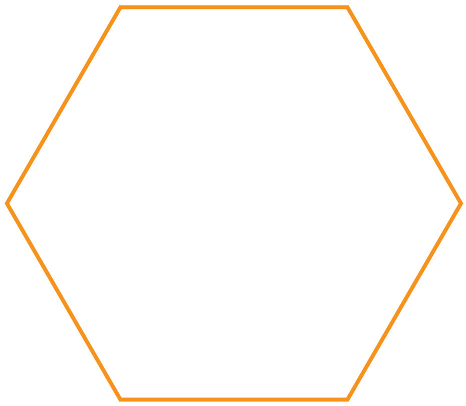
- The n points that are connected are called the **vertices** of the polygon, and the n connecting line segments are called the **sides** of the polygon.
- Every polygon can be divided into $(n-2)$ non-overlapping triangles. Hence, the sum of the interior angles of a polygon is $(n-2)\cdot\pi$ or $(n-2)\cdot 180°$.
- Line segments connecting two vertices not adjacent to the same side of the polygon are called the **diagonals** of the polygon.

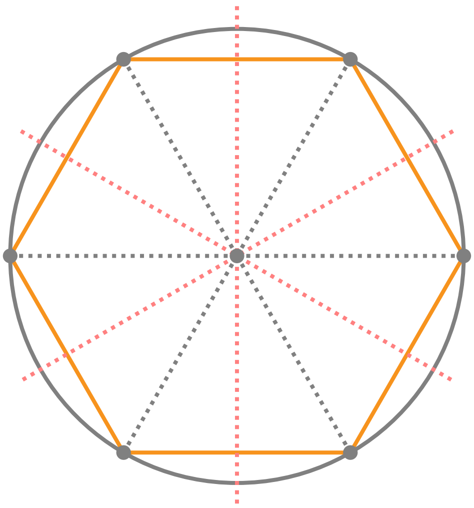
Further statements hold for polygons with sides of equal length and equal interior angles. For $n = 3$, these are equilateral triangles, end for quadrilaterals these are squares.

Regular Polygons5.4.9

A polygon that is equilateral (all sides have the same length) and equiangular (all angles are equal in measure) is called a **regular polygon** or a **regular n-gon**.

Honeycombs are – when seen from above – approximately regular hexagons.





Regular polygons have various symmetry properties. All lines perpendicular to the sides, passing through the midpoint of the respective side, intersect in a point M. Reflecting a polygon across such a line maps it onto itself.

Furthermore, regular polygons have rotational symmetry, i.e. a n-gon maps onto itself if it is rotated around M by an angle of $\frac{2\pi}{n}$.

The vertices of a regular polygon have all the same distance from M and thus lie all on a circle around M.

5.4.4 Circumference

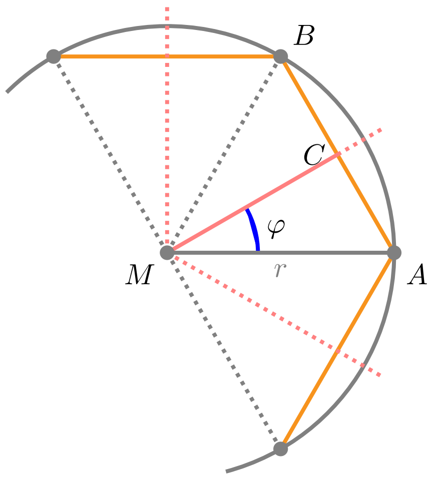
The circumference of a polygon is the sum of the lengths of all its line segments. If a polygon has further properties concerning the side lengths, more statements concerning the circumference can hold.

First, quadrilaterals are considered. If a and b are adjacent sides of a parallelogram, then its circumference is $U = a + b + a + b = 2 \cdot a \cdot b$.

For a rhombus (and therefore for a square), all four sides have the same length a such that its circumference is $U = 4 \cdot a$.

Likewise, for every regular polygon, all sides have the same length. If n is the number of vertices and a is the length of a side, then the circumference U_n can simply be calculated by $U_n = n \cdot a$.

As an outlook to trigonometric functions described in Section 5.6 auf Seite 201 the circumference of a regular polygon shall now be calculated in another way.



The vertices of a regular polygon all lie on a common circle with radius r. The angle φ between the line segments connecting the centre of the circle to the vertices A and B of a side is the n-th part of the complete angle: $\varphi = \frac{2\pi}{n}$. The centre of the circle and the midpoint C of the line segment $\overline{AB}$ form a right triangle MAC with the angle $\measuredangle(AMC) = \frac{1}{2} \cdot \varphi = \frac{\pi}{n}$. If the value of a is calculated by

$$\sin(\measuredangle(AMC)) = \frac{\frac{1}{2}a}{r}$$

and is inserted in $U = n \cdot a$, then we obtain the formula

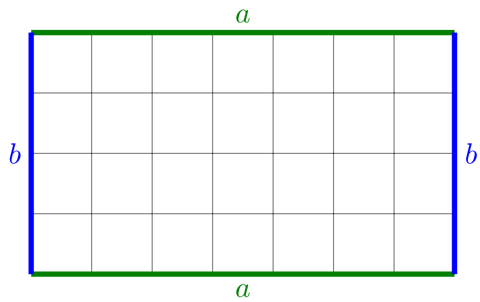
$$U_n = n \cdot a = 2 \cdot r \cdot n \cdot \sin\left(\frac{\pi}{n}\right)$$

for the circumference of a regular polygon. For example, $U_6 = 2 \cdot r \cdot 6 \cdot \frac{1}{2} = 6 \cdot r$. The larger n is, the closer the circumference is to the value $2 \cdot r \cdot \pi \approx 6.283 \cdot r$ describing the circumference of a circle with radius r. This can be shown by means of more advanced methods of calculus, the basic ideas of which are introduced in Chapter 7 auf Seite 295. The approach described here is based on the following idea: it is difficult to calculate the value of the circumference of a circle. Therefore, one looks for similar objects, in this case the regular polygons, with two properties: their circumference can be calculated easily, and if the number of vertices is sufficiently large, then the circumference of the polygon differs from the circumference of a circle less than any given positive number (here, one thinks of “small” numbers). This approach can also be used to calculate the area of surfaces that are not bounded by line segments (see Chapter 8 auf Seite 337). For this purpose, it will be illustrated here how to calculate the area of polygons, which is in this respect relatively easy. Further, this can be used as the starting point of an approximation, as the figure above showing a circle inscribed into a hexagon suggests.

5.4.5 Area

The area of a surface equals the number of unit squares required to cover this surface completely.

Let us first consider rectangles. If the sides of the rectangle are of lengths a and b, then the rectangle contains b rows with a unit squares, i.e. $b \cdot a$ unit squares.

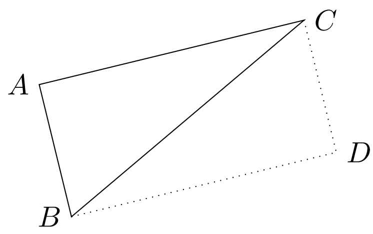


Area of a Rectangle5.4.10

The area F of a rectangle with sides of lengths a and b is

$$F = b \cdot a = a \cdot b\,.$$

With this, the area of a right triangle can be calculated easily. Let ABC be a right triangle rotated by an angle of 180°. If the original and the rotated triangle are merged along the hypotenuse, one obtains a rectangle.

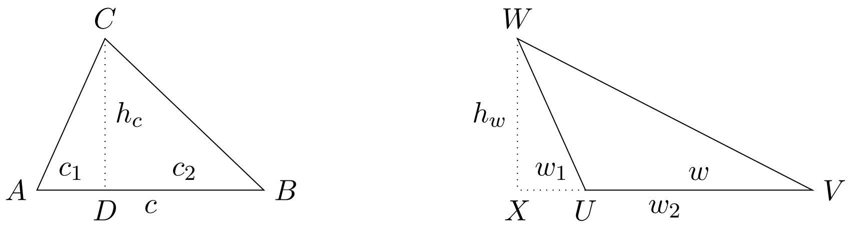


The area of the right triangle is then half the area of the rectangle, i.e. $F = \frac{1}{2} \cdot a \cdot b$.

And how is the area calculated if the triangle is not right-angled?

Every triangle can be divided into two right triangles by drawing a line from one vertex to the opposite side such that this line is perpendicular to the side. This line is called the **altitude** h_i of a triangle on a specific side i, where i is the index of the side a, b, or c.

Depending on whether the new line is interior or exterior to the triangle, the area of the triangle equals the sum or the difference of the areas of the two resulting right triangles:



Thus, on the left, we have (if F_Δ is the area of the triangle Δ)

$$F_{ABC} = F_{DBC} + F_{ADC} = \tfrac{1}{2} \cdot h_c \cdot c_2 + \tfrac{1}{2} \cdot h_c \cdot c_1 = \tfrac{1}{2} \cdot h_c \cdot (c_2 + c_1) = \tfrac{1}{2} \cdot h_c \cdot c \,.$$

On the right, we have

$$F_{UVW} = F_{XVW} - F_{XUW} = \tfrac{1}{2} \cdot h_w \cdot w_2 - \tfrac{1}{2} \cdot h_w \cdot w_1 = \tfrac{1}{2} \cdot h_w \cdot (w_2 - w_1) = \tfrac{1}{2} \cdot h_w \cdot w \,.$$

Thus the area can always be calculated from the length of one side and the length of the altitude perpendicular to the corresponding side.

Area of a Triangle5.4.11

The area F_{ABC} of a triangle equals half the product of the length of a side and the length of the corresponding altitude of the triangle:

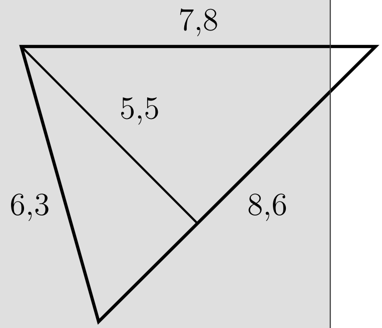
$$F_{ABC} = \frac{1}{2} \cdot a \cdot h_a = \frac{1}{2} \cdot b \cdot h_b = \frac{1}{2} \cdot c \cdot h_c \,.$$

Here, the **altitude of a triangle on a side** denotes the line segment from the vertex opposite the side to the line containing the side itself, perpendicular to this side.

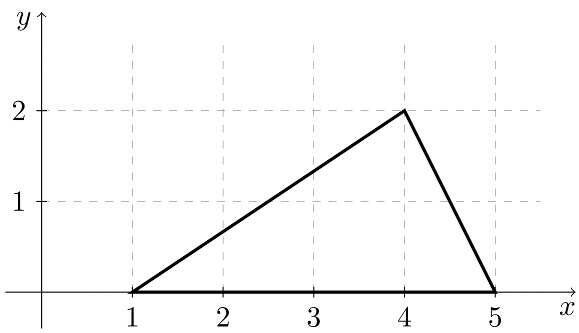
Example 5.4.12

For the triangle to the right, the altitude corresponding to the side of length 8.6 is given. The given values are rounded numerical values. Hence, the area F of the triangle is approximately

$$F = \frac{8.6 \cdot 5.5}{2} = 23.65 \ .$$



Exercise 5.4.1
Calculate the area of the triangle below.

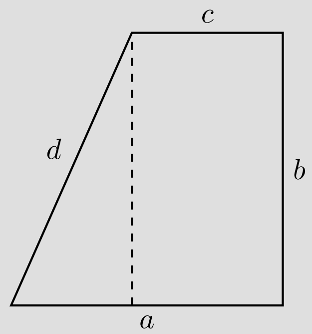


Solution:
For this triangle, one altitude can be read off easily, namely the altitude perpendicular the side on the x-axes. The length h of this altitude is $h = 2$, and the length of the corresponding side is $c = 5 - 1 = 4$. Hence, the area F of the triangle is $F = \frac{1}{2} \cdot c \cdot h = \frac{1}{2} \cdot 4 \cdot 2 = 4$.

Using the formula for the area of triangles, areas of polygons can also be calculated. This is due to the fact that every polygon can be divided into triangles by adding diagonals to the polygon until all subareas are triangles. However, the considerations will remain restricted here to a few simple shapes. In the following example, the polygon can be divided into a triangle and a rectangle. As a result, the calculation will be particularly easy.

Example 5.4.13

Consider the polygon to the right, namely a trapezoid. In this example, the polygon can be divided into a right triangle with the legs $(a - c)$ and b and the hypotenuse d as well as a rectangle with sides of length b and c.

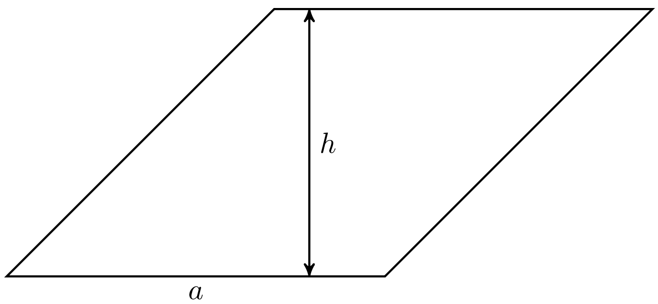


Then, the area of the polygon is:

$$F = F_{\text{triangle}} + F_{\text{rectangle}} = \frac{1}{2}(a-c)\cdot b + b\cdot c = \frac{1}{2}ab - \frac{1}{2}bc + bc = \frac{1}{2}(a+c)\cdot b\ .$$

Exercise 5.4.2

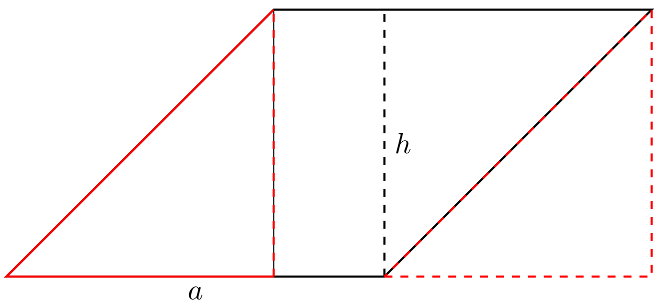
Calculate the area of the **parallelogram** to the right for $a = 4$ and $h = 5$.



Hint:
Divide the parallelogram appropriately and look at the resulting triangles carefully!

Solution:

The parallelogram can be divided into the left red rectangle, a rectangle, and the right triangle. Shifting the left red triangle to the right one obtains a rectangle with sides of lengths a and h. Then, the area of the parallelogram is



$$F = a \cdot h = 4 \cdot 5 = 20\ .$$

Finally, we will calculate the area of a circle. Info Box 5.2.4 auf Seite 159 introduced the number π describing the ratio of the circumference of a circle to its radius. The formula for the area of the circle also involves π.

Area of a Circle5.4.14

The area of a circle with radius r is

$$F = \pi \cdot r^2 .$$

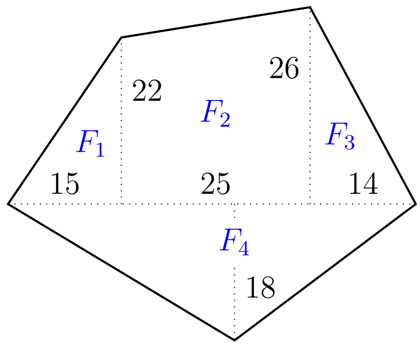
Example 5.4.15

Let the area of a circle with radius $r = 2$ be 12.566. This fact can be used to calculate an approximate value of the number π: We have $F = \pi \cdot r^2$, hence $\pi = \frac{F}{r^2}$. Inserting the given values results in the approximate value

$$\pi = \frac{F}{r^2} \approx \frac{12.566}{4} = 3.1415 .$$

5.4.6 Exercises

Exercise 5.4.3
Calculate the area of the polygon to the right.



Solution:
The values of the indicated subareas are calculated separately.

- F_1 is a triangle: $F_1 = \frac{15 \cdot 22}{2} = 165$.
- F_2 is a trapezoid that can be divided into two triangles with the altitude 25: $F_2 = \frac{22 \cdot 25}{2} + \frac{26 \cdot 25}{2} = 275 + 325 = 600$.
- F_3 is a triangle: $F_3 = \frac{14 \cdot 26}{2} = 182$.
- The surface F_4 is also a triangle: $F_4 = \frac{(15+25+14) \cdot 18}{2} = 486$.

Finally, we obtain the area of the entire polygon by summing up all these subareas: $F_1 + F_2 + F_3 + F_4 = 165 + 600 + 182 + 486 = 1433$.

5.5 Simple Geometric Solids

5.5.1 Introduction

The shapes of common objects such as a notepad or a mobile phone as well as of technical structures such as tunnels can be described by simple basic solids, apart form the "rounded vertices". Why is that?

If a broom is moved straight across a plane floor covered with dust, a rectangular section of the clean floor becomes visible. Geometrically idealised, a rectangle is formed by shifting a line segment (the broom). If the broom is rotated, a circle can be created. In this way, from simple objects more complex objects are constructed that can nevertheless be described easily.

5.5.2 Simple Geometric Solids

Points are the simplest basic geometric objects. Translations of points result in line segments, and transformations such as translation or rotation of line segments result in simple geometric figures. For example, polygons and circles are obtained in a way described above.

If figures are shifted or rotated out of their plane, then new objects are created that are denoted as solids. In this section some simple solids will be described, whose shapes can be identified easily in many everyday objects and technical constructions.

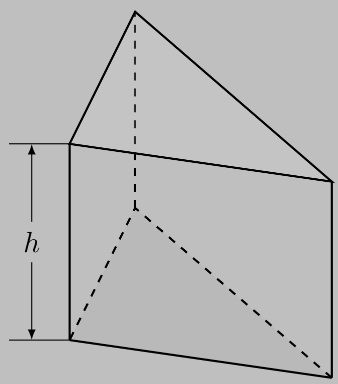
Example 5.5.1

Let us consider a rectangle and shift it perpendicular to the drawing plane. In this way, a rectangular cuboid (or informally a rectangular box) is constructed. Its surface consists of the given rectangle and a copy of that (two faces). Four further rectangles (faces) are formed by the four sides of the given rectangle.

Taking any polygon and shifting it perpendicular to the drawing plane, results in a solid that is called a prism. The term also denotes a transparent optical element of this shape used to refract light waves. Because the angle of refraction depends on the wavelength (i.e. the colour of the light), the different wave lengths of seemingly white light are refracted differently. In this way, the different colours of white light become visible.

Prism5.5.2

Let a polygon G be given. A **prism** is a solid resulting from a perpendicular translation of a polygon G by a line segment of length h. The two faces, i.e. the given polygon and the shifted copy of this polygon, are then called the base faces. They are parallel to each other. All other faces together form the lateral surface M.



The figure above shows a prism with a triangle as its base. The other faces adjacent to the base face are rectangles.

The volume V of the prism is the product of the area of the polygon G and the height h: We have $V = G \cdot h$.

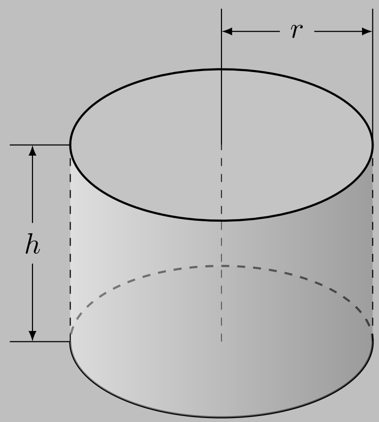
The area O of the surface is the sum of twice the area of the base face G and the area of the lateral surface M. If U is the circumference of the given polygon, we have $O = 2 \cdot G + M = 2 \cdot G + U \cdot h$.

In the introductory example a rectangular cuboid was described. Using the definition above, it can be considered as a special case of a prism, namely a prism with a rectangle as its base face. If all faces are squares, the prism is called a cube.

The construction principle can be varied in different ways. For example, the polygon can be replaced by a disk that is shifted. By a perpendicular translation of the disk a solid is created that is especially symmetric, namely a cylinder. A tunnel drilling machine creates – if considered in a simplified manner – a cylindrical tube.

Cylinder5.5.3

Let a disk G be given. A **cylinder** is a solid created by a perpendicular translation of a disk G by a line segment h. The two faces, i.e. the given disk and its copy, are then called the base faces of the cylinder. They are parallel to each other. The curved part of the surface between the two disks forms the lateral face M of the cylinder.



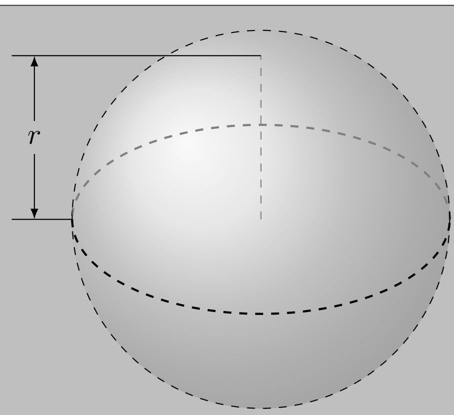
The volume V of the cylinder is the product of the area of the disk G with radius r and the height h of the cylinder: $V = G \cdot h = \pi \cdot r^2 \cdot h$.

The area O of the surface is the sum of twice the area of the disk G and the area of the lateral surface M. With the circumference $U = 2 \cdot \pi \cdot r$ of the disk we have $O = 2 \cdot G + M = 2 \cdot \pi \cdot r^2 + 2 \cdot \pi \cdot r \cdot h = 2 \cdot \pi \cdot r \cdot (r + h)$.

If the disk is not translated but rotated, where the axis of rotation passes trough the centre of the disk and one of its boundary points, then the resulting solid is a sphere.

Sphere5.5.4

Let a disk with centre M and radius r be given. If the disk M is rotated around an axis trough M and a boundary point of the disk, the resulting solid is a sphere with radius r.



The volume V of the sphere is $V = \frac{4}{3} \cdot \pi \cdot r^3$.

The area O of the surface is given by $O = 4 \cdot \pi \cdot r^2$.

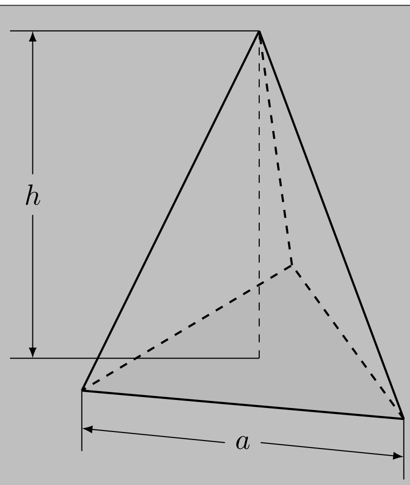
A sphere can also be described as a solid consisting of all points that have a distance less than or equal to r from M (see also Chapter 10 auf Seite 452).

In this approach, a prism is a solid consisting of all points that lie on a connecting line between the base face and its copy.

In the following, two variations of this approach will be considered. We start again with a polygon as base face. Moreover, instead of a copy of the base face, only a point is given.

Pyramid5.5.5

Let a polygon G and a point S with distance $h > 0$ from G be given. A pyramid with the base G and the apex S is a solid consisting of all points lying on a line segment between S and a point of the base face G.



The figure above shows a pyramid with a triangular base face.

The volume V of the pyramid is proportional to the area of the base face G and the height h: $V = \frac{1}{3} \cdot G \cdot h$.

The area O of the surface is the sum of the area of the base face G and the area of the lateral surface M, where the area of the lateral surfaces is the sum of the areas of its triangular faces D_k $(1 \leq k \leq n)$. Thus, we have $O = G + M = G + D_1 + \ldots + D_n$.

In special situations one obtains simple formulas that can be used to calculate the volume and the surface area of the solid. One example is the pyramid shown above. There, the base face is an equilateral triangle. The following exercise illustrates how to derive a formula for the surface area of a special case of a pyramid from the properties of equilateral triangles.

Exercise 5.5.1
Calculate the surface area O of a pyramid whose faces are all equilateral triangles with sides of length a.

Answer: O = [ ]

Solution:
A pyramid whose faces are all equilateral triangles has in total four faces: a triangular base face and three further adjacent faces. Since all faces are equal, the surface area of this pyramid is given by $O = 4 \cdot F$, where F is the area if a single equilateral triangle.

The height (altitude) ℓ of an equilateral triangle with sides of length a is, according to Pythagoras' theorem,

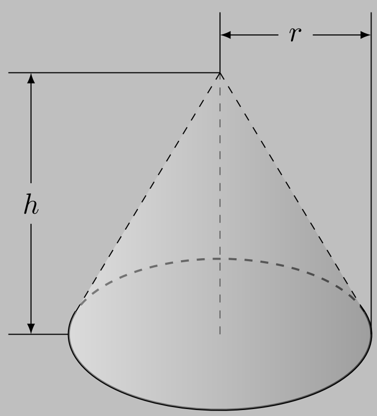
$$a^2 = \ell^2 + \left(\frac{a}{2}\right)^2 ,$$

equal to $\ell = \sqrt{a^2 - \frac{1}{4} \cdot a^2} = \frac{1}{2} \cdot a \cdot \sqrt{3}$. Hence, $F = \frac{1}{2} \cdot a \cdot \ell = \frac{1}{4} \cdot a^2 \cdot \sqrt{3}$, so $O = 4 \cdot F = a^2 \cdot \sqrt{3}$.

The considerations above, that a prism and a cylinder share the same constructing principle for different base faces, can be applied to the new situation of a pyramid as well. One obtains another solid if instead of a polygon (as for the case of a pyramid) a disk is now used as base face.

Cone5.5.6

Let a disk G with radius r and a point S with distance $h > 0$ from G be given. A cone with the base face G and the apex S is the solid consisting of all points lying on a line segment between S and a point of the base face G.



The volume V of the cone is proportional to the area of the disk G and its height h. We have $V = \frac{1}{3} \cdot G \cdot h = \frac{1}{3} \cdot \pi \cdot r^2 \cdot h$.

A cone whose apex is perpendicularly above the centre of the disk is called a **right circular cone**.

The area of the surface of a right circular cone is the sum of the area of the disk G and the area of the lateral surface M. If ℓ is the distance of the apex from the boundary of the disk, then with the circumference of a circle $U = 2\pi r$ we have $O = G + M = \pi \cdot r^2 + \pi \cdot r \cdot \ell = \pi \cdot r \cdot (r + \ell)$.

In the figure above the height h of the cone is given. However, the formulas given above contain the distance between the apex and the boundary of the disk. How these quantities are related?

Exercise 5.5.2
Let's imagine that the right circular cone is divided into two equal pieces by a plane trough the apex perpendicular to the base face. Then, one face of the two pieces is a triangle whose sides are defined by the height h, the distance ℓ, and the radius r or the diameter of the circle, respectively. Here, r and h are given.

a. Describe ℓ as a function of h and r:

ℓ = [ ]

Solution:
The length of ℓ is the hypotenuse of the triangle whose legs are the height h of the circular cone and the radius r of the disk. Form Pythagoras' theorem we have $\ell = \sqrt{r^2 + h^2}$.

b. Describe the surface area O of the cone as a function of h and r:

O = [ ]

Solution:
The result for $\ell = \sqrt{r^2 + h^2}$ obtained in the first part of the exercise is inserted in the formula for the surface area O given above. Thus, we have

$$O = \pi \cdot r \cdot (r + \ell) = \pi \cdot r \cdot \left(r + \sqrt{r^2 + h^2}\right) = \pi \cdot r^2 \cdot \left(1 + \sqrt{1 + \left(\frac{h}{r}\right)}\right).$$

The last formula shows how the relation between surface area of the cone and the area $G = \pi \cdot r^2$ of the disk depends on the ratio of the height h to the radius r.

5.5.3 Exercises

Exercise 5.5.3
Calculate the volume of a prism of height $h = 8\,\text{cm}$ with a triangle as its base. Two sides of this triangle are of length $5\,\text{cm}$, and one side is of length $6\,\text{cm}$.

Answer: [ ] cm^3

Exercise 5.5.4
The surface area of a cylinder of height $h = 6\,\text{cm}$ is to be covered with a coloured sheet. The surface area shall be $O = 200\,\text{cm}^2$. Calculate the diameter d of the disk and the volume of the cylinder. Use the approximate value 3.1415 for π and round off your result to the nearest millimetre.

Answers:

a. d = [ ] cm

b. V = [ ] cm^3

Exercise 5.5.5
Consider a piece of wood with the shape of a rectangular cuboid with the volume V. The height of the cuboid is $h = 120\,\text{cm}$, and the base face is a square with sides of length $s = 40\,\text{cm}$. From the piece of wood, a cylindrical hole of height g with a diameter $d = 20\,\text{cm}$ is drilled "centrically" (i.e. the intersection point of the diagonals of the quadratic base face is the centre of the base disk of the cylinder). Use the approximate value 3.1415 for π and round off your result to integers. Calculate

a. the volume V_Z of the drilled hole:

   V_Z = [ ] cm^3
   Solution:
   The volume of the cylinder is

$$V_Z = \pi \cdot \left(\frac{d}{2}\right)^2 \cdot h = 3.1415 \cdot \left(\frac{20\,\text{cm}}{2}\right)^2 \cdot 120\,\text{cm} = 3.1415 \cdot 12000\,\text{cm}^3 = 37698\,\text{cm}^3$$

b. the percentage of the volume V_1 of the new piece of wood remaining after drilling of the volume V_0:

   Answer: [ ] %
   Solution:
   The volume V of the piece of wood is

$$V = s^2 \cdot h = (40\,\text{cm})^2 \cdot 120\,\text{cm} = 1600 \cdot 120\,\text{cm}^3 = 16 \cdot 12000\,\text{cm}^3$$

The percentage p_Z of the drilled cylinder is

$$p_Z = \frac{V_Z}{V} = \frac{\pi \cdot 12000\,\mathrm{cm}^3}{16 \cdot 12000\,\mathrm{cm}^3} \approx \frac{3.1415}{16} \approx 19\%$$

and thus, $p = (100 - 19)\% = 81\%$ is the percentage original wooden rectangular cuboid which forms the new piece of wood.

5.6 Trigonometric Functions: Sine, et cetera

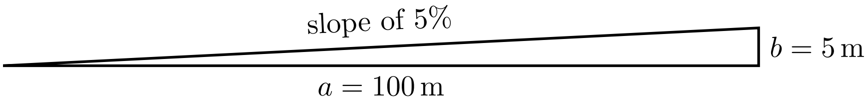
5.6.1 Introduction

On mountain roads, warning signs are put up if the road goes steeply downhill. The percentage describes how steep the terrain slopes compared to a horizontal movement. Questions for the conditions of movements on an inclined plane in physics have been investigated by Galileo Galilei. The results are also relevant for technical constructions.

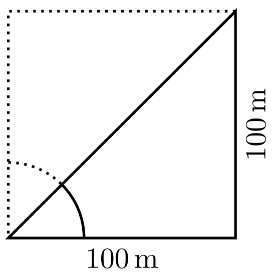
Trigonometric functions serve as a mathematical tool: they describe a geometric situation by means of a mathematical expression. This section describes how the relation between the percentage of the slope and the corresponding angle can be expressed. A first investigation of the properties of the trigonometric functions gives an idea of the various possible applications far beyond geometry, which will be revisited repeatedly in the later sections.

5.6.2 Trigonometry in Triangles

If one drives downhill on a road with a slope of five percent, then the height falls five metres for every 100 metres travelled horizontally. Here, the difference in height is considered in comparison to the horizontal line.



Accordingly, the slope is 100% if the difference in height between two positions with a horizontal distance of 100 m is 100 m. Geometrically, the connecting line segment between the two points is a diagonal of a square. Hence, the angle between the horizontal line and the diagonal, i.e. the road on which ones moves, has a degree measure of 45°.

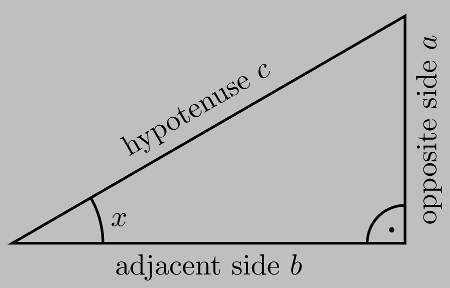


In other words: An angle of 45° corresponds to a slope of $\frac{100\,\mathrm{m}}{100\,\mathrm{m}} = 1$, i.e. the ratio of the horizontal line segment to the vertical line segment is 1. According to the intercept theorem, this ratio does not depend on the lengths of the individual segments. It only depends on the position of the two rays with respect to each other, i.e. the measure of the angle they enclose. If this assignment of a ratio of the line segments to an angle is also known for other angles, many constructive problems can be solved. For example, for a given angle the height can be determined.

Even the question for the ratio that corresponds to an angle of 30° shows, however, that in general it is not that simple to determine the assignment of a ratio of line segments to an angle. Therefore, the time-consumingly determined values that we considered initially were listed in mathematical tables such that they could be looked up later again easily. Now, these values are available practically everywhere, provided by calculators and computers. The most common assignments of an angle to a ratio of line segments are presented below. They are called circular functions or trigonometric functions. The branch of mathematics dealing with their properties is called **trigonometry**.

Trigonometric Functions in the Right Triangle5.6.1

Here, the most common **circular functions** are described as assignments of ratios of the sides in a right triangle to an angle. The circular functions are also called the **trigonometric functions**. Here, x denotes an angle in a right triangle that is not a right angle. The **opposite (side)** is the side opposite the angle x, and the other leg is called the **adjacent (side)**.



- The ratio of the opposite side a to the adjacent side b to an angle is called the tangent function:
$$\tan(x) := \frac{\text{opposite side}}{\text{adjacent side}} = \frac{a}{b}$$
- The ratio of the adjacent side b to the hypotenuse c to an angle is called the cosine function:
$$\cos(x) := \frac{\text{adjacent side}}{\text{hypotenuse}} = \frac{b}{c}$$

- The ratio of the opposite side a to the hypotenuse c to an angle is called the sine function:
$$\sin(x) := \frac{\text{opposite side}}{\text{hypotenuse}} = \frac{a}{c}$$

The tangent function describes the assignment of the ratio of height to width to the angle of inclination, i.e. the slope. In Chapter 8 auf Seite 337 this is also relevant in the context to the geometrical interpretation of the derivative.

According to the definition, the tangent function of the angle α is

$$\tan(\alpha) = \frac{a}{b} = \frac{a}{b} \cdot \frac{c}{c} = \frac{a}{c} \cdot \frac{c}{b} = \frac{\sin(\alpha)}{\cos(\alpha)} .$$

Thus, it suffices to know the values of sine and cosine to be able to calculate the tangent function.

Example 5.6.2

Let a triangle with a right angle $\gamma = \frac{\pi}{2} = 90°$ be given. The side c is of length 5 cm, and the side a is of length 2.5 cm. Calculate the sine, cosine and tangent function of the angle α.

The sine can be calculated immediately from the given values:

$$\sin(\alpha) = \frac{a}{c} = \frac{2.5\,\text{cm}}{5\,\text{cm}} = 0.5 .$$

To calculate the cosine the length of the side b is required - it can be obtained by means of Pythagoras' theorem:

$$b^2 = c^2 - a^2$$

Hence,

$$\cos(\alpha) = \frac{b}{c} = \frac{\sqrt{c^2 - a^2}}{c} = \frac{\sqrt{(5\,\text{cm})^2 - (2.5\,\text{cm})^2}}{5\,\text{cm}} = 0.866 .$$

Thus, the tangent of the angle α is

$$\tan(\alpha) = \frac{\sin(\alpha)}{\cos(\alpha)} = \frac{0.5}{0.866} = 0.5773 .$$

Exercise 5.6.1
Determine some approximate values of the trigonometric functions sine, cosine and tangent graphically. Let a right triangle with the hypotenuse $c = 5$ be given. Use Thales' circle to draw right triangles for the angles

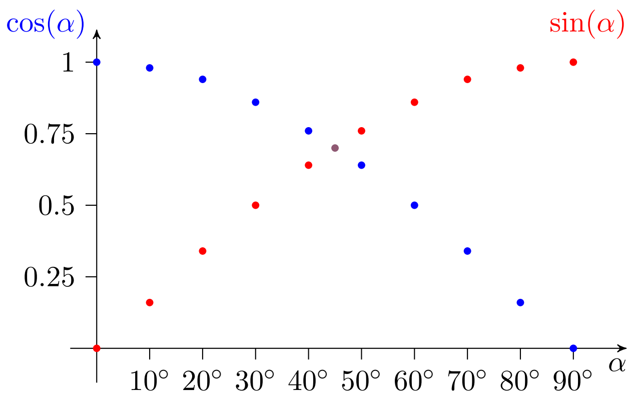
$$\alpha \in \{10°; 20°; 30°; 40°; 45°; 50°; 60°; 70°; 80°\} \ .$$

Use a drawing scale of 1 unit length $\hat{=}$ 2 cm, and fill in the measured values for the sides a and b in a table. From the measured values, calculate the sine, cosine, and tangent of each angle and decide for which functions also values for $\alpha = 0°$ and $\alpha = 90°$ exist. After that, plot the calculated values of sine and cosine against the angle α.

Solution:
In the process of measurement, errors will always occur! Therefore, the values in your table will be slightly different from the ones given in the table below. The table could look as follows:

α	a	b	$\sin(\alpha)$	$\cos(\alpha)$	$\tan(\alpha)$
0	0.0	5.0	0.0	1.0	0.0
10°	0.8	4.9	0.160	0.98	0.1633
20°	1.7	4.7	0.34	0.94	0.3617
30°	2.5	4.3	0.5	0.86	0.5814
40°	3.2	3.8	0.64	0.76	0.8421
45°	3.5	3.5	0.7	0.7	1.0
50°	3.8	3.27	0.76	0.64	1.1875
60°	4.3	2.5	0.86	0.5	1.7200
70°	4.7	1.7	0.94	0.34	2.7647
80°	4.9	0.8	0.98	0.160	6.1250
90°	5.0	0.0	1.0	0.0	–

Then, the corresponding diagram looks as follows:



If we once again look closer at the results obtained in the last exercise, we can find different ways to interpret them, and then identify some relations.

- With increasing angle α the opposite side a increases and the adjacent side b decreases.

  Likewise, $\sin(\alpha) \sim a$ and $\cos(\alpha) \sim b$.

- With increasing angle α the opposite side a increases to the same extent as the adjacent side b decreases with the angle α decreasing from $90°$. In the Thales circle, the two triangles with the opposite values of a and b are two solutions for the construction of a right triangle with a given hypotenuse and a given altitude (see also Example 5.3.6 auf Seite 168).

- In the right triangle the adjacent side of the angle $\beta = 90° - \alpha$ is the same side as the opposite side of the angle α (and vice versa). Thus,

  $$\sin(\alpha) = \cos(90° - \alpha) = \cos\left(\frac{\pi}{2} - \alpha\right)$$

  and

  $$\cos(\alpha) = \sin(90° - \alpha) = \sin\left(\frac{\pi}{2} - \alpha\right) .$$

- For $\alpha = 45°$ the opposite side and adjacent side are equal, and thus sine and cosine are equal as well. This observation was used at the beginning of this section for the determination of the slope.

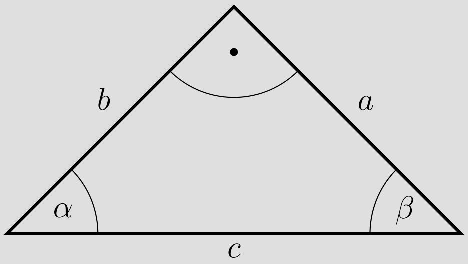
- The tangent function, i.e. the ratio of a to b, increases with increasing angle α from zero to “infinity”.

In the following example we will continue our considerations from the beginning of this section and use a triangle with an angle of $45°$ to calculate the value of the corresponding sine value exactly.

Example 5.6.3

Calculate the sine of the angle $\alpha = 45°$ now exactly, i.e. unlike as in Exercise 5.6.1 auf Seite 203, where the sine was calculated from measured (and hence error-prone) values.

If in a right triangle with $\gamma = 90°$ the angle α is equal to $45°$, then, because of the formula for the sum of interior angles in a right triangle, $\alpha + \beta + \gamma = \pi = 180°$, the angle β also needs to be equal to $45° = \pi/4$, and the two legs a and b are of equal length. A triangle with two sides of equal length is called an **isosceles**.



We have:

$$\sin(\alpha) = \sin(45°) = \frac{a}{c} \, .$$

Moreover:

$$a^2 + b^2 = 2a^2 = c^2 \quad \Rightarrow \quad c = \sqrt{2} \cdot a$$

$$\Rightarrow \quad \sin(45°) = \sin(\pi/4) = \frac{a}{\sqrt{2} \cdot a} = \frac{1}{2} \cdot \sqrt{2} \, .$$

In Exercise 5.6.1 auf Seite 203 the value of the sine of 45° was approximated by a value of 0.7 which is quite close to the actual value of $\frac{1}{2} \cdot \sqrt{2}$.

In the next example we will calculate the sine of the angle $\alpha = 60°$. For this purpose, we first do not consider a right triangle but an equilateral triangle. By a clever decomposition of the triangle and by using another "auxiliary quantity" we will obtain the required result.

Example 5.6.4

Consider a **equilateral** triangle to calculate $\sin(60°)$. As the name implies, the sides of this triangle are all of equal length, and the angles are also all of the same magnitude, namely $\alpha = \beta = \gamma = \frac{180°}{3} = 60° = \frac{\pi}{3}$. According to the theorem for congruent triangles "sss", the triangle is defined uniquely by the specification of a side a. This triangle is constructed by drawing the side a and then drawing a circle with radius r around both endpoints of the side. Now, the intersection point of the two circles is the third vertex.

This triangle is not right-angled. If an altitude h is drawn on one of the sides a, the triangle can be divided into two congruent right triangles.
We have:

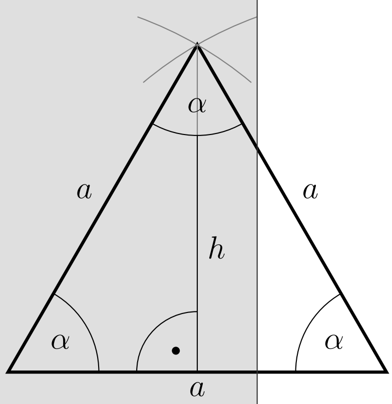
$$\sin(\alpha) = \sin(60°) = \frac{h}{a} \, .$$

According to Pythagoras' theorem we have

$$\left(\frac{a}{2}\right)^2 + h^2 = a^2 \, .$$

Therefore,

$$h^2 = \frac{3}{4}a^2 \quad \text{and hence} \quad h = \frac{1}{2}\sqrt{3} \cdot a \, .$$



As a result we obtain the required value

$$\sin(60^\circ) = \sin\left(\frac{\pi}{3}\right) = \frac{h}{a} = \frac{1}{2}\cdot\sqrt{3}\,.$$

From this triangle the sine of another angle can also be calculated: the altitude h bisects the above angle such that in the two congruent smaller triangles the above angle is $30^\circ = \frac{\pi}{6}$. Now we have

$$\sin(30^\circ) = \sin\left(\frac{\pi}{6}\right) = \frac{a/2}{a} = \frac{1}{2}\,.$$

Exercise 5.6.2
Calculate the exact value of the cosine of the angles $\alpha_1 = 30^\circ$, $\alpha_2 = 45^\circ$, and $\alpha_3 = 60^\circ$. To do this, use the results obtained in the example above and in Exercise 5.6.1 auf Seite 203.

Solution:
From Exercise 5.6.1 auf Seite 203 it is known that $\cos(\alpha) = \sin(90^\circ - \alpha)$. With the results obtained in the example above it follows

$$\begin{aligned}
\cos(30^\circ) &= \sin(90^\circ - 30^\circ) = \sin(60^\circ) = \frac{1}{2}\cdot\sqrt{3}\,,\\
\cos(45^\circ) &= \sin(90^\circ - 45^\circ) = \sin(45^\circ) = \frac{1}{2}\cdot\sqrt{2}\,,\\
\cos(60^\circ) &= \sin(90^\circ - 60^\circ) = \sin(30^\circ) = \frac{1}{2}\,.
\end{aligned}$$

The following small table lists the values for frequently used angles: In the first row denoted by x the angle is given in degree measure, and in the last row denoted by α the angle is given in radian measure.

x	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
sin	$0 = \frac{1}{2}\cdot\sqrt{0}$	$\frac{1}{2} = \frac{1}{2}\cdot\sqrt{1}$	$\frac{1}{2}\cdot\sqrt{2}$	$\frac{1}{2}\cdot\sqrt{3}$	$\frac{1}{2}\cdot\sqrt{4} = 1$
cos	$1 = \frac{1}{2}\cdot\sqrt{4}$	$\frac{1}{2}\cdot\sqrt{3}$	$\frac{1}{2}\cdot\sqrt{2}$	$\frac{1}{2}\cdot\sqrt{1} = \frac{1}{2}$	$\frac{1}{2}\cdot\sqrt{0} = 0$
tan	0	$\frac{\sqrt{3}}{3}$	1	$\sqrt{3}$	–
α	0°	30°	45°	60°	90°

You should learn these values by heart. The values of the trigonometric functions for other angles are listed in tables or saved in your calculator.

Hence, a height can be calculated very easily from an angle and a distance. Namely, if s is the distance of a building with a flat roof, which is observed at an angle of x, then from $\tan(x) = \frac{h}{s}$ we have $h = s \cdot \tan(x)$. Likewise, sine and cosine can be used to calculate lengths. This relation between angles and lengths is often used.

For example, an area can be calculated in this way even if the required length is not given directly. In the following example, the altitude of a triangle is to be calculated. Since the h starting in a vertex C is perpendicular to the line of the opposite side $c = \overline{AB}$, the vertices of h and A or B, respectively, form a right triangle. If an angle and the adjacent side are given, then the altitude can be calculated from $\sin(\alpha) = \frac{h}{b}$ or from $\sin(\beta) = \frac{h}{a}$, using standard notation.

Exercise 5.6.3
Calculate the area F of a triangle with the sides $c = 7$, $b = 3$, and the angle $\alpha = 30°$ between the two sides c and b.

Result: F = ______

Solution:
The area F can be calculated from $F = \frac{1}{2} \cdot c \cdot h_c$, where we still need to determine h_c. From $\sin(\alpha) = \frac{h_c}{b}$ we have

$$h_c = b \cdot \sin(\alpha) = 3 \cdot \sin(30°) = 3 \cdot \frac{1}{2} .$$

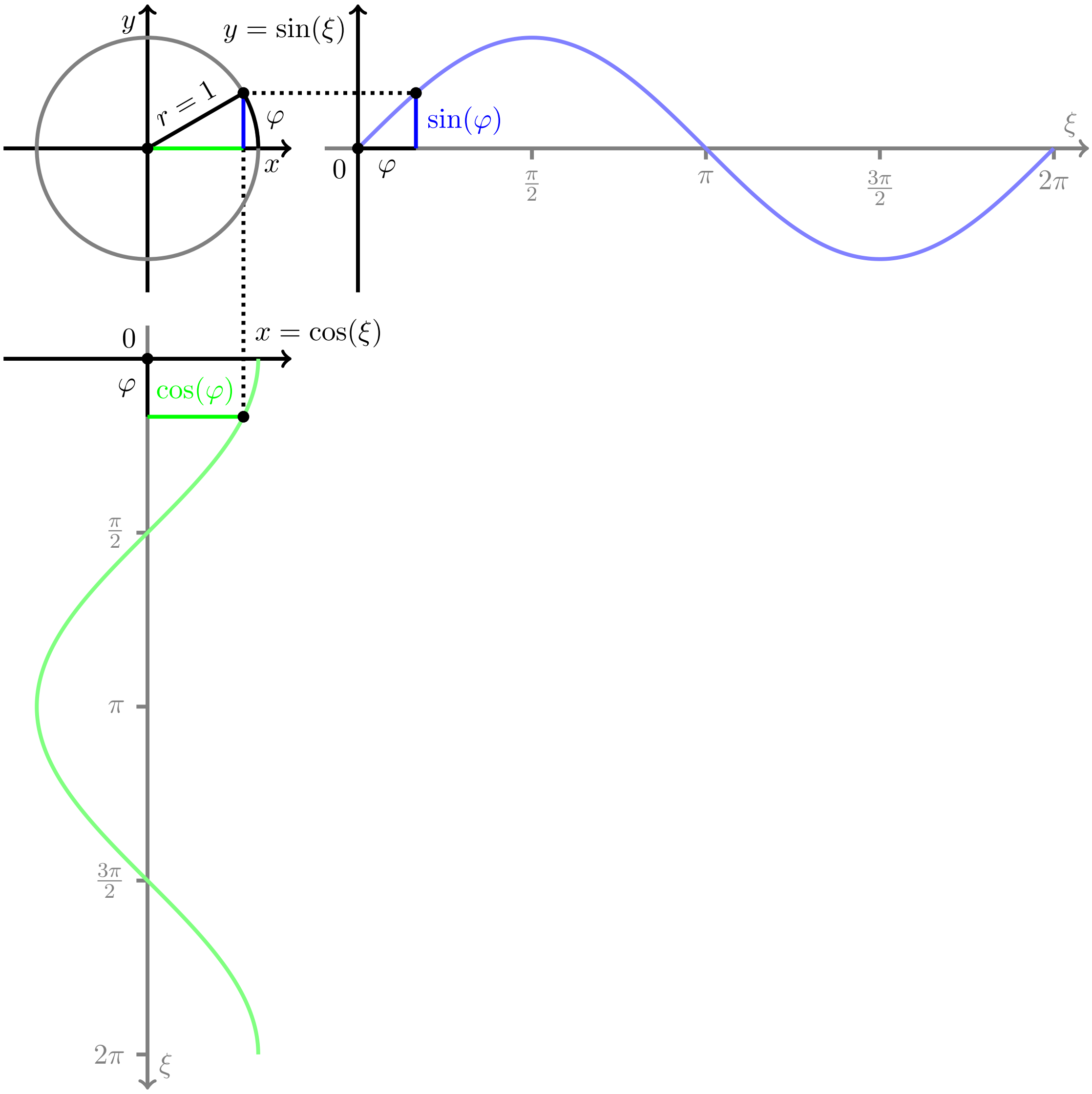
Hence,

$$F = \frac{1}{2} \cdot c \cdot b \cdot \sin(\alpha) = \frac{1}{2} \cdot 7 \cdot 3 \cdot \frac{1}{2} = \frac{21}{4} .$$

5.6.3 Trigonometry in the Unit Circle

In the previous section the trigonometric functions were introduced by means of a right triangle. Hence, the properties described above are valid for an angle ranging from 0° to 90° or from 0 to $\frac{\pi}{2}$, respectively.

To extend the acquired insights to angles greater than $\pi/2$, it is particularly useful to investigate the so-called unit circle.



The unit circle is a circle with a radius of 1. Its centre is positioned at the origin in the Cartesian coordinate system. Consider a line segment of length 1 starting from the centre. From its horizontal initial position on the positive x-axis, this segment is now rotated counter-clockwise, i.e. in the mathematical positive direction, around its centre. In this process, its rotating end point is sweeping the unit circle enclosing the angle φ with the positive x-axis. During rotation, the angle φ increases from 0 to 2π or $360°$, respectively. Thus, to any angle φ there corresponds a point with the coordinates x_φ and y_φ on the unit circle.

For φ from 0 to $\frac{\pi}{2}$, the line segment, the corresponding segment on the x-axis, and the the corresponding segment on the y-axis can be regarded as a right triangle. The hypotenuse is the line segment of length 1, the x-intercept is the adjacent side, and the y-intercept is the opposite side. This matches the situation described in the previous section.

Hence, the sine of the angle φ is

$$\sin(\varphi) = \frac{y_\varphi}{1} = y_\varphi$$

and the cosine is

$$\cos(\varphi) = \frac{x_\varphi}{1} = x_\varphi \ .$$

Based on the description above, these definitions now are also valid for angles $\varphi > \pi/2$. Here, the values of x_φ and y_φ can be negative as well, hence also sine and cosine can be negative. If the y-values are plotted against the angle φ, one obtains for the sine function the blue curve. Plotting y-values against the angle φ one obtains for the cosine function the green curve. If the line segment is rotated in the opposite direction, values for negative angles can be defined accordingly.

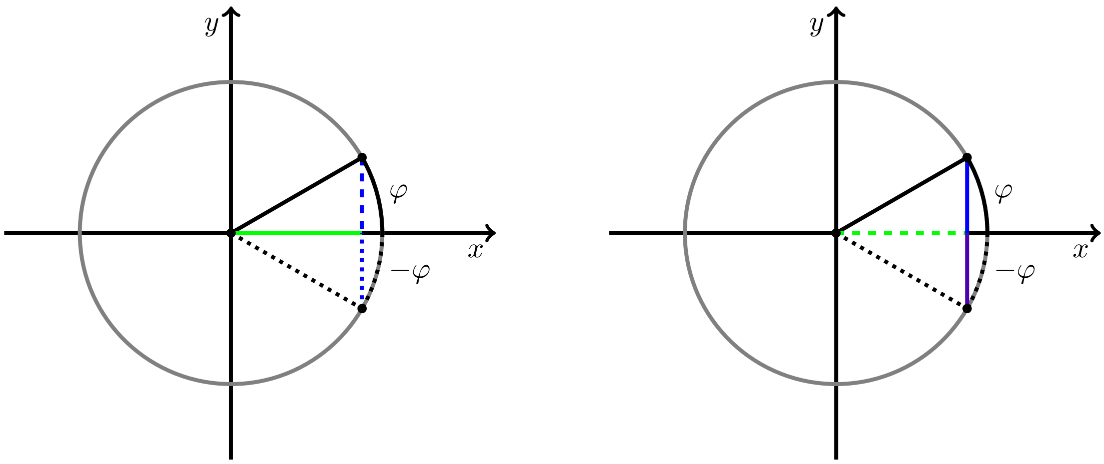
Furthermore, using Pythagoras' theorem, we have

$$x_\varphi^2 + y_\varphi^2 = 1 \ .$$

Replacing x_φ and y_φ by the corresponding relations to the trigonometric functions results for any φ in the important relation

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

Additionally, from the description of the sine and cosine function, it can be seen that the values of the cosine function do not change if the line segment is reflected across the x-axis. Hence, the cosine value of the angle φ is equal to the cosine value of the angle $-\varphi$ (indicated in the figure below by the green line). For the sine function, a reflection across the x-axis results in a change of sign of the sine value (indicated in the figure below by the blue line and the violet line, respectively)



Expressed in formulas, this is

$$\cos(-\varphi) = \cos(\varphi) \qquad \text{and} \qquad \sin(-\varphi) = -\sin(\varphi)$$

for every angle φ. These symmetry properties are useful for many calculations. An elementary example is the calculation of the angle between the x-axis and the connecting line from the origin to a point in the Cartesian coordinate system (see also Exercise 5.6.4 auf der nächsten Seite).

Example 5.6.5

Find the values of the sine, cosine, and tangent function of the angle $\alpha = 315°$.

For $\alpha = 315°$, the point P_α lies in the fourth quadrant. On the unit circle it es also described by the negative angle $\varphi = 315° - 360° = -45°$. Therefore, we have $\sin(315°) = \sin(-45°) = -\sin(45°) = \frac{1}{2}\sqrt{2}$ and $\cos(315°) = \cos(-45°) = \cos(45°) = \frac{1}{2}\sqrt{2}$ as well as $\tan(315°) = \tan(-45°) = -1$.

5.6.4 Exercises

Exercise 5.6.4
What is the degree measure of the angle φ between the x-axis and the connecting line from the origin in the Cartesian coordinate system to the point $P_\varphi = (-0.643; -0.766)$ on the unit circle? Use a calculator, but do not trust it blindly!

Result: φ = [   ] °

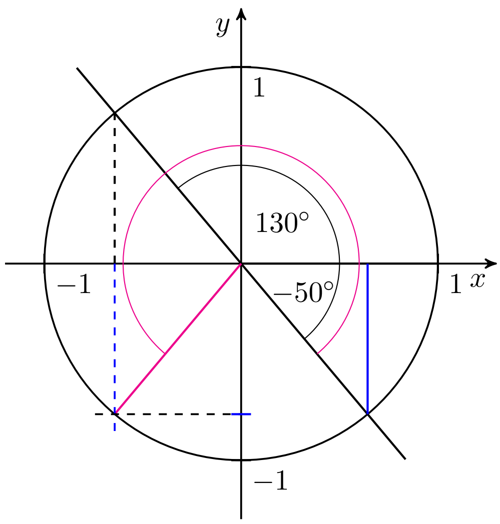
Solution:
From the coordinates of the point P_φ we have

$$\cos(\alpha) = -0.643 \quad \text{and} \quad \sin(\alpha) = -0.766\ .$$

If you enter

- `invers(cos(-0.643))` or $\cos^{-1}$(-0.643) in the calculator, you obtain approximately 130°
- `invers(sin(-0.766))` or $\sin^{-1}$(-0.766) in the calculator, you obtain approximately $-50°$.

Moreover, you know that the point lies in the third quadrant. Thus, the angle must be in the range from 180° to 270°.



The figure to the left shows that the negative cosine value corresponds to the angle $-130°$ and to the angle $\varphi = -130° = -130° + 360° = 230°$.
Likewise, the negative sine value can correspond to the angle $-50°$ and to the angle $\varphi = -(-50°) + 180° = 230°$.
Since this last value lies in the range stated above, the required value of the angle is $\varphi = 230°$ indicated in the figure by a pink line.

Exercise 5.6.5 1. Let a right triangle with the right angle at the vertex C and the sides $b = 2.53\,\text{cm}$ and $c = 3.88\,\text{cm}$ be given. Calculate the values of $\sin(\alpha)$, $\sin(\beta)$, and a.

Results:

- $\sin(\alpha)$ = [ ]
- $\sin(\beta)$ = [ ]
- a = [ ] cm

Solution:
We have

$$a = \sqrt{c^2 - b^2} = \sqrt{(3.88\,\text{cm})^2 - (2.53\,\text{cm})^2} = \sqrt{15.0544\,\text{cm}^2 - 6.4009\,\text{cm}^2} = \sqrt{8.6535}\,\text{cm}\ ,$$

and

$$\sin(\alpha) = \frac{a}{c} = \frac{\sqrt{8.6535}\,\text{cm}}{3.88\,\text{cm}} = \frac{\sqrt{86535}}{388} \quad \text{and} \quad \sin(\beta) = \frac{b}{c} = \frac{2.53\,\text{cm}}{3.88\,\text{cm}} = \frac{253}{388}\ .$$

Numerically, we obtain $a \approx 2.9417\,\text{cm}$, $\sin(\alpha) \approx 0.7587$, and $\sin(\beta) \approx 0.65201$.

2. Calculate the area F of a triangle with the sides $a = 4\,\text{m}$, $c = 60\,\text{cm}$, and the angle $\beta = \measuredangle(a, c) = \frac{11\pi}{36}$.

   Result: F = [ ] m^2

   Solution:

$$\frac{(a \cdot \sin(\beta)) \cdot c}{2} = \sin\left(\tfrac{11\pi}{36}\right) \cdot 1.2\,\text{m}^2 \approx 0.98298\,\text{m}^2\ .$$

5.7 Final Test

5.7.1 Final Test Module 7

Exercise 5.7.1
Identify the figures below as precisely as possible by specifying the name of the type (preceded by an adjective if necessary) and describing as many properties of the figure as possible.

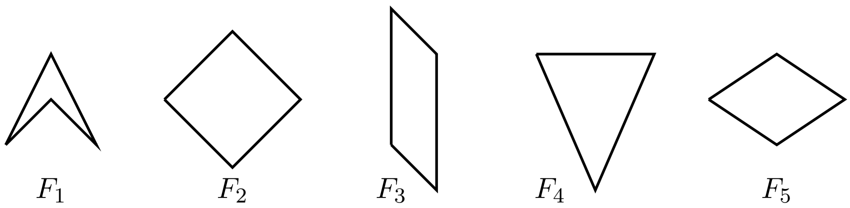


Figure:	Description of the Type:
F_1	
F_2	
F_3	
F_4	
F_5	

Exercise 5.7.2
Are the following results and statements right or wrong?

right	wrong	
☐	☐	Every rectangle is a rhombus.
☐	☐	Every square is a parallelogram.
☐	☐	It exists exactly one square with a diagonal of 5 cm.
☐	☐	A triangle with the angles $36°$ and $54°$ is right-angled.
☐	☐	In a rectangle the sum of all (interior) angles in radian measure is equal to 4π.

Exercise 5.7.3
Consider a triangle ABC with side lengths $a = 5\,\text{cm}$, $b = 6\,\text{cm}$, and $c = 9\,\text{cm}$. On the side c a point P and on the side b a point Q are chosen such that PQ is parallel to the side a and $[\overline{PQ}] = 0.50\,\text{cm}$. Calculate the lengths of the line segments $[\overline{PB}]$ and $[\overline{QC}]$ specified in centimetre.

a. $[\overline{PB}]$ = ☐ cm

b. $[\overline{QC}]$ = ☐ cm

Exercise 5.7.4

Consider a square with sides of length a. Find the formulas for the area and the circumference for the largest circle inscribed to the square as well as for the smallest circle containing the square completely:

a. Circumference of the circle within the square as a function of the side length a:

b. Area of the circle within the square as a function of the side length a:

c. Circumference of the circle around the square as a function of the side length a:

d. Area of the circle around the square as a function of the side length a:

Do not enter any brackets or radical terms. Enter, for example, $2^{0.5}$ instead of $\sqrt{2}$ to avoid the radical.

6 Elementary Functions

Module Overview

6.1 Basics of Functions

6.1.1 Introduction

From Module 1 we already know that real numbers are sets and intervals are subsets of real numbers.

Example 6.1.1

All real numbers $\mathbb{R}$, excluding the number $0 \in \mathbb{R}$, are to be collected in a set. How is this set of numbers described? For this, the notation

$$\mathbb{R} \setminus \{0\}$$

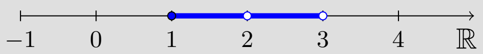
is used. This reads as "$\mathbb{R}$ without 0". Alternatively, this set can be described as a union of two open intervals:

$$\mathbb{R} \setminus \{0\} = (-\infty; 0) \cup (0; \infty) \ .$$

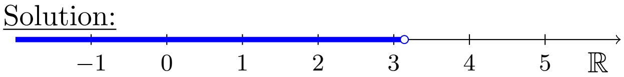
In the same way, single numbers can be removed from any other sets. So, for example, the set

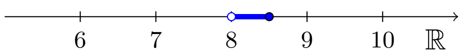
$$[1; 3) \setminus \{2\}$$

contains all numbers of the half-open interval $[1; 3)$, excluding the number 2:



Exercise 6.1.1
Indicate the intervals $(-\infty; \pi)$ and $(8; 8.5]$ on the number line.
Solution:





For doing and applying mathematics, it is not sufficient to just consider sets and equations and inequalities for numbers of these sets, as done in the previous modules (for example, in Module 1). We also need functions (which are often also denoted as maps).

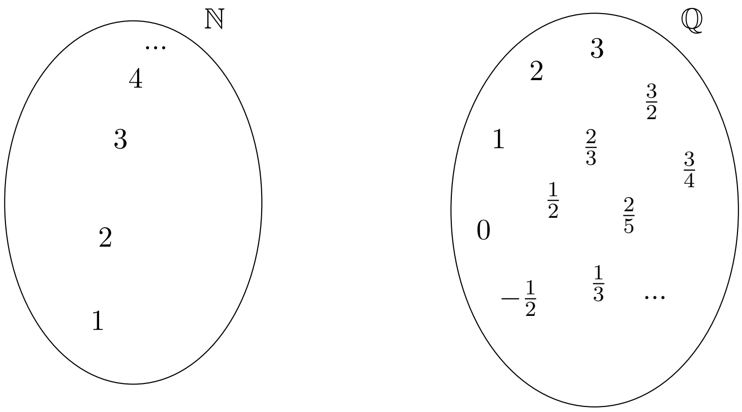
Info6.1.2

Functions (or **maps**) are assignments between elements of two sets such that there is exactly one element in the second set assigned to each element in the first set.

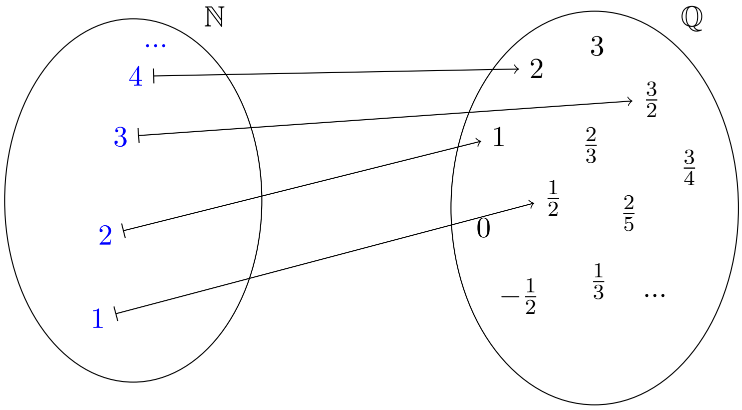
We will focus on the basic mathematical term of an assignment between numbers in Section 6.1.2. In Section 6.1.3 we will refer to applications of mathematics in other sciences, and we will illustrate how useful the mathematical notion of a function as a tool for the formalisation of relations between dependent quantities is. Finally, we will study the graphical representation of functions using graphs in Section 6.1.4. Later in this chapter we will consider the most relevant elementary functions together with their graphs. It is fundamental to know the behaviour of the graphs of the elementary functions.

6.1.2 Assignments between Sets

We start with a first example of a function as an assignment between two sets. For this purpose, we consider the set of natural numbers $\mathbb{N}$ and the set of rational numbers $\mathbb{Q}$, and visualise these two sets as two “containers” filled with numbers.



Now, we want to create an assignment between the elements of these two sets as follows. To every number $n \in \mathbb{N}$ half of this number $\frac{n}{2} \in \mathbb{Q}$ is assigned, i.e. to the number $1 \in \mathbb{N}$ we assign the number $\frac{1}{2} \in \mathbb{Q}$, to the number $2 \in \mathbb{N}$ we assign the number $1 \in \mathbb{Q}$, etc. In the figure below, this is illustrated by arrows that indicate which numbers in $\mathbb{Q}$ are assigned to which numbers in $\mathbb{N}$.



For the assignment of single elements of the sets, described above in words, we use the so-called assignment arrow. This is an arrow that has a bar at its tail: $\longmapsto$. It says that to the number on the side with the bar the number on the side with the arrow is assigned:

$$\mathbb{N} \ni 1 \longmapsto 0.5 \in \mathbb{Q}\,, \quad \mathbb{N} \ni 2 \longmapsto 1 \in \mathbb{Q}\,, \quad \text{etc.}$$

With these assignments, we have now constructed a function from the natural numbers $\mathbb{N}$ to the rational numbers $\mathbb{Q}$. In mathematics, this assignment is named, i.e. a symbol is allocated (often this is f for function), that shall describe exactly this assignment. For this purpose, the sets of numbers from which and to which the assignment will be done must be noted. In our case, to the elements of the set of natural numbers $\mathbb{N}$ the rational numbers are assigned. Mathematically, this is expressed by a so-called mapping arrow $\longrightarrow$, i.e. an arrow that has the target set at its head and the set, whose elements are assigned to, at its tail. Thus, in our case we have

$$f\colon \mathbb{N} \longrightarrow \mathbb{Q}\,.$$

This reads as "the function f maps $\mathbb{N}$ to $\mathbb{Q}$".

Furthermore, we can now ask whether the assignments of this function, $1 \longmapsto \frac{1}{2}$, $2 \longmapsto 1$, etc., can be described in a more compact way. To do this, we recall the beginning of this example. We decided to assign to every natural number n its half $\frac{n}{2}$. So we can write this arbitrary natural number n and the corresponding rational number $\frac{n}{2}$ left and right to the assignment arrow, respectively:

$$n \longmapsto \frac{n}{2}\,.$$

This reads as "n is mapped to $\frac{n}{2}$". This notation is also called mapping rule of the function. Another notation for the mapping rule uses the name of the function:

$$f(n) = \frac{n}{2}\,.$$

This reads as "f of n equals $\frac{n}{2}$". Altogether we can describe this function f as follows:

$$f : \begin{cases} \mathbb{N} & \longrightarrow & \mathbb{Q} \\ n & \longmapsto & \frac{n}{2}\,. \end{cases}$$

Finally, this reads as "the function f maps $\mathbb{N}$ to $\mathbb{Q}$, each $n \in \mathbb{N}$ is mapped to $\frac{n}{2} \in \mathbb{Q}$". Throughout this module, we will continue to use this summarising notation of functions.

Let us consider a few further simple examples for functions.

Example 6.1.3

- A function g assigns to every real number x its square: $x \cdot x = x^2$. This results in the so-called standard parabola (see Section 6.2.6):

$$g : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{cases}.$$

  Hence, the mapping rule of g is $g(x) = x^2$. Then, assignments can be calculated for specific numbers. For example, $g(2) = 2^2 = 4$ or $g(-\pi) = (-\pi)^2 = \pi^2$, etc.

- A function φ shall assign to every real number y between 0 and 1 three times its value plus 1. This is an example for a so-called linear affine function (see Section 6.2.4):

$$\varphi : \begin{cases} (0;1) & \longrightarrow & \mathbb{R} \\ y & \longmapsto & 3y+1 \end{cases}.$$

  Hence, the mapping rule of φ is $\varphi(y) = 3y + 1$. Thus, for example, $\varphi(\frac{1}{3}) = 3 \cdot \frac{1}{3} + 1 = 2$, etc. However, in this case the values $\varphi(8)$ or $\varphi(1)$ cannot be calculated since 8 and 1 do not belong to the set $(0;1)$.

Exercise 6.1.2 (i) Specify a function h that assigns to every positive real number x its reciprocal. Calculate the values $h(2)$ and $h(1)$. Complete the two assignments

$$3 \longmapsto\ ? \quad \text{and} \quad ?\ \longmapsto 2$$

of h.

(ii) Describe in words the assignment that is done by the following function:

$$w : \begin{cases} [4;9] & \longrightarrow & \mathbb{R} \\ \alpha & \longmapsto & \sqrt{\alpha} \end{cases}.$$

Calculate $w(9)$ and $w(5)$. Can $w(10)$ also be calculated?

Solution:

(i) The reciprocal of x is $\frac{1}{x}$. The set of positive real numbers is $(0;\infty)$. Hence, the function h can be written as

$$h : \begin{cases} (0;\infty) & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{1}{x} \end{cases}.$$

This is an example for a function of hyperbolic type that will be described in more detail in Section 6.3. The mapping rule of h is $h(x) = \frac{1}{x}$, hence $h(2) = \frac{1}{2}$ and $h(1) = \frac{1}{1} = 1$. Moreover, we have $h(3) = \frac{1}{3}$, hence $3 \longmapsto \frac{1}{3}$. The observation $h(\frac{1}{2}) = \frac{1}{\frac{1}{2}} = 2$ gives us $\frac{1}{2} \longmapsto 2$.

(ii) The function w assigns to every real number α greater or equal to 4 and less or equal to 9 its square root $\sqrt{\alpha}$. The mapping rule is $w(\alpha) = \sqrt{\alpha}$, hence $w(9) = \sqrt{9} = 3$ and $w(5) = \sqrt{5}$. The value $w(10)$ cannot be calculated since $10 \notin [4;9]$.

The examples above show some basic properties of functions for which we will now introduce specific terminology.

Info6.1.4

For the definition of a function, a set of numbers is specified whose elements are to be assigned to other numbers by the function. This set is called the **domain** or set of definition of the function. If the function has a name, e.g. f, then the domain is denoted by the symbol D_f. For example, the domain of the function

$$h : \begin{cases} (0;\infty) & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{1}{x} \end{cases}$$

in Exercise 6.1.2 is the set $D_h = (0;\infty)$. There is also a specific term for the elements of the domain. In this exercise, the numbers $x \in D_h$ are assigned by the mapping rule $h(x) = \frac{1}{x}$. Here, the variable x is called the **independent variable** of the function h.

Exercise 6.1.3

Specify the domain of function w in Exercise 6.1.2 and function g in Example 6.1.3.

Solution:

We have

$$w : \begin{cases} [4;9] & \longrightarrow & \mathbb{R} \\ \alpha & \longmapsto & \sqrt{\alpha} \end{cases}$$

and

$$g : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{cases},$$

hence $D_w = [4; 9]$ and $D_g = \mathbb{R}$.

If we consider the mapping rule $h(x) = \frac{1}{x}$ of the function, we see that there is no reason not to insert in $\frac{1}{x}$ any real number x, excluding the value $x = 0$, since the operation "$\frac{1}{0}$" has no solution. Therefore, in specifying the domain, we can distinguish between numbers that are excluded because they are not allowed to be inserted in the mapping rule at all and those that are excluded because the function is just defined accordingly. This now leads to the term of the maximal domain of a function, i.e. the maximum subset of real numbers $\mathbb{R}$ that can be used as the domain of a function with a known mapping rule.

Example 6.1.5

The maximum domain $D_h \subset \mathbb{R}$ of the function

$$h : \begin{cases} D_h & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{1}{x} \end{cases},$$

is $D_h = \mathbb{R} \setminus \{0\}$.

Exercise 6.1.4
Find the maximum domain of the function

$$w : \begin{cases} D_w & \longrightarrow & \mathbb{R} \\ \alpha & \longmapsto & \sqrt{\alpha} \end{cases}.$$

Solution:
The square root has a real number as its result for all non-negative real numbers. Hence, $D_w = [0; \infty)$.

For the definition of a function, a second set (beside the domain) is required, the set that is the target of the assignment by the function. This set is called target set or codomain. Let us again consider the function

$$\varphi : \begin{cases} (0; 1) & \longrightarrow & \mathbb{R} \\ y & \longmapsto & 3y + 1 \end{cases}$$

in Example 6.1.3. Its target set is the set of real numbers $\mathbb{R}$. The target set of the function

$$f : \begin{cases} \mathbb{N} & \longrightarrow & \mathbb{Q} \\ n & \longmapsto & \frac{n}{2} \end{cases}$$

in the first example of this section is the set of rational numbers $\mathbb{Q}$. Here, we see an important difference between the domain and target set of a function. The domain contains all numbers, and only those numbers, that are allowed to be inserted (and one wishes to insert) in the mapping rule of the function. However, the target set can contain all numbers that can potentially appear as a result of the mapping function.

In this context the question arises: What is the smallest target set that can be used for a function with given domain and known mapping rule? The smallest target set is the set of numbers that – for a given domain and mapping rule – indeed appear as targets of the assignment. This set is called **range** or image, and its elements are called values of the function. For a function f, the symbol W_f is used for the image. For the values of a function f of an independent variable x we write generally $f(x) \in W_f$, as in the mapping rule, or we introduce another variable, e.g. $y = f(x) \in W_f$.

Example 6.1.6

Let us consider again the example

$$\varphi : \begin{cases} (0;1) & \longrightarrow & \mathbb{R} \\ y & \longmapsto & 3y + 1 \ . \end{cases}$$

The range of this function is

$$W_\varphi = (1;4) \ .$$

This can be seen by inserting some values from $D_\varphi = (0;1)$ in the mapping rule and calculating the results. In this way, a so-called **table of values** is obtained:

y	0.1	0.3	0.5	0.7	0.9
$\varphi(y)$	1.3	1.9	2.5	3.1	3.7

Such tables of values are useful to get an overview of the values of a function. However, from a mathematical point of view, they are not sufficient to be completely sure what the actual range of a function is. One method to determine the range of a function is based on solving inequalities:

Example 6.1.7

For the function

$$\varphi : \begin{cases} (0;1) & \longrightarrow & \mathbb{R} \\ y & \longmapsto & 3y+1 , \end{cases}$$

due to the domain $D_\varphi = (0;1)$ for the independent variable y we have

$$0 < y < 1 .$$

Now, we use equivalent transformations to create the mapping rule $\varphi(y) = 3y+1$ in this inequality:

$$0 < y < 1 \,|\cdot 3 \quad \Leftrightarrow \quad 0 < 3y < 3 \,|+1 \quad \Leftrightarrow \quad 1 < 3y+1 < 4 \quad \Leftrightarrow \quad 1 < \varphi(y) < 4 .$$

Hence, we have for the values of the function $\varphi(y) \in (1;4)$ and therefore $W_\varphi = (1;4)$.

6.1.3 Functions in Mathematics and Applications

Mathematical functions often describe relations between quantities that arise from other sciences or everyday life. For example, the volume V of a cube depends on the side length a of the cube. The volume can be considered as a mathematical function that assigns the corresponding volume $V(a) = a^3$ to every possible side length $a > 0$:

$$V : \begin{cases} (0;\infty) & \longrightarrow & \mathbb{R} \\ a & \longmapsto & V(a) = a^3 . \end{cases}$$

The result is the cubic standard parabola (see Section 6.2.6) for the relation between side length and volume. In this way, many more examples can be found arising from sciences and everyday life: the position as a function of time in physics, the reaction rate as a function of concentration in chemistry, the amount of flour needed as a function of the desired amount of dough in a cake recipe, etc.

To this end, let us consider an example.

Example 6.1.8

The intensity of nuclear radiation is inversely proportional to the square of the distance to the source. This is called the inverse square law. Using a physical proportionality factor $c > 0$, the relation between intensity I of the radiation and distance

$r > 0$ from the source can be formulated mathematically as follows:

$$I : \begin{cases} (0;\infty) & \longrightarrow & \mathbb{R} \\ r & \longmapsto & \frac{c}{r^2} \end{cases} .$$

Hence, for the intensity we have the mapping rule $I(r) = \frac{c}{r^2}$ that describes the relation between the quantities I and r.

Exercise 6.1.5

In the construction of wind turbines it is known that the wind turbine power is proportional to the cube of the wind velocity. Under the condition that the proportionality factor satisfies the relation $\rho > 0$, which of the following mathematical functions describes this relation of physical quantities correctly?

a)

$$P : \begin{cases} (0;\infty) & \longrightarrow & \mathbb{R} \\ v & \longmapsto & P(v) = \frac{\rho}{v^3} \end{cases}$$

b)

$$P : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ v & \longmapsto & P(v) = \rho v^3 \end{cases}$$

c)

$$P : \begin{cases} [0;\infty) & \longrightarrow & \mathbb{R} \\ v & \longmapsto & P(v) = \rho v^3 \end{cases}$$

d)

$$x : \begin{cases} [0;\infty) & \longrightarrow & \mathbb{R} \\ f & \longmapsto & x(f) = \rho f^3 \end{cases}$$

Solution:

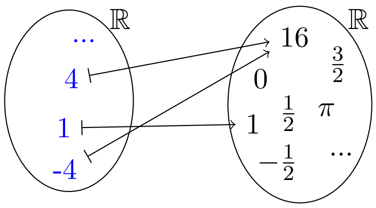
c) and d) are correct. a) is wrong since it describes an inverse proportionality. b) is not adequate to the problem. Only non-negative wind velocities make sense in this context and hence, negative values should be excluded from the domain. c) is correct; in this case we have the relation $P(v) = \rho v^3$ between the power P and the wind velocity v. We see that d) is equally correct. In this case we have a function that is completely identical to case c), except for the fact that in this case the power is denoted by x and the wind velocity is denoted by f. This again clarifies that the letters used to denote the function and the variables are mathematically completely arbitrary. However, in the sciences there are conventions, which letters are generally used to denote certain quantities. So, here it is more common to denote the velocity by v and the power by P.

6.1.4 Invertability

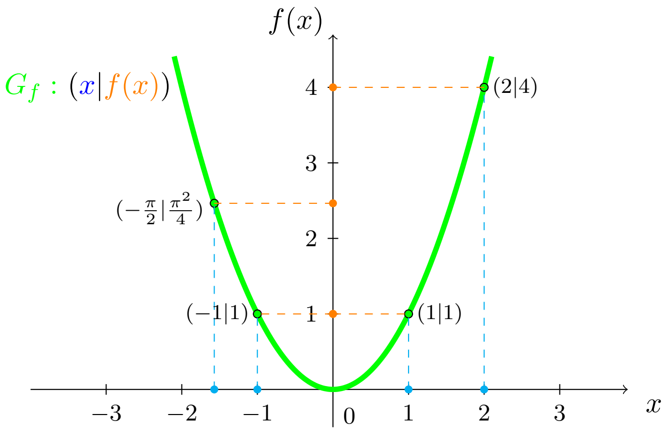
The visualisation of a function, as for example

$$f: \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{cases},$$

in form of a so-called Venn diagram (see Section 6.1.2)



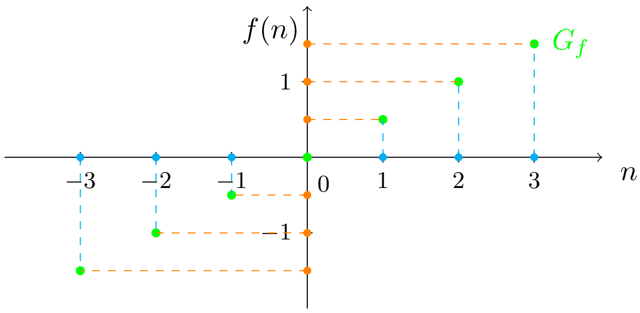
is in fact useful to understand what a function really is, but it says nothing about the special properties of the function. For this purpose, another kind of graphical representation exists, namely the representation as the graph of the function. For this representation, we draw a two-dimensional coordinate system (see Module 9) in which the numbers of the domain of the function are indicated on the horizontal axis and the numbers from the range are indicated on the vertical axis. In such a coordinate system we mark all points $(x|f(x))$ arising by the assignment of the function $x \longmapsto f(x)$. In this case, these are all points $(x|x^2)$, i.e. $(1|1)$, $(-1|1)$, $(-\frac{\pi}{2}|\frac{\pi^2}{4})$, etc. This results in a curve that is called graph of f and that is denoted by the symbol G_f:



If we consider the function

$$f: \begin{cases} \mathbb{N} & \longrightarrow & \mathbb{Q} \\ n & \longmapsto & \frac{n}{2} \end{cases}$$

from Section 6.1.2 and its graph

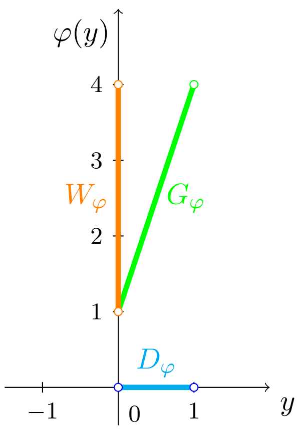


then we see that graphs not always have to be continuous curves, but also can consist, as in this case, only of single points.

Now, from the graph we can see many basic properties of the function. Recall the function

$$\varphi : \begin{cases} (0;1) & \longrightarrow & \mathbb{R} \\ y & \longmapsto & 3y+1 \end{cases}$$

with domain $D_\varphi = (0;1)$ and range $W_\varphi = (1;4)$ from Section 6.1.2. If we draw its graph, then we see that the domain and the range appear on the horizontal and on the vertical axes, respectively:

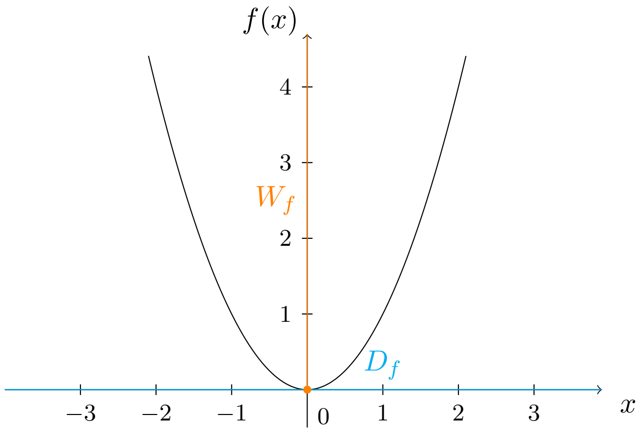


Exercise 6.1.6
Consider again the graph of the function

$$f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{cases},$$

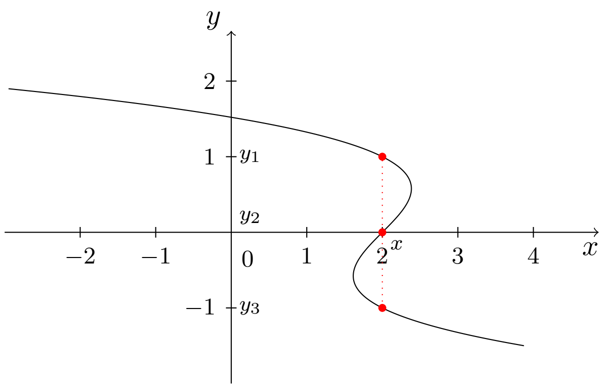
indicate domain and range on the the horizontal axis and vertical axes and label them.

<u>Solution:</u>



$D_f = \mathbb{R}$, $W_f = [0; \infty)$

Furthermore, the property of uniqueness of functions can be seen from the graph. To convince ourselves of this, we realise that a curve as shown in the figure below

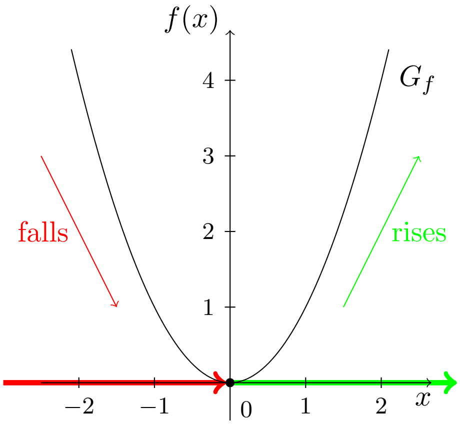


cannot be a graph of a function. For a single x-value in the domain, several values y_1, y_2, y_3 had to exist in the range. Graphs of functions indicate the uniqueness by the fact that they “cannot reverse in the horizontal direction”.

Another important property of a graph is its growing behaviour. Let us consider the function

$$f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{cases}$$

and its graph.

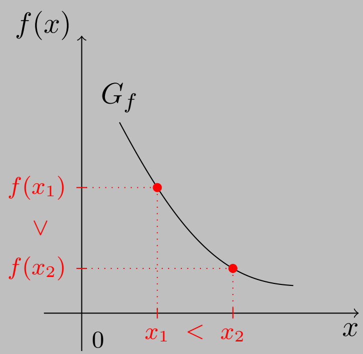


On the horizontal axis in this graph we see two regions in which the growing behaviour of the graph is different. In the region of x-values with $x \in (-\infty; 0)$ the graph falls, i.e. if the x values increase, then the corresponding function values on the vertical axis decrease. In the region of x-values with $x \in (0; \infty)$ we observe the contrary behaviour. For increasing x-values the corresponding function values also increase: the graph rises. At the x-value $0 \in \mathbb{R}$ the region of decreasing growing behaviour merges into the region of increasing growing behaviour. Such values will be of particular importance for the study of vertices in Section 6.2.7 and for the determination of extreme values in Module 7.

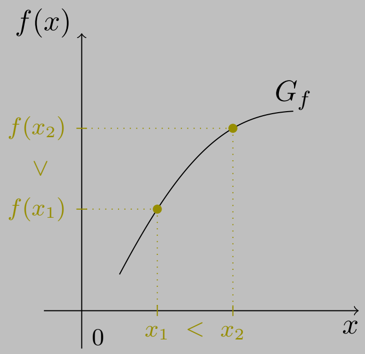
These two properties are referred to as strictly decreasing and strictly increasing, respectively, and they are defined mathematically as follows:

Info6.1.9

- If for $x_1 < x_2$ in a subset of the domain of a function f we always have $f(x_1) > f(x_2)$, the function f is said to be strictly decreasing for every x in this subset.



- If for $x_1 < x_2$ in a subset of the domain of a function f we always have $f(x_1) < f(x_2)$, the function f is said to be strictly increasing for every x in this subset.



This is true for all functions that will be investigated throughout this module. Often, the described monotonicity properties are only true for certain regions in the functions's domain, as we have seen for the case of the standard parabola. However functions also exist which only have one of the monotonicity properties for the entire domain (see Example 6.1.10 below). In this case the entire function is called strictly increasing or strictly decreasing. A function that is either strictly decreasing or strictly increasing is simply called strictly monotonic.

A further example shows how strict monotonicity can be checked explicitly for a function by solving inequalities from Module 3 auf Seite 82.

Example 6.1.10

Consider the function

$$h : \left\{ \begin{array}{rcl} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & -\frac{1}{2}x + 1 \end{array} \right. .$$

Determine whether the function h is either strictly increasing or strictly decreasing.

First, we choose two arbitrary numbers $x_1, x_2 \in D_h = \mathbb{R}$ for which

$$x_1 < x_2 \ .$$

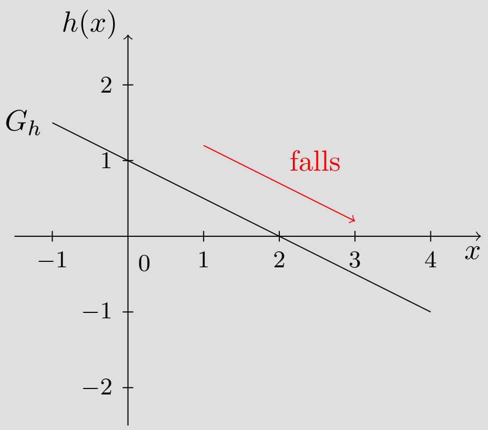
By equivalent transformations of inequalities, we can transform $x_1 < x_2$ either to $h(x_1) < h(x_2)$ or to $h(x_1) > h(x_2)$, and from this we can conclude that h is strictly increasing or strictly decreasing. We apply the following equivalent transformations to the inequality $x_1 < x_2$:

$$x_1 < x_2 \,|\cdot(-\frac{1}{2}) \quad \Leftrightarrow \quad -\frac{1}{2}x_1 > -\frac{1}{2}x_2 \,.$$

Furthermore, we add $+1$ to the mapping rule. Thus, we obtain

$$-\frac{1}{2}x_1 > -\frac{1}{2}x_2 \,|+1 \quad \Leftrightarrow \quad -\frac{1}{2}x_1 + 1 > -\frac{1}{2}x_2 + 1 \quad \Leftrightarrow \quad h(x_1) > h(x_2)\,.$$

Since $h(x_1) > h(x_2)$, the function h is strictly decreasing. This can also be seen from the graph of h:



Exercise 6.1.7
Use equivalent transformations to check explicitly that the function

$$\eta : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & 2x+2 \end{cases}$$

is strictly increasing.
<u>Solution:</u>
We have

$$x_1 < x_2 \,|\cdot 2 \quad \Leftrightarrow \quad 2x_1 < 2x_2 \,|+2 \quad \Leftrightarrow \quad 2x_1 + 2 < 2x_2 + 2 \quad \Leftrightarrow \quad \eta(x_1) < \eta(x_2)\,,$$

hence the function η is strictly increasing.

6.2 Linear Functions and Polynomials

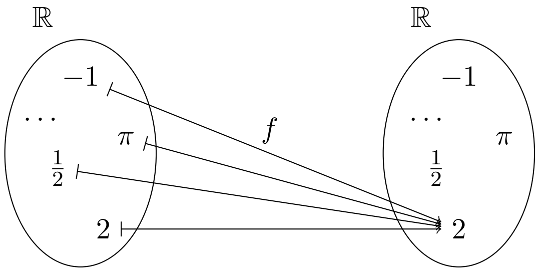
6.2.1 Introduction

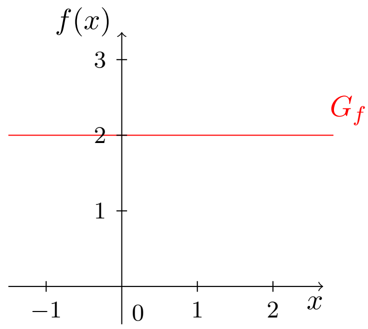
In this section we study the following types of functions: constant functions, linear functions, linear affine functions, monomials and polynomials.

6.2.2 Constant Functions and the Identity

so-called constant functions assign to every number in the domain $\mathbb{R}$ exactly the same constant number in the target set $\mathbb{R}$, e.g. the constant number 2, in the following way:

$$f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & 2\,. \end{cases}$$

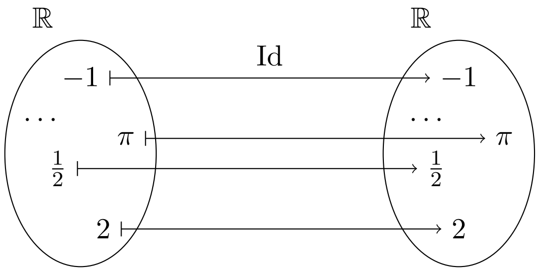


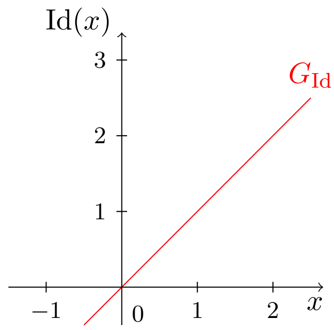


Here, we then have $f(x) = 2$ for all $x \in \mathbb{R}$. Hence, the range of this function f consists only of the set $W_f = \{2\} \subset \mathbb{R}$.

The identity function on $\mathbb{R}$ is the function that assigns each real number to itself. This is written as follows:

$$\mathrm{Id} : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x\,. \end{cases}$$





Here, we then have $\mathrm{Id}(x) = x$ for all $x \in \mathbb{R}$. Hence, the range of Id is the set of real numbers ($W_{\mathrm{Id}} = \mathbb{R}$). Furthermore, the identity function is (obviously) a strictly increasing function.

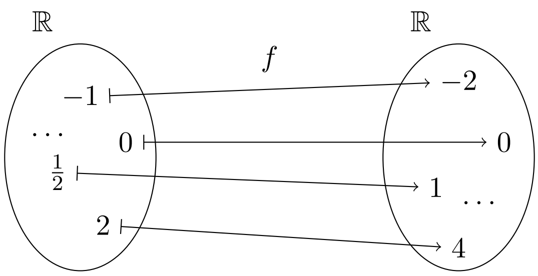
6.2.3 Linear Functions

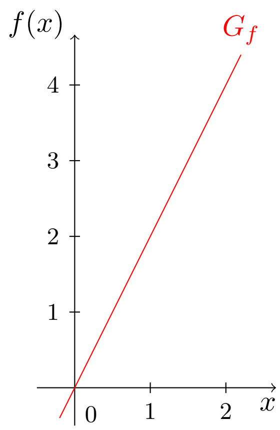
Starting from the identity function, more complex functions - so-called linear functions - can be constructed. So, for example, one can think of a function that assigns to every real number twice its value or π times its value, etc., e.g.

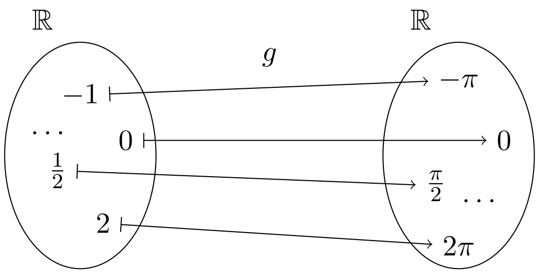
$$f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & 2x \end{cases}$$

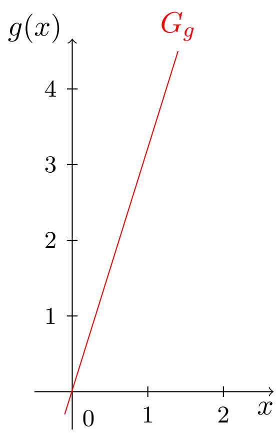
or

$$g : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \pi x\,. \end{cases}$$



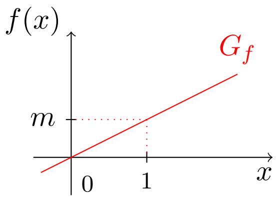






Hence, all linear functions (except for the zero function, see below) also have the entire set of real numbers as their ranges ($W_f, W_g = \mathbb{R}$). The factor that multiplies each real number in such a linear function is called slope of the linear function. Often, also for linear functions one does not like to specify a certain function with a specific slope, but an arbitrary function with an arbitrary slope $m \in \mathbb{R}$:

$$f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & mx \,. \end{cases}$$



Where does the term slope of a linear function come from? If the difference in height by which the graph is rising vertically is divided by the corresponding length in the horizontal direction, then one obtains the slope m. So $m = \frac{f(x_2)-f(x_1)}{x_2-x_1}$ for $x_1 < x_2$.

Info6.2.1

A linear function

$$f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & mx \end{cases}$$

is strictly increasing if and only if its slope is positive, i.e. $m > 0$; and it is strictly decreasing if and only if its slope is negative, i.e. $m < 0$.

Exercise 6.2.1
Which of the linear functions above has the slope $m = 1$?

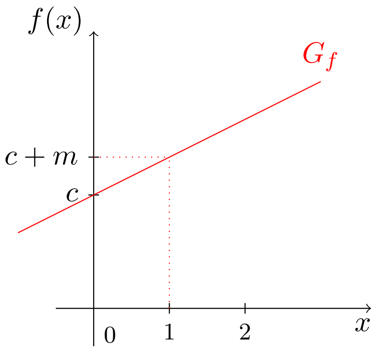
Solution:
We have $f(x) = 1 \cdot x = x = \mathrm{Id}(x)$, i.e. the identity function.

Exercise 6.2.2
Which of the linear functions above has the slope $m = 0$?
Solution:
We have $f(x) = 0 \cdot x = 0$, i.e. the constant function that is always 0.

6.2.4 Linear Affine Functions

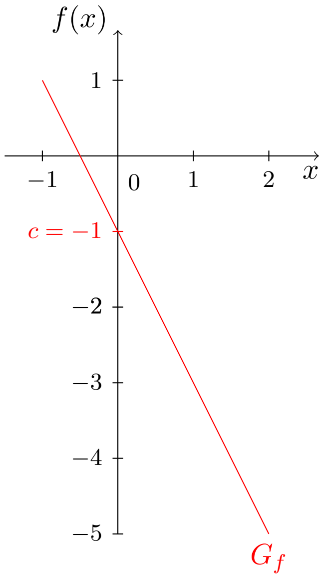
Combining linear functions with constant functions results in so-called linear affine functions. These are the sum of a linear function and a constant function. Generally, without any specification for the slope ($m \in \mathbb{R}$) this is written as follows:

$$f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & mx + c\,. \end{cases}$$



The graphs of linear affine functions are also called lines. For linear affine functions, the constant m is still called slope, and the constant $c \in \mathbb{R}$ is called y-intercept. The reason for this term is as follows: if the intersection point of the graph of the linear affine function with the vertical axis is considered, then this point has the distance c from the origin (see figure above). So, for the linear affine function shown in the figure below

$$f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & -2x - 1 \end{cases}$$



we have the slope $m = -2$ and the y-intercept $c = -1$. The y-intercept is the value of the function at $x = 0$ and hence given by

$$c = f(0) = -2 \cdot 0 - 1 = -1 \ .$$

Exercise 6.2.3
Find the slope and the y-intercept of the function

$$f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \pi x - 42 \ . \end{cases}$$

Solution:
Slope: π and y-intercept: -42

Exercise 6.2.4
Which functions are the linear affine functions that have slope $m = 0$, and which are the ones with y-intercept $c = 0$?
Solution:
If $m = 0$, then $f(x) = 0 \cdot x + c = c$. Hence, the constant functions are the ones that have slope 0. A zero y-intercept, i.e. $c = 0$, implies $f(x) = mx + 0 = mx$. Hence, in this case we obtain exactly the linear functions.

6.2.5 Absolute Value Functions

In Module 2 the absolute value of a real number x was introduced in the following way:

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

In the context of this module, the absolute value can be regarded as a function. This results in the **absolute value function**:

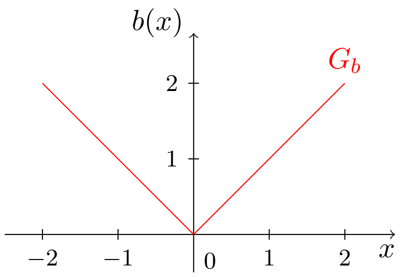
$$b : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & |x| \,. \end{cases}$$

Exercise 6.2.5
What is the range W_b of the absolute value function b?
Solution:
Since $b(x) = |x| \geq 0$ for all numbers x in $D_b = \mathbb{R}$, we have $W_b = [0; \infty)$.

Due to the definition by cases

$$b(x) = |x| = \begin{cases} x & \text{for } x \geq 0 \\ -x & \text{for } x < 0 \,, \end{cases}$$

the absolute value function is an example of a piecewise defined function. If absolute values are defined according to different cases, then it is also said that the absolute value is resolved. Then, the graph of the absolute value function b looks as follows:



One property of the graph of the absolute value function, which most of the more general functions involving absolute values have in common, is the kink at $x = 0$. The absolute value function b defined above is only the simplest case of a function involving an absolute value. More complicated examples of functions can be constructed, involving one or several absolute values, e.g.

$$f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & |2x - 1| \,. \end{cases}$$

For such functions, it is a relevant task to get an idea of how the graph of the function looks. To do this, we use the piecewise definition of the absolute value, and the approach is similar to the one for the solution of absolute value equations and inequalities. Here, we demonstrate this approach for the example of the function f defined above:

Example 6.2.2

Consider the function

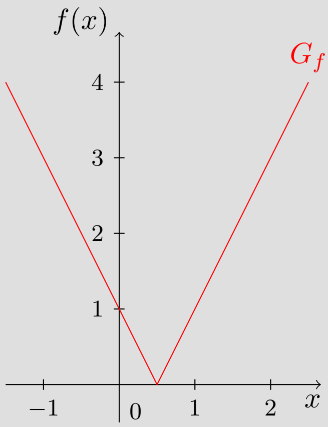
$$f : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & |2x-1| \end{array} \right. .$$

What does the graph look like?

We calculate:

$$f(x) = |2x-1| = \left\{ \begin{array}{rl} 2x-1 & \text{for } 2x-1 \geq 0 \\ -(2x-1) & \text{for } 2x-1 < 0 \end{array} \right. = \left\{ \begin{array}{rl} 2x-1 & \text{for } x \geq \frac{1}{2} \\ -2x+1 & \text{for } x < \frac{1}{2} \end{array} \right. .$$

Thus, we obtain a piecewise defined function whose graph is an increasing line with the slope 2 and the y-intercept -1 for x in the region $x \geq \frac{1}{2}$ and a decreasing line with the slope -2 and the y-intercept 1 for x in the region $x < \frac{1}{2}$. With this information we can draw the graph of f:



Info6.2.3

CAUTION! If absolute values are resolved as in the calculation in the example above, two important calculation rules have to be observed:

1. The regions for the cases are defined by inequalities for the entire expression

between the absolute value bars, here $2x - 1 \geq 0$ and $2x - 1 < 0$, and not only by $x \geq 0$ and $x < 0$. It is always that way if absolute values are resolved.

2. For the case < 0 the entire expression gets a minus sign. Here, care has to be taken that the expression is bracketed. In the example above, we therefore have $-(2x-1) = -2x+1$ and not $-2x-1$. This is always the case if absolute values are resolved.

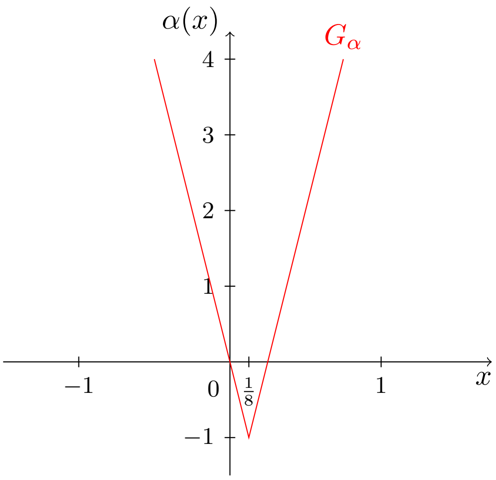
Exercise 6.2.6
Sketch the graph of the function

$$\alpha : \left\{ \begin{array}{ccl} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & |-8x+1| - 1 \; . \end{array} \right.$$

Moreover, specify its range W_α.
Solution:
We have

$$\alpha(x) = |-8x+1|-1 = \left\{ \begin{array}{rl} -8x+1-1 & \text{for } -8x+1 \geq 0 \\ -(-8x+1)-1 & \text{for } -8x+1 < 0 \end{array} \right. = \left\{ \begin{array}{rl} -8x & \text{for } x \leq \frac{1}{8} \\ 8x-2 & \text{for } x > \frac{1}{8} \end{array} \right. ,$$

hence:

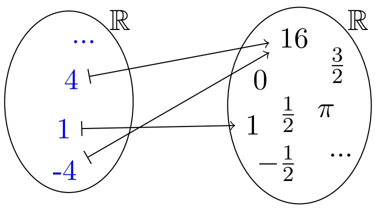


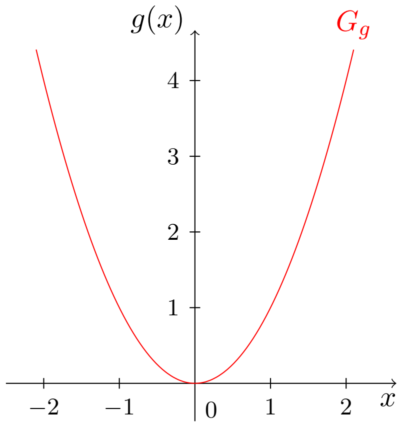
Since $|-8x+1| \geq 0$, it follows $|-8x+1|-1 \geq -1$ which implies $W_\alpha = [-1; \infty)$.

6.2.6 Monomials

In addition to the linear affine functions studied in the previous section, we can also think of functions that assign to every real number a non-negative integer power of the number. An example is the function

$$g : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{array} \right. .$$





This works for every non-negative integer exponent, and generally this function is written as

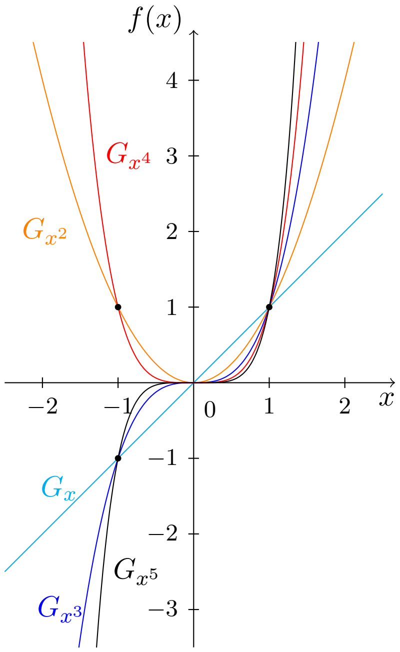
$$f : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^n \end{array} \right.$$

with $n \in \mathbb{N}_0$, and it is called a monomial. The exponent n of a monomial is called the degree of the monomial. For example, the function g described at the beginning of this section is a monomial of degree 2.

Exercise 6.2.7
Which functions are the monomials of degree 1 and 0?

Solution:
Since $x^1 = x$ for all $x \in \mathbb{R}$, the identity function Id is the monomial of degree 1. Likewise, $x^0 = 1$ for all $x \in \mathbb{R}$, and thus, the constant function $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = 1$ is the monomial of degree 0.

The monomial of degree 2 is called the standard parabola. The monomial of degree 3 is called the cubic standard parabola. The figure below shows the graphs of a few monomials.



On the basis of these graphs, we now summarise some conclusions on monomials: There is a fundamental difference between monomials (with the mapping rule $f(x) = x^n$, $n \in \mathbb{N}$) of even and odd degree. The range of monomials of an even non-zero degree is always the set $[0;\infty)$, while monomials of odd degree have the range $\mathbb{R}$. Furthermore, we always have

$$f(1) = 1^n = 1\ ,$$
$$f(0) = 0^n = 0$$

and

$$f(-1) = \begin{cases} 1 & \text{for } n \text{ even} \\ -1 & \text{for } n \text{ odd}\ . \end{cases}$$

Moreover, we have

$$\begin{cases} x > x^2 > x^3 > x^4 > \ldots & \text{for } x \in (0;1) \\ x < x^2 < x^3 < x^4 < \ldots & \text{for } x \in (1;\infty)\ . \end{cases}$$

Exercise 6.2.8
How can our conclusions concerning monomials be seen immediately from the exponent rules?

Solution:

From the exponent rules, we know that $1^n = 1$ and $0^n = 0$ for arbitrary non-negative integers n. We have $(-1)^n = 1$ if n is an even number and $(-1)^n = -1$ if n is an odd number. From this, the described conclusions result for all monomials. $x > x^2 > x^3 > x^4 > \ldots$ for positive x less than 1 and $x < x^2 < x^3 < x^4 < \ldots$ for x greater than 1 result from the exponent rules since higher powers of positive numbers less than 1 will always have decreasing values while, in contrast, higher powers of positive numbers greater than 1 will always have increasing values.

6.2.7 Polynomials and Their Roots

The monomials considered until now always involved only exactly one power of the independent variable. From these monomials we can easily construct more complex functions involving several different powers of the independent variable. These are sums of multiples of monomials. They are called polynomials. A few examples and their graphs are given below.

$$f_1 : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & 2x^3 + 4x^2 - 3x + 42 \end{cases} \qquad \text{(degree: 3)}$$

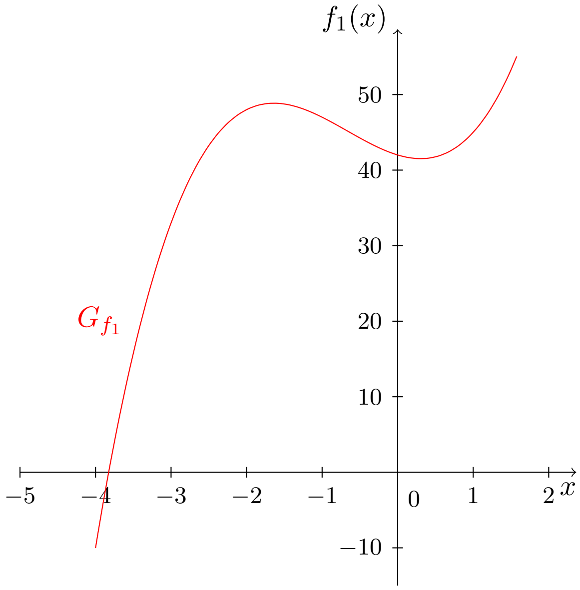
$$f_2 : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & -x^{101} + 3x - 14 \end{cases} \qquad \text{(degree: 101)}$$

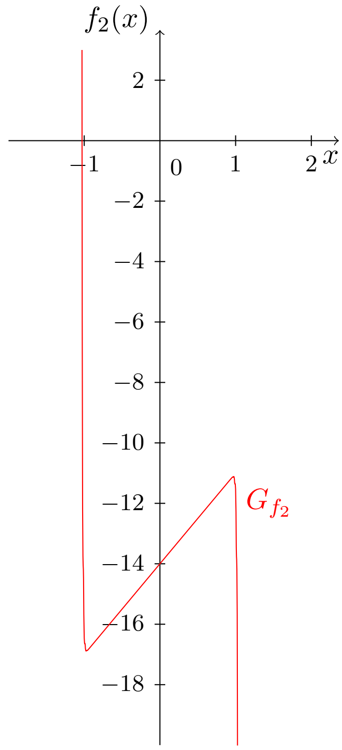
$$f_3 : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & 9x^4 + 9x^3 - 2x^2 - 19x \end{cases} \qquad \text{(degree: 4)}$$

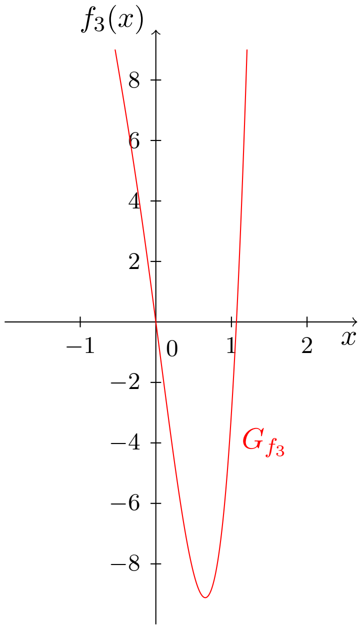
$$f_4 : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 + 2x + 2 \end{cases} \qquad \text{(degree: 2)}$$

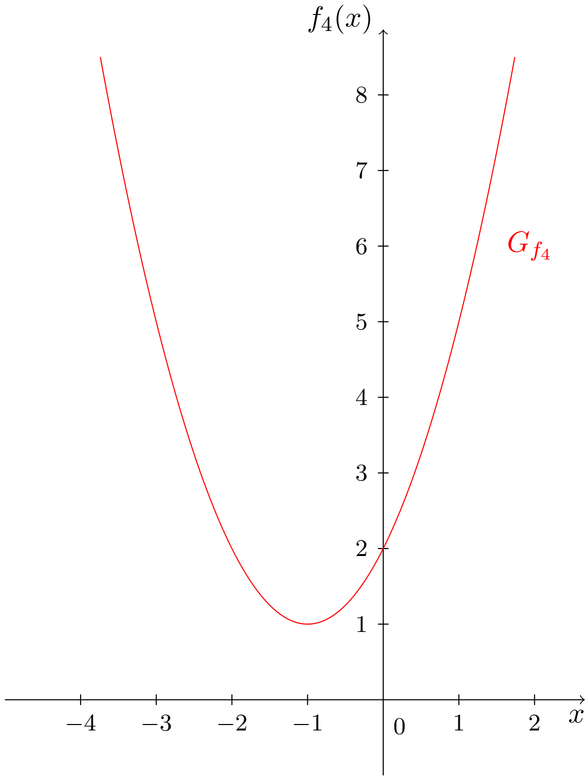
$$f_5 : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & 8x - 2 \end{cases} \qquad \text{(degree: 1)}$$

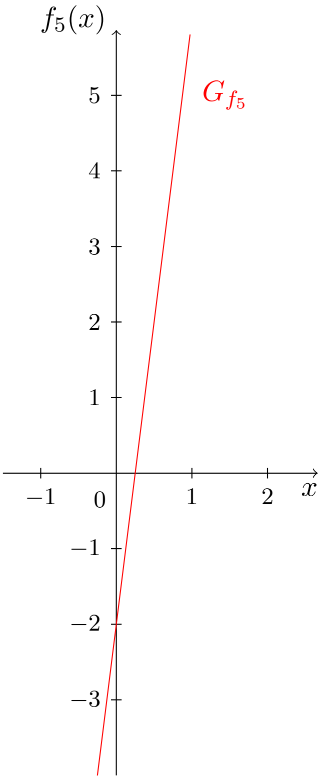
$$f_6 : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & 13 \end{cases} \qquad \text{(degree: 0)}$$

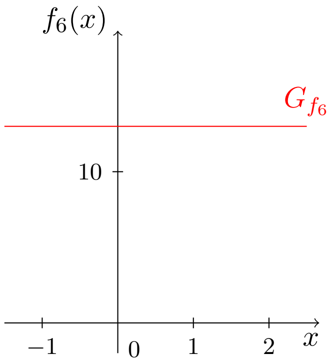

$f_1(x)$
G_{f_1}
50
40
30
20
10
−10
−5
−4
−3
−2
−1
0
1
2
x


$f_2(x)$
G_{f_2}
2
−2
−4
−6
−8
−10
−12
−14
−16
−18
−1
0
1
2
x


$f_3(x)$
G_{f_3}
8
6
4
2
−2
−4
−6
−8
−1
0
1
2
x


$f_4(x)$
G_{f_4}
8
7
6
5
4
3
2
1
−4
−3
−2
−1
0
1
2
x





Obviously, the degree of a polynomial is determined by the monomial with highest degree. Moreover, we see that all types of functions (constant functions, linear functions and linear affine functions) studied so far – as well as the monomials – occur naturally as special cases of polynomials. Thus, the polynomials include all types of functions considered so far.

An unspecific polynomial of degree $n \in \mathbb{N}$ is written as follows:

$$f : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} \cdots + a_2 x^2 + a_1 x + a_0 \end{array} \right. .$$

Here, $a_0, a_1, \ldots, a_n$ with $a_n \neq 0$ are real prefactors of the individual monomials, which are called coefficients of the polynomial.

Exercise 6.2.9
What is the polynomial $f(x)$ with the coefficients $a_0 = -4$, $a_2 = \pi$, and $a_4 = 9$, and what is its range?

The polynomial is $f(x)$ = ______ ,
its range is W_f = ______ .

Solution:
The polynomial is

$$f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & 9x^4 + \pi x^2 - 4 \end{cases},$$

and its range is $W_f = [-4; \infty)$ since the even powers of x can only take non-negative values.

For general polynomials, the roots are of particular interest. The roots of a polynomial can be found by solving equations of nth degree. In the case of degree polynomials of degree $n = 2$ (which are also called general parabolas), this is possible by solving a quadratic equation. In Module 2 the relevant terms and methods, i.e. completing the square, the *pq* formula, and the vertex form of quadratic expressions are explained in detail.

Example 6.2.4

Consider the parabola

$$\zeta : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ y & \longmapsto & 2y^2 - 8y + 6 \end{cases}.$$

We find the roots and the vertex and then sketch the graph.

We complete the square in the mapping rule $\zeta(y) = 2(y^2 - 4y + 3)$:

$$y^2 - 4y + 3 = y^2 - 4y + 4 - 1 = (y - 2)^2 - 1 .$$

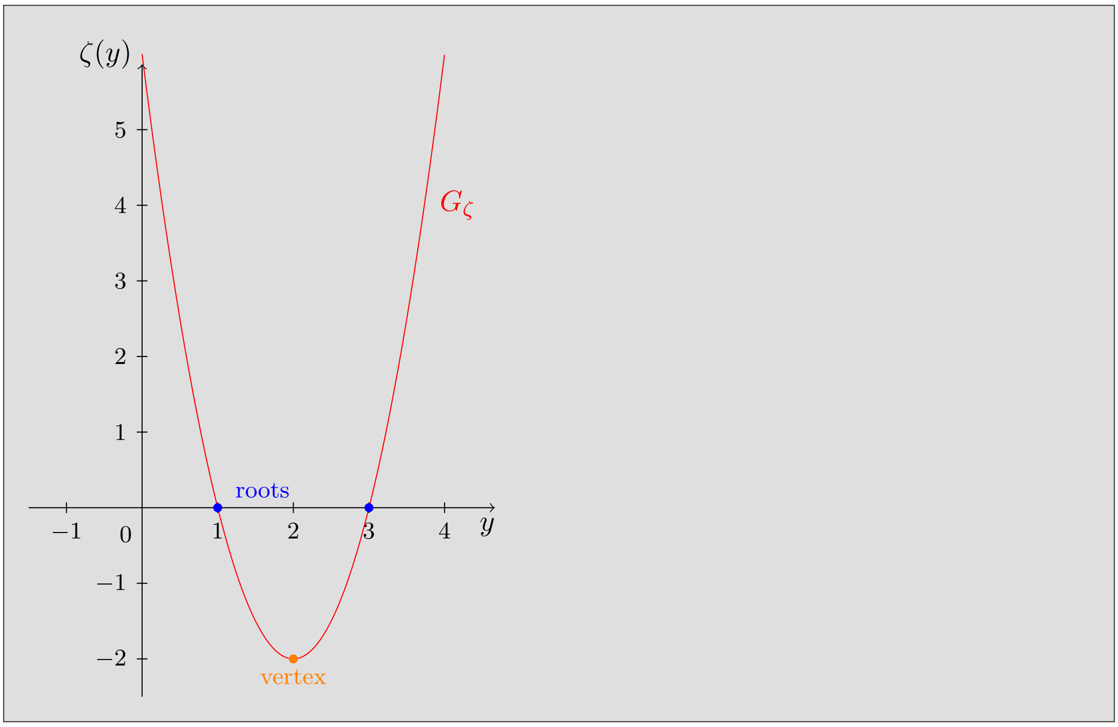
Thus, the mapping rule can be written as

$$\zeta(y) = 2(y - 2)^2 - 2 .$$

We see that the parabola is shifted with respect to the standard parabola by 2 units to the right and 2 units downwards. It can be seen that the vertex is at $(2, -2)$. The roots can be calculated according to:

$$\zeta(y) = 2((y-2)^2-1) = 0 \quad \Leftrightarrow \quad (y-2)^2 = 1 \quad \Leftrightarrow \quad y_{1,2}-2 = \begin{cases} 1 \\ -1 \end{cases} \quad \Leftrightarrow \quad y_{1,2} = \begin{cases} 3 \\ 1 \end{cases}.$$

Finally, the graph of the function is as shown in the figure below.



6.2.8 Hyperbolas

We consider functions which have a reciprocal relation in their mapping rule. For the determination of the maximum domain of such a function, note that the denominator must be non-zero.

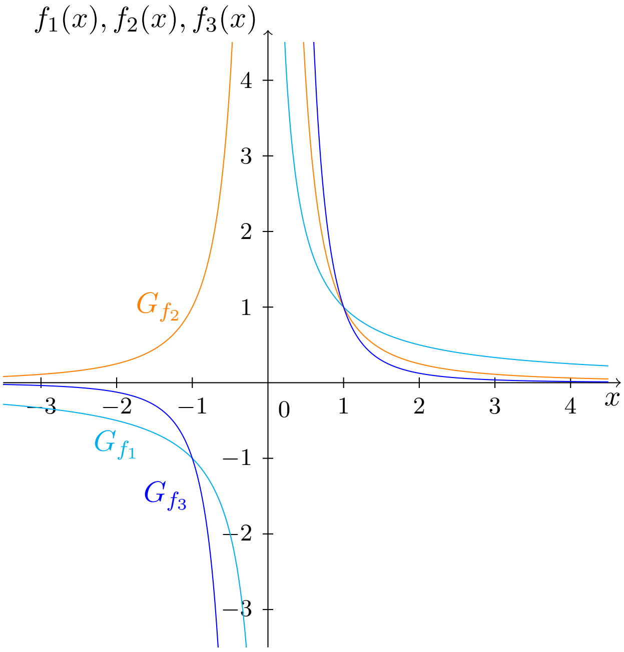
A few examples of reciprocal functions are listed below; these are reciprocals of monomials, and they are also called functions of hyperbolic type.

$$f_1 : \left\{ \begin{array}{rcl} \mathbb{R} \setminus \{0\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{1}{x} \end{array} \right. ,$$

$$f_2 : \left\{ \begin{array}{rcl} \mathbb{R} \setminus \{0\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{1}{x^2} \end{array} \right. ,$$

$$f_3 : \left\{ \begin{array}{rcl} \mathbb{R} \setminus \{0\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{1}{x^3} \end{array} \right. ,$$

etc. Their graphs are as follows.



In particular, the graph of the function

$$f_1 : \begin{cases} \mathbb{R} \setminus \{0\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{1}{x} \end{cases}$$

is called the hyperbola.

Generally, for the reciprocal of an arbitrary monomial of degree $n \in \mathbb{N}$ a corresponding function of hyperbolic type can be specified.

$$f_n : \begin{cases} \mathbb{R} \setminus \{0\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{1}{x^n} \end{cases} .$$

Exercise 6.2.10
What is the range W_{f_n} of the function f_n for even or odd $n \in \mathbb{N}$?

Solution:
We always have $\frac{1}{x^n} \neq 0$ since a quotient can only be zero if the numerator is zero. Thus, the range never contains $0 \in \mathbb{R}$. Since $x^n \geq 0$ for even $n \in \mathbb{N}$, we have $\frac{1}{x^n} > 0$ for even $n \in \mathbb{N}$. However, for odd $n \in \mathbb{N}$ we can also have $\frac{1}{x^n} < 0$. This results in

$$W_{f_n} = \begin{cases} \mathbb{R} \setminus \{0\} & \text{for } n \text{ odd} \\ (0; \infty) & \text{for } n \text{ even} . \end{cases}$$

This can also be seen from the graphs of the functions of hyperbolic type.

Further examples for functions of hyperbolic type were already considered in Example 6.1.8 and in Exercise 6.1.5 in Section 6.1.3.

6.2.9 Rational Functions

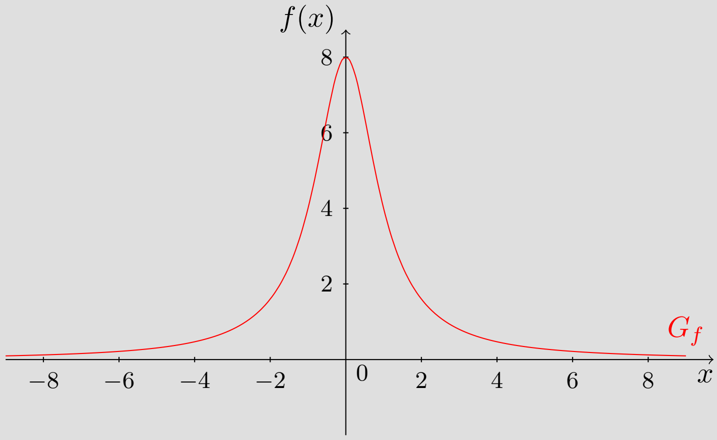
A general rational function has a mapping rule that is the quotient of two polynomials. Some examples with their graphs are given below. Of course, for these functions numbers for which the denominator in the mapping rule equals zero must also be excluded from the domain.

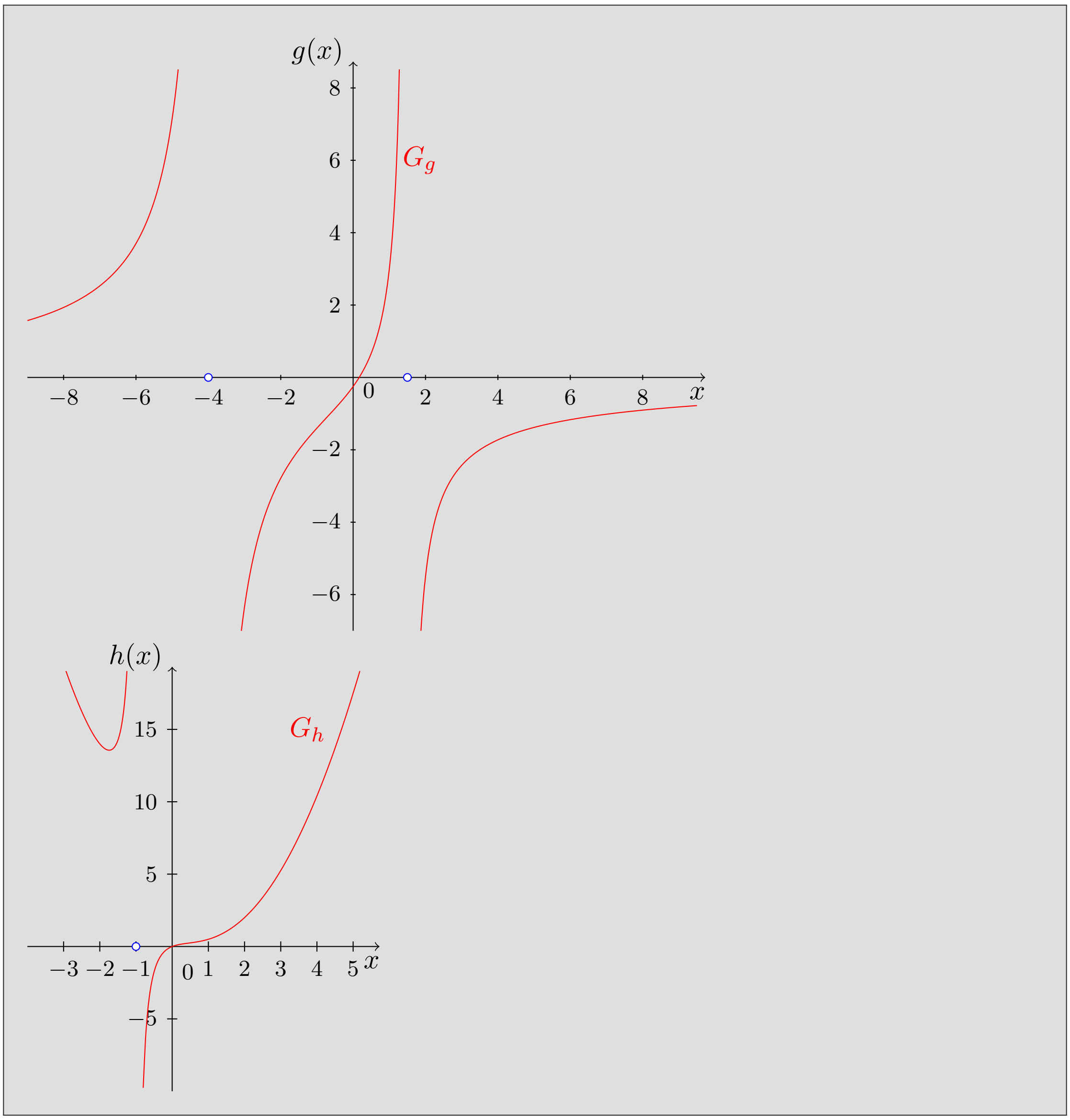
Example 6.2.5

$$f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{8}{x^2+1} \end{cases},$$

$$g : \begin{cases} \mathbb{R} \setminus \{-4, \frac{3}{2}\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{-18x+3}{2x^2+5x-12} \end{cases},$$

$$h : \begin{cases} \mathbb{R} \setminus \{-1\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{x^3-x^2+x}{x+1} \end{cases}.$$





Exercise 6.2.11
For the function

$$\psi : \begin{cases} D_\psi & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{-42x}{x^2-\pi} \end{cases}$$

find the maximum domain $D_\psi \subset \mathbb{R}$ of ψ.

Solution:

The roots of the denominator are

$$x^2 - \pi = 0 \quad \Leftrightarrow \quad x^2 = \pi \quad \Leftrightarrow \quad x = \pm\sqrt{\pi}\,.$$

Thus, we have $D_\psi = \mathbb{R} \setminus \{-\sqrt{\pi}, \sqrt{\pi}\}$.

Exercise 6.2.12

For the rational functions in Example 6.2.5, specify the degree of the polynomials in the numerator and the denominator and find their roots.

Solution:

The function f has the a numerator of degree 0 and denominator of degree 2. The numerator does not have a root ($8 \neq 0$), nor does the denominator ($x^2 + 1 = 0$ has no solution).

The function g has the a numerator of degree 1 and denominator of degree 2. The root of the numerator is at $x = \frac{1}{6}$ ($-18x+3 = 0 \Leftrightarrow x = \frac{3}{18}$), and the roots of the denominator at $x_1 = -4$, $x_2 = \frac{3}{2}$ are obtained by solving the quadratic equation $2x^2 + 5x - 12 = 0$, e.g. by means of the quadratic formula.

The function h has a numerator of degree 3 and denominator of degree 1. The root of the denominator is simply at $x = -1$ ($x + 1 = 0 \Leftrightarrow x = -1$). To find the roots of the numerator, the equation $x^3 - x^2 + x = 0$ has to be solved. By factoring out x one obtains $x(x^2 - x + 1) = 0$, and it can be seen immediately that one root is at $x = 0$. Finally, the quadratic equation $x^2 - x + 1 = 0$ has to be solved by means of the quadratic formula. However, the discriminant $\Delta = 1^2 - 4 \cdot 1 \cdot 1 = -3$ is negative, so no other real solution of the equation – and hence no other root of the numerator – exists.

The roots of a rational function are the roots of the numerator. For example, the function

$$j : \begin{cases} \mathbb{R} \setminus \{-1; 3\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{x-1}{x^2-2x-3} \end{cases}$$

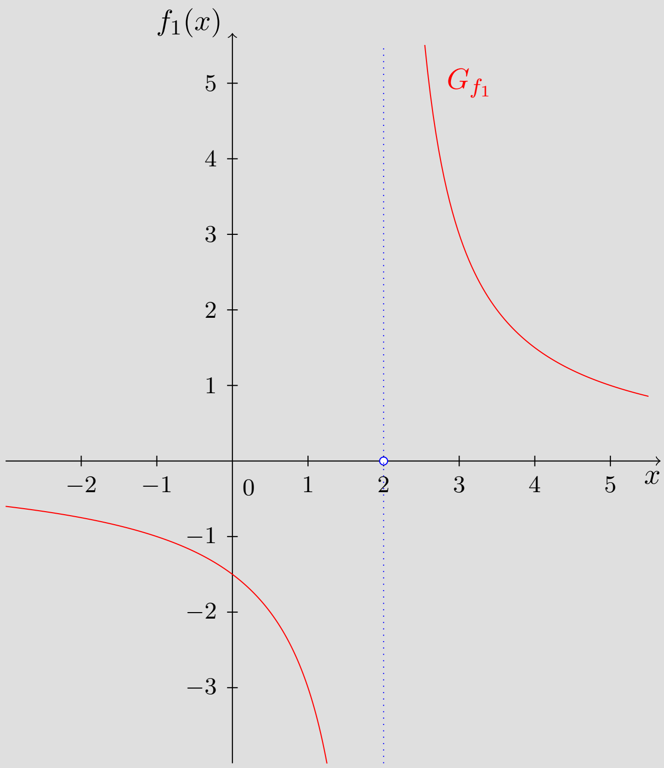
has a single root at $x = 1$. The roots of the denominator of rational functions that must be excluded from the domain often have to be investigated further. It is of particular interest how the graphs of functions behave in the neighbourhood of gaps in the domain. The roots of the denominator are also called **poles**. The next examples will illustrate the different types of poles that can occur.

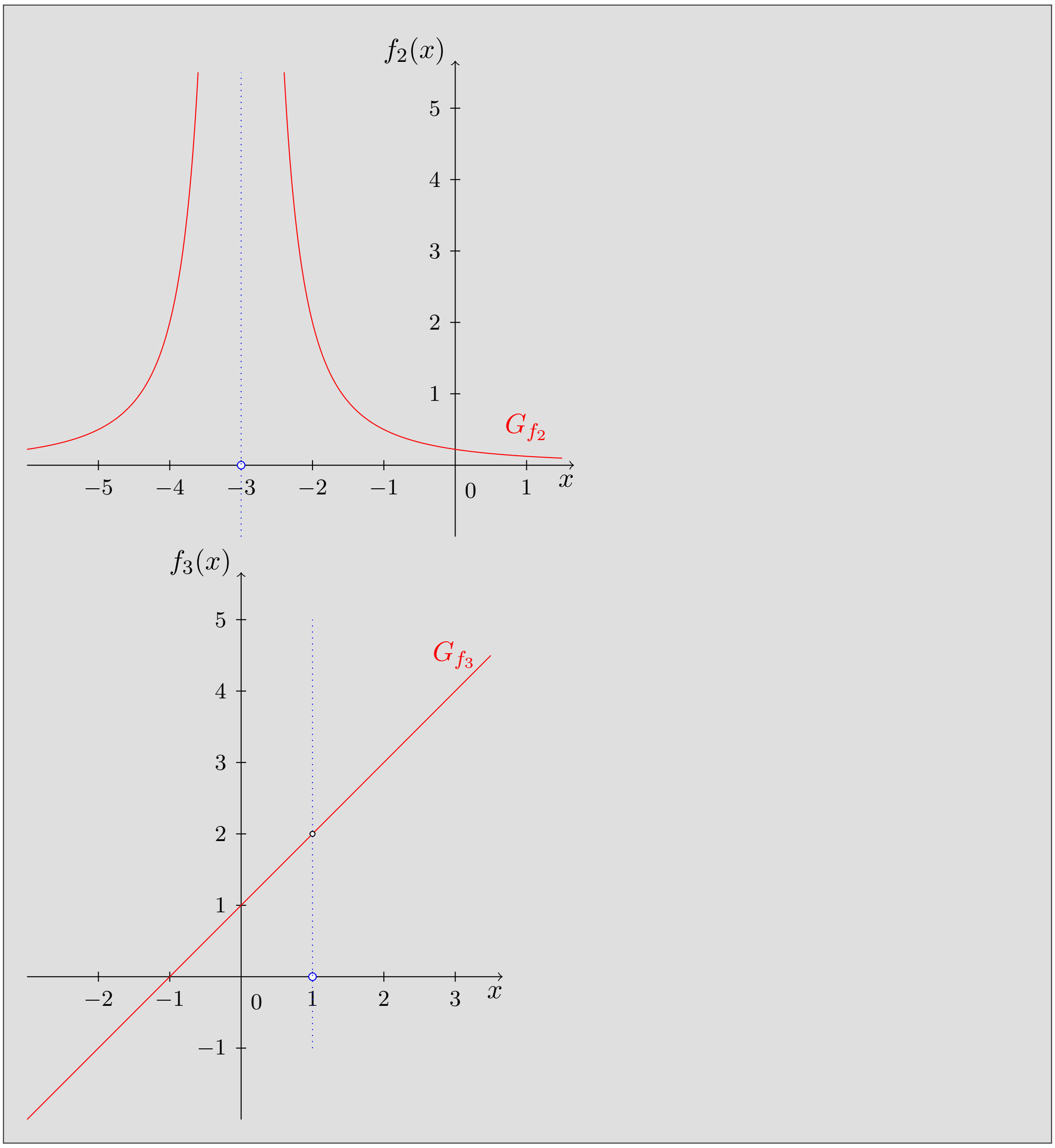
Example 6.2.6

$$f_1 : \begin{cases} \mathbb{R} \setminus \{2\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{3}{x-2} \end{cases}$$

$$f_2 : \begin{cases} \mathbb{R} \setminus \{-3\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{2}{(x+3)^2} \end{cases}$$

$$f_3 : \begin{cases} \mathbb{R} \setminus \{1\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{x^2-1}{x-1} \end{cases}$$





At $x = 2$ and $x = -3$ the functions f_1 and f_2, respectively, have so-called (proper) poles, and at $x = 1$ the function f_3 has a so-called removable singularity. Looking at the graphs, the difference between these types of poles becomes clear. For (proper) poles, the graph rises or falls unboundedly in the neighbourhood of the pole, and for removable singularities the graph ends left and right in the “gap”.

In the mapping rules of the three functions, this difference is expressed as follows: the values $x = 2$ and $x = -3$ are roots of the denominator, but they are not roots of the numerator of the functions f_1 and f_2, respectively. Actually, the functions f_1 and f_2 do not have any roots in the numerator. In such cases, the roots of the denominator are always (proper) poles.

Exercise 6.2.13
Is the denominator's root of the function

$$q : \begin{cases} \mathbb{R} \setminus \{\frac{1}{2}\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{x^4-1}{2x-1} \end{cases}$$

a proper pole? If so, give reasons for your answer.

Solution:
The point $x = \frac{1}{2}$ is a root of the denominator:

$$2x - 1 = 0 \quad \Leftrightarrow \quad 2x = 1 \quad \Leftrightarrow \quad x = \frac{1}{2} \ .$$

However, for the numerator, we have:

$$x^4 - 1 = 0 \quad \Leftrightarrow \quad x^4 = 1 \quad \Leftrightarrow \quad x = \pm 1 \ ,$$

and thus, the roots of the numerator are at $x = -1$ and $x = 1$. Hence, $x = \frac{1}{2}$ is not a root of the numerator. Thus, $x = \frac{1}{2}$ is a proper pole.

Between the two poles of f_1 and f_2 there is a further difference. At the pole $x = 2$ of f_1, the function has a change of sign. The graph of f_1 falls left to the pole unboundedly to minus infinity and rises right to the pole (coming from the right) unboundedly to plus infinity.
The graph of f_2 rises on both sides of the pole at $x = -3$ (while approaching the pole) unboundedly to plus infinity, and hence, there is no change of sign in the function values.

However, in the mapping rule of f_3, the term responsible for the pole at $x = 1$ can be cancelled out. For rational functions that have a removable singularity, this is always possible.

Exercise 6.2.14
Find all poles/singularities of the function

$$\gamma : \begin{cases} D_\gamma & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{3x+6}{x^2-x-6} \end{cases}$$

and determine their type. Specify the maximum domain $D_\gamma \subset \mathbb{R}$ of the function.

Solution:
The roots of the denominator are the solutions of the quadratic equation $x^2 - x - 6 = 0$, thus

$$x_{1,2} = \frac{-(-1) \pm \sqrt{(-1)^2 - 4 \cdot (-6) \cdot 1}}{2} = \frac{1 \pm 5}{2} = \begin{cases} 3 \\ -2 \, . \end{cases}$$

Hence, the maximum domain is

$$D_\gamma = \mathbb{R} \setminus \{-2; 3\} \, .$$

The root of the numerator results from $3x + 6 = 0$, i.e. the root of the numerator is also at $x = -2$. Thus, we can transform the mapping rule of γ for $x \in D_\gamma$ as follows:

$$\gamma(x) = \frac{3x + 6}{x^2 - x - 6} = \frac{3(x + 2)}{(x - 3)(x + 2)} = \frac{3}{x - 3} \, .$$

Hence, the function can also be written in the form

$$\gamma : \begin{cases} \mathbb{R} \setminus \{-2; 3\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{3}{x-3} \, , \end{cases}$$

and thus, the function has a continuously removable singularity at $x = -2$ and a (proper) pole with change of sign at $x = 3$.

6.2.10 Asymptotes

In this section we will study how rational functions behave as x tends to infinity if the degree of the polynomial in the numerator is less than or equal to the degree of the polynomial in the denominator. An example of this is the function

$$f : \begin{cases} \mathbb{R} \setminus \{-\pi\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{x}{x+\pi} \, . \end{cases}$$

In the function f, the degree of the polynomial in the numerator is 1, and degree of the polynomial in the denominator is 1. Other examples were investigated in the previous Section 6.2.9.

Example 6.2.7

Let us consider the function

$$g : \begin{cases} \mathbb{R} \setminus \{1\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & 1 + \frac{1}{x-1} \, . \end{cases}$$

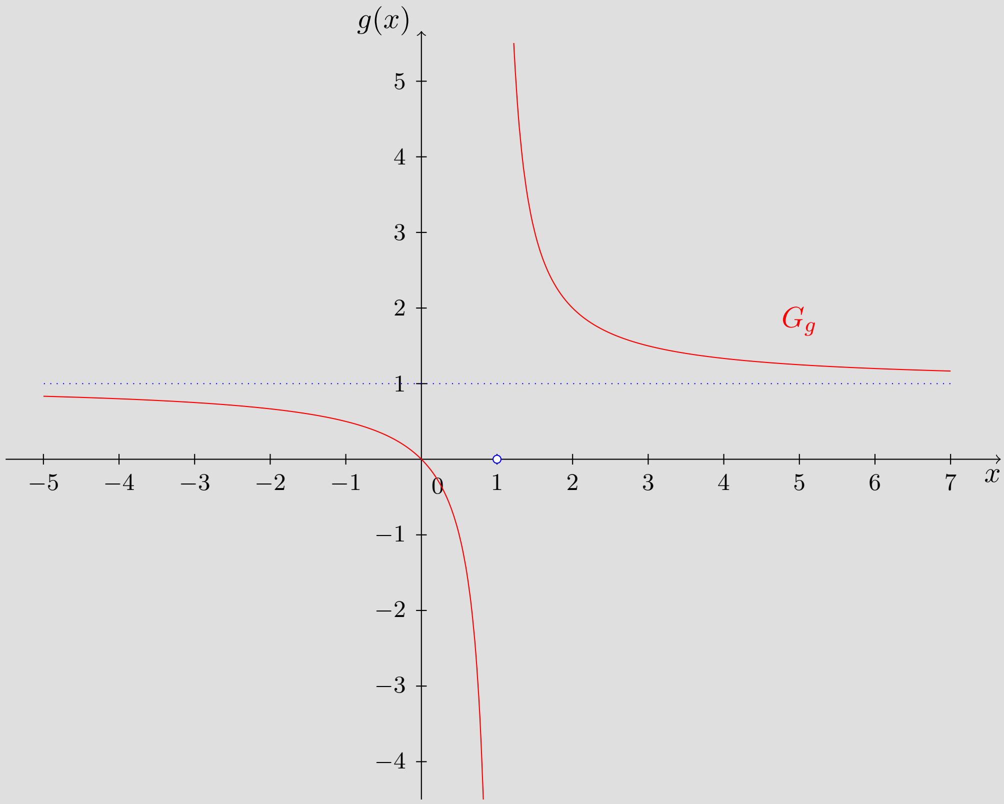
Its mapping rule is a sum of a polynomial (of degree 0) and a rational term. By finding the common denominator it is easy to transform $g(x)$ into a rational form in which the degree of the polynomial in the numerator equals the degree of the polynomial in the denominator:

$$g(x) = 1 + \frac{1}{x-1} = \frac{x-1}{x-1} + \frac{1}{x-1} = \frac{x-1+1}{x-1} = \frac{x}{x-1} \, .$$

Thus, we can rewrite g in the form

$$g : \begin{cases} \mathbb{R} \setminus \{1\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{x}{x-1} \, , \end{cases}$$

and we now consider the corresponding graph.



Besides the pole and the singularity at $x = 1$, we see that the value $y = 1$ is of specific relevance. Obviously, this value is never taken by the function g. Thus, the range of g is $W_g = \mathbb{R} \setminus \{1\}$. Instead, for "very large" values (x tends to plus infinity) and "very small" values (x tends to minus infinity) of the independent variable x, the function g approaches its limiting value 1 indefinitely without ever reaching it for any real number x.

This can be seen from the mapping rule $g(x) = 1 + \frac{1}{x-1}$ as follows. For "very large" (50, 100, 1000, etc.) or "very small" (50, 100, 1000, etc.) values of x, the rational part $\frac{1}{x-1}$ approaches 0 since x occurs in the denominator. In general, for these values of x, only the polynomial part 1 of the mapping rule remains. This part can be described by a – in this case constant – function that is called an asymptote g_{As} of the function g:

$$g_{As} : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & 1 \, . \end{cases}$$

Since this is a constant function, it is also called a horizontal asymptote.

Exercise 6.2.15
Identify the asymptotes of the function

$$i : \begin{cases} \mathbb{R} \setminus \{-2\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & 3 - \frac{6}{x+2} \end{cases}$$

and the asymptote of the hyperbola described in Section 6.2.8.
Solution:
We have

$$i(x) = 3 - \frac{6}{x+2}$$

with the rational part $\frac{6}{x+2}$. Hence, the horizontal asymptote of i has the mapping rule $i_{As}(x) = 3$. The hyperbola

$$f : \begin{cases} \mathbb{R} \setminus \{0\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{1}{x} \end{cases}$$

also has an asymptote. We can write the mapping rule as

$$f(x) = 0 + \frac{1}{x} \, ,$$

thus, we have $f_{As}(x) = 0$ for the asymptote, i.e. the asymptote is the function that is constantly 0: the zero function or the horizontal axis of the coordinate system.

Info6.2.8

A rational function f with the polynomial $p(x)$ of degree $z \geq 0$ in the numerator and the polynomial $q(x)$ of degree $n \geq 0$ in the denominator of the form

$$f : \begin{cases} \mathbb{R} \setminus \{\text{denominator's roots}\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & f(x) = \frac{p(x)}{q(x)} \end{cases}$$

has a constant function (or a horizontal line) as an asymptote if $z \leq n$. In particular, the zero function is the asymptote in the case $z < n$.

6.3 Power Functions

6.3.1 Introduction

In Section 6.2.6 and Section 6.2.8 we studied monomials and functions of hyperbolic type. In summary, these can be described as the following type of functions:

$$f : \begin{cases} D_f & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^k \end{cases},$$

where $k \in \mathbb{Z} \setminus \{0\}$ and $D_f = \mathbb{R}$ for $k \in \mathbb{N}$ as well as $D_f = \mathbb{R} \setminus \{0\}$ for $k \in \mathbb{Z}$ with $k < 0$. In this section, we will allow arbitrary rational for the exponent in the mapping rule. This results in so-called power functions that again include monomials and functions of hyperbolic type as special cases. We will collect their fundamental properties and see some applications.

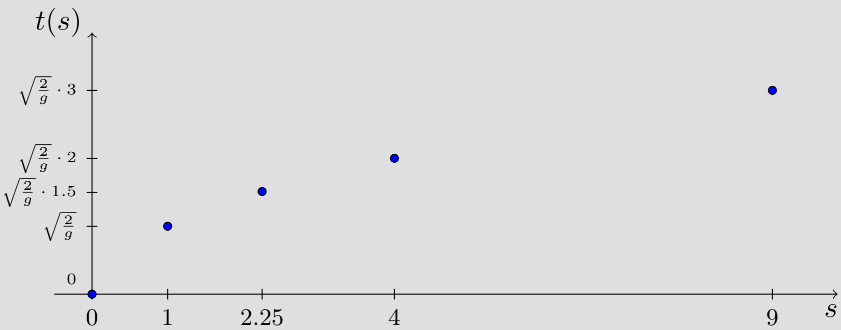
6.3.2 Radical Functions

Example 6.3.1

If an object falling in the homogeneous gravitational field of the Earth is observed, then the following relation between the falling time and the travelled distance can be found:

Falling time t in seconds	0	$\sqrt{\frac{2}{g}}$	$\sqrt{\frac{2}{g}} \cdot 1.5$	$\sqrt{\frac{2}{g}} \cdot 2$	$\sqrt{\frac{2}{g}} \cdot 3$
Travelled distance s in metres	0	1	2.25	4	9

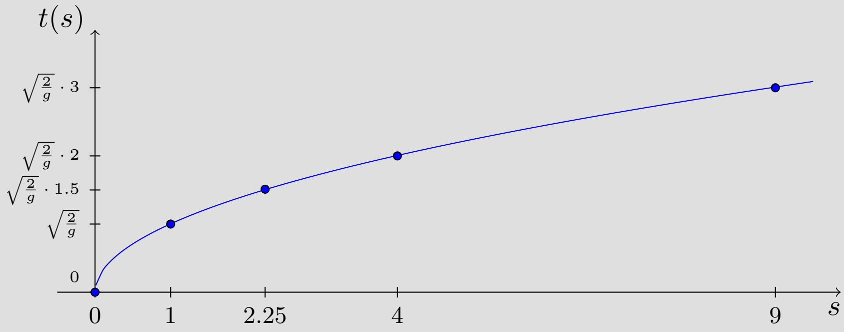
Here, $g \approx 9.81 \frac{\mathrm{m}}{\mathrm{s}^2}$ is the physical constant of the gravitational acceleration. Now, plotting these values in a diagram with the horizontal axis s and the vertical axis t results in the figure below.



This suggests that the relation between t and s can be described mathematically by the function

$$t: \begin{cases} [0;\infty) & \longrightarrow & \mathbb{R} \\ s & \longmapsto & \sqrt{\frac{2}{g}} \cdot \sqrt{s} \end{cases}$$

with s being the independent variable. This is a function, in whose mapping rule a root (more specifically, a square root) of the independent variable occurs. Then, the graph of this function contains the measurement points listed above:



This example shows that functions with mapping rules that contain roots of the independent variables occur naturally in applications of mathematics.

For natural numbers $n \in \mathbb{N}$, $n > 1$, the functions

$$f_n: \begin{cases} D_{f_n} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \sqrt[n]{x} = x^{\frac{1}{n}} \end{cases}$$

are called radical functions. Obviously, these include the square root $f_2(x) = \sqrt{x}$, the cube root $f_3(x) = \sqrt[3]{x}$, the fourth root $f_4(x) = \sqrt[4]{x}$, etc., as mapping rules of functions (see exponent rules).

Exercise 6.3.1
Transform the mapping rule of the radical functions using exponent rules such that only exponents still occur in the mapping rule.

<u>Solution:</u>
According to the exponent rules, we have

$$f_n(x) = \sqrt[n]{x} = x^{\frac{1}{n}}$$

for all natural numbers n. Hence, for example

$$f_2(x) = \sqrt{x} = x^{\frac{1}{2}} \ , \quad f_3(x) = \sqrt[3]{x} = x^{\frac{1}{3}} \ , \quad f_4(x) = \sqrt[4]{x} = x^{\frac{1}{4}} \ , \quad \ldots$$

Exercise 6.3.2
What is the function f_n with $n = 1$?

Solution:

According to the exponent rules, we have for $n = 1$

$$f_1(x) = \sqrt[1]{x} = x^{\frac{1}{1}} = x = \mathrm{Id}(x) \ .$$

This is the identity function. Generally, this function is excluded from the class of radical functions.

Of great relevance is now the maximum domain D_{f_n} that a radical function can have. Obviously, it depends on the exponent n of the root which values of x are allowed to be inserted in the mapping rule to obtain real values as a result. So we see that the square root $\sqrt{\ }$ has a real value as a result only for a non-negative number. However, if we consider the cube root $\sqrt[3]{\ }$, then we see that the cube root has a real value as a result for all real numbers, for example, $\sqrt[3]{-27} = -3$. Generally, we have:

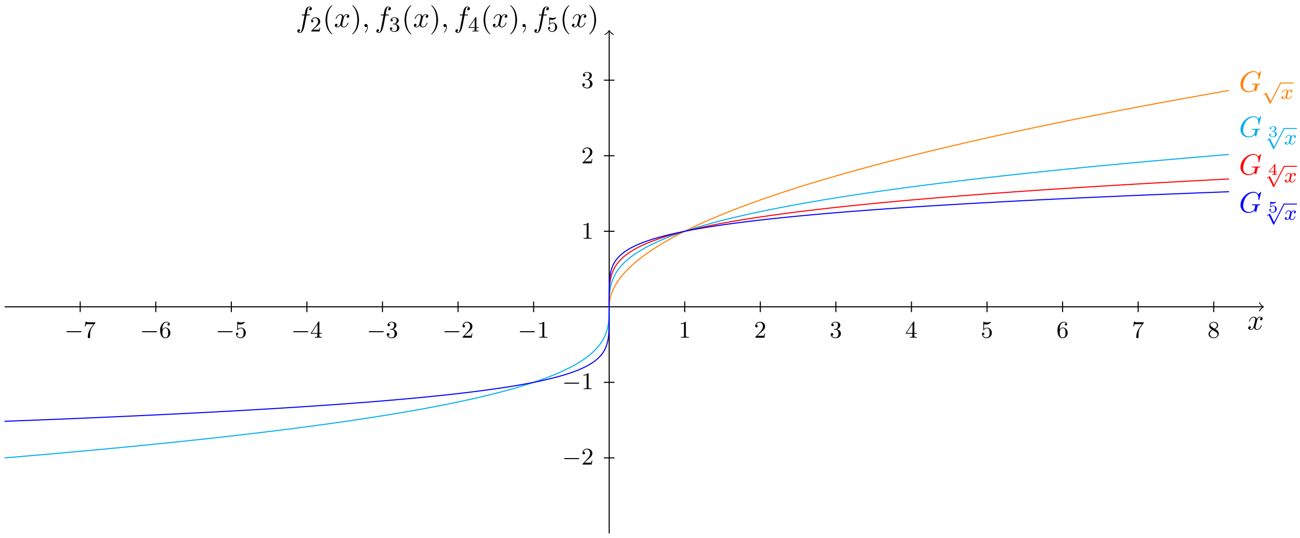
Info6.3.2

The radical functions

$$f_n : \begin{cases} D_{f_n} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \sqrt[n]{x} \end{cases}$$

with $n \in \mathbb{N}$ and $n > 1$ have the maximum domain $D_{f_n} = [0; \infty)$ if n is even and the maximum domain $D_{f_n} = \mathbb{R}$ if n is odd.

Thus, the graphs of the first four radical functions, f_2, f_3, f_4, f_5, look like as in the figure below.



From the graphs, it can be seen that all radical functions are strictly increasing.

Exercise 6.3.3
For the radical functions

$$f_n : \begin{cases} D_{f_n} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \sqrt[n]{x} \end{cases}$$

with $n \in \mathbb{N}$, $n > 1$, find the range W_{f_n} depending on whether n is even or odd.

Solution:
For radical functions with n even, obviously only non-negative numbers can occur as results since, according to the exponent rules, the roots $\sqrt{x}$, $\sqrt[4]{x}$, $\sqrt[6]{x}$ are always non-negative for $x \geq 0$. In contrast, for radical functions with n odd, all negative real numbers can occur as results. In fact, we have $\sqrt[3]{x} < 0$, $\sqrt[5]{x} < 0$ if and only if $x < 0$. In summary, we then have, regarding the strict monotonicity of the radical functions, $W_{f_n} = \mathbb{R}$ if n is odd, and $W_{f_n} = [0; \infty)$ if n is even.

6.4 Exponential and Logarithmic Functions

6.4.1 Introduction

In contrast to power functions, in **exponential functions** the independent variable does not occur in the base of the exponential term but in the exponent. Accordingly, we will consider mapping rules such as

$$x \longmapsto 2^x \text{ or } x \longmapsto 10^x .$$

Exponential functions are relevant in many different fields, for example, for the description of biological growth processes – including diverse population models –, the processes of radioactive decay, or a certain kind of interest calculation. Let us consider an example.

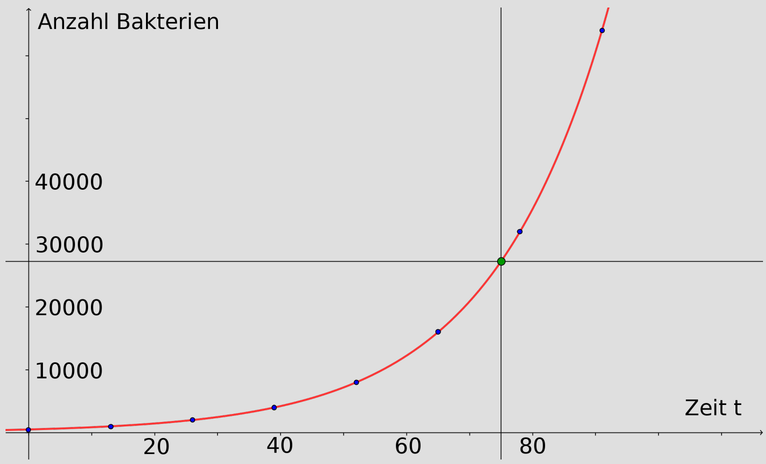
Example 6.4.1

A bacterial culture starts with 500 bacteria and doubles in size every 13 minutes. We would like to know how many bacteria will be in the culture after 1 hour and 15 minutes (i.e. after 75 minutes).

As a first try we can create a simple table of values that lists the number of bacteria in the population at the beginning ($t = 0$ min), after $t = 13$ min, after $t = 26$ min, etc., i.e. at multiples of the 13 minutes duplication time:

Time t in min	0	13	26	39	52	65	78	91	etc.
Number of Bacteria	500	1 000	2 000	4 000	8 000	16 000	32 000	64 000	etc.

From the table we can estimate that the answer to our question will be between 16 000 and 32 000, probably closer to 32 000. What about a (more) precise answer? For this, we need to know the functional relation between the values of t and the number of bacteria. In the figure below the graph of a function p is shown; so to speak, this graph closes the gaps between the isolated points that correspond to the pairs of values in the table and are plotted as well. The corresponding mapping rule assigns to every real valued point in time a number of bacteria. As we will see, the corresponding function is an exponential function.



From the graphical representation, the required number of bacteria can be read off a bit more precisely. However, for the exact specification we need the mapping rule underlying the graph, which we will first simply state here:

$$p : [0; \infty) \longrightarrow (0; \infty) \text{ with } t \longmapsto p(t) = 500 \cdot 2^{(t/13)} .$$

(In Exercise 6.4.2 auf Seite 266 we will give a reason for this functional relation.) For $t = 75$ (measured in minutes) we obtain the function value

$$p(75) = 500 \cdot 2^{(75/13)} \approx 500 \cdot 54.539545 \approx 27\,270 .$$

Hence, after 75 minutes approx. 27 270 bacteria live in the considered population.

6.4.2 Contents

In the previous example, an exponential function with base $a = 2$ occurs, and the independent variable – in this example this is the variable t – occurs in the exponent. We will now specify the general mapping rule for an exponential function with an arbitrary base a; however, we here assume $a > 0$:

$$f : \begin{cases} \mathbb{R} & \longrightarrow & (0; \infty) \\ x & \longmapsto & f(x) = f_0 \cdot a^{\lambda x} \end{cases}$$

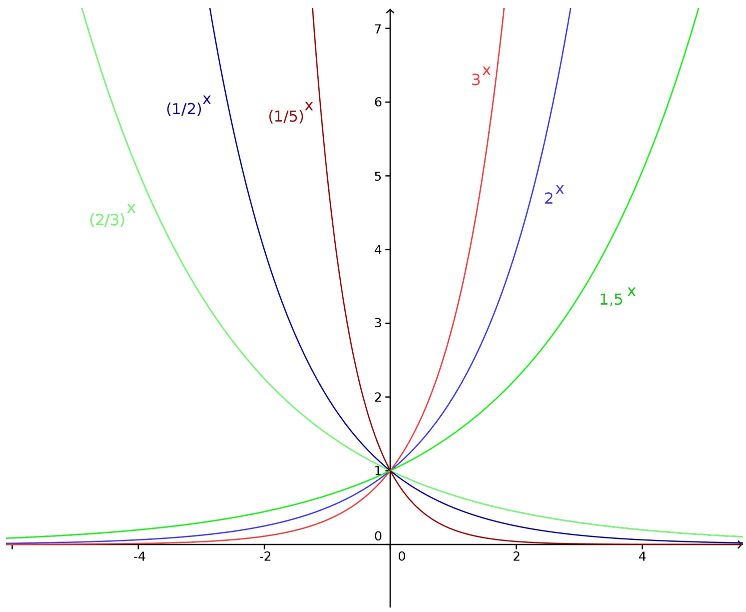
Here, f_0 and λ denote so-called parameters of the exponential function that will be introduced below.

The domain of all exponential functions is the set of all real numbers, i.e. $D_f = \mathbb{R}$, whereas the range only consists of the positive real numbers, i.e. $W_f = (0; \infty)$, since every power of a positive number can only be positive.

Exercise 6.4.1
Why it is assumed that the base a of the exponential function is greater than zero?

Solution:
An exponential function will be defined not only for certain, specific, or isolated values of the variable x but, if possible, for all real numbers. If negative bases $a < 0$ were allowed, then problems would immediately arise in extracting roots – referring to $a^{(1/2)} = \sqrt{a}, a^{1/4}, a^{1/12}$, etc. For example, square roots of negative numbers are not defined, see Section 6.3 auf Seite 259.

Some general properties can be seen from the figure below showing exponential functions $g : \mathbb{R} \to (0; \infty)$, $x \longmapsto g(x) = a^x$ for different values of a:



- All these exponential functions pass trough the point $(x = 0, y = 1)$, since $g(x = 0) = a^0$ and $a^0 = 1$ for every number a.
- If $a > 1$, then the graph of g rises from left to right (i.e. for increasing x-values); one also says that the function g is strictly increasing. The greater the value of a, the steeper the graph of g rises for positive values of x. Moving towards ever larger negative values of x (i.e. approaching from right to left) the negative x-axis is an asymptote of the graph.

- If $a < 1$, then the graph of g falls from left to right (i.e. for increasing x-values); one also says that the function d is strictly decreasing. The greater the value of a, the slower the graph of g falls for negative values of x. Moving towards ever larger positive values of x (i.e. approaching from left to right) the positive x-axis is an asymptote of the graph.

What are the parameters f_0 and λ? The parameter f_0 is easily explained: if the value $x = 0$ is inserted in the general exponential function f, resulting in

$$f(x = 0) = f_0 \cdot a^{\lambda \cdot 0} = f_0 \cdot a^0 = f_0 \cdot 1 = f_0 \ ,$$

then it can be seen that f_0 is a kind of starting point or **initial value** (at least if the variable x is taken for a time); the exponential progression $a^{\lambda x}$ is generally multiplied by the factor f_0 and thus weighted accordingly, i.e. stretched (for $|f_0| > 1$) or compressed (for $|f_0| < 1$).

The parameter λ that occurs in the exponent is called **growth rate**; it determines how strong the exponential function – with the same base – increases (for $\lambda > 0$) or decreases (for $\lambda < 0$). The expression $a^{\lambda x}$ is called growth factor.

Exercise 6.4.2
Explain the form of the exponential function $f(t) = 500 \cdot 2^{(t/13)}$ that occurs in Example 6.4.1 auf Seite 263.

Solution:
In every duplication period of 13 minutes – as the name suggests – the population of bacteria is doubled. Compared to the initial value (500 bacteria), the number of bacteria doubles after a period of 13 minutes, quadruples after two such periods, increases by a factor of eight after $3 \cdot 13$ minutes (always compared to the initial value, as already said), etc. From that fact we see that a growth process involves powers of 2; accordingly, we take as the base of the functional relation $a = 2$.

This consideration also determines the exponent of the required exponential function: Our measurement of time has to refer to the duplication period of 13 minutes, hence, the exponent is $\frac{t}{13}$. After 13 minutes, the exponent has the value $\frac{13}{13} = 1$. Then, the growth factor is $2^{(13/13)} = 2$. After two duplication periods (26 minutes), the exponent has the value $\frac{26}{13} = 2$, and thus the growth factor is altogether $2^{(26/13)} = 2^2 = 4$, etc.

Finally, we have to weight our exponential function $2^{(t/13)}$ by the correct initial value (500 bacteria); this is done by the factor 500.

6.4.3 The Natural Exponential Function

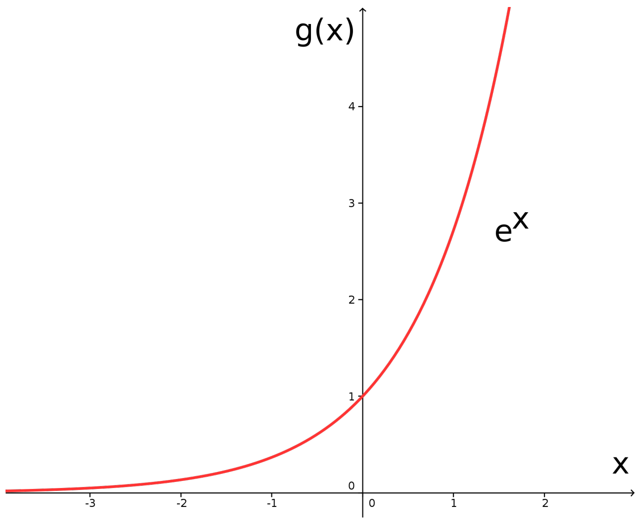
There is a very special exponential function, sometimes also called *the* exponential function, that we will study now. In fact, all other exponential functions can be reduced to this special exponential function. It has **Euler's number** e as its base. Its value is (approximately) equal to

$$e = 2.718281828459045235\ldots\ .$$

So, let us consider the graph of the exponential function – for the time being without any additional parameters –

$$g : \begin{cases} \mathbb{R} & \longrightarrow & (0;\infty) \\ x & \longmapsto & g(x) = e^x \end{cases}$$

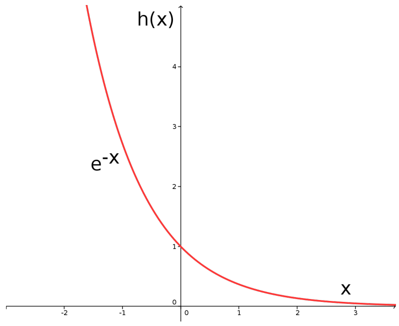
which is, because of its base e, also called the e **function** or **natural exponential function**:



Unsurprisingly, the natural exponential function shows the typical behaviour of exponential functions $x \mapsto a^x$ $(a > 1)$ already discussed in Section 6.4.2 auf Seite 264, after all we have only chosen a special value for the base, namely $a = e$. In particular, we note again that the natural exponential function is strictly increasing, i.e. for large negative values of x, it approaches the negative x-axis, and for $x = 0$, it takes the value 1.

Exercise 6.4.3
What does the graph of the function $h : \mathbb{R} \to (0;\infty)$, $x \mapsto h(x) = e^{-x}$ look like, and which general properties has this function?

Solution:



The function h is strictly decreasing, for large positive values of x the graph of h approaches the positive x-axis, and for $x = 0$ the function h takes the value $h(x = 0) = 1$.

At the beginning of this subsection we claimed that the exponential functions described above can be reduced to the natural exponential function. This is done by means of the identity

$$a^x = e^{x \cdot \ln(a)}$$

that is valid for any real number $a > 0$ and any real number x. Here, ln denotes the natural logarithmic function that will be studied in detail in the following Section 6.4.4 auf Seite 270.

Exercise 6.4.4
Explain why the identity $a^x = e^{x \cdot \ln(a)}$ is valid.
Solution:
According to the exponent rule $(b^r)^s = b^{r \cdot s}$ the right hand side of the identity in question can be rewritten as $e^{x \cdot \ln(a)} = (e^{\ln(a)})^x$. Since the natural logarithmic function ln is the inverse of the natural exponential function, we have $e^{\ln(a)} = a$. This implies $(e^{\ln(a)})^x = a^x$ which is indeed the left hand side of the identity.

In general natural exponential functions, the parameters f_0 and λ occur that were already introduced in Section 6.4.2 auf Seite 264; thus, its functional description is as follows:

$$f : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & (0; \infty) \\ x & \longmapsto & f(x) = f_0 \cdot e^{\lambda x} \end{array} \right. .$$

Again, the parameter f_0 describes initial values different from 1, and the factor λ in the exponent allows for different (positive or negative) growth rates. This shall be finally illustrated by means of an example.

Example 6.4.2

A series of experiments with radioactive iodine atoms (^{131}I) results in the following mean data:

Number of Iodine Atoms	10 000	5 000	2 500	1 250	etc.
Number of Days Elapsed	0	8.04	16.08	24.12	etc.

In other words: every 8.04 days the number of iodine atoms halves due to radioactive decay. For this reason one says in this context that the half-life h of ^{131}I equals 8.04 days.

The radioactive decay follows an exponential law:

$$N(t) = N_0 \cdot e^{\lambda t} \ .$$

Our exponential function is here denoted by N; it describes the number of remaining iodine atoms. Accordingly, N_0 denotes the number of iodine atoms at the beginning, i.e. $N_0 = 10000$. The independent variable is in this case the time t (measured in days). We expect the parameter λ to be negative since the exponential function describes a decay process, i.e. a process with a negative growth rate. We will determine λ from the measurement data.

After $h = 8.04$ days only 5000 iodine atoms are still present, i.e. $N(t = 8.04) = 5000 = \frac{N_0}{2}$. Using the exponential law for the radioactive decay, we obtain:

$$\frac{N_0}{2} = N_0 \cdot e^{\lambda \cdot h} \ .$$

Now, we can cancel N_0 on both sides of the equation and subsequently take the natural logarithm of the equation (see Section 6.4.4 auf der nächsten Seite):

$$\ln\left(\frac{1}{2}\right) = \ln(e^{\lambda \cdot h}) \ .$$

We transform the left hand side according to the calculation rules for logarithmic functions (see Section 6.4.4 auf der nächsten Seite): $\ln(1/2) = \ln(1) - \ln(2) = 0 - \ln(2) = -\ln(2)$. For the right hand side we note that the natural logarithmic function is the inverse function of the natural exponential function, i.e. $\ln(e^{\lambda \cdot h}) = \lambda \cdot h$; thus we have:

$$\begin{aligned} -\ln(2) &= \lambda \cdot h \\ \Leftrightarrow \qquad \lambda &= -\frac{\ln(2)}{h} \ . \end{aligned}$$

Inserting the half-life $h = 8.04$ days of ^{131}I results in this case in

$$\lambda \approx -0.0862 \, \frac{1}{\text{day}} \ .$$

Other radioactive substances have different half-lifes, e.g. ^{239}Pu has a half-life of 24 000 years, and hence they result in different values of the parameter λ in the exponential law for the radioactive decay.

6.4.4 Logarithmic Function

In Section 6.4.3 auf Seite 267 we studied the natural exponential function

$$g : \begin{cases} \mathbb{R} & \longrightarrow & (0;\infty) \\ x & \longmapsto & g(x) = e^x \end{cases}$$

In particular, we pointed out a very important property of the natural exponential function: it is strictly increasing. If the graph of this function is reflected about the angle bisector between the first and the third quadrant (see Chapter 9), one obtains the graph of the natural logarithmic function that has the symbol ln:

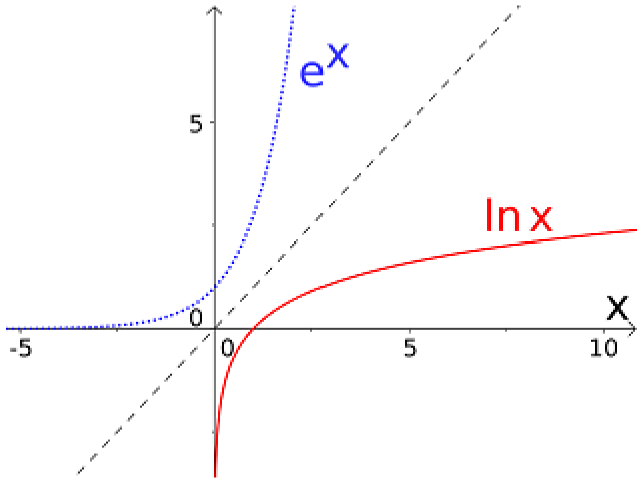
Info6.4.3

The function defined by the equation $e^{\ln(x)} = x$

$$\ln : \begin{cases} (0;\infty) & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \ln(x) \end{cases}$$

is called the **natural logarithmic function**.

Here, the equation shall be read in such a way that $\ln(x) = a$ is just the value a with $e^a = x$. The construction mentioned above is shown in the figure below.



The following properties of the natural logarithmic function can be seen from the graph:

- The function ln is strictly increasing.
- Approaching zero from the right on the x-axis, $\ln(x)$ takes ever larger negative values: We note that the graph of ln gets arbitrarily close to the negative vertical axis (y-axis).
- At the point $x = 1$ the natural logarithmic function takes the value 0, i.e. $\ln(1) = 0$.

As well as the natural logarithmic function there are other logarithmic functions, which each correspond to a certain exponent.

Info6.4.4

If $b > 0$ is an arbitrary exponent, then the function defined by the equation $b^{\log_b(x)} = x$ (read as: "$\log_b(x) = a$ is the exponent a with $b^a = x$")

$$\log_b : \begin{cases} (0;\infty) & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \log_b(x) \end{cases}$$

is called the **general logarithmic function** with base b.

Normally, the logarithmic function cannot be calculated directly. Since it is defined as the inverse function of the exponential function, one generally tries to rewrite its input as a power and reads off the exponent.

Example 6.4.5

Typical calculations for the natural logarithmic function are

$$\ln(e^5) \;=\; 5\;,\;\; \ln(\sqrt{e}) \;=\; \ln(e^{\frac{1}{2}}) \;=\; \frac{1}{2}$$

and for the general logarithmic function:

$$\log_5(25) \;=\; \log_5(5^2) \;=\; 2\;,\;\; \log_3(81) \;=\; \log_3(3^4) \;=\; 4\;.$$

Here, the base of the logarithmic function has to be observed, for example, we have

$$\log_2(64) \;=\; \log_2(2^6) \;=\; 6\;,\;\; \text{but} \;\; \log_4(64) \;=\; \log_4(4^3) \;=\; 3\;.$$

Exercise 6.4.5
Calculate the values of the following logarithmic functions:

a. $\ln(\sqrt[3]{e}) \;=$ ☐ .
Solution:
We have $\ln(\sqrt[3]{e}) = \ln(e^{\frac{1}{3}}) = \frac{1}{3}$.

b. $\log_2(256) \;=$ ☐ .
Solution:
We have $\log_2(256) = \log_2(2^8) = 8$.

c. $\log_9(3) \;=$ ☐ .
Solution:
We have $\log_9(3) = \log_9(9^{\frac{1}{2}}) = \frac{1}{2}$.

In mathematics and in the sciences the following logarithmic functions are frequently used and thus have their own dedicated names:

- Logarithmic function to the base 10: denoted by $\log_{10}(x) = \lg(x)$ or sometimes only by $\log(x)$. This logarithmic function is associated to the powers of ten and is used, for example, in chemistry for the calculation of the pH level.
- Logarithmic function to the base 2: denoted by $\log_2(x) = \operatorname{ld}(x)$. This logarithmic function is relevant in computer science.

- Logarithmic function to the base e: denoted by $\log_e(x) = \ln(e)$. The natural logarithmic function is mostly inadequate for practical calculations (unless the expression is a power of e). It is called natural since the exponential function with base e, from a mathematical point of view, is simpler than the general exponential function (e.g. because e^x is its own derivative, but b^x for $b \neq e$ is not).

There are several calculation rules for the logarithmic function, which will be explained in the next section.

6.4.5 Logarithm Rules

For calculations involving logarithmic functions certain rules apply that can be derived form the exponent rules.

Info6.4.6

The following rules are called **logarithm rules**:

$$\begin{aligned} \log(u \cdot v) &= \log(u) + \log(v) \quad (u, v > 0)\ , \\ \log\left(\frac{u}{v}\right) &= \log(u) - \log(v) \quad (u, v > 0)\ , \\ \log(u^x) &= x \cdot \log(u) \quad (u > 0, x \in \mathbb{R})\ . \end{aligned}$$

These rules do not only apply to natural logarithmic functions but also to all other logarithmic functions. They can be used to transform a given expression in such a way that the power occurs only in the logarithmic terms.

Example 6.4.7

For example, the value $\mathrm{ld}(4^5)$ can be calculated applying the logarithm rules:

$$\mathrm{ld}(8^5) = \log_2(8^5) = 5 \cdot \log_2(8) = 5 \cdot \log_2(2^3) = 5 \cdot 3 = 15\ .$$

Products in logarithmic functions can be split into sums outside the logarithmic functions:

$$\lg\left(100 \cdot \sqrt{10} \cdot \frac{1}{10}\right) = \lg(100) + \lg(\sqrt{10}) - \lg(10) = 2 + \frac{1}{2} - 1 = \frac{3}{2}\ .$$

Importantly, the splitting rule $\log(u \cdot v) = \log(u) + \log(v)$ transforms products into sums. The other way round is impossible for logarithmic functions: the logarithm of a sum cannot be transformed any further.

6.5 Trigonometric Functions

6.5.1 Introduction

Trigonometry, according to its origin in the Greek words trigonon for "triangle" and metron for "measure", is the measuring theory of (angles and sides of) triangles. In this theory, the **trigonometric functions** sine function, cosine function and tangent function, play a central role.

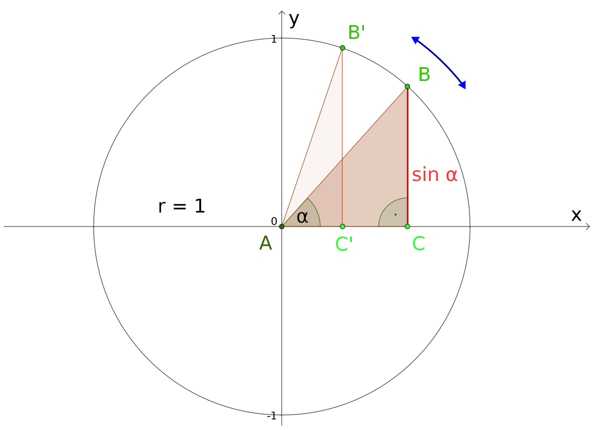
However, the field of applications of sine, cosine, tangent etc is not restricted to "simple" triangle calculations. In fact, the trigonometric functions show their real potential in the manifold fields of applications. The most relevant of those are probably in the description of oscillation processes and wave phenomena in physics and engineering. However, they are also applied in many other fields as, for example, in geodesy and astronomy.

6.5.2 Sine Function

In Module 5 auf Seite 146 the trigonometric functions were introduced in Section 5.6 auf Seite 201 on right triangles, for example by the relation

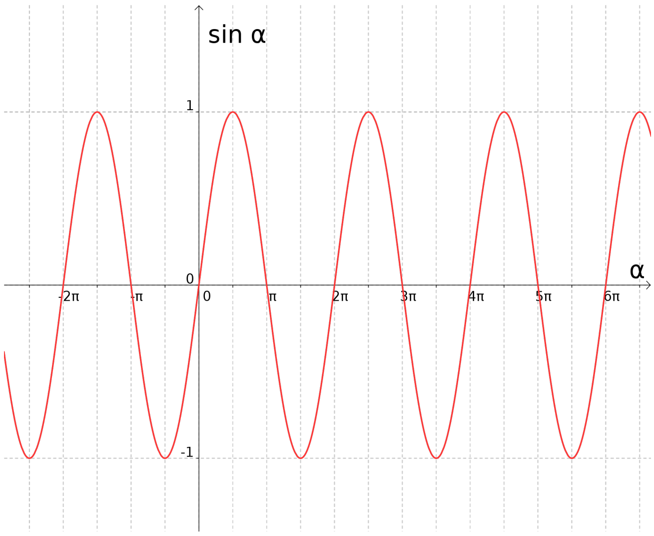
$$\sin(\alpha) = \frac{\text{opposite side}}{\text{hypotenuse}} ,$$

and explained using the unit circle. Starting from this definition of $\sin(\alpha)$ we arrive at the **sine function** if we declare the angle α to be the independent variable of a function named sin. This can be illustrated by means of a family of right triangles ABC that are inscribed the **unit circle**, i.e. a circle with radius $r = 1$, in a certain way.



If we start with the angle $\alpha = 0°$, corresponding to a triangle degenerated to a line segment, then the length of the line segment $\overline{BC}$ equals 0. If we now rotate the point

B counterclockwise around the circle, then the angle α – and $\sin(\alpha)$ as well – increases until for $\alpha = 90°$ a maximum value ($\sin(90°) = 1$) is reached. Afterwards, the angle α continues to increase while $\sin(\alpha)$ is starting to decrease again. For $\alpha = 180°$, the triangle ABC is again degenerated to a line segment and $\sin(180°) = 0$. If α further increases, the triangle "flips downwards" and the line segment $\overline{BC}$ is oriented parallel to the negative vertical line (y-axis), hence, its length is negative. For $\alpha = 270°$ the maximum negative value occurs, before α approaches 0 again. At $\alpha = 360°$ the game starts again.



The figure above shows the graph of the sine function

$$sin : \begin{cases} \mathbb{R} & \longrightarrow & [-1;+1] \\ \alpha & \longmapsto & \sin(\alpha) \end{cases} .$$

In contrast to the discussion so far, the angle α is plotted on the horizontal axis (α-axis) in radian measure, used more commonly in this context, and not in degree measure.

Let us specify some of the most relevant properties of the sine function:

- The sine function is defined on the entire set of real numbers $\mathbb{R}$. Hence, $D_{\sin} = \mathbb{R}$. In contrast, the range only consists of the interval between -1 and $+1$, including these two endpoints: $W_{\sin} = [-1;+1]$.
- After certain measures, the graph of the since function repeats its shape exactly; in this context this is called the periodicity of the sine function. The **period** of the sine function is 360° or 2π. Mathematically, this relation can be expressed as

$$\sin(\alpha) = \sin(\alpha + 2\pi) .$$

Just a quick look at the graph of the simple sine function suggests that we could use this function for the description of wave phenomena. However, to be able to use the full potential of the sine function a few parameters will be introduced. For example, the amplitude of the sine function can be amplified or damped by a so-called amplitude factor A, the frequency of the oscillation can be affected by a frequency-like factor a, and the entire path of the graph can be shifted to the left or to the right by a shifting constant b. Thus, the general sine function has the following form:

$$f : \begin{cases} \mathbb{R} & \longrightarrow & [-A;+A] \\ x & \longmapsto & f(x) = A\sin(ax+b) \end{cases} .$$

Example 6.5.1

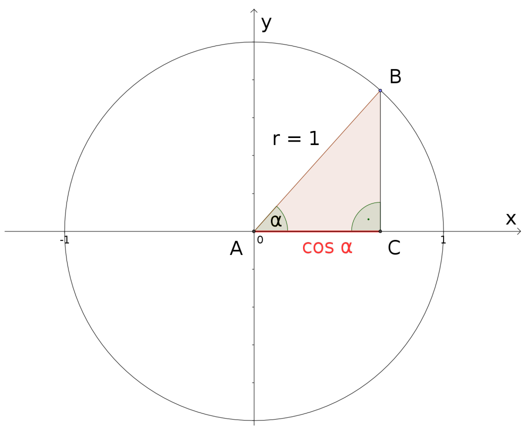
Let us consider a simple pendulum. A small, heavy weight swings freely in the gravitational field of the Earth at a long, very thin cord that is, for example, fixed at the ceiling of a (high) room. Under certain idealised conditions and for small values of the displacement angle φ from the rest position (the vertical axis), the relation between the angle φ and the independent variable t, i.e. the time, is described by a general sine function:

$$\varphi(t) = A\ \sin(\frac{2\pi}{T}t + b)\ .$$

Here, T denotes the so-called period of the pendulum, i.e. the period of time, required by the pendulum for a full oscillation.

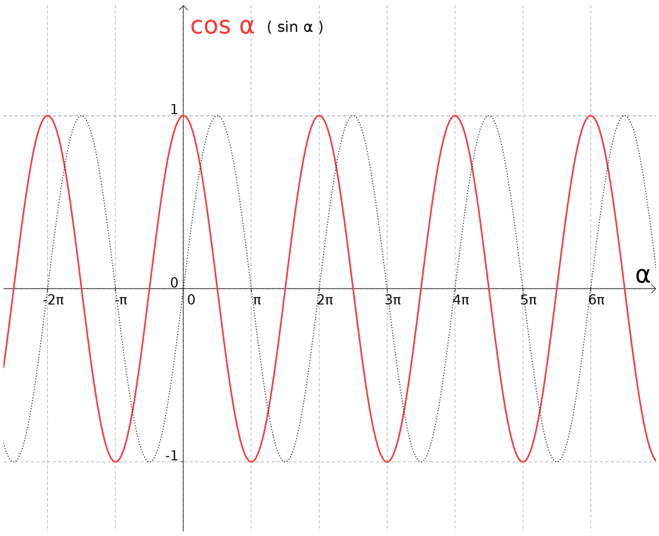
6.5.3 Cosine and Tangent Function

Essentially, for the cosine function and the tangent function we have to do the same considerations as for the sine function that we already know from Subsection 6.5.2 auf Seite 275. As we have some experience, we can shorten the discussion a bit. We start with the **cosine function** and consider again our triangles inscribed the unit circle.



Again, all hypotenuses of these right triangles have the length 1, such that the cosines of the angles α occur as the lengths of the line segments $\overline{AC}$ in the figure. If we uniformly rotate the point B counterclockwise around the circle, varying the angle α, we finally obtain the cosine function:

$$\cos : \begin{cases} \mathbb{R} & \longrightarrow & [-1;+1] \\ \alpha & \longmapsto & \cos(\alpha) \end{cases} .$$



The figure above shows the graph of the cosine function(red line) and the graph of the sine function (grey line) side by side for comparison. We see a very strong relationship, which we will discuss later.

What are the relevant properties of the cosine function?

- The cosine function is also a periodic function. The period is again 2π or $360°$.
- The domain of the cosine function consists of the entire set of the real numbers. Hence, $D_{\cos} = \mathbb{R}$. The range is the interval between -1 and $+1$, including the endpoints: $W_{\cos} = [-1;+1]$.
- From the figure above, showing $\cos(\alpha)$ and $\sin(\alpha)$, we immediately see that

$$\cos(\alpha) = \sin\left(\alpha + \frac{\pi}{2}\right)$$

for all real values of α. Also true, but not as obvious, is the relation

$$\cos(\alpha) = -\sin\left(\alpha - \frac{\pi}{2}\right) .$$

Exercise 6.5.1
At which points does the cosine function take its maximum positive value 1, and at which points does it take its maximum negative value -1? What are its roots (at which points is the function 0)?
Solution:
We have $\cos(0) = 1$. Because of the periodicity with period 2π this is also true for all α equal to $\pm 2\pi, \pm 2 \cdot 2\pi, \pm 3 \cdot 2\pi, \ldots$. Hence, the cosine function takes its maximum value 1 for all integer multiples of 2π (or all even multiples of π). This can also be written as:

$$\cos(\alpha) = 1 \Leftrightarrow \alpha \in \{2k \cdot \pi \,:\, k \in \mathbb{Z}\} \,.$$

The value -1 is taken by the cosine function at the points $\ldots, -3\pi, -\pi, \pi, 3\pi, 5\pi, \ldots$, i.e. for all odd multiples of π:

$$\cos(\alpha) = -1 \Leftrightarrow \alpha \in \{(2k+1) \cdot \pi \,:\, k \in \mathbb{Z}\} \,.$$

Roots occur at $\ldots, -\frac{3}{2}\pi, -\frac{1}{2}\pi, \frac{1}{2}\pi, \frac{3}{2}\pi, \ldots$, i.e. at half-integer multiples of π:

$$\cos(\alpha) = 0 \Leftrightarrow \alpha \in \{\tfrac{2k+1}{2} \cdot \pi; k \in \mathbb{Z}\} \,.$$

As in the case of the sine, the cosine has also a general cosine function. In its definition additional degrees of freedom occur in form of parameters (amplitude factor B, frequency factor c, and shifting constant d). In this way, it is possible to fit the function's graph to different situations (in application examples):

$$g : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & [-A; +A] \\ x & \longmapsto & B\cos(cx+d) \end{array} \right. \,.$$

Exercise 6.5.2
In Example 6.5.1 auf Seite 277 we briefly discussed the simple pendulum. In particular, the displacement angle φ of the pendulum can be determined as a function of time under the condition that the period T equals π seconds and that the pendulum at $t = 0$ is started with an initial displacement angle of 30°:

$$\varphi(t) = \frac{\pi}{6} \cdot \sin\left(2t + \frac{\pi}{2}\right) \,.$$

Can this situation also be described using the (general) cosine function (instead of the sine function), and if so, what form does $\varphi(t)$ take in this case?

Solution:
The answer to the first question is: "Yes, it is possible to describe the present situation using the cosine function" (as we will see in a moment).

In principle, we could start with the general cosine function g given above and determine the parameters B, c and d for the present situation using the same consideration as in Example 6.5.1 auf Seite 277. However, it is easier to apply the relation $\cos(\alpha) = \sin(\alpha + \frac{\pi}{2})$ between cosine and sine functions since then, it immediately follows that

$$\sin(2t + \frac{\pi}{2}) = \cos(2t) ,$$

and thus:

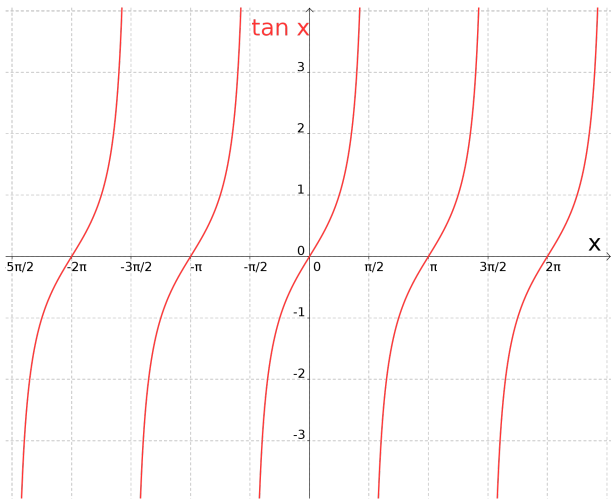
$$\varphi(t) = \frac{\pi}{6} \cdot \cos(2t) .$$

The tangent is the ratio of sine to cosine: $tan(\alpha) = \frac{\sin(\alpha)}{\cos(\alpha)}$. Thus, it follows immediately that the **tangent function** cannot be defined for all real numbers since finally the cosine function has an infinite number of roots. This can be seen, for example, in Exercise 6.5.1 auf Seite 278. In Exercise 6.5.1 auf Seite 278 also the positions of the roots of the cosine function are determined, namely $\cos(\alpha) = 0 \Leftrightarrow \alpha \in \{\frac{2k+1}{2} \cdot \pi; k \in \mathbb{Z}\}$. Thus, the domain of the tangent function is $D_{\tan} = \mathbb{R} \setminus \{\frac{2k+1}{2} \cdot \pi; k \in \mathbb{Z}\}$.

And what about the range? At the roots of the cosine function the tangent function tends to infinite positive or negative values and has a pole, and at the root of the sine function the ratio of sine and cosine is zero. In between, all values can be taken by the tangent function, and hence $W_{\tan} = \mathbb{R}$. All in all, for the tangent function we have

$$tan : \begin{cases} \mathbb{R} \setminus \{\frac{2k+1}{2} \cdot \pi; k \in \mathbb{Z}\} & \longrightarrow & \mathbb{R} \\ \alpha & \longmapsto & \tan(\alpha) \end{cases}$$

The graph of the function is shown in the figure below.



In addition, the tangent function is periodic, however, the period is π or 180°.

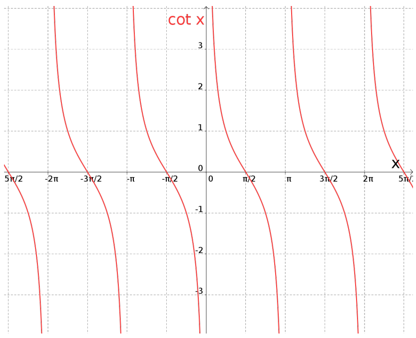
Exercise 6.5.3
The so-called cotangent function (abbreviated to cot) is defined by $\cot(\alpha) = \frac{1}{\tan(\alpha)} = \dfrac{\cos(\alpha)}{\sin(\alpha)}$.

Specify the domain and the range of the cotangent function.

<u>Solution:</u>
The poles of the cotangent functions are at the points where the sine function is 0. This is the case if α is an integer multiple of π. Thus, we have to exclude these points in the definition of the cotangent function:

$$D_{\text{cot}} = \mathbb{R} \setminus \{k \cdot \pi \,:\, k \in \mathbb{Z}\} \,.$$

For the determination of the range, the considerations are very similar to the ones for the range of the tangent function. We thus have $W_{\text{cot}} = \mathbb{R}$.



6.6 Properties and Construction of Elementary Functions

6.6.1 Introduction

In this section, we will consider a further property of elementary functions that we have not discussed yet in previous sections: the symmetry of functions. Furthermore, we will investigate how to construct new functions from known elementary functions. For this purpose sums, products and compositions of functions are introduced.

6.6.2 Symmetry

Info6.6.1

A function $f\colon \mathbb{R} \longrightarrow \mathbb{R}$ is called even or axially symmetric if, for all $x \in \mathbb{R}$, we have

$$f(x) = f(-x)\ .$$

Analogously, a function is called odd or centrally symmetric if, for all $x \in \mathbb{R}$, we have

$$f(x) = -f(-x)\ .$$

These two symmetry conditions for functions allow us to make conclusions about the behaviour of the graphs. For even functions, a reflection across the vertical axis does not change the graph, and for odd functions, a reflection across the origin does not change the graph. A few illustrative examples are listed below.

Example 6.6.2

- The functions

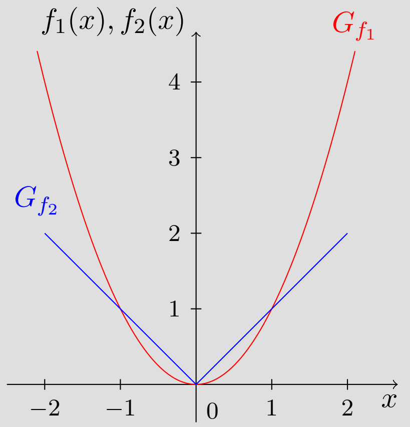
  $$f_1 : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{array} \right.$$

  and

  $$f_2 : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & |x|\ , \end{array} \right.$$

  i.e. the standard parabola (see Section 6.2.6) and the absolute value function (see Section 6.2.5), are examples of even functions. We have $f_1(-x) = (-x)^2 =$

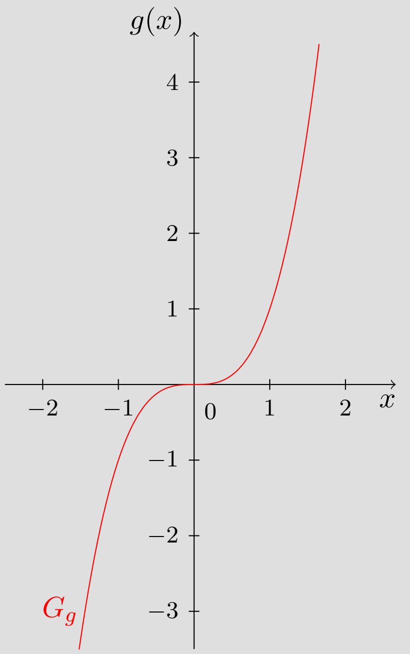
$x^2 = f_1(x)$ and $f_2(-x) = |-x| = |x| = f_2(x)$ for all $x \in \mathbb{R}$. The graphs of these two functions are symmetric under reflection across the vertical axis.



- The function

$$g : \begin{cases} \mathbb{R} \longrightarrow \mathbb{R} \\ x \longmapsto x^3 \end{cases},$$

i.e. the cubic parabola (see Section 6.2.6), is an example of an odd function. We have $g(-x) = (-x)^3 = -x^3 = -g(x)$ for all $x \in \mathbb{R}$. The graph of the function is centrally symmetric with respect to the origin.



Of course, the symmetry properties of functions can also be used if the domain of the function is not the entire set of real numbers. However, then the domain must contain the number 0 in the middle of the interval. An example of this case is the tangent function in the exercise below.

Exercise 6.6.1
Specify whether the following functions are even, odd or non-symmetric.

a)
$$f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & e^x \end{cases}$$

b)
$$g : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ y & \longmapsto & \sin(y) \end{cases}$$

c)
$$h : \begin{cases} (-\frac{\pi}{2}; \frac{\pi}{2}) & \longrightarrow & \mathbb{R} \\ \alpha & \longmapsto & \tan(\alpha) \end{cases}$$

d)
$$i : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ u & \longmapsto & \cos(u) \end{cases}$$

e)
$$j : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & 42 \end{cases}$$

Solution:
a) non-symmetric, b) odd, c) odd, d) even, e) even

6.6.3 Sums, Products, Compositions

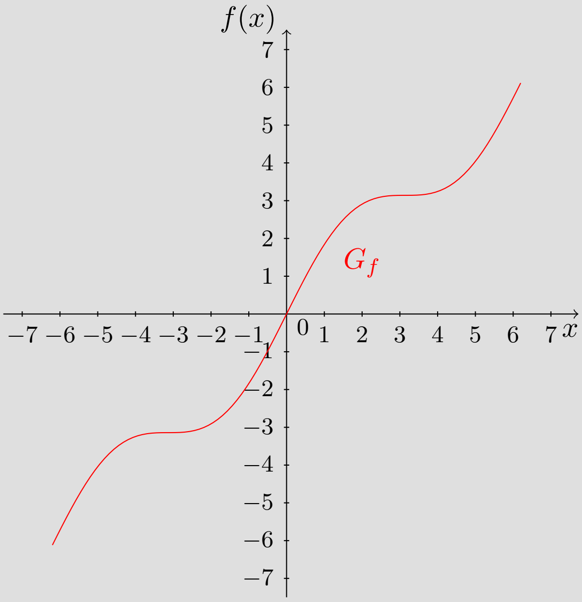
In this section, we will now use the large collection of elementary functions that we acquired in this module to create new more complex functions out of this elementary functions. At different places throughout this module we already studied functions those mapping rules were composed of sums and products of simpler mapping rules. Of course, one can also take differences and, under certain conditions, quotients from mapping rule. The example below lists a few such combined functions.

Example 6.6.3

- The function

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

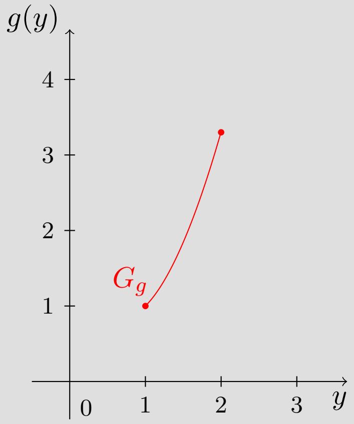
is the sum of the identity function (see Section 6.2.3) and the sine function (see Section 6.5). The graph of this function is shown in the figure below.



- The function

$$g : \begin{cases} [1;2] & \longrightarrow & \mathbb{R} \\ y & \longmapsto & y^2 - \ln(y) \end{cases}$$

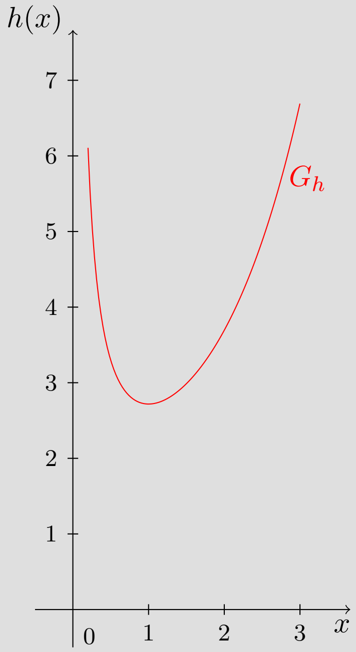
is the difference of the standard parabola (see Section 6.2.6) and the natural logarithmic function (see Section 6.4.4). The graph of this function is shown in the figure below.



- The function

$$h : \begin{cases} (0;\infty) & \longrightarrow & \mathbb{R} \\ x & \longmapsto & e^x \frac{1}{x} \end{cases}$$

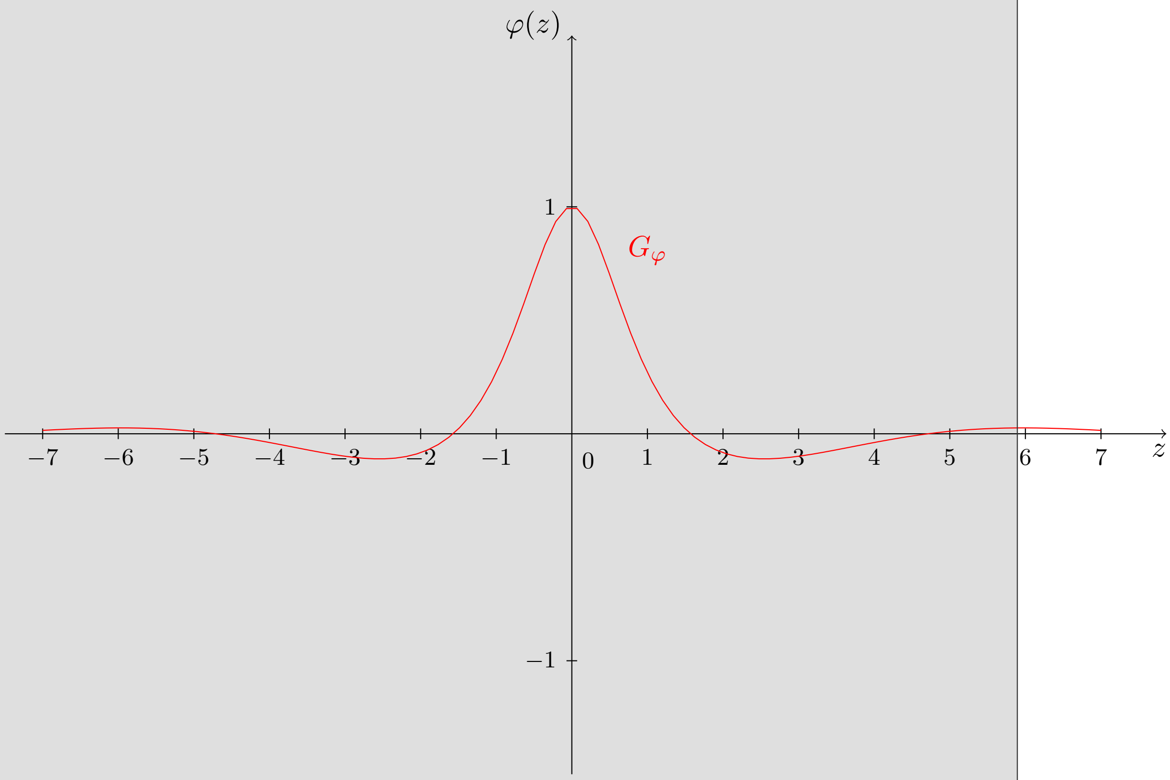
  is the product of the natural exponential function with the mapping rule e^x (see Section 6.4.3) and the hyperbola with the mapping rule $\frac{1}{x}$ (see Section 6.2.8). The graph of this function is shown in the figure below.



- The function

$$\varphi : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ z & \longmapsto & \frac{cos(z)}{z^2+1} \end{cases}$$

is the quotient of the cosine function (see Section 6.5.3) and the polynomial of degree 2 (see Section 6.2.7) with the mapping rule $z^2 + 1$. The graph of this function is shown in the figure below.



Exercise 6.6.2
Find further examples for elementary functions we studied in this module that were constructed from simpler elementary functions by combining sums, differences, product, or quotients.

<u>Solution:</u>
Further examples are:

- Functions of hyperbolic type (see Section 6.2.8) are all quotients of the constant function 1 and a monomial.
- Monomials (see Section 6.2.6) are all multiple products of the identity function $\mathrm{Id}(x) = x$.
- Linear functions (see Section 6.2.3) are products of constant functions describing

the slope and the identity function.

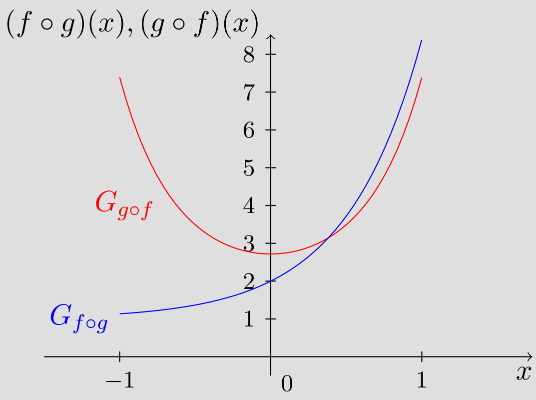
- Polynomials (see Section 6.2.7) are sums and differences of functions that are themselves products of constant functions and monomials.

Finally, there is another way to combine elementary functions to obtain new functions. This is the so-called **composition** of functions.

Let us consider a few examples.

Example 6.6.4

- The functions
$$f:\begin{cases}\mathbb{R} & \longrightarrow & \mathbb{R}\\ x & \longmapsto & f(x)=x^2+1\end{cases}$$
and
$$g:\begin{cases}\mathbb{R} & \longrightarrow & \mathbb{R}\\ x & \longmapsto & g(x)=e^x\end{cases}$$
can be composed in two ways. We can compose the function $f\circ g\colon \mathbb{R}\longrightarrow\mathbb{R}$ or the function $g\circ f\colon\mathbb{R}\longrightarrow\mathbb{R}$. We obtain
$$(f\circ g)(x)=f(g(x))=f(e^x)=(e^x)^2+1=e^{2x}+1\ ,$$
i.e.
$$f\circ g:\begin{cases}\mathbb{R} & \longrightarrow & \mathbb{R}\\ x & \longmapsto & e^{2x}+1\ ,\end{cases}$$
and
$$(g\circ f)(x)=g(f(x))=g(x^2+1)=e^{x^2+1}\ ,$$
i.e.
$$g\circ f:\begin{cases}\mathbb{R} & \longrightarrow & \mathbb{R}\\ x & \longmapsto & e^{x^2+1}\ .\end{cases}$$
If we look at the graphs, we see that these are two completely different functions, i.e. the order of the composition is relevant.



- If two functions such as

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

and

$$w : \begin{cases} [0; \infty) & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \sqrt{x} \end{cases}$$

are composed, however, the domains of the functions have to be observed. For example, if we want to consider the composed function $w \circ h$, then we have

$$(w \circ h)(x) = w(h(x)) = w(\sin(x)) = \sqrt{\sin(x)}\ .$$

Since the values of the sine function can also be negative but the square root only accepts non-negative values, the domain of the sine function has to be restricted accordingly such that the corresponding function values of the sine function are always non-negative, for example, by the restriction $x \in [0; \pi] = D_{w \circ h}$. Thus, we have

$$w \circ h : \begin{cases} [0; \pi] & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \sqrt{\sin(x)}\ . \end{cases}$$

Exercise 6.6.3
Let the functions

$$f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & 2x - 3 \end{cases}$$

$$g : \begin{cases} \mathbb{R} \setminus \{0\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{1}{x} \end{cases}$$

and

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

be given. Specify the compositions $f \circ g$, $g \circ f$, $h \circ f$, $h \circ g$, $f \circ f$, and $g \circ g$. Restrict the domains, if necessary, such that the composition is allowed. However, for the composed function always use the maximum domain.

<u>Solution:</u>

$$f \circ g : \begin{cases} \mathbb{R} \setminus \{0\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{2}{x} - 3 \end{cases}$$

$$g \circ f : \begin{cases} \mathbb{R} \setminus \{\frac{3}{2}\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{1}{2x-3} \end{cases}$$

$$h \circ f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \sin(2x-3) \end{cases}$$

$$h \circ g : \begin{cases} \mathbb{R} \setminus \{0\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \sin(\frac{1}{x}) \end{cases}$$

$$f \circ f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & 4x - 9 \end{cases}$$

$$g \circ g : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x \end{cases}$$

6.7 Final Test

6.7.1 Final Test Module 7

Exercise 6.7.1
Specify the maximum domains D_f and D_g of the two functions

$$f : \begin{cases} D_f & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{9x^2-\sin(x)+42}{x^2-2} \end{cases}$$

and

$$g : \begin{cases} D_g & \longrightarrow & \mathbb{R} \\ y & \longmapsto & \frac{\ln(y)}{y^2+1} \end{cases} .$$

Exercise 6.7.2
Specify the range W_i of the function

$$i : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 - 4x + 4 + \pi \end{cases} .$$

Exercise 6.7.3
Find the parameters $A, \lambda \in \mathbb{R}$ in the exponential function

$$c : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & A \cdot e^{\lambda x} - 1 \end{cases} ,$$

such that $c(0) = 1$ and $c(4) = 0$.

Answer: A = ☐ , λ = ☐ .

Exercise 6.7.4
Specify the composition $h = f \circ g \colon \mathbb{R} \to \mathbb{R}$ (note: $h(x) = (f \circ g)(x) = f(g(x))$) of the functions

$$f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & C \cdot \sin(x) \end{cases}$$

and

$$g : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & B \cdot x + \pi \end{cases} .$$

Answer: $h(x)$ = ☐ .

Find the parameters such that the sine wave described by the function h has the graph shown below.

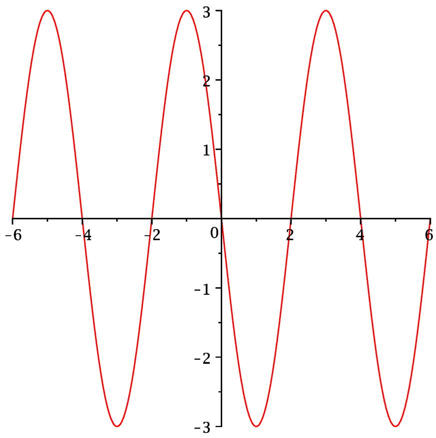


Abbildung 1: A sine wave.

Answer: $h(x) =$ [ ] .

Exercise 6.7.5

Specify the inverse function $f = u^{-1}$ of

$$u : \begin{cases} (0;\infty) & \longrightarrow \mathbb{R} \\ y & \longmapsto -\log_2(y)\,. \end{cases}$$

The function $f = u^{-1}$ has

a. the domain $D_f =$ [ ] .

b. the range $W_f =$ [ ] .

c. the mapping rule $f(y) = u^{-1}(y) =$ [ ] .

Exercise 6.7.6

Please indicate whether the following statements are right or wrong:

The function

$$f : \begin{cases} [0;3) & \longrightarrow \mathbb{R} \\ x & \longmapsto 2x+1 \end{cases}$$

- [ ] ... can be also written for short as $f(x) = 2x + 1$.
- [ ] ... is a linear affine function.
- [ ] ... has the range $\mathbb{R}$.
- [ ] ... has the slope 2.
- [ ] ... can only take values greater or equal 1 and less than 7.
- [ ] ... has a graph that is a piece of a line.
- [ ] ... has at $x = 0$ the value 1.
- [ ] ... has the domain $\mathbb{R}$.

Exercise 6.7.7
Calculate the following logarithms:

a. $\ln(e^5 \cdot \frac{1}{\sqrt{e}}) =$ [ ] .

b. $\log_{10}(0.01) =$ [ ] .

c. $\log_2(\sqrt{2 \cdot 4 \cdot 16 \cdot 256 \cdot 1024}) =$ [ ] .

7 Differential Calculus

Module Overview

7.1 Derivative of a Function

7.1.1 Introduction

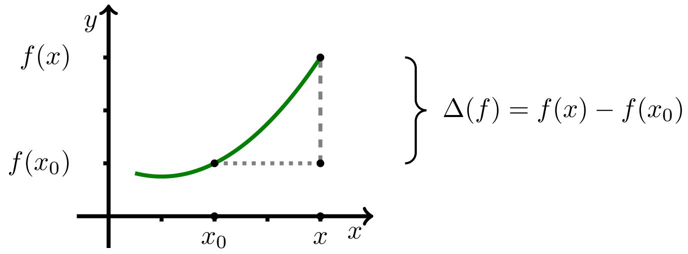
A family is going on holiday by car. The car is moving through roadworks with a velocity of 60 km[]/h. The sign at the end of the roadworks says that the speed limit is, as of now, 120 km[]/h. Even though the car driver puts the pedal to the metal, the velocity of the car will not jump up immediately but increase as a function of time. If the velocity increases from 60 km[]/h to 120 km[]/h in 5 seconds at a constant rate of change, then the *acceleration* (= change of velocity per time) equals this constant (in this case) rate of velocity change: the acceleration is the quotient of the velocity change and the time required for this change. Thus, its value is here 12 kilometre per hour per second. In reality, the velocity of the car will not increase at a constant rate but at a *time-dependent* rate. If the velocity v is described as a function of time t, then the acceleration is the slope of this function. This does not depend on the fact whether this slope is constant (in time) or not. On other words: The acceleration is the *derivative* of the velocity *function* v with respect to the time t.

Similar relations can also be found in other technical fields such as, for example, the calculation of internal forces acting in steel frames of buildings, the forecast of atmospheric and oceanic currents, or in the modelling of financial markets, which is currently highly relevant.

This chapter reviews the basic ideas underlying these calculations, i.e. it deals with **differential calculus**. In other words: we will take derivatives of functions to find their slopes or rates of change. Even thought these calculations will be carried out here in a strictly mathematical way, their motivation is not purely mathematical. Derivatives, interpreted as rates of change of different functions, play an important role in many scientific fields and are often investigated as special quantities.

7.1.2 Relative Rate of Change of a Function

Consider a function $f : [a; b] \to \mathbb{R}$, $x \mapsto f(x)$ and a sketch of the graph of f (shown in the figure below). We would like to describe the rate of change of this function at an arbitrary point x_0 between a and b. This will lead us to the notion of a derivative of a function. Generally, calculation rules are to be applied that are as simple as possible.



If x_0 and the corresponding function value $f(x_0)$ are fixed and another arbitrary, but variable point x between a and b as well as the corresponding function value $f(x)$ are chosen, then through these two points, i.e. the points $(x_0; f(x_0))$ and $(x; f(x))$, a line can be drawn that is characterised by its slope and its y-intercept. For the slope of this line one obtains the so-called **difference quotient**

$$\frac{\Delta(f)}{\Delta(x)} = \frac{f(x) - f(x_0)}{x - x_0}$$

that describes how the function values of f between x_0 and x change **on average**. Thus, an average rate of change of the function f on the interval $[x_0; x]$ is found. This quotient is also called **relative change**.

If we let the variable point x approach the point x_0, then we see that the line that intersects the graph of the function in the points $(x_0; f(x_0))$ and $(x; f(x))$ gradually becomes a tangent line to the graph in the point $(x_0; f(x_0))$. In this way, the rate of change of the function f – or the **slope** of the graph of f – at the point x_0 **itself** can be determined. If the approaching process of x to x_0 described above leads to, figuratively speaking, a unique tangent line (i.e. a line with a unique slope that, in particular, must not be infinity), then in mathematical terms one says that the **limit** of the difference quotient does **exist**. This limiting process, i.e. letting x approach x_0, is described here and in the following by the symbol

$$\lim_{x \to x_0} ,$$

where lim is an abbreviation for the Latin word *limes*, meaning "border" or "boundary". If the limit of the difference quotient exists, then

$$f'(x_0) = \lim_{x \to x_0} \frac{\Delta(f)}{\Delta(x)} = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

denotes the value of the **derivative** of f at x_0. The function f is then said to be **differentiable** at the point x_0.

Example 7.1.1

For the function $f(x) = \sqrt{x}$ the relative change at the point $x_0 = 1$ is given by

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{\sqrt{x} - \sqrt{1}}{x - 1} = \frac{\sqrt{x} - 1}{(\sqrt{x} - 1)(\sqrt{x} + 1)} = \frac{1}{\sqrt{x} + 1} .$$

If x approaches $x_0 = 1$, this results in the limit

$$\lim_{x \to x_0} \frac{\Delta(f)}{\Delta(x)} = \frac{1}{2} .$$

The value of the derivative of the function f at the point $x_0 = 1$ is denoted by $f'(1) = \frac{1}{2}$.

Exercise 7.1.1
Consider the function $f : \mathbb{R} \to \mathbb{R}$ with $x \mapsto f(x) = x^2$ and a point $x_0 = 1$. At this point, the relative change for a real value of x equals $\dfrac{f(x) - f(1)}{x - 1} = \square$.

If x approaches $x_0 = 1$, this results in the slope $\square$ of the graph of the function f at the point $x_0 = 1$.

Solution:
For $f(x) = x^2$, the relative change at the point $x_0 = 1$ is given by

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

Then, if x approaches x_0, this results in the limit

$$\lim_{x \to 1} \frac{\Delta(f)}{\Delta(x)} = 2 .$$

This is the slope of the tangent line to the graph of f at the point $(x_0; f(x_0)) = (1; 1)$. The value of the derivative of f at the point $x_0 = 1$ is denoted by $f'(1) = 2$.

Using the formula for the relative rate of change, calculating the derivative can be very cumbersome and also only works for very simple functions. Typically, the derivative is determined by applying calculation rules and inserting known derivatives for the individual terms.

7.1.3 Derivative

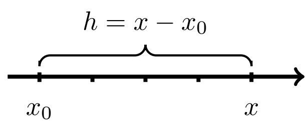
Notation of the Derivative7.1.2

In mathematics, sciences and engineering, different but equivalent notations for derivatives are used:

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

These different notations all denote the derivative of the function f at the point x_0.

If the derivative is to be calculated using the difference quotient $\frac{f(x)-f(x_0)}{x-x_0}$, then it is often convenient to rewrite the difference quotient in another way. Denoting the difference of x and x_0 by $h := x - x_0$ (see figure below),



the difference quotient can be rewritten as

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

where $x = x_0 + h$. There is no statement about whether x has to be greater or less than x_0. Hence, the quantity h can take positive or negative values. To determine the derivative of the function f, the limit for $h \to 0$ has to be calculated:

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

If this limit exists **for all** points x_0 in a function's domain, then the function is said to be **differentiable** (everywhere). Many of the common functions are differentiable. However, a simple example of a function that is not differentiable everywhere is the absolute value function $f : \mathbb{R} \to \mathbb{R}$ with $x \mapsto f(x) := |x|$.

Example 7.1.3

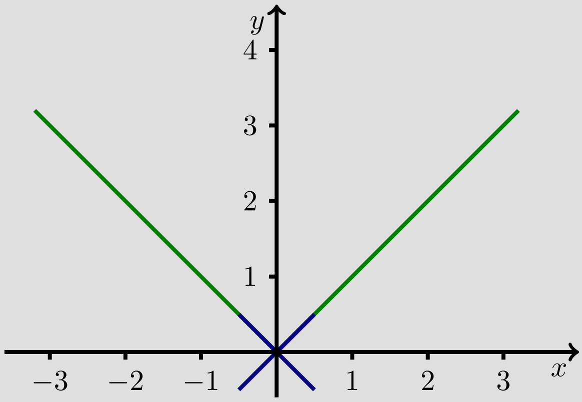
The absolute value function (see Module 6, Section 6.2.5 auf Seite 238) is not differentiable at the point $x_0 = 0$. The difference quotient of f at the point $x_0 = 0$

is:

$$\frac{f(0+h)-f(0)}{h} = \frac{|h|-|0|}{h} = \frac{|h|}{h} .$$

Since h can be greater or less than 0, two cases are to be distinguished: For $h > 0$, we have $\frac{|h|}{h} = \frac{h}{h} = 1$, and for $h < 0$, we have $\frac{|h|}{h} = \frac{-h}{h} = -1$. In these two cases, the limiting process, i.e. h approaching 0, results in two different values (1 and -1). Thus, **the** limit of the difference quotient at the point $x_0 = 0$ does not exist. Hence, the absolute value function is not differentiable at the point $x_0 = 0$.

The graph changes its direction at the point $(0;0)$ abruptly: Casually speaking, one says that the graph of the function has a kink at the point $(0;0)$.



Likewise, if a function has a jump at a certain point, a unique tangent line to the graph at this point does not exist and thus, the function has no derivative at this point.

7.1.4 Exercises

Exercise 7.1.2
Using the difference quotient, calculate the derivative of $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x) := 4 - x^2$ at the points $x_1 = -2$ and $x_2 = 1$.

Answer:

a. The difference quotient of f at the point $x_1 = -2$ is [ ] and has for $x \to -2$ the limit $f'(-2) =$ [ ] .

b. The difference quotient of f at the point $x_2 = 1$ is [ ] and has for $x \to 1$ the limit $f'(1) =$ [ ] .

Solution:

a. At the point $x_1 = -2$, we have for the difference quotient

$$\frac{\Delta(f)}{\Delta(x)} = \frac{f(x) - f(x_1)}{x - x_1} = \frac{f(x) - f(-2)}{x - (-2)} = \frac{4 - x^2 - 0}{x + 2} = \frac{(2 - x)(2 + x)}{2 + x} = 2 - x \,.$$

For $x \to x_1$, i.e. for $x \to -2$, this difference quotient tends to $2 - (-2) = 4$; hence, $f'(-2) = 4$.

b. At the point $x_2 = 1$, we have for the difference quotient

$$\frac{\Delta(f)}{\Delta(x)} = \frac{f(x) - f(x_2)}{x - x_2} = \frac{f(x) - f(1)}{x - 1} = \frac{4 - x^2 - 3}{x - 1} = \frac{1 - x^2}{x - 1} = -\frac{(x - 1)(x + 1)}{x - 1} = -x - 1 \,.$$

For $x \to x_2$, i.e. for $x \to 1$, this difference quotient has the limit $-1 - 1 = -2$; hence, $f'(1) = -2$.

Exercise 7.1.3
Explain why the functions

a. $f : [-3; \infty[\to \mathbb{R}$ with $f(x) := \sqrt{x + 3}$ at $x_0 = -3$ and

b. $g : \mathbb{R} \to \mathbb{R}$ with $g(x) := 6 \cdot |2x - 10|$ at $x_0 = 5$

are not differentiable.

Answer:

a. The derivative of the function f at the point $x_0 = -3$ does not exist since the difference quotient [ ] does not converge for $h \to 0$.

b. The derivative of the function g at the point $x_0 = 5$ does not exist since the difference quotient for $h < 0$ has the value [ ] and for $h > 0$ has the value [ ]. Thus, the limit for $h \to 0$ does not exist.

Solution:

a. The difference quotient of the function f at the point $x_0 = -3$ is

$$\frac{\Delta(f)}{\Delta(x)} = \frac{f(x_0+h) - f(x_0)}{h} = \frac{\sqrt{-3+h+3} - \sqrt{-3+3}}{h} = \frac{\sqrt{h} - 0}{h} = \frac{1}{\sqrt{h}}\,.$$

For $h \to 0$ ($h > 0$), this difference quotient increases infinitely, i.e. the limit of the difference quotient does not exist.

b. The difference quotient of the function g at the point $x_0 = 5$ is

$$\frac{\Delta(g)}{\Delta(x)} = \frac{g(x_0+h) - g(x_0)}{h} = \frac{6 \cdot |2(5+h) - 10| - 6 \cdot |2 \cdot 5 - 10|}{h} = \frac{12|h| - 0}{h} = \frac{12|h|}{h}\,.$$

For $h < 0$, since $|h| = -h$, the difference quotient has the value -12. In contrast, for $h > 0$, since $|h| = h$, it has the value 12. Thus, the limit of the difference quotient does not exist. (The limit has always to be unique.)

7.2 Standard Derivatives

7.2.1 Introduction

Most of the common functions, such as polynomials, trigonometric functions, and exponential functions (see Module 6) are differentiable. In the following, the differentiation rules for these functions are repeated.

7.2.2 Derivatives of Power Functions

In the last section, the derivative was introduced as the limit of the difference quotient. Accordingly, for a linear affine function (see Module 6, Section 6.2.4 auf Seite 236) $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x) = mx + b$, where m and b are given numbers, we obtain for the derivative at the point x_0 the value $f'(x_0) = m$. (Readers are invited to verify that fact themselves.)

For monomials x^n with $n \geq 1$, it is easiest to determine the derivative using the difference quotient. Without any detailed calculation or any proof we state the following rules:

Derivative of x^n7.2.1

Let a natural number n and a real number r be given.

The constant function $f : \mathbb{R} \to \mathbb{R}$ with $x \mapsto f(x) := r = r \cdot x^0$ has the derivative $f' : \mathbb{R} \to \mathbb{R}$ with $x \mapsto f'(x) = 0$.

The function $f : \mathbb{R} \to \mathbb{R}$ with $x \mapsto f(x) := r \cdot x^n$ has the derivative

$$f' : \mathbb{R} \to \mathbb{R} \text{ with } x \mapsto f'(x) = r \cdot n \cdot x^{n-1} .$$

This differentiation rule is true for all $n \in \mathbb{R} \setminus \{0\}$.

Again, we leave the verification of these statements to the reader.

Example 7.2.2

Let us consider the function $f : \mathbb{R} \to \mathbb{R}$ with $x \mapsto f(x) = 5x^3$. According to the notation above, this is a function with $r = 5$ and $n = 3$. Thus for the value of the derivative at the point x , we have

$$f'(x) = 5 \cdot 3x^{3-1} = 15x^2 \ .$$

For root functions, an equivalent statement holds. However, it should be noted that root functions are only differentiable for $x > 0$ since the tangent line to the graph of the function at the point $(0;0)$ is parallel to the y-axis and thus, it is not a graph of a function.

Derivative of $x^{\frac{1}{n}}$ 7.2.3

For $n \in \mathbb{Z}$ with $n \neq 0$, the function $f : [0;\infty[\to \mathbb{R}$, $x \mapsto f(x) := x^{\frac{1}{n}}$ is differentiable for $x > 0$, and we have

$$f' :]0;\infty[\to \mathbb{R} \ , \quad x \mapsto f'(x) = \frac{1}{n} \cdot x^{\frac{1}{n}-1} \ .$$

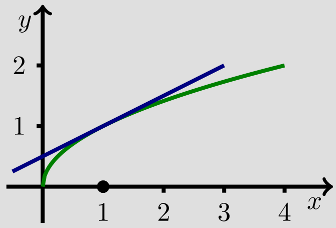
For $n \in \mathbb{N}$, root functions are described by $f(x) = x^{\frac{1}{n}}$. Of course, the differentiation rule given here also holds for $n = 1$ or $n = -1$.

Example 7.2.4

The root function $f : [0;\infty[\to \mathbb{R}$ with $x \mapsto f(x) := \sqrt{x} = x^{\frac{1}{2}}$ is differentiable for $x > 0$. The value of the derivative at an arbitrary point $x > 0$ is given by

$$f'(x) = \frac{1}{2} \cdot x^{\frac{1}{2}-1} = \frac{1}{2} \cdot x^{-\frac{1}{2}} = \frac{1}{2 \cdot \sqrt{x}} \ .$$

The derivative at the point $x_0 = 0$ does not exist since the slope of the tangent line to the graph of f would be infinite there.



The tangent line to the graph of the given root function at the point $(1;1)$ has the slope $\frac{1}{2\sqrt{1}} = \frac{1}{2}$.

For $x > 0$, the statements above can be extended to exponents $p \in \mathbb{R}$ with $p \neq 0$: The value $f'(x)$ of the derivative of the function f with the mapping rule $f(x) = x^p$ is, for $x > 0$,

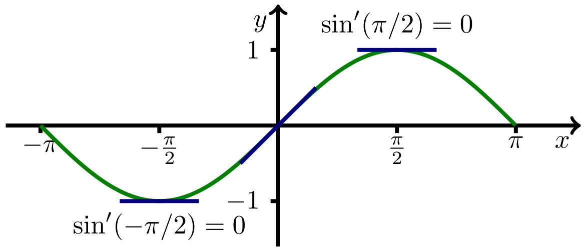
$$f'(x) = p \cdot x^{p-1} .$$

7.2.3 Derivatives of Special Functions

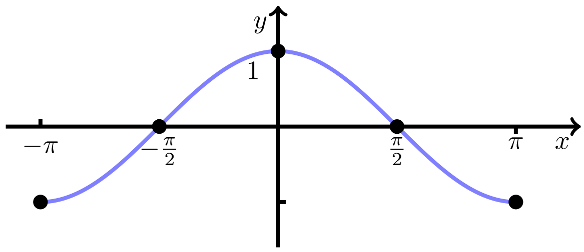
Derivatives of Trigonometric Functions

The sine function $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x) = \sin(x)$ is periodic with period 2π. Thus, it is sufficient to consider the function on an interval of length 2π. A section of the graph for $-\pi \leq x \leq \pi$ is shown in the figure below:

Sine function



Derivative



As we see from the figure above, the slope of the sine function at $x_0 = \pm\frac{\pi}{2}$ is $f'(\pm\frac{\pi}{2}) = 0$. The tangent line to the graph of the sine function at $x_0 = 0$ has the slope $f'(0) = 1$. At $x_0 = \pm\pi$, the tangent line has the same slope as the tangent line at $x_0 = 0$, but the sign is opposite. Hence, the slope at $x_0 = \pm\pi$ is $f'(\pm\pi) = -1$. Thus, the derivative of the sine function is a function that exhibits exactly these properties. A detailed investigation of the regions between these specially chosen points shows that the derivative of the sine function is the cosine function:

Derivatives of Trigonometric Functions7.2.5

For the sine function $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x) := \sin(x)$, we have

$$f' : \mathbb{R} \to \mathbb{R}\ , \quad x \mapsto f'(x) = \cos(x)\ .$$

For the cosine function $g : \mathbb{R} \to \mathbb{R}$, $x \mapsto g(x) := \cos(x)$, we have

$$g' : \mathbb{R} \to \mathbb{R}\ , \quad x \mapsto g'(x) = -\sin(x)\ .$$

For the tangent function $h : \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi\ :\ k \in \mathbb{Z}\} \to \mathbb{R}$, $x \mapsto h(x) := \tan(x)$, we have

$$h' : \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi\ :\ k \in \mathbb{Z}\} \to \mathbb{R}\ , \quad x \mapsto h'(x) = 1 + (\tan(x))^2 = \frac{1}{\cos^2(x)}\ .$$

This last result comes from the calculation rules explained (explained below) and the definition of the tangent function as the quotient of the sine function and the cosine function.

Derivative of the Exponential Function

Info7.2.6

The exponential function $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x) := e^x = \exp(x)$ has the special property that its derivative f' is also the exponential function, i.e. $f'(x) = e^x = \exp(x)$.

Derivative of the Logarithmic Function

The derivative of the logarithmic function is given here without proof. For $f :]0; \infty[\to \mathbb{R}$ with $x \mapsto f(x) = \ln(x)$ one obtains $f' :]0; \infty[\to \mathbb{R}$, $x \mapsto f'(x) = \frac{1}{x}$.

7.2.4 Exercises

Exercise 7.2.1
Find the following derivatives by simplifying the terms of the functions and then applying your knowledge of the differentiation of common functions ($x > 0$):

a. $f(x) := x^6 \cdot x^{\frac{7}{2}} =$ [ ] .

b. $g(x) := \frac{x^{-\frac{3}{2}}}{\sqrt{x}} =$ [ ] .

Thus, we have:

a. $f'(x) =$ [ ] .

b. $g'(x) =$ [ ] .

Solution:

a. We have $f(x) = x^6 \cdot x^{\frac{7}{2}} = x^{6+\frac{7}{2}} = x^{\frac{19}{2}}$, and hence $f'(x) = \frac{19}{2} x^{\frac{19}{2}-1} = \frac{19}{2} x^{\frac{17}{2}}$.

b. We have $g(x) = \frac{x^{-\frac{3}{2}}}{\sqrt{x}} = x^{-\frac{3}{2}} \cdot x^{-\frac{1}{2}} = x^{-\frac{3}{2}-\frac{1}{2}} = x^{-2}$, and hence $g'(x) = (-2) \cdot x^{-2-1} = -2 \cdot x^{-3} = -\frac{2}{x^3}$.

Exercise 7.2.2
Simplify the terms of the functions and find their derivatives:

a. $f(x) := 2 \sin\left(\frac{x}{2}\right) \cdot \cos\left(\frac{x}{2}\right) =$ [ ] .

b. $g(x) := \cos^2(3x) + \sin^2(3x) =$ [ ] .

Thus, we have:

a. $f'(x) =$ [ ] .

b. $g'(x) =$ [ ] .

Solution:

a. Generally, we have

$$\sin(u) \cdot \cos(v) = \frac{1}{2} \left(\sin(u - v) + \sin(u + v)\right) .$$

Thus, in the present case we have $f(x) = 2 \cdot \frac{1}{2}(\sin(0) + \sin(x)) = \sin(x)$, and hence

$f'(x) = \cos(x)$.

b. Since $\sin^2(u) + \cos^2(u) = 1$, we have $g(x) = 1$, and thus $g'(x) = 0$.

Exercise 7.2.3
Simplify the terms of the functions and find the derivatives (for $x > 0$ in the first part of this exercise):

a. $f(x) := 3\ln(x) + \ln\left(\frac{1}{x}\right) =$ [ ] .

b. $g(x) := (e^x)^2 \cdot e^{-x} =$ [ ] .

Thus, we have:

a. $f'(x) =$ [ ] .

b. $g'(x) =$ [ ] .

Solution:

a. We have

$$f(x) = 3\ln(x) + \ln\left(\frac{1}{x}\right) = \ln(x^3) + \ln\left(\frac{1}{x}\right) = \ln\left(x^3 \cdot \frac{1}{x}\right) = \ln\left(x^2\right) .$$

For the value of the derivative at the point x ($x > 0$), it follows from the chain rule $f'(x) = \frac{1}{x^2} \cdot 2x = \frac{2}{x}$. (The chain rule is explained in detail in Section 7.3.4 auf Seite 313.)

b. We have

$$g(x) = (e^x)^2 \cdot e^{-x} = e^x \cdot e^x \cdot e^{-x} = e^{x+x-x} = e^x .$$

Hence, it follows that $g'(x) = e^x$.

7.3 Calculation Rules

7.3.1 Introduction

Using a few calculation rules and the derivatives presented in the last section, a variety of functions can be differentiated.

7.3.2 Multiples and Sums of Functions

In the following, $u, v : D \to \mathbb{R}$ will denote two arbitrary differentiable functions, and r denotes an arbitrary real number.

Sum Rule and Constant Factor Rule7.3.1

Let two differentiable functions u and v be given. Then, the sum $f := u + v$ with $f(x) = (u + v)(x) := u(x) + v(x)$ is also differentiable, and we have

$$f'(x) = u'(x) + v'(x) \ .$$

Likewise, a function multiplied by a factor r, i.e. $f := r \cdot u$ with $f(x) = (r \cdot u)(x) := r \cdot u(x)$, is also differentiable, and we have

$$f'(x) = r \cdot u'(x) \ .$$

Using these two rules together with the differentiation rules for monomials x^n, any arbitrary polynomial can be differentiated. Here are some examples.

Example 7.3.2

The polynomial f with the mapping rule $f(x) = \frac{1}{4}x^3 - 2x^2 + 5$ is differentiable, and we have

$$f'(x) = \frac{3}{4}x^2 - 4x \ .$$

The derivative of the function $g :]0;\infty[\to \mathbb{R}$ with $g(x) = x^3 + \ln(x)$ is

$$g' :]0;\infty[\to \mathbb{R} \text{ with } g'(x) = 3x^2 + \frac{1}{x} = \frac{3x^3 + 1}{x} \ .$$

Differentiating the function $h : [0; \infty[\to \mathbb{R}$ with $h(x) = 4^{-1} \cdot x^2 - \sqrt{x} = \frac{1}{4}x^2 + (-1) \cdot x^{\frac{1}{2}}$ results, for $x > 0$, in

$$h'(x) = \frac{1}{2}x - \frac{1}{2}x^{-\frac{1}{2}} = \frac{x^{\frac{3}{2}} - 1}{2\sqrt{x}} .$$

7.3.3 Product and Quotient of Functions

Product and Quotient Rule7.3.3

Likewise, the product of functions, i.e. $f := u \cdot v$ with $f(x) = (u \cdot v)(x) := u(x) \cdot v(x)$, is differentiable, and the following **product rule** applies:

$$f'(x) = u'(x) \cdot v(x) + u(x) \cdot v'(x) .$$

The quotient of functions, i.e. $f := \frac{u}{v}$ with $f(x) = \left(\frac{u}{v}\right)(x) := \frac{u(x)}{v(x)}$, is defined and differentiable for all x with $v(x) \neq 0$, and the following **quotient rule** applies:

$$f'(x) = \frac{u'(x) \cdot v(x) - u(x) \cdot v'(x)}{(v(x))^2} .$$

These calculation rules shall be illustrated by means of a few examples.

Example 7.3.4

Find the derivative of $f : \mathbb{R} \to \mathbb{R}$ with $f(x) = x^2 \cdot e^x$. The product rule can be applied choosing, for example, $u(x) = x^2$ and $v(x) = e^x$. The corresponding derivatives are $u'(x) = 2x$ and $v'(x) = e^x$. Combining these terms according to the product rule results in the derivative of the function f:

$$f' : \mathbb{R} \to \mathbb{R} , \quad x \mapsto f'(x) = 2xe^x + x^2e^x = (x^2 + 2x)e^x .$$

Next, we investigate the tangent function g with $g(x) = \tan(x) = \frac{\sin(x)}{\cos(x)}$ $(\cos(x) \neq 0)$.In order to use the quotient rule we set $u(x) = \sin(x)$ and $v(x) = \cos(x)$. The corresponding derivatives are $u'(x) = \cos(x)$ and $v'(x) = -\sin(x)$. Combining these terms and applying the quotient rule results in the derivative of the function g:

$$g'(x) = \frac{\cos(x) \cdot \cos(x) - \sin(x) \cdot (-\sin(x))}{\cos^2(x)} .$$

This result can be transformed into any of the following expressions:

$$g'(x) = 1 + \left(\frac{\sin(x)}{\cos(x)}\right)^2 = 1 + \tan^2(x) = \frac{1}{\cos^2(x)} .$$

For the last transformation, the relation $\sin^2(x) + \cos^2(x) = 1$ was used, which was given in Module 5 (see Section 5.6.2 auf Seite 201).

Exercise 7.3.1
Calculate the derivative of $f : \mathbb{R} \to \mathbb{R}$ with $f(x) = \sin(x) \cdot x^3$ by factorising the product into two factors, taking the derivatives of each single factor, and finally combining the results according to the product rule.

a. The derivative of the left factor $u(x) =$ [ ] is $u'(x) =$ [ ] .
b. The derivative of the right factor $v(x) =$ [ ] is $v'(x) =$ [ ] .
c. Thus, applying the product rule to f results in $f'(x) =$ [ ] .

Solution:
The four terms are

$$u(x) = \sin(x) , \quad u'(x) = \cos(x) , \quad v(x) = x^3 , \quad v'(x) = 3x^2 ,$$

and combining them according to the product rule results in

$$f'(x) = \cos(x) \cdot x^3 + \sin(x) \cdot 3x^2 .$$

Exercise 7.3.2
Calculate the derivative of $f :]0;\infty[\to \mathbb{R}$ with $f(x) = \frac{\ln(x)}{x^2}$ by splitting the quotient up into numerator and denominator, taking the derivatives of both, and combining them according to the quotient rule.

a. The derivative of the numerator $u(x) =$ [ ] is $u'(x) =$ [ ] .

b. The derivative of the denominator $v(x) =$ [ ] is $v'(x) =$ [ ] .

c. Thus, applying the quotient rule to f results in $f'(x) =$ [ ] .

Solution:
The four terms are

$$u(x) = \ln(x)\ ,\ \ u'(x) = \frac{1}{x}\ ,\ \ v(x) = x^2\ ,\ \ v'(x) = 2x\ ,$$

and combining them according to the quotient rule results in

$$f'(x) = \frac{\frac{1}{x} \cdot x^2 - \ln(x) \cdot 2x}{x^4} = \frac{1 - 2\ln(x)}{x^3}\ ,$$

where the last transformation step (cancelling x) is just to simplify the expression.

7.3.4 Composition of Functions

Finally, we investigate composition of functions (see Module 6, Section 6.6.3 auf Seite 284): what happens if a function u (the inner function) is substituted into another function v (the outer function)? In mathematics, such a composition is denoted by $f := v \circ u$ with $f(x) = (v \circ u)(x) := v(u(x))$. That is, first the value of a function u is determined depending on the variable x. The value $u(x)$ calculated this way is then used as an argument of the function v. This results in the final function value $v(u(x))$.

Chain Rule7.3.5

The derivative of the function $f := v \circ u$ with $f(x) = (v \circ u)(x) := v(u(x))$ can be calculated applying the **chain rule**:

$$f'(x) = v'(u(x)) \cdot u'(x)\ .$$

Here, the expression $v'(u(x))$ is considered in such a way that v is a function of u and thus, the derivative is taken with respect to u; then $v'(u)$ is evaluated for $u = u(x)$. The following phrase is a useful summary: the derivative of a composite function is the product of the outer derivative and the inner derivative.

This differentiation rule shall be illustrated by a few examples.

Example 7.3.6

Find the derivative of the function $f : \mathbb{R} \to \mathbb{R}$ with $f(x) = (3-2x)^5$. To apply the chain rule, inner and outer functions must be identified. If we take the function $u(x) = 3-2x$ as the inner function u, then the outer function v is given by $v(u) = u^5$. With this, we have the required form $v(u(x)) = f(x)$.

Taking the derivative of the inner function u with respect to x results in $u'(x) = -2$. For the outer derivative, the function v is differentiated with respect to u, which results in $v'(u) = 5u^4$. Inserting these terms into the chain rule results in the derivative f' of the function f with

$$f'(x) = 5(u(x))^4 \cdot (-2) = 5(3-2x)^4 \cdot (-2) = -10(3-2x)^4 \; .$$

As a second example, let's calculate the derivative of $g : \mathbb{R} \to \mathbb{R}$ with $g(x) = e^{x^3}$. For the inner function u the assignment $x \mapsto u(x) = x^3$ and for the outer function v the assignment $u \mapsto v(u) = e^u$ is appropriate. Taking the inner and the outer derivative results in $u'(x) = 3x^2$ and $v'(u) = e^u$. Inserting these terms into the chain rule results in the derivative of the function g:

$$g' : \mathbb{R} \to \mathbb{R} \; , \quad x \mapsto g'(x) = e^{u(x)} \cdot 3x^2 = e^{x^3} \cdot 3x^2 = 3x^2 e^{x^3} \; .$$

7.3.5 Exercises

Exercise 7.3.3
Calculate the derivatives of the functions f, g, and h defined by the following mapping rules:

a. The derivative of $f(x) := 3 + 5x$ is $f'(x) =$ [ ] .

b. The derivative of $g(x) := \frac{1}{4x} - x^3$ is $g'(x) =$ [ ] .

c. The derivative of $h(x) := 2\sqrt{x} + 4x^{-3}$ is $h'(x) =$ [ ] .

Solution:

a. We have $f'(x) = 0 + 5 \cdot 1 \cdot x^0 = 0 + 5 = 5$.

b. Since $g(x) = \frac{1}{4x} - x^3 = \frac{1}{4}x^{-1} - x^3$, we have $g'(x) = \frac{1}{4} \cdot (-1) \cdot x^{-2} - 3 \cdot x^2 = -\frac{1}{4x^2} - 3x^2$.

c. Since $h(x) = 2\sqrt{x} + 4x^{-3} = 2x^{\frac{1}{2}} + 4x^{-3}$, we have $h'(x) = 2 \cdot \frac{1}{2} \cdot x^{-\frac{1}{2}} + 4 \cdot (-3) \cdot x^{-4} = \frac{1}{\sqrt{x}} - \frac{12}{x^4}$.

Exercise 7.3.4
Calculate the derivatives of the functions f, g, and h described by the following mapping rules, and simplify the results.

a. The derivative of $f(x) := \cot x = \frac{\cos x}{\sin x}$ is $f'(x) =$ [ ] .

b. The derivative of $g(x) := \sin(3x) \cdot \cos(3x)$ is $g'(x) =$ [ ] .

c. The derivative of $h(x) := \frac{\sin(3x)}{\sin(6x)}$ is $h'(x) =$ [ ] .

Solution:

a. From the quotient rule, we find

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

b. From the product rule and the chain rule, we find

$$g'(x) = \cos(3x) \cdot 3 \cdot \cos(3x) + \sin(3x) \cdot (-\sin(3x)) \cdot 3 = 3\left(\cos^2(3x) - \sin^2(3x)\right) .$$

Since generally $\cos^2(u) - \sin^2(u) = \cos(2u)$, we have $g'(x) = 3\cos(6x)$.

c. According to the general relation $\sin(2u) = 2\sin(u)\cos(u)$, we have

$$h(x) = \frac{\sin(3x)}{\sin(6x)} = \frac{\sin(3x)}{2\sin(3x)\cos(3x)} = \frac{1}{2\cos(3x)} = \frac{1}{2} \cdot (\cos(3x))^{-1} \ .$$

Applying the chain rule several times results in

$$h'(x) = \frac{1}{2} \cdot (-1) \cdot (\cos(3x))^{-2} \cdot (-\sin(3x)) \cdot 3 = \frac{3\sin(3x)}{2\cos^2(3x)} = \frac{3\tan(3x)}{2\cos(3x)} \ .$$

Exercise 7.3.5
Calculate the derivatives of the functions f, g, and h defined by the following mapping rules:

a. The derivative of $f(x) := e^{5x}$ is $f'(x) =$ [ ] .

b. The derivative of $g(x) := x \cdot e^{6x}$ is $g'(x) =$ [ ] .

c. The derivative of $h(x) := (x^2 - x) \cdot e^{-2x}$ is $h'(x) =$ [ ] .

Solution:

a. From the chain rule, we immediately find $f'(x) = 5e^{5x}$.

b. From the product rule and the chain rule, we find $g'(x) = 1 \cdot e^{6x} + x \cdot e^{6x} \cdot 6 = e^{6x}(1 + 6x)$.

c. From the product rule and the chain rule, we find $h'(x) = (2x - 1) \cdot e^{-2x} + (x^2 - x) \cdot e^{-2x} \cdot (-2) = -(2x^2 - 4x + 1)e^{-2x}$.

Exercise 7.3.6
Calculate the first four derivatives of $f : \mathbb{R} \to \mathbb{R}$ with $f(x) := \sin(1 - 2x)$.

Answer: The kth derivative of f is denoted by $f^{(k)}$. Here, $f^{(1)} = f'$, $f^{(2)}$ is the derivative of $f^{(1)}$, $f^{(3)}$ is the derivative of $f^{(2)}$, etc. Thus, we have:

- $f^{(1)}(x) =$ [ ] .
- $f^{(2)}(x) =$ [ ] .
- $f^{(3)}(x) =$ [ ] .
- $f^{(4)}(x) =$ [ ] .

Solution:
From the chain rule, we find successively:

$$\begin{aligned}
f^{(1)}(x) &= \cos(1-2x)\cdot(-2) = -2\cos(1-2x)\ ,\\
f^{(2)}(x) &= -2\cdot(-\sin(1-2x))\cdot(-2) = -4\sin(1-2x)\ ,\\
f^{(3)}(x) &= -4\cdot\cos(1-2x)\cdot(-2) = 8\cos(1-2x)\ ,\\
f^{(4)}(x) &= 8\cdot(-\sin(1-2x))\cdot(-2) = 16\sin(1-2x)\ .
\end{aligned}$$

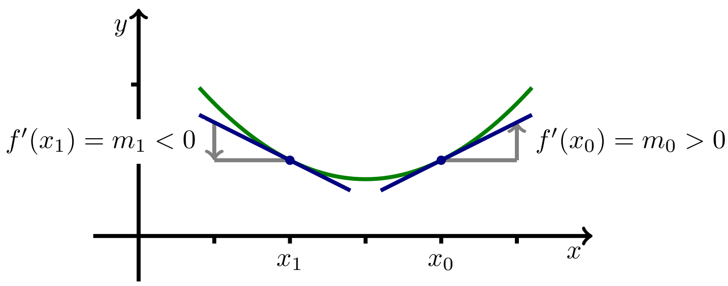
7.4 Properties of Functions

7.4.1 Introduction

The derivative was introduced above by means of a tangent line to a graph of a function. This tangent line describes the given function "approximately" in a certain region. The properties of this tangent line give also information on the properties of the approximated function in this region.

7.4.2 Monotony

The derivative of a function can be used to study the growth behaviour, i.e. whether the function values increase or decrease for increasing values of x. For this purpose, we consider a function $f : D \to \mathbb{R}$ that is differentiable on $]a; b[\subseteq D$:



If $f'(x) \leq 0$ for all x between a and b, then f is monotonically decreasing on the interval $]a; b[$.

If $f'(x) \geq 0$ for all x between a and b, then f is monotonically increasing on the interval $]a; b[$.

Thus, it is sufficient to determine the sign of the derivative f' to decide whether a function is monotonically increasing or decreasing on the interval $]a; b[$.

Example 7.4.1

The function $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^3$ is differentiable with $f'(x) = 3x^2$. Since $x^2 \geq 0$ for all $x \in \mathbb{R}$, we have $f'(x) \geq 0$, and therefore f is monotonically increasing.

For $g : \mathbb{R} \to \mathbb{R}$ with $g(x) = 2x^3+6x^2-18x+10$, the function $g'(x) = 6x^2+12x-18 = 6(x+3)(x-1)$ has the roots $x_1 = -3$ and $x_2 = 1$. If the monotony of the function g is investigated, then three regions are to be distinguished in which g' has a different sign.

The following table is used to determine in which region the derivative of g is positive or negative. These regions correspond to the monotony regions of g. The entry "+" says that the considered term is positive on the given interval. If the term is negative, then "−" is entered.

x	$x < -3$	$-3 < x < 1$	$1 < x$
$x + 3$	−	+	+
$x - 1$	−	−	+
$g'(x)$	+	−	+
g is monotonically	increasing	decreasing	increasing

For the function $h : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ with $h(x) = \frac{1}{x}$, we have $h'(x) = -\frac{1}{x^2}$, that is $h'(x) < 0$ for all $x \neq 0$.

Even though the function h exhibits the same monotony behaviour for the two subregions $x < 0$ and $x > 0$, it is not monotonically decreasing on the entire region. As a counterexample let us consider the function values $h(-2) = -\frac{1}{2}$ and $h(1) = 1$. Here, we have $-2 < 1$ but also $h(-2) < h(1)$. This corresponds to an increasing growth behaviour if we change from one subregion to the other. The statement that the function h is monotonically decreasing on $]-\infty; 0[$ thus means that the restriction of h on this interval is monotonically decreasing. Moreover, the function h is also monotonically decreasing for all $x > 0$.

7.4.3 Second Derivative and Bending Properties (Curvature)

Let us consider a function $f : D \to \mathbb{R}$ that is differentiable on the interval $]a; b[\subseteq D$. If its derivative f' is also differentiable on the interval $]a; b[\subseteq D$, then f is called **twice-differentiable**. The derivative of the first derivative of f ($(f')' = f''$) is called the **second derivative** of the function f.

The second derivative of the function f can be used to investigate the bending behaviour (curvature) of the function:

Bending Properties (Curvature)7.4.2

If $f''(x) \geq 0$ for all x between a and b, then f is called **convex** (**left curved** or **concave up**) on the interval $]a; b[$.

If $f''(x) \leq 0$ for all x between a and b, then f is called **concave** (**right curved** or

concave down) on the interval $]a;b[$.

Thus, it is sufficient to determine the sign of the second derivative f'' to decide whether a function is convex (left curved) or concave (right curved).

Comment on the Notation7.4.3

The second derivative and further "higher" derivatives are often denoted using superscript natural numbers in round brackets: $f^{(k)}$ then denotes the kth derivative of f. In particular, this notation is used in generally written formulas even for the (first) derivative ($k = 1$) and for the function f itself ($k = 0$).

Hence,

- $f^{(0)} = f$ denotes the function f,
- $f^{(1)} = f'$ denotes the (first) derivative,
- $f^{(2)} = f''$ the second derivative,
- $f^{(3)}$ the third derivative, and
- $f^{(4)}$ the fourth derivative of f.

This list can be continued as long as the derivatives of f exist.

The following example shows that a monotonically increasing function can be convex on one region and concave on another.

Example 7.4.4

Certainly, the function $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^3$ is at least twice-differentiable. Since $f'(x) = 3x^2 \geq 0$ for all $x \in \mathbb{R}$, the function f is monotonically increasing on its entire domain. Moreover, we have $f''(x) = 6x$. Thus, for all $x < 0$, we also have $f''(x) < 0$ and hence, the function f is concave (right curved) on this region. For $x > 0$, we have $f''(x) > 0$. Hence, for $x > 0$, the function f is convex (left curved).

7.4.4 Exercises

Exercise 7.4.1
Specify the (maximum) open intervals on which the function f with $f(x) := \frac{x^2-1}{x^2+1}$ is monotonically increasing or decreasing.

Answer:

- f is monotonically [ ] on $]-\infty;0[$.
- f is monotonically [ ] on $]0;\infty[$.

Solution:
The derivative f' of the function f is given by

$$f'(x) = \frac{2x\cdot(x^2+1)-(x^2-1)\cdot 2x}{(x^2+1)^2} = \frac{2x\left(x^2+1-x^2+1\right)}{(x^2+1)^2} = \frac{4x}{(x^2+1)^2}\ .$$

Since the denominator of $f'(x)$ is always positive, the sign of $f'(x)$ is determined solely by the numerator. For all negative $x \in \mathbb{R}$, we have $f'(x) < 0$, and hence the function f is monotonically decreasing on this region. In contrast, for all positive $x \in \mathbb{R}$, we have $f'(x) > 0$, and hence the function f is monotonically increasing on this region.

Exercise 7.4.2
Specify the (maximum) open intervals $]c;d[$ on which the function f with $f(x) := \frac{x^2-1}{x^2+1}$ for $x > 0$ is convex or concave. Answer:

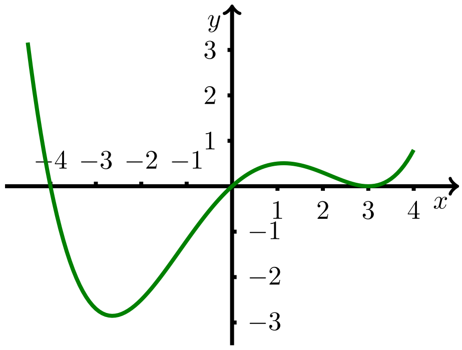
- The function f is convex on [ ] .
- The function f is concave on [ ] .

Solution:
From the quotient rule, we have for the first and the second derivative of f:

$$\begin{aligned} f'(x) &= \frac{4x}{(x^2+1)^2}\ , \\ f''(x) &= \frac{4\cdot(x^2+1)^2-4x\cdot 2(x^2+1)\cdot 2x}{(x^2+1)^4} = \frac{4x^2+4-16x^2}{(x^2+1)^3} = \frac{4(1-3x^2)}{(x^2+1)^3}\ . \end{aligned}$$

Since $4/(x^2+1)^3$ is always positive, the sign of $f''(x)$ is only determined by the factor $(1-3x^2)$. The roots of $f''(x)$ are at $x_0 = \pm\frac{1}{\sqrt{3}}$. Thus, for $x > 0$, the derivative $f''(x)$ is greater than 0 on the open interval $\left]0;\frac{1}{\sqrt{3}}\right[$, so the function f is convex (left curved) on this region. On the interval $\left]\frac{1}{\sqrt{3}};\infty\right[$, we have $f''(x) < 0$; the function f is concave (right curved) on this region.

Exercise 7.4.3
Consider the function $f : [-4.5; 4] \to \mathbb{R}$ with $f(0) := 2$. Its derivative f' has the graph shown in the figure below:



a. Where is the function f monotonically increasing and where it is monotonically decreasing? Find the maximum open intervals $]c; d[$ on which f has this property.

b. What can you say about the maximum and minimum points of the function f?

Answer:

- The function f is monotonically [ ] on $]-4.5;$ [ ] $[$.
- The function f is monotonically [ ] on $]$ [ ] $;0[$.
- The function f is monotonically [ ] on [ ] .
- The function f is monotonically [ ] on $]3;4[$.

The maximum point of f is at [ ] . The minimum point of f is at [ ]
.

Solution:
The monotony behaviour is determined by the derivative f' of the function f. Since the graph of the derivative f' is given in the exercise, we only have to read off on which intervals the graph lies above (or below) the x-axis: On the intervals $[-4.5; -4[$, $]0; 3[$, and $]3; 4[$, we have $f'(x) > 0$, and hence, the function f is monotonically increasing there. However, on the interval $]-4; 0[$, we have $f'(x) < 0$, and hence the function f is monotonically decreasing there.
At an extremal point x_e (maximum or minimum point) of a function f (which does not lie on the boundary of the domain) the first derivative is zero: $f'(x_e) = 0$. Graphically, this means that the tangent line to the graph of f is a horizontal line. According to the graph, the zeros of $f'(x)$ are at $x_1 = -4$, $x_2 = 0$, and $x_3 = 3$. Since f is monotonically increasing on $]-4.5; -4[$ and monotonically decreasing on $]-4; 0[$, $x_1 = -4$ is a maximum

point. This tells us that the function has a minimum point at $x_2 = 0$. (At $x_3 = 3$ the function has a saddle point.)

7.5 Applications

7.5.1 Curve Analysis

Curve analysis is an established part of the German mathematics syllabus that consists of curve sketching for a function together with the gathering of certain standardized qualitative and quantitative information about the function graph.

Let a differentiable function $f :]a; b[\to \mathbb{R}$ with the mapping rule $x \mapsto y = f(x)$ for $x \in]a; b[$ be given. In this course, a complete curve analysis of f includes the following information:

- maximum domain
- x- and y-intercepts of the graph
- symmetry of the graph
- limiting behaviour/asymptotes
- first derivatives
- extremal values
- monotony behaviour
- inflexion points
- bending behaviour (curvature)
- sketch of the graph

Many of these points were already discussed in Module 6 auf Seite 217. Therefore, in this section we shall only briefly repeat what is meant by the different steps of curve analysis. Subsequently, we will discuss one example of a curve analysis in detail.

The first part of the curve analysis involves the algebraic and geometric aspects of f:

Maximum Domain All real numbers x for which $f(x)$ exists are determined. The set D of these numbers is called the maximum domain.

x- and y-Intercepts

- x-axis: All zeros of f are determined.
- y-axis: The function value $f(0)$ (if $0 \in D$) is calculated.

Symmetry of the Graph The graph of the function is symmetrical with respect to the y-axis if $f(-x) = f(x)$ for all $x \in D$. Then the function f is called **even**. If $f(-x) = -f(x)$ for all $x \in D$, the graph is centrally symmetric with respect to the origin $(0;0)$ of the coordinate system. In this case, the function is called **odd**.

Asymptotic Behaviour at the Domain Boundary The limits of the function f at the boundaries of its domain are investigated.

In the second part, the function is investigated analytically by means of conclusions from the first derivatives. Of course, the first and the second derivative have to be calculated first, provided they exist.

Derivatives Calculation of the first and the second derivative (if they exist).

Extremal Values and Monotony The necessary condition for x to be an extremum (if $x \in D$ is not a boundary point of D), is $f'(x) = 0$.
Thus, we calculate the points x_0 at which the derivative f' takes the value zero. If at these points also the second derivative exists, then we have:

- $f''(x_0) > 0$: x_0 is a minimum point of f.
- $f''(x_0) < 0$: x_0 is a maximum point of f.

The function f is monotonically increasing on that intervals of the domain on which we have $f'(x) \geq 0$. It is monotonically decreasing on that intervals where $f'(x) \leq 0$.

Inflexion Points and Curvature Properties The necessary condition for x to be an inflexion point (if the second derivative f'' exists), is $f''(x) = 0$
If $f''(w_0) = 0$ and $f^{(3)}(w_0) \neq 0$, then w_0 is an inflexion point, i.e. the bending behaviour of f changes at this point.
The function f is convex (left curved) on that intervals of the domain in which we have $f^{(2)}(x) \geq 0$. It is concave (right curved) on that intervals on which we have $f^{(2)}(x) \leq 0$.

Sketch of the Graph A sketch of the graph is drawn based on the information gained during the previous steps.

7.5.2 Detailed Example

We investigate a function f defined by the mapping rule

$$f(x) = \frac{4x}{x^2+2} .$$

Maximum Domain
The maximum domain of this function is $D_f = \mathbb{R}$ since the denominator $x^2 + 2 \geq 2$, i.e. it is always non-zero, and hence no points have to be excluded.

x- and y-Intercepts

The zeros of the function are the zeros of the numerator. Hence, the graph of f intersects the x-axis only at the origin $(0;0)$ since the numerator is only zero for $x = 0$. This is also the only intersection point with the y-axis since there we have $f(0) = 0$.

Symmetry
To investigate the symmetry we replace the argument x by $(-x)$. We have

$$f(-x) = \frac{4\cdot(-x)}{(-x)^2+2} = -\frac{4x}{x^2+2} = -f(x)$$

for all $x \in \mathbb{R}$. Hence, the graph of the function f is centrally symmetric with respect to the origin.

Limiting Behaviour
The function is defined on the entire set of real numbers $\mathbb{R}$, so only the limiting behaviour for $x \to \infty$ and $x \to -\infty$ has to be investigated. Since $f(x)$ is a fraction of two polynomials and the denominator has a greater power than the numerator, the x-axis is a horizontal asymptote in both directions:

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

Derivatives
The first two derivatives of the function are calculated from the quotient rule. For the first derivative, we have:

$$f'(x) = 4\cdot\frac{1\cdot(x^2+2) - x\cdot 2x}{(x^2+2)^2} = 4\cdot\frac{-x^2+2}{(x^2+2)^2}\ .$$

Taking the derivative again and simplifying the terms results in:

$$\begin{aligned} f''(x) &= 4\cdot\frac{-2x(x^2+2)^2 - (-x^2+2)\cdot 2(x^2+2)\cdot 2x}{(x^2+2)^4} \\ &= 4\cdot\frac{-2x(x^2+2) - (-x^2+2)\cdot 4x}{(x^2+2)^3} \\ &= 4\cdot\frac{-2x^3 - 4x + 4x^3 - 8x}{(x^2+2)^3} \\ &= 4\cdot\frac{2x^3 - 12x}{(x^2+2)^3} \\ &= 8\cdot\frac{x(x^2-6)}{(x^2+2)^3}\ . \end{aligned}$$

Extremal Values
The necessary condition for an extremum at x is $f'(x) = 0$, in this case $-x^2+2 = 0$. Thus,

we obtain $x_1 = \sqrt{2}$ and $x_2 = -\sqrt{2}$. In addition, we have to investigate the behaviour of the second derivative at these points:

$$f''(x_1) \;=\; 8\frac{\sqrt{2}\cdot(2-6)}{(2+2)^3} < 0\,, \quad f''(x_2) \;=\; 8\frac{(-\sqrt{2})\cdot(2-6)}{(2+2)^3} > 0\,.$$

Hence, x_1 is a maximum point and x_2 is a minimum point of f. Inserting these values for x into f results in the maximum point $\left(\sqrt{2};\sqrt{2}\right)$ and the minimum point $\left(-\sqrt{2};-\sqrt{2}\right)$ of f.

Monotony Behaviour

Since f is defined on the entire set of real numbers $\mathbb{R}$, the monotony behaviour can be derived from the position of the extremal points and their types: f is monotonically decreasing on $\left]-\infty;-\sqrt{2}\right[$, monotonically increasing on $\left]-\sqrt{2};\sqrt{2}\right[$, and monotonically decreasing on $\left]\sqrt{2};\infty\right[$. Monotony intervals are always given as open intervals.

Inflexion Points

From the necessary condition $f''(x) = 0$ for x to be an inflexion point, we have the equation $8x(x^2-6) = 0$. Thus, $w_0 = 0$, $w_1 = \sqrt{6}$, and $w_2 = -\sqrt{6}$ are the only solutions. The polynomial in the denominator of f'' is always greater than zero. Since the polynomial in the numerator has only single roots, the second derivative $f''(x)$ changes its sign at all these points. Hence, these points are inflexion points of f. The coordinates of the inflexion points $(0;0)$, $\left(\sqrt{6};\frac{1}{2}\sqrt{6}\right)$, $\left(-\sqrt{6};-\frac{1}{2}\sqrt{6}\right)$ are determined by inserting the corresponding values for x into f.

Bending Behaviour

The twice-differentiable function f is convex if the second derivative is greater or equal to zero. It is concave if the second derivative is less or equal to zero. Since the polynomial in the denominator of $f''(x)$ is always positive, it is sufficient to examine the sign of the polynomial $p(x) = 8x(x-\sqrt{6})(x+\sqrt{6})$ in the numerator. For $0 < x < \sqrt{6}$, it is negative (f is concave there). For $x > \sqrt{6}$ it is positive (f is convex there). Since f is centrally symmetric, it follows that f is convex on the intervals $\left]-\sqrt{6};0\right[$ and $\left]\sqrt{6};\infty\right[$ and concave on the intervals $\left]-\infty;-\sqrt{6}\right[$ and $\left]0;\sqrt{6}\right[$.

Sketch of the Graph

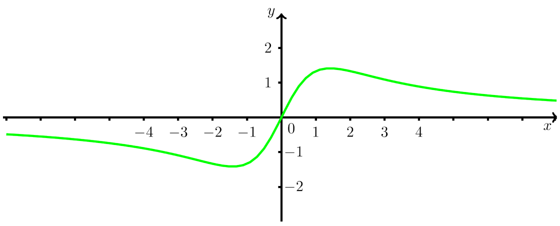


Abbildung 1: The graph of the function f, sketched on the interval $[-8;8]$.

7.5.3 Exercises

With the following exercise the elements of the curve analysis method can be trained:

`In the online version, exercises from an exercise list will be shown here`

Exercise 7.5.1
Carry out a complete curve analysis for the function f with $f(x) = (2x - x^2)e^x$ and enter your results into the input fields.

Maximum domain: [ ] .

Set of intersection points with the x-axis (zeros of $f(x)$): [ ] .

The y-intercept is at y = [ ] .

Symmetry: The function is
[ ] axially symmetric with respect to the y-axis,
[ ] centrally symmetric with respect to the origin.

Limiting behaviour: For $x \to \infty$, the functions values $f(x)$ tend to [ ] , and for $x \to -\infty$, they tend to [ ] .

Derivatives: We have $f'(x) =$ [ ] and $f''(x) =$ [ ] .

Monotony behaviour: The function is monotonically increasing on the interval [ ] and monotonically decreasing otherwise.

Extremal values: The point $x_1 =$ [ ] is a minimum point and the point $x_2 =$ [ ] is a maximum point.

Inflexion points: The set of inflexion points consists of [ ] .

Sketch the graph and compare your result to the sample solution.

Solution:
Maximum Domain
We have $f(x) = -x(x-2)e^x$ and $e^x > 0$ for all $x \in \mathbb{R}$; thus, $D_f = \mathbb{R} =]-\infty; \infty[$ is the maximum domain.

x- and y-Intercepts
The intersection points with the x-axis (roots of the function) lie at $x_1 = 0$ and $x_2 = 2$, i.e. the coordinates of the points are $(0;0)$ and $(2;0)$. The y-intercept is the point $(0;0)$.

Symmetry
The function f is neither even nor odd, and hence the graph of f is neither axially symmetric with respect to the y-axis nor centrally symmetric with respect to the origin.

Limiting Behaviour
Since the function is defined for all real numbers, only the asymptotes for $\pm\infty$ have to be investigated:

$$\lim_{x\to\infty} -x(x-2)e^x \;=\; -\infty \text{ and } \lim_{x\to-\infty} -x(x-2)e^x \;=\; 0\;.$$

Hence, $y = 0$ is an asymptote for $x \to -\infty$.

Derivatives
The first two derivatives of f are

$$\begin{aligned} f'(x) &= (2-2x)e^x + (2x-x^2)e^x = (2-x^2)e^x = -(x^2-2)e^x\;, \\ f''(x) &= -2xe^x + (2-x^2)e^x = -(x^2+2x-2)e^x\;. \end{aligned}$$

Monotony Behaviour and Extremal Values
The solutions of $f'(x) = 0$ are $x_1 = -\sqrt{2}$ and $x_2 = \sqrt{2}$. Furthermore, we have $x_1 < x_2$ and

$$f'(x) \;=\; -(x+\sqrt{2})(x-\sqrt{2})e^x\;.$$

On $\left]-\infty;-\sqrt{2}\right[$ the first derivative f' is negative, so f is monotonically decreasing there. On $\left]-\sqrt{2};\sqrt{2}\right[$ the first derivative f' is positive, hence f is monotonically increasing there. On $\left]\sqrt{2};\infty\right[$ the first derivative f' is negative, so f is monotonically decreasing there. Thus, $x_1 = -\sqrt{2}$ is a minimum point, and $x_2 = \sqrt{2}$ is a maximum point.

Inflexion Points
The necessary condition $f''(x) = 0$ for x to be an inflexion point results in the quadratic equation $x^2 + 2x - 2 = 0$. It has the solutions $w_1 = \frac{-2-\sqrt{4+8}}{2} = -1-\sqrt{3}$ and $w_2 = \frac{-2+\sqrt{4+8}}{2} = -1+\sqrt{3}$.

Sketch of the Graph

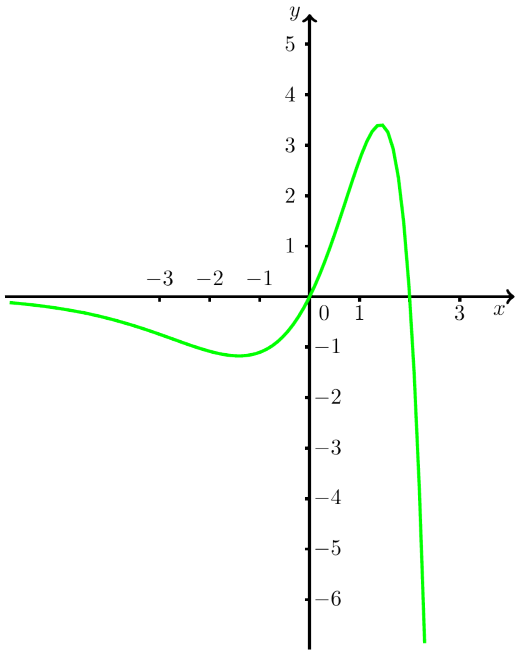


Abbildung 2: Graph of the function f, sketched on the interval $]-6.2;3[$.

7.5.4 Optimisation Problems

In many applications in engineering and business, solutions to problems can be found which are not unique. They often depend on variable conditions. To find an ideal solution, additional properties (constraints) are defined that are to be satisfied by the solution. This very often results in so-called **optimisation problems**, in which one solution has to be selected from a family of solutions such that it best satisfies a previously specified property.

As an example, we consider the problem of constructing a cylindrical can. This can must satisfy the additional condition of having a capacity (volume) V of one litre (a.k.a. one cubic decimetre, $1\,\mathrm{dm}^3$). Thus, if V is specified in dm^3 and r is the radius and h the height of the can, each measured in decimetre (dm), then the volume is $V = \pi r^2 \cdot h = 1$. The can with the least surface area $O = 2 \cdot \pi r^2 + 2\pi r h$ is required in order to save material. Here, the surface area O, measured in square decimetres (dm^2), is a function of the radius r and the height h of the can.

In mathematical terms, our question results in the problem of finding a minimum of the surface function O, where the minimum has to be found for values of r and h that also satisfy the additional condition for the volume: $V = \pi r^2 \cdot h = 1$. In the context of finding extrema, such an additional condition is also called a **constraint**.

Optimisation Problem7.5.1

In an **optimisation problem**, we search for an extremum x_{ext} of a function f satisfying a given equation $g(x_{\text{ext}}) = b$.

If we search for a minimum point, this problem is called a **minimisation problem**. If we search for a maximum point, this problem is called a **maximisation problem**.

The function f is called the **target function**, and the equation $g(x) = b$ is called the **constraint** of the optimisation problem.

7.5.5 Example

Let us consider the example above in more detail. Obviously, the problem is to minimise the surface area of a cylindrical can with a given volume (base multiplied by height):

$$V = \pi r^2 h = 1 \ ,$$

where r is the radius of the base and h is the height of the can. The surface area consists of the lid and the base (both with an area of πr^2) and the lateral surface (with an area of $2\pi r h$), which results in the equation $O = 2\pi r^2 + 2\pi r h$ for the surface area of the can. The surface area of the can is a function of the radius r and the height h. In contrast, a fixed volume (constraint) is assigned to the volume. Thus, it can be written:

$$O\left(r, h\right) = 2\pi r^2 + 2\pi r h \ .$$

Due to the constraint $V = \pi r^2 h = 1$, this problem that initially involves two variables (r and h) can be reduced to a problem that only involves one variable. Solving the constraint for the height of the can results in:

$$\begin{aligned} \pi r^2 h &= 1 \\ \Leftrightarrow \quad h &= \frac{1}{\pi r^2} \ . \end{aligned}$$

Substituting this formula into the function $O(r, h)$ results in a function that only depends on one variable. This function is also called O **for simplicity**:

$$O\left(r, h\right) = 2\pi r^2 + 2\pi r h = 2\pi r^2 + 2\pi r \frac{1}{\pi r^2} = 2\left(\pi r^2 + \frac{1}{r}\right) = O\left(r\right) \ .$$

After this manipulation, the problem of finding the cans minimal surface area can be solved analogously to normal extremal value problems of functions. Thus, we take the first derivative of the function O with respect to the variable r and set this derivative equal to zero:

$$\begin{aligned} O'(r) = 2\left(2\pi r - \frac{1}{r^2}\right) &= 0 \\ \Leftrightarrow \quad 2\pi r &= \frac{1}{r^2} \\ \Leftrightarrow \quad 2\pi r^3 &= 1 \\ \Leftrightarrow \quad r^3 &= \frac{1}{2\pi} \\ \Leftrightarrow \quad r &= \sqrt[3]{\frac{1}{2\pi}}\ . \end{aligned}$$

The last equivalent transformation used the fact that the radius r cannot take negative values. Substituting this result into the second derivative of O shows whether a minimum was actually found ($O''(r) = 4\pi + 4/r^3$):

$$O''\left(\sqrt[3]{\frac{1}{2\pi}}\right) = 4\pi + \frac{4}{\left(\sqrt[3]{\frac{1}{2\pi}}\right)^3} = 12\pi > 0\ .$$

For the radius $r = \sqrt[3]{\frac{1}{2\pi}}$, the surface area of the cylindrical can with the given volume $V = 1$ is a minimum. The corresponding height of the can is $h = \frac{1}{\pi\left(\sqrt[3]{\frac{1}{2\pi}}\right)^2} = \sqrt[3]{\frac{4}{\pi}}$. If a can of these dimensions is manufactured, the material usage for the given volume is minimised.

7.6 Final Test

7.6.1 Final Test Module 6

Exercise 7.6.1
In a container at 9 a.m. a temperature of $-10°C$ is measured. At 3 p.m. the measured temperature is $-58°C$. After a period of 14 hours, the temperature has fallen to $-140°C$.

a. What is the average rate of temperature change between the first and second measurements?

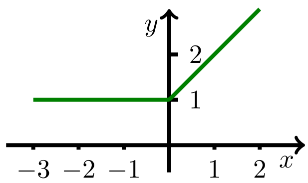
   Answer: [ ]

b. The 'falling' property of the temperature shows in the fact that the rate of change is [ ] .
   Hint:
   Enter an adjective.

c. Calculate the average rate of temperature change for the whole measuring period.

   Answer: [ ]

Exercise 7.6.2
A function $f : [-3; 2] \to \mathbb{R}$, $x \mapsto f(x)$ has a first derivative f' whose graph is shown in the figure below.



The function values of f between -3 and 0

- [ ] are constant,
- [ ] increase by 3,
- [ ] decrease.

At the point $x = 0$ the function f has

- [ ] a jump,
- [ ] no derivative,
- [ ] a derivative of 1.

Exercise 7.6.3
Calculate for the function

a. $f : \{x \in \mathbb{R} : x > 0\} \to \mathbb{R}$ with $f(x) := \ln\left(x^3 + x^2\right)$ the value of the first derivative f' at x:
$f'(x) =$ [ ] .

b. $g : \mathbb{R} \to \mathbb{R}$ with $g(x) := x \cdot e^{-x}$ the value of the second derivative g'' at x:
$g''(x) =$ [ ] .

Exercise 7.6.4
Consider the function $f :]0;\infty[\to \mathbb{R}$, $x \mapsto f(x)$ with $f'(x) = x \cdot \ln x$. On which regions is f monotonically decreasing, and on which regions is f concave? Specify the regions as open intervals $]c;d[$ that are as large as possible:

a. f is monotonically decreasing on [ ] .

b. f is concave on [ ] .

8 Integral Calculus

Module Overview

8.1 Antiderivatives

8.1.1 Introduction

In the previous module we studied derivatives of functions. As for every other arithmetic operation,the question of finding an inverse operation arises. For example, subtraction is the inverse operation of addition, and division is the inverse operation of multiplication. The question of the inverse operation of differentiation leads to the introduction of integral calculus and the definition of an antiderivative. The relation between derivative and antiderivative can be explained very easily. If a derivative f' can be assigned to a function f, and the derivative f' is also considered as a function, then the function f could be assigned to this function f' by inverting the operation of "differentiation". Thus, in this chapter the question is: for a given function f, can we find another function with f as its derivative?

The applications of integral calculus are as varied as the applications of differential calculus. In physics, for example, the force F acting on a mass m may be investigated. From the well-known relation $F = ma$ (a being the acceleration of the object), the acceleration $a = F/m$ can be calculated from the force. If the acceleration is interpreted as the rate of velocity change, i.e. $a = \frac{dv}{dt}$, then the velocity can be determined subsequently from the inverse of the derivative – from integral calculus. Similar relations can be found in many fields of science and engineering, and also in economics. Integral calculus is used for the calculation of areas, centres of mass, bending properties of beams or the solutions of so-called differential equations, which are used so frequently in science and engineering.

8.1.2 Antiderivatives

In the context of this course we will discuss integral calculus for functions on "connected domains", which are of particular significance for many practical applications. In mathematical terms, the domains of the functions will be intervals. As the inverses of derivatives, antiderivatives will be also defined on intervals.

Antiderivative8.1.1

Let an interval $D \subseteq \mathbb{R}$ and a function $f : D \to \mathbb{R}$ be given. If there exists a differentiable function $F : D \to \mathbb{R}$ that has f as its derivative, i.e. $F'(x) = f(x)$ for all $x \in D$, then F is called an **antiderivative** of f.

Let us first consider a few examples.

Example 8.1.2

The function F with $F(x) = -\cos(x)$ has the derivative

$$F'(x) = -(-\sin(x)) = \sin(x) .$$

Thus, F is an antiderivative of f with $f(x) = \sin(x)$.

Example 8.1.3

The function G with $G(x) = \frac{1}{3}e^{3x+7}$ has the derivative

$$G'(x) = \frac{1}{3} \cdot 3 \cdot e^{3x+7} .$$

Hence, G is an antiderivative of g with $g(x) = e^{3x+7}$.

Next we will consider another very simple example, which illustrates an important point to note when calculating antiderivatives.

Example 8.1.4

Let a constant function H with the function value $H(x) = 18$ be given on an interval. Then the function H has the derivative

$$H'(x) = 0 .$$

Hence, H is an antiderivative of h with $h(x) = 0$.

The last example is a little surprising because the derivative of a constant function is the zero function. Thus, *every* constant function F is an antiderivative of f with $f(x) = 0$ on an interval, i.e. $F(x)$ is equal to any number C for every value of x. However, the antiderivative $F(x)$ cannot be any other function than a *constant* one if f is defined on an interval.

All Derivatives of the Zero Function8.1.5

The function F is an antiderivative of f with $f(x) = 0$ on an interval if and only if F is a constant function, i.e. if a real number C exists such that $F(x) = C$ for all values of x in the interval.

If the functions F and G have the same derivative, i.e. $f = F' = G'$, then we have $G'(x) - F'(x) = 0$. Taking the antiderivative on an interval on both sides of the equation results in the relation $G(x) - F(x) = C$. Thus, we have $G(x) = F(x) + C$. Therefore, if F is an antiderivative of f, then G with $G(x) = F(x) + C$ is also an antiderivative of f.

Statement on Antiderivatives8.1.6

If F and G are antiderivatives of $f : D \to \mathbb{R}$ on an interval D, then there exists a real number C such that

$$F(x) = G(x) + C \qquad \text{for all } x \in D\,.$$

This is also written as

$$\int f(x)\,dx = F(x) + C,$$

to express how all antiderivatives of f look.

The set of all antiderivatives is also called the **indefinite integral** and is written according to the statement above as

$$\int f(x)\,dx = F(x) + C\ ,$$

where F is any antiderivative of f.

This notation of the indefinite integral emphasises that it is a function F with $F' = f$ that is calculated for a given function f. How this expression is used to calculate the (definite) integral of a continuous function f is described by the **fundamental theorem of calculus** discussed in the next section in Info Box 8.2.3 auf Seite 350.

How do we know the value of this constant C? If we only look for an antiderivative of f with $f(x) = 0$ on an interval without knowing any other conditions, then the constant C is indefinite. C is only definite if an additional function value $y_0 = F(x_0)$ of F at a point x_0 is given.

Example 8.1.7

For example, for f with $f(x) = 2x + 5$, we have

$$\int (2x+5)\,dx = x^2 + 5x + C\ .$$

If we look for the antiderivative F of f with $F(0) = 6$, then we set $6 = F(0) = 0^2 + 5 \cdot 0 + C = C$ and hence, $C = 6$. Thus, the antiderivative is in this case $F(x) = x^2 + 5x + 6$.

If the relation between the derivative $f = F'$ and the antiderivative F is written in the way discussed above for the types of functions considered so far, then one obtains the following table:

A Small Table of Antiderivatives8.1.8

The functions f are considered on an interval. The antiderivatives of these functions are given as an indefinite integral:

Function f	Antiderivatives F
$f(x) = 0$	$F(x) = \int 0\,dx = C$
$f(x) = x^n$	$F(x) = \int x^n\,dx = \frac{1}{n+1} \cdot x^{n+1} + C$
$f(x) = \sin(x)$	$F(x) = \int \sin(x)\,dx = -\cos(x) + C$
$f(x) = \sin(kx)$	$F(x) = \int \sin(kx)\,dx = -\frac{1}{k}\cos(kx) + C$
$f(x) = \cos(x)$	$F(x) = \int \cos(x)\,dx = \sin(x) + C$
$f(x) = \cos(kx)$	$F(x) = \int \cos(kx)\,dx = \frac{1}{k}\sin(kx) + C$
$f(x) = e^x$	$F(x) = \int e^x\,dx = e^x + C$
$f(x) = e^{kx}$	$F(x) = \int e^{kx}\,dx = \frac{1}{k}e^{kx} + C$
$f(x) = x^{-1} = \frac{1}{x}$	$F(x) = \int \frac{1}{x}\,dx = \ln\lvert x\rvert + C$ for $x \neq 0$

Here, k and C denote arbitrary real numbers with $k \neq 0$, and n is an integer with $n \neq -1$.

The next example shows how the table is used.

Example 8.1.9

Find the indefinite integral of the function f with $f(x) = 10x^2 - 6 = 10x^2 - 6x^0$.

From the table above we read off the antiderivatives of g with $g(x) = x$ and h with $h(x) = x^0 = 1$: The function G with $G(x) = \frac{1}{1+1} \cdot x^{1+1} = \frac{1}{2} \cdot x^2$ is an antiderivative of g, and the function H with $H(x) = \frac{1}{0+1} \cdot x^{0+1} = x$ is an antiderivative of h. Thus, the function $F : \mathbb{R} \to \mathbb{R}$ with

$$F(x) = 10 \cdot \frac{1}{2}x^2 - 6 \cdot x = 5x^2 - 6x$$

is an antiderivative of f. We see that

$$\int (10x - 6)\, dx = 5x^2 - 6x + C$$

describes the set of antiderivatives of $f : \mathbb{R} \to \mathbb{R}$ with $f(x) = 10x - 6$, where C is an arbitrary real number.

The notation using the constant C expresses that, for example, $G : \mathbb{R} \to \mathbb{R}$ with $G(x) := 5x^2 - 6x - 7$ is also an antiderivative of f, where $C = -7$, since $G'(x) = 5 \cdot 2x - 6 = f(x)$ for all $x \in \mathbb{R}$.

In table books the antiderivatives are generally listed neglecting the constants. However, for calculations it is necessary to state that several functions differing by a constant can exist. In solving problems of applied mathematics, the constant C is often determined by additional conditions, such as a given function value of the antiderivative.

Practical Note8.1.10

It is very easy to check whether the antiderivative of a given function f was found correctly. Take the derivative of the found antiderivative and compare it to the initially given function f. If the functions coincide, then the calculation was correct. If the result does not coincide with the function f, then the antiderivative has to be checked again.

8.1.3 Exercises

Exercise 8.1.1
Specify an antiderivative:

a. $\int \left(12x^2 - 4x^7\right)\, dx \; = \;$ [ ] .

Solution:
We have
$$\int \left(12x^2 - 4x^7\right)\, dx = 12 \cdot \frac{1}{3} \cdot x^3 - 4 \cdot \frac{1}{8} \cdot x^8 + C = 4 \cdot x^3 - \frac{1}{2} \cdot x^8 + C\,.$$

b. $\int (\sin(x) + \cos(x))\, dx \; = \;$ [ ] .

Solution:
We have
$$\int (\sin(x) + \cos(x))\, dx = -\cos(x) + \sin(x) + C = \sin(x) - \cos(x) + C\,.$$

c. $\int \frac{1}{6\sqrt{x}}\, dx \; = \;$ [ ] .

Solution:
We have $\frac{1}{6\sqrt{x}} = \frac{1}{6} \cdot \frac{1}{\sqrt{x}} = \frac{1}{6} \cdot x^{-\frac{1}{2}}$. Hence,
$$\int \frac{1}{6\sqrt{x}}\, dx = \frac{1}{6} \cdot \int x^{-\frac{1}{2}}\, dx = \frac{1}{6} \cdot \frac{1}{-\frac{1}{2}+1} \cdot x^{-\frac{1}{2}+1} + C = \frac{1}{6} \cdot 2 \cdot x^{\frac{1}{2}} + C = \frac{1}{3} \cdot \sqrt{x} + C\,.$$

Exercise 8.1.2
Find an antiderivative:

a. $\int e^{x+2}\, dx \; = \;$ [ ] .

Solution:
We have $\int e^{x+2}\, dx = \int e^x \cdot e^2\, dx = e^2 \cdot \int e^x\, dx = e^2 \cdot e^x + C = e^x \cdot e^2 + C = e^{x+2} + C.$

b. $\int 3x \cdot \sqrt[4]{x}\, dx \; = \;$ [ ] .

Solution:
We have $\int 3x \cdot \sqrt[4]{x}\, dx = 3 \cdot \int x \cdot x^{\frac{1}{4}}\, dx = 3 \cdot \int x^{\frac{5}{4}}\, dx = 3 \cdot \frac{4}{9} \cdot x^{\frac{9}{4}} + C = \frac{4}{3} \cdot x^{\frac{9}{4}} + C.$

Exercise 8.1.3
Decide whether the following statements about real-valued functions are true.

True:	False:	Statement:
☐	☐	F with $F(x) = -\frac{\cos(\pi x)+2}{\pi}$ is an antiderivative of f with $f(x) = \sin(\pi x) + 2$.
☐	☐	F with $F(x) = -\frac{\cos(\pi x)+2}{\pi}$ is an antiderivative of f with $f(x) = \sin(\pi x)$.
☐	☐	F with $F(x) = -7$ is an antiderivative of f with $f(x) = -7x$ for $x \in \mathbb{R}$.
☐	☐	F with $F(x) = (\sin(x))^2$ is an antiderivative of f with $f(x) = 2\sin(x)\cos(x)$.
☐	☐	If F is an antiderivative of f, and G is an antiderivative of g, then $F + G$ is an antiderivative of $f + g$.

Solution:

- The derivative of F with $F(x) = -\frac{\cos(\pi x)+2}{\pi}$ is $F'(x) = \sin(\pi x) \neq \sin(\pi x) + 2 = f(x)$. Therefore F is not an antiderivative of f.
- The derivative of F with $F(x) = -\frac{\cos(\pi x)+2}{\pi}$ is $F'(x) = \sin(\pi x) = f(x)$. F is an antiderivative of f.
- The derivative of F with $F(x) = -7$ is $F'(x) = 0 \neq -7x = f(x)$ (for $x \neq 0$). Hence, F is not an antiderivative of f.
- The derivative of F with $F(x) = (\sin(x))^2$ is, using the chain rule, $F'(x) = 2 \cdot \sin(x) \cdot \cos(x) = f(x)$. Thus, F is an antiderivative of f.
- If F is an antiderivative of f, and G is an antiderivative of g, then F and G are differentiable, where $F' = f$ and $G' = g$. Thus, $F + G$ is differentiable, and we have $(F + G)'(x) = F'(x) + G'(x) = f(x) + g(x) = (f + g)(x)$, i.e. $F + G$ is an antiderivative of $f + g$.

Exercise 8.1.4
Find an antiderivative of

a. $f(x) := \dfrac{8x^3 - 6x^2}{x^4}$

b. $g(x) := \dfrac{18x^2}{3\sqrt{x^5}}$

c. $h(x) := \dfrac{x + 2\sqrt{x}}{4x}$

for $x > 0$, after rewriting the terms as reduced sums of fractions:

a. With the simplification $f(x) =$ ☐
we have the antiderivative $F(x) =$ ☐ for $x > 0$.

Solution:
The function can be rewritten as $f(x) = \frac{8}{x} - \frac{6}{x^2} = \frac{8}{x} - 6x^{-2}$ which results in the

antiderivative

$$F(x) = 8 \cdot \ln(x) + \frac{6}{x} + C$$

for $x > 0$ and $C \in \mathbb{R}$.

b. With the simplification $g(x) =$ [ ]
we have the antiderivative $G(x) =$ [ ] for $x > 0$.

Solution:
The function can be rewritten as $g(x) = \frac{6}{\sqrt{x}} = 6 \cdot x^{-\frac{1}{2}}$ which results in the antiderivative

$$G(x) = 6 \cdot 2 \cdot \sqrt{x} + C = 12 \cdot \sqrt{x} + C$$

for $x > 0$ and $C \in \mathbb{R}$.

c. With the simplification $h(x) =$ [ ]
we have the antiderivative $H(x) =$ [ ] for $x > 0$.

Solution:
The function can be rewritten as $h(x) = \frac{1}{4} + \frac{1}{2 \cdot \sqrt{x}}$ which results in the antiderivative

$$H(x) = \frac{1}{4} \cdot x + \sqrt{x} + C$$

for $x > 0$ and $C \in \mathbb{R}$.

Exercise 8.1.5
Consider a function f with $f(x) = \frac{1}{x}$ for $x > 0$. Moreover, the functions F_1 and F_2 with $F_1(x) = \ln(7x)$ or $F_2(x) = \ln(x + 7)$ for $x > 0$ are given. Calculate the derivatives of F_1 and F_2, and state whether F_1 and F_2 are antiderivatives of f:

We have $F_1'(x) =$ [ ] and $F_2'(x) =$ [ ]
.

Check the correct answer(s).

- [ ] F_1 is an antiderivative of f.
- [ ] F_2 is an antiderivative of f.

Solution:
The derivative of F_1 with $F_1(x) = \ln(7x)$ for $x > 0$ is $F_1'(x) = \frac{1}{7x} \cdot 7 = \frac{1}{x} = f(x)$. Thus, F_1 is an antiderivative of f.

The function F_2 with $F_2(x) = \ln(x+7)$ for $x > 0$ has the derivative $F_2'(x) = \frac{1}{x+7} \cdot 1 = \frac{1}{x+7} \neq \frac{1}{x}$ for all $x > 0$. Hence, F_2 is not an antiderivative of f.

Exercise 8.1.6
Assume that F is an antiderivative of f with $f(x) = 1 + x^2$, and F has the function value $F(0) = 1$. $F(3)$ equals [ ] .
Solution:
By assumption, F is an antiderivative of f with $f(x) = 1 + x^2$. Thus, we have $F'(x) = f(x) = 1 + x^2 = x^0 + x^2$, from which

$$F(x) = \frac{1}{0+1} \cdot x^{0+1} + \frac{1}{2+1} \cdot x^{2+1} + C = x + \frac{1}{3} \cdot x^3 + C$$

follows, for a real number C. Moreover, we have $F(0) = 1$ which tells us $1 = F(0) = 0 + \frac{1}{3} \cdot 0^3 + C = C$. This results in $F(x) = \frac{1}{3} \cdot x^3 + x + 1$. Substituting $x = 3$ gives the required value $F(3) = \frac{1}{3} \cdot 3^3 + 3 + 1 = 13$.

8.2 Definite Integral

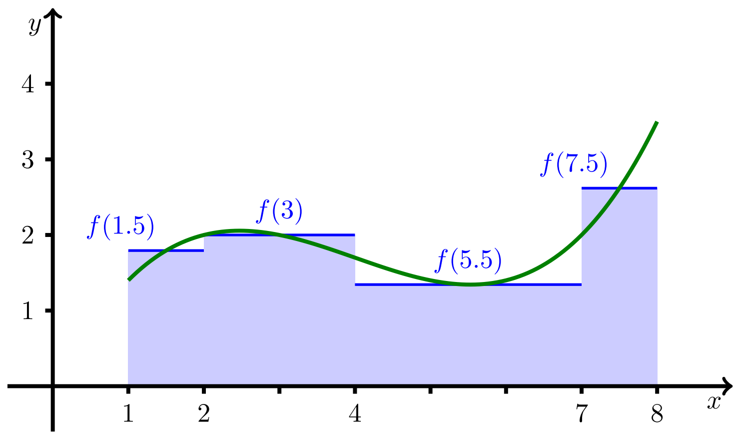
8.2.1 Introduction

The derivative $f'(x_0)$ of a differentiable function f describes how the function values change “in the vicinity of” a point x_0. If, for example, the derivative is positive, then the function f is monotonically increasing. Geometrically, this means that the slope of a tangent line to the graph at the point x_0 is positive. The derivative provides a local view of the function at every point x_0. This way, a lot of detailed information can be collected.

Conversely, a “global characteristic” is obtained if a “summary” of the function is generated by summing up the weighted function values. In mathematics, this sum is called the integral or integral value of the function. Geometrically, this concept provides a way to calculate the area under the graph of a function. It was Bernhard Riemann who specified this approach and who gave his name to the so-called **Riemann integral**.

8.2.2 Integral

The integral of a function f with $f(x) \geq 0$ can be interpreted as the “area under the graph” of the function. In the so-called Riemann integral, the graph of the function is approximated by a step function, and the values of this step function are summed up, weighted by the corresponding interval length, i.e. the “width of a step”. This approach is illustrated in the figure below.



Definition of the Riemann integral. The function is approximated by a step function that is divided here into four subintervals.

We can see here that the area under the graph of the function is initially approximated by rectangles. The length of the (horizontal) side of these rectangles is determined by the length of an interval on the x-axis, while the length of the second (vertical) side is determined by a function value $f(z_k)$ at the point z_k in the corresponding x-interval. Then the areas of these rectangles are calculated, and finally summed up. The smaller the intervals on the x-axis, the more the sum calculated this way approaches the "real" value of the area under the graph (the correct integral value of the function).

Formally, a sum S_n of the form

$$S_n = \sum_{k=0}^{n-1} f(z_k) \cdot \Delta(x_k) \qquad \text{with } \Delta(x_k) = x_{k+1} - x_k$$

is calculated. In the example considered here, the interval $[0; 8]$ is divided into four parts, where $x_0 = 1$, $x_1 = 2$, $x_2 = 4$, and $x_3 = 7$. If the areas of these four parts are summed according to the formula above, this results in

$$\begin{aligned} S_4 &= f(z_0) \cdot (x_1 - x_0) + f(z_1) \cdot (x_2 - x_1) + f(z_2) \cdot (x_3 - x_2) + f(z_3) \cdot (x_4 - x_3) \\ &= f(z_0) \cdot (2 - 1) + f(z_1) \cdot (4 - 2) + f(z_2) \cdot (7 - 4) + f(z_3) \cdot (8 - 7) \\ &= f(z_0) \cdot 1 + f(z_1) \cdot 2 + f(z_2) \cdot 3 + f(z_3) \cdot 1 \,. \end{aligned}$$

To get precise values for the area, it is obviously not sufficient to subdivide the interval into just a few subintervals. In general, it will be necessary to decrease the maximal length of the subintervals $x_{k+1} - x_k$ gradually, which in turn requires more summands $f(z_k) \cdot (x_{k+1} - x_k)$ to be calculated and added. Hence, the limit is considered as the maximal length of the subintervals tends to zero.

In principle, the approach discussed above can also be applied to functions with negative function values. We will explain in Section 8.3 auf Seite 364 how the area is calculated then. However, note that in the definition of the integral a few aspects have to be observed that go beyond the scope of this course. Therefore, we refer to advanced textbooks for details concerning the assumptions in the definitions below.

Integral8.2.1

Let a function $f : [a; b] \to \mathbb{R}$ on a real interval $[a; b]$ be given. If the number of subintervals is increased such that $x_{k+1} - x_k$ approaches 0, then

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \sum_{k=0}^{n-1} f(z_k) \cdot (x_{k+1} - x_k) \quad \text{with } x_k \le z_k \le x_{k+1} \qquad (8.2.1)$$

is called the **definite integral** of f with the lower limit of integration a and the upper limit of integration b (if the limit exists and does not depend on the respective partition). The function f is then referred to as **integrable**. In this context the function f is also called the **integrand**.

In principle, this approach can also result in an indefinite value, i.e. the integral does not exist. However, advanced considerations show that the integral exists for every continuous function. As an example, we calculate the integral of $f : [0;1] \to \mathbb{R}$, $x \mapsto x$, where we focus on the calculation of the limit.

Example 8.2.2

Calculate the integral of $f : [0;1] \to \mathbb{R}$, $x \mapsto x$.

For this purpose, the interval $[0;1]$ is subdivided into subintervals $[x_k; x_{k+1}]$ of equal length with $x_0 := 0$ and $x_k := x_{k-1} + \frac{1}{n}$. Thus, the length of a subinterval is $\Delta(x_k) = x_{k+1} - x_k = \frac{1}{n}$.

Investigating the length of the interval for n tending to infinity shows that $\Delta(x_k)$ is getting smaller and approaches zero. Thus, the assumption for the calculation of a definite integral is satisfied.

Furthermore, for the values of x_k, we obtain from the length of an interval $x_k = \frac{k}{n}$. If we set $z_k = x_k$ for the intermediate points, we obtain $f(z_k) = f(x_k) = x_k = \frac{k}{n}$.

Substituting these terms into the formula (8.2.1) gives the equation

$$\begin{aligned} S_n &= \sum_{k=0}^{n-1} f(x_k) \cdot \Delta(x_k) = \sum_{k=0}^{n-1} x_k \cdot \frac{1}{n} = \sum_{k=0}^{n-1} \frac{k}{n} \cdot \frac{1}{n} = \frac{1}{n^2} \cdot \sum_{k=0}^{n-1} k = \frac{1}{n^2} \cdot \sum_{k=1}^{n-1} k \\ &= \frac{1}{n^2} \frac{n(n-1)}{2} = \frac{1}{2} \cdot \frac{n-1}{n} = \frac{1}{2} \cdot \left(1 - \frac{1}{n}\right) , \end{aligned}$$

where we used the formula $\sum_{k=1}^{n-1} k = \frac{1}{2} n(n-1)$ of Carl Friedrich Gauss. And with $\lim\limits_{n\to\infty} \frac{1}{n} = 0$, we find for the integral

$$\int_0^1 x \, dx = \lim_{n\to\infty} S_n = \frac{1}{2} .$$

The large class of integrable functions includes all polynomials, rational functions, trigonometric functions, exponential functions, logarithmic functions and their compositions.

To simplify the calculations, rules for integrating functions are required that are as easy as possible. An important result is provided by the so-called **fundamental theorem of calculus**. It describes the relation between the antiderivative of a continuous function and its integral.

Fundamental Theorem of Calculus8.2.3

Let a continuous function $f : [a;b] \to \mathbb{R}$ on a real interval $[a;b]$ be given. Then the function f has an antiderivative, and for every antiderivative F of f, we have

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

As a simple example, we will calculate the definite integral of the function f with $f(x) = x^2$ between $a = 1$ and $b = 2$. Using the rules for the determination of antiderivatives and the fundamental theorem of calculus this problem can be solved easily.

Example 8.2.4

The function $f : [1;2] \to \mathbb{R}$ with $f(x) := x^2$ has the antiderivative F with $F(x) = \frac{1}{3}x^3 + C$ for a real number C. From the fundamental theorem, we have

$$\int_1^2 x^2\,dx = \left[\frac{1}{3}x^3 + C\right]_1^2 = \left(\frac{1}{3}\,2^3 + C\right) - \left(\frac{1}{3}\,1^3 + C\right) = \frac{7}{3}\ .$$

The calculation shows that the constant is cancelled after substituting the lower and upper limits of integration such that, in practice, it can already be “suppressed” in taking the antiderivative for the definite integral, so for the calculation of the definite integral one can choose $C = 0$.

The equation in the fundamental theorem also applies to every intermediate value $z \in [a;b]$ such that all function values $F(z)$ can be calculated from

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

if the derivative $F' = f$ and a function value, for example the value $F(a)$, are known. One also says that the antiderivative F is reconstructed from the derivative $F' = f$.

Application examples for the reconstruction of a function F from its derivative $F' = f$ are discussed at the end of Section 8.3 auf Seite 364.

8.2.3 Calculation Rules

Partition of the Interval of Integration8.2.5

Let $f : [a; b] \to \mathbb{R}$ be an integrable function. Then for every number z between a and b, we have

$$\int_a^b f(x)\,dx = \int_a^z f(x)\,dx + \int_z^b f(x)\,dx \ .$$

With the definition

$$\int_d^c f(x)\,dx := -\int_c^d f(x)\,dx$$

the rule above applies to all real numbers z for which the two integrals on the right-hand side of the equation exist, even if z does not lie between a and b. Before we demonstrate this calculation with an example, we will examine the definition above in more detail.

Exchanging the Limits of Integration8.2.6

Let $f : [a; b] \to \mathbb{R}$ be an integrable function. The integral of the function f from a to b is calculated according to the rule

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

The calculation rule described above is convenient when integrating functions that involve absolute values, or piecewise-defined functions.

Example 8.2.7

The integral of the function $f : [-4; 6] \to \mathbb{R}, x \mapsto |x|$ is

$$\begin{aligned}\int_{-4}^{6} |x|\, dx &= \int_{-4}^{0} (-x)\, dx + \int_{0}^{6} x\, dx \\ &= \left[-\frac{1}{2}x^2\right]_{-4}^{0} + \left[\frac{1}{2}x^2\right]_{0}^{6} \\ &= (0-(-8)) + (18-0) \\ &= 26\,.\end{aligned}$$

The integration over a sum of two functions can also be split up into two integrals:

Sum and Constant Multiple Rule8.2.8

Let f and g be integrable functions on $[a; b]$, and let r be a real number. Then

$$\int_a^b (f(x) + g(x))\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx\,. \tag{8.2.2}$$

For constant multiples of a function, we have

$$\int_a^b r \cdot f(x)\, dx = r \cdot \int_a^b f(x)\, dx\,. \tag{8.2.3}$$

Further material.:

There is also a calculation rule for the integration of a product of two functions, which results from the product rule for the derivative.

Integration by Parts8.2.9

Let u and v be differentiable functions on $[a;b]$ with the continuous derivatives u' and v'. For the integral of the function $u \cdot v'$, we have

$$\int_a^b u(x) \cdot v'(x)\,dx = [u(x) \cdot v(x)]_a^b - \int_a^b u'(x) \cdot v(x)\,dx\ ,$$

where u' is the derivative of u and v is an antiderivative of v'. This calculation rule is called **integration by parts or partial integration**.

This rule is also illustrated by an example.

Example 8.2.10

Calculate the integral

$$\int_0^{\pi} x \sin(x)\,dx$$

by means of integration by parts. For this purpose, we choose the functions u and v' such that

$$u(x) = x \quad \text{and} \quad v'(x) = \sin(x)\ .$$

Thus, we have

$$u'(x) = 1 \quad \text{and} \quad v(x) = -\cos x\ .$$

The required integral can now be calculated using integration by parts:

$$\begin{aligned} \int_0^{\pi} x \sin x\,dx &= [x \cdot (-\cos(x))]_0^{\pi} - \int_0^{\pi} 1 \cdot (-\cos x)\,dx \\ &= \pi \cdot (-\cos(\pi)) - 0 + \int_0^{\pi} \cos(x)\,dx \\ &= \pi \cdot (-(-1)) + [\sin(x)]_0^{\pi} = \pi\ . \end{aligned}$$

The assignments of the functions u and v' have to be appropriate. This becomes obvious if in this example the assignments of u and v' are exchanged. Readers are invited to calculate this integral with exchanged assignments of u and v'!

In the following two exercises we practice using the rule of integration by parts.

Exercise 8.2.1

Calculate the integral $I = \int_1^4 x \cdot e^x \, dx$: $I =$ [ ] .

Solution:

The integrand f with $f(x) = x \cdot e^x$ is a product of a polynomial u with $u(x) = x$ and an exponential function. The derivative of u is $u'(x) = 1$, i.e. a constant function. Moreover, the antiderivative of v' with $v'(x) = e^x$ is v with $v(x) = e^x$. Thus, integration by parts results in

$$\begin{aligned} \int_1^4 x \cdot e^x \, dx &= [x \cdot e^x]_1^4 - \int_1^4 1 \cdot e^x \, dx \\ &= [x \cdot e^x]_1^4 - [e^x]_1^4 \\ &= [x \cdot e^x - e^x]_1^4 \\ &= [(x-1) \cdot e^x]_1^4 \\ &= 3e^4 \, . \end{aligned}$$

Exercise 8.2.2

Calculate the integral $I = \int_1^8 x \cdot \ln(x) \, dx$: $I =$ [ ] .

Solution:

The integrand f with $f(x) = x \cdot \ln(x) = \ln(x) \cdot x$ for $1 \leq x \leq 8$ is a product of a polynomial and a logarithmic function. The derivative of the logarithmic function u with $u(x) = \ln(x)$ is $u'(x) = \frac{1}{x}$. Thus, the function u' is a "simple" rational function. Moreover, the antiderivative of the polynomial v' with $v'(x) = x$ is known, namely v with $v(x) = \frac{1}{2} \cdot x^2$.

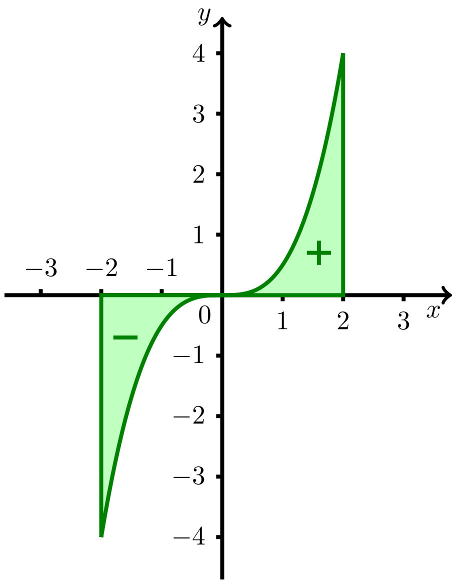
The product $u' \cdot v$ with $u'(x) \cdot v(x) = \frac{1}{x} \cdot \frac{1}{2}x^2 = \frac{1}{2}x$ for $1 \leq x \leq 8$ is a continuous function for which an antiderivative is known. Thus, the required integral can be calculated by means of integration by parts:

$$\begin{aligned} \int_1^8 x \cdot \ln(x) \, dx = \int_1^8 \ln(x) \cdot x \, dx &= \left[\ln(x) \cdot \frac{1}{2}x^2\right]_1^8 - \int_1^8 \frac{1}{x} \cdot \frac{1}{2}x^2 \, dx \\ &= \left[\ln(x) \cdot \frac{1}{2}x^2\right]_1^8 - \int_1^8 \frac{1}{2}x \, dx \\ &= \left[\frac{1}{2}x^2 \cdot \ln(x) - \frac{1}{4}x^2\right]_1^8 \\ &= (32\ln(8) - 16) - \left(\frac{1}{2} \cdot \ln(1) - \frac{1}{4}\right) \\ &= 96\ln(2) - 16 + \frac{1}{4} \, , \end{aligned}$$

where $\ln(8) = \ln(2^3) = 3 \cdot \ln(2)$ and $\ln(1) = 0$ was used.

8.2.4 Properties of the Integral

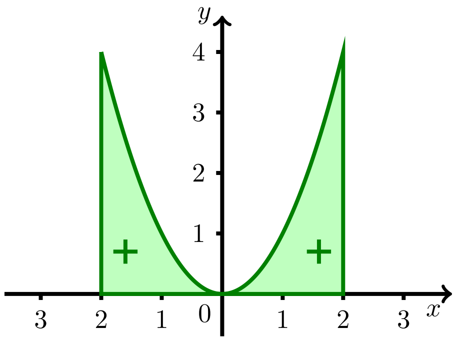
For odd functions $f : [-c; c] \to \mathbb{R}$, the integral is zero. For example, consider the function f on $[-2; 2]$ with $f(x) = x^3$ shown in the figure below.



The graph of f is subdivided into two parts, namely in a part between -2 and 0 and a part between 0 and 2, and the two regions bounded by the graph and the x-axis are investigated. The two regions can be transferred into each other by a point reflection (central inversion). Both regions are equal in size. However, forming the Riemann sum of both regions, one finds that the area of the region below the x-axis takes a negative value. Thus, if the integrals for the two regions are added to calculate the integral over the interval from -2 to 2, the area of the region above the positive x-axis is positive. The area of the region below the negative x-axis is equal in size but has a negative sign. Thus, the sum of the two areas equals zero. Thus, for odd functions f, we have the following rule:

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

For an even function $g : [-c; c] \to \mathbb{R}$, the graph is symmetric with respect to the y-axis.



The region between the graph of g and the x-axis is here symmetric with respect to the y-axis. Thus, the region to the left of the y-axis is the mirror image of the region to the right. The sum of the areas of the two regions is

$$\int_{-c}^{c} g(x)\,dx = 2 \cdot \int_{0}^{c} g(x)\,dx\ .$$

This rule for the integral applies to every integrable function g that is even, even if the function takes negative values. Due to the calculation rule above, it is then sufficient to calculate the integral for non-negative values of x with the lower limit 0 and the upper limit c.

Often, the calculation of an integral is easier if the integrand is first transformed into a known form. Examples of possible transformations shall be considered below. In the first example, power functions are investigated.

Example 8.2.11

Calculate the integral

$$\int_{1}^{4} (x-2) \cdot \sqrt{x}\,dx\ .$$

First, the integrand is transformed to simplify the calculation:

$$(x-2) \cdot \sqrt{x} = x\sqrt{x} - 2\sqrt{x} = x^{\frac{3}{2}} - 2x^{\frac{1}{2}}\ .$$

Now the integral can be calculated more easily:

$$\begin{aligned}\int_1^4 (x-2)\cdot\sqrt{x}\,dx &= \int_1^4 \left(x^{\frac{3}{2}} - 2x^{\frac{1}{2}}\right) dx = \left[\frac{2}{5}x^{\frac{5}{2}} - \frac{4}{3}x^{\frac{3}{2}}\right]_1^4 \\ &= \left(\frac{2}{5}\left(\sqrt{4}\right)^5 - \frac{4}{3}\left(\sqrt{4}\right)^3\right) - \left(\frac{2}{5}\cdot 1 - \frac{4}{3}\cdot 1\right) \\ &= \left(\frac{64}{5} - \frac{32}{3}\right) - \left(\frac{2}{5} - \frac{4}{3}\right) \\ &= \frac{62}{5} - \frac{28}{3} \\ &= 3 + \frac{1}{15}\ .\end{aligned}$$

The next example demonstrates a transformation of an integrand involving exponential functions.

Example 8.2.12

Calculate the integral

$$\int_{-2}^3 \frac{8e^{3+x} - 12e^{2x}}{2e^x}\,dx\ .$$

According to the calculation rule for exponential functions one obtains

$$\frac{8e^{3+x} - 12e^{2x}}{2e^x} = \frac{8e^{3+x}}{2e^x} - \frac{12e^{2x}}{2e^x} = 4e^{3+x-x} - 6e^{2x-x} = 4e^3 - 6e^x\ ,$$

such that the integral can finally be calculated easily:

$$\begin{aligned}\int_{-2}^3 \frac{8e^{3+x} - 12e^{2x}}{2e^x}\,dx = \int_{-2}^3 \left(4e^3 - 6e^x\right) dx &= \left[4e^3\cdot x - 6e^x\right]_{-2}^3 \\ &= \left(4e^3\cdot 3 - 6e^3\right) - \left(4e^3\cdot(-2) - 6e^{-2}\right) \\ &= 14e^3 + \frac{6}{e^2}\ .\end{aligned}$$

Consider a rational function. If the degree of the numerator polynomial is greater or equal to the degree of the denominator polynomial, a polynomial long division is carried out first (see Module 6 auf Seite 217). Depending on the situation, further transformations (e.g. partial fraction decomposition) may be appropriate. These can be found in advanced textbooks and formularies. In the following example, a polynomial long division is carried out to integrate a rational function.

Example 8.2.13

Calculate the integral

$$\int_{-1}^{1} \frac{4x^2 - x + 4}{x^2 + 1}\, dx \,.$$

First, we transform the integrand using polynomial long division:

$$4x^2 - x + 4 = (x^2 + 1) \cdot 4 - x$$

so

$$\frac{4x^2 - x + 4}{x^2 + 1} = 4 - \frac{x}{x^2 + 1} \,.$$

Thus, we have

$$\begin{aligned} \int_{-1}^{1} \frac{4x^2 - x + 4}{x^2 + 1}\, dx &= \int_{-1}^{1} \left(4 - \frac{x}{x^2 + 1}\right) dx \\ &= \int_{-1}^{1} 4\, dx - \int_{-1}^{1} \frac{x}{x^2 + 1}\, dx = [4x]_{-1}^{1} - 0 = 8 \,. \end{aligned}$$

The integrand in the second integral is an odd function and centrally symmetric on the interval $[-1; 1]$, so the second integral equals zero.

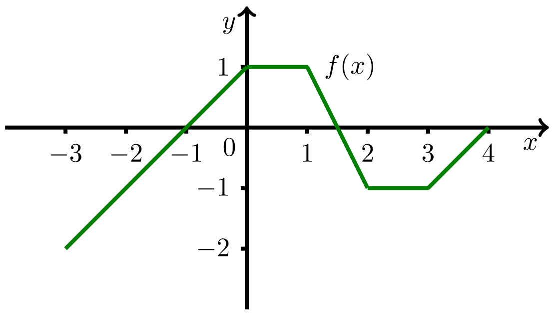
Here, a specific example was given to provide a first impression of the integration of rational functions. In advanced mathematics lectures and in the literature this approach is described in general terms.

8.2.5 Exercises

The first exercise picks up the idea (from the definition of the integral) of calculating the value of the integral based on an appropriate partition. For this purpose, in this first exercise more general shapes are used besides the standard rectangles, such as triangles.

Exercise 8.2.3

Calculate the integral $\int_{-3}^{4} f(x)\,dx$ of the function $f : [-3; 4] \to \mathbb{R}$ with the graph shown in the figure below applying methods of the elementary geometry, i.e. by subdividing the "area below the graph" into elementary geometrical shapes such as triangles or rectangles that either lie above or below the x-axis. In this case, you can subsequently calculate the single areas using the formulas for triangles and rectangles.



The value of the integral is the sum of the areas of regions lying above the x-axis minus the sum of the areas of regions lying below the x-axis. In this sense, the value of the integral can be considered as the sum of signed areas.

The value of the integral $\int_{-3}^{4} f(x)\,dx$ is [ ] .

Solution:

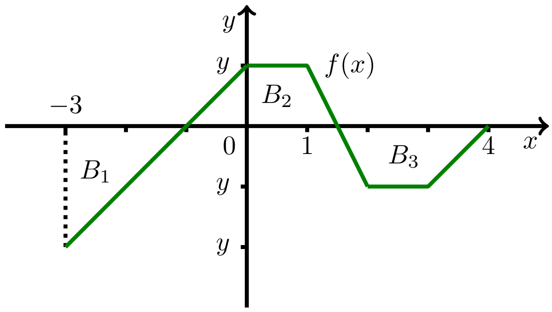
The area is subdivided by vertical lines at $x_0 = -3$, $x_1 = -1$, $x_2 = 0$, $x_3 = 1$, $x_4 = \frac{3}{2}$, $x_5 = 2$, $x_6 = 3$, and $x_7 = 4$ into regions that are bounded each by the graph of f, the x-axis and the lines at x_{k-1} and x_k for $1 \leq k \leq 7$.

Then the corresponding signed areas are summed up resulting in the value of the integral:

$$\int_{-3}^{4} f(x)\,dx = -\frac{2 \cdot 2}{2} + \frac{1 \cdot 1}{2} + 1 \cdot 1 + \frac{\frac{1}{2} \cdot 1}{2} - \frac{\frac{1}{2} \cdot 1}{2} - 1 \cdot 1 - \frac{1 \cdot 1}{2} = -2\,.$$

Of course, the area can also be subdivided into other regions. If, for example, the area is subdivided by vertical lines at $z_0 = -3$, $z_1 = -1$, $z_2 = \frac{3}{2}$, and $z_3 = 4$ into three regions, the area of the region B_2 between the lines at z_1 and z_2 equals the area of the region B_3

between the lines at z_2 and z_3. However, the signs of the areas of B_2 and B_3 are opposite since B_2 lies above the x-axis and B_3 below.



Thus, the value of the integral equals the negative area of the region B_1 between the lines z_0 and z_1, which lies below the x-axis.

Exercise 8.2.4
Calculate the following integrals:

a. $\displaystyle\int_0^5 3\,dx \;=$ ☐

b. $\displaystyle\int_0^5 -4\,dx \;=$ ☐

c. $\displaystyle\int_0^4 2x\,dx \;=$ ☐

d. $\displaystyle\int_0^4 (4-x)\,dx \;=$ ☐

<u>Solution:</u>
Using the fundamental theorem of calculus, we obtain

1. $\displaystyle\int_0^5 3\,dx = [3x]_0^5 = 15$

2. $\displaystyle\int_0^5 -4\,dx = [-4x]_0^5 = -20$

3. $\displaystyle\int_0^4 2x\,dx = \left[x^2\right]_0^4 = 16$

4. $\displaystyle\int_0^4 (4-x)\,dx = \left[4x - \frac{1}{2}\cdot x^2\right]_0^4 = 8$

Exercise 8.2.5
The value of the integral $\int_{-\pi}^{\pi} \left(5x^3 - 4\sin(x)\right)\,dx$ is $\boxed{}$.

Solution:
From the fact that the integrand is odd and the interval of integration is symmetric with respect to the origin $(0,0)$, we can deduce that the value of the integral is 0. Alternatively, we can calculate the integral using the fundamental theorem of calculus:

$$\int_{-\pi}^{\pi} \left(5x^3 - 4\sin(x)\right)\,dx = \left[\frac{5}{4}\cdot x^4 + 4\cos(x)\right]_{-\pi}^{\pi} = \left[\frac{5}{4}\cdot x^4 + 4\cos(x)\right]_{-\pi}^{\pi} = 0\,.$$

Exercise 8.2.6
Calculate a real number z such that the value of the integral

$$\int_0^2 \left(x^2 + z\cdot x + 1\right)\,dx$$

equals zero. The required value of z is $z = \boxed{}$.

Solution:
If we consider z as an unknown constant, we have

$$\int_0^2 \left(x^2 + z\cdot x + 1\right)\,dx = \left[\frac{1}{3}x^3 + \frac{1}{2}zx^2 + x\right]_0^2 = \frac{8}{3} + 2z + 2\,.$$

Hence, $z = -\frac{14}{6} = -\frac{7}{3}$ is the required value.

Exercise 8.2.7
Calculate the following integrals:

a. $\int_{-3}^{2} \left(1 + 6x^2 - 4x\right)\,dx = \boxed{}$

b. $\int_1^9 \frac{5}{\sqrt{4x}}\,dx = \boxed{}$

Solution:
The integrand f with $f(x) = 1 + 6x^2 - 4x = 6x^2 - 4x + 1$ is a polynomial. Thus, F with $F(x) = 2x^3 - 2x^2 + x$ is an antiderivative of f. From the fundamental theorem, we have

$$\int_0^1 \left(1 + 6x^2 - 4x\right)\,dx = \left[2x^3 - 2x^2 + x\right]_{-3}^{2} = 2(8-4) + 2 - (2(-27-9) - 3) = 85\,.$$

In the second part of the exercise, the integrand $f(x) = \frac{5}{\sqrt{4x}} = \frac{5}{2}x^{-1/2}$ is a product of a root function and a constant factor. Thus, F with $F(x) = 5x^{1/2} = 5\sqrt{x}$ is an antiderivative of f. From the fundamental theorem, we have

$$\int_1^9 \frac{5}{\sqrt{4x}}\,dx = \left[5\sqrt{x}\right]_1^9 = 5(3-1) = 10\ .$$

Exercise 8.2.8

The value of the integral $\displaystyle\int_{-24}^{-6} \frac{1}{2x}\,dx$ is ☐ .

Solution:

The integrand f with $f(x) = \frac{1}{2x} = \frac{1}{2}\cdot\frac{1}{x}$ for $x < 0$ has an antiderivative F with $F(x) = \frac{1}{2}\ln|x|$. From the fundamental theorem, we have

$$\int_{-24}^{-6} \frac{1}{2x}\,dx = \left[\frac{1}{2}\ln|x|\right]_{-24}^{-6} = \frac{1}{2}\left(\ln|-6| - \ln|-24|\right) = \frac{1}{2}\ln\left(\frac{6}{24}\right) = \frac{1}{2}\ln\left(2^{-2}\right) = -\ln(2)\ .$$

Exercise 8.2.9

Calculate the following integrals

a. $\displaystyle\int_0^3 (2x-1)\,dx =$ ☐

b. $\displaystyle\int_{-3}^0 (1-2x)\,dx =$ ☐

Solution:

The integrand f with $f(x) = 2x-1$ is a polynomial. F with $F(x) = x^2 - x$ is an antiderivative of f. From the fundamental theorem, we have

$$\int_0^3 (2x-1)\,dx = \left[x^2 - x\right]_0^3 = 9 - 3 - 0 = 6\ .$$

In the second part of the exercise, the integrand f with $f(x) = 1-2x$ is also a polynomial. Thus, F with $F(x) = x - x^2$ is an antiderivative of f. From the fundamental theorem, we have

$$\int_{-3}^0 (1-2x)\,dx = \left[x - x^2\right]_{-3}^0 = 0 - (-3-9) = 12\ .$$

Exercise 8.2.10

Calculate the integral

$$\int_\pi^{3\pi} \left(\frac{3\pi}{x^2} - 4\sin(x)\right) dx =$$ ☐ .

Solution:
The integrand f with $f(x) = \frac{3\pi}{x^2} - 4\sin(x)$ has an antiderivative F with $F(x) = -\frac{3\pi}{x} + 4\cos(x)$. From the fundamental theorem, we have

$$\int_{\pi}^{3\pi} \left(\frac{3\pi}{x^2} - 4\sin(x) \right) dx = \left[-\frac{3\pi}{x} + 4\cos(x) \right]_{\pi}^{3\pi} = (-1-4) - (-3-4) = 2 \; .$$

Remark: The integral over a period 2π of the two periodic functions sin and cos equals zero. However, for other periodic functions, as for example $f(x) = \sin(x) + \frac{1}{2}$, the value of the integral over a period can differ from zero.

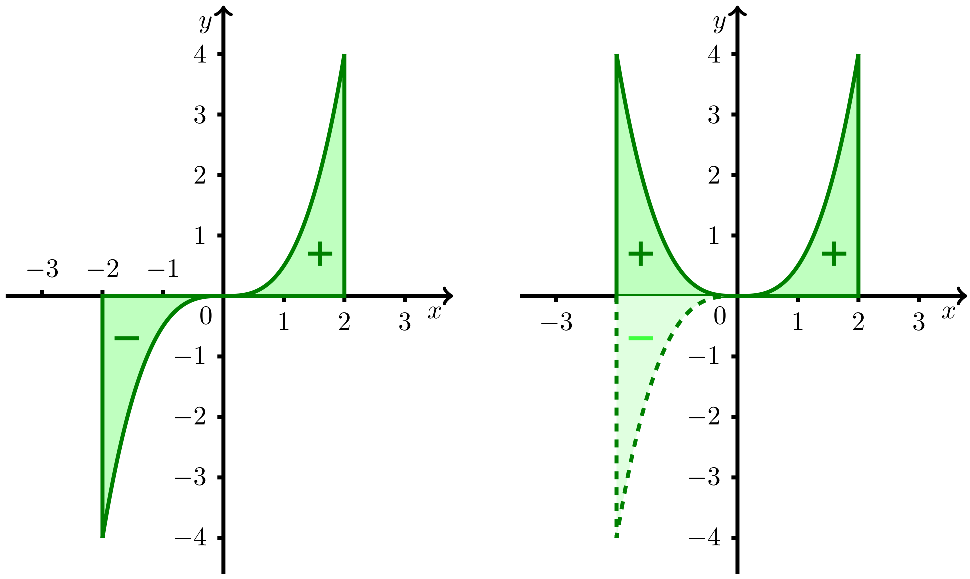
8.3 Applications

8.3.1 Introduction

Integral calculus has many varied applications, in particular in science and engineering. Here, the calculation of areas of regions with boundaries described by mathematical functions shall be studied first. This application is not purely mathematical, but is used in the determination of centres of mass, rotational properties of rigid bodies or the bending properties of beams or girders. At the end of this section a few more physical applications are considered.

8.3.2 Calculation of Areas

A first application of the integral calculus is the calculation of **areas** of regions with boundaries that can be described by mathematical functions. For illustration purposes, the figure below (left part) shows the function $f(x) = \frac{1}{2}x^3$ on the interval $[-2; 2]$. The goal is to calculate the area that is bounded by the graph of the function and the x-axis. From our previous investigations, we know that the integral over this odd function in the limits from -2 to 2 equals zero since the area of the left region equals the area of the right region but during integration they are assigned different signs. Thus, the value of the integral does not match the actual value of the area. However, if the "negative" area is reflected across the x-axis, i.e. the function is assigned a positive sign (right part of the figure). Now the area is determined correctly by the integral. In mathematical terms, it is not the integral of the function f that is calculated but the integral of its absolute value $|f|$.



Since the integral is now taken over the absolute value of the function, a partition of the integral depending on the sign of the function is required: the interval of integration is subdivided into subintervals on which the function values have the same sign. For continuous functions, these subintervals are determined by the zeros of the function.

Calculation of Areas8.3.1

Let a continuous function $f : [a; b] \to \mathbb{R}$ on an interval $[a; b]$ be given. Moreover, let $x_1, \ldots, x_m$ be the zeros of f with $x_1 < x_2 < \ldots < x_m$. We set $x_0 := a$ and $x_{m+1} := b$.

Then, the area bounded by the graph of f and the x-axis equals

$$\int_a^b |f(x)|\, dx = \sum_{k=0}^{m} \left| \int_{x_k}^{x_{k+1}} f(x)\, dx \right| .$$

This shall be explained in more detail for the example above.

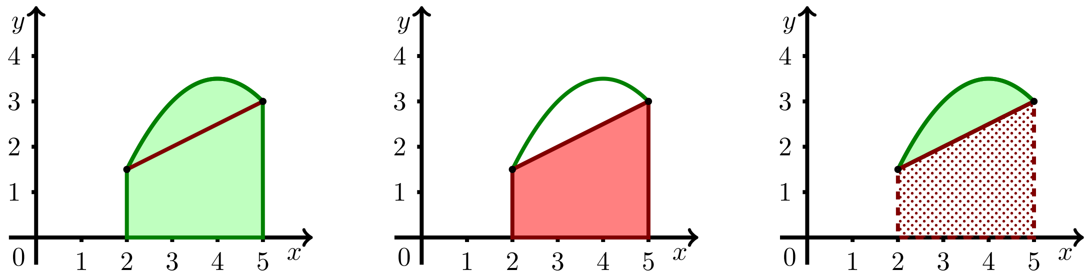
Example 8.3.2

We have to calculate the area I_A of the region that is bounded by the continuous function f with $f(x) = \frac{1}{2}x^3$ and the x-axis on the interval $[-2; 2]$. The only zero of the given function is at $x_0 = 0$. The interval of integration is subdivided into the two subintervals $[-2; 0]$ and $[0; 2]$. Thus, for the area of the region between the graph of the function and the x-axis we calculate

$$\begin{aligned} I_A = \int_{-2}^{2} \left| \frac{1}{2}x^3 \right| dx &= \left| \int_{-2}^{0} \frac{1}{2}x^3\, dx \right| + \left| \int_{0}^{2} \frac{1}{2}x^3\, dx \right| \\ &= \left| \left[\frac{1}{8}x^4 \right]_{-2}^{0} \right| + \left| \left[\frac{1}{8}x^4 \right]_{0}^{2} \right| \\ &= |0 - 2| + |2 - 0| \\ &= 4 , \end{aligned}$$

to obtain the value $I_A = 4$.

However, not only areas of regions between a graph of a function and the x-axis can be calculated but also areas of regions that are bounded by two graphs of functions as illustrated in the figure below. The right-hand part of this figure indicates the required area that is calculated as the difference of the area in the left part of the figure and the area in the middle.



Again, this principle shall first be presented formally, then explained by means of an example.

Calculation of Areas of Regions between the Graphs of two Functions8.3.3

Let two continuous functions $f, g : [a; b] \to \mathbb{R}$ on an interval $[a; b]$ be given. Moreover, let $x_1, \ldots, x_m$ be the zeros of $f - g$ with $x_1 < x_2 < \ldots < x_m$. We set $x_0 := a$ and $x_{m+1} := b$.

Then the area of the region between the graphs of f and g can be calculated from

$$\int_a^b |f(x) - g(x)|\, dx = \sum_{k=0}^{m} \left| \int_{x_k}^{x_{k+1}} (f(x) - g(x))\, dx \right| .$$

Let us now consider an example.

Example 8.3.4

Calculate the area I_A of the region between the graphs of f and g with $f(x) = x^2$ and $g(x) = 8 - \frac{1}{4}x^4$ for $x \in [-2; 2]$.

First, we find the zeros of the function $f-g$. From

$$\begin{aligned} f(x)-g(x) &= \frac{1}{4}x^4+x^2-8 \\ &= \frac{1}{4}\left(x^4+4x^2-32\right) \\ &= \frac{1}{4}\left(x^4+4x^2+2^2-2^2-32\right) \\ &= \frac{1}{4}\left(\left(x^2+2\right)^2-36\right) \end{aligned}$$

the real zeros of $f-g$ can be calculated:

$$\begin{aligned} & \left(x^2+2\right)^2-36=0 \\ \Leftrightarrow\quad & \left(x^2+2\right)^2=36 \\ \Leftrightarrow\quad & x^2+2=6 \\ \Leftrightarrow\quad & x^2=4 \\ \Leftrightarrow\quad & x=\pm 2\ . \end{aligned}$$

Alternatively, from the third binomial formula, we have:

$$\begin{aligned} 0 &= \left(x^2+2\right)^2-36=\left(x^2+2\right)^2-6^2 \\ &= \left(x^2+2-6\right)\cdot\left(x^2+2+6\right)=\left(x^2-4\right)\cdot\left(x^2+2+6\right)=(x-2)\cdot(x+2)\cdot\left(x^2+8\right)\ . \end{aligned}$$

In the first calculation, after taking the first root we did not consider the case $x^2+2=-6$ any further since the zeros obtained from the resulting equation $x^2=-8$ are not real. Thus, the real zeros of $f-g$ are -2 and 2. These are also the boundary points of the interval of integration. Thus, a partition of the integral into different parts is not necessary. On the interval of integration, the function values of f are smaller than the function values of g. For the area of the region, we obtain

$$\begin{aligned} I_A &= \int_{-2}^{2}|f(x)-g(x)|\,dx \\ &= \int_{-2}^{2}(g(x)-f(x))\,dx \\ &= \int_{-2}^{2}\left(-\frac{1}{4}x^4-x^2+8\right)dx \\ &= 2\int_{0}^{2}\left(-\frac{1}{4}x^4-x^2+8\right)dx\ , \qquad \text{since the integrand is even} \\ &= \left[-\frac{1}{20}x^5-\frac{1}{3}x^3+8x\right]_0^2 \\ &= 2\cdot\left(-\frac{32}{20}-\frac{8}{3}+16\right) \\ &= \frac{352}{15}\ . \end{aligned}$$

8.3.3 Applications in the Sciences

The velocity v of an object describes the instantaneous rate of change of position at the time t. Thus, we have $v = \frac{ds}{dt}$ if $v = v(t)$ and $s = s(t)$ are considered as functions of t. The current position $s(T)$ results from the inversion of the derivative, i.e. by integration of the velocity over the time t. With the initial value $s(t = 0) = s_0$ at the time $t = 0$, we have

$$\begin{aligned} \int_0^T \frac{ds}{dt}\,dt &= \int_0^T v\,dt \\ \Leftrightarrow \quad [s(t)]_0^T &= \int_0^T v\,dt \\ \Leftrightarrow \quad s(T) - s(0) &= \int_0^T v\,dt \\ \Leftrightarrow \quad s(T) &= s_0 + \int_0^T v(t)\,dt\,. \end{aligned}$$

In mathematical terms, this situation can be described as follows: if the *derivative* f' of a function f and a single function value $f(x_0)$ are known, then the function can by calculated by means of the integral. In this context one says that the function values are reconstructed from the derivative.

If, for example, a population of bacteria increases approximately according to B' with $B'(t) = 0.6t$ for $t \geq 0$ and initially the population consists of $B(0) = 100$ bacteria, then the number B of bacteria in the population at time T is described by

$$B(T) - B(0) = \int_0^T 0.6t\,dt$$

and hence by

$$B(T) = B(0) + \int_0^T 0.6t\,dt = 100 + 0.6\int_0^T t\,dt = 100 + 0.3\left(T^2 - 0^2\right) = 100 + 0.3T^2\,.$$

Thus, the fundamental theorem of calculus provides an important tool for reconstructing a function if its derivative is known (and continuous). However, in practical applications the functions will often be more sophisticated, for example consisting of combinations of exponential functions.

A further example from physics, which may be familiar, is the determination of the work as a product of force and displacement: $W = F_s \cdot s$. Here, F_s is the projection of the force onto the direction of the travelled path. However, if the force depends on the path, then this law does not apply in its simple form. For example, to calculate the work done by moving a massive body along a path, the force has to be integrated along the path from the initial point s_1 to the end point s_2:

$$W = \int_{s_1}^{s_2} F_s(s)\,ds \;.$$

These are only three examples from the sciences of how the notion of an integral is useful. Depending on the subject of your study you will encounter a whole series of further applications of integration.

8.3.4 Exercises

Exercise 8.3.1
Calculate the area I_A of the region A that is bounded by the graph of the function $f : [-2\pi; 2\pi] \to \mathbb{R}, x \mapsto 3\sin(x)$ and the x-axis.

Answer: $I_A =$ ______ .

Solution:
The function f with $f(x) = 3\sin(x)$ has, on the interval $[-2\pi; 2\pi]$, the zeros -2π, $-\pi$, 0, π, and 2π. Since the graph of f is centrally symmetric with respect to the origin, for the area we obtain:

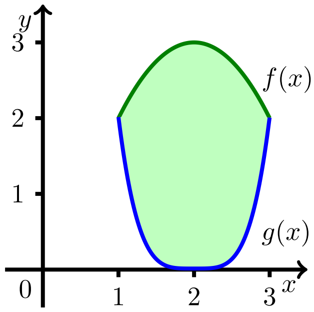
$$\begin{aligned}
\int_{-2\pi}^{2\pi} |f(x)|\,dx &= 3 \cdot \int_{-2\pi}^{2\pi} |\sin(x)|\,dx \\
&= 3 \cdot 2 \cdot \int_{0}^{2\pi} |\sin(x)|\,dx, \qquad \text{since the function } |\sin| \text{ is even,} \\
&= 6 \cdot \left(\int_0^{\pi} \sin(x)\,dx + \int_{\pi}^{2\pi} (-\sin(x))\,dx \right) \\
&= 6 \cdot \left([-\cos(x)]_0^{\pi} + [\cos(x)]_{\pi}^{2\pi} \right) \\
&= 6 \cdot ((-(-1) + 1) + (1 - (-1))) \\
&= 24\,.
\end{aligned}$$

Of course, the integral can also be calculated without noting that the graph of f is centrally symmetric with respect to the origin.

Exercise 8.3.2
Calculate the area I_A of the region A bounded by the graphs of the functions $f : [1; 3] \to \mathbb{R}, x \mapsto 3 - (x-2)^2$ and $g : [1; 3] \to \mathbb{R}, x \mapsto 2 \cdot (x-2)^4$. Draw the graphs of the functions before calculating the area.

Answer: $I_A =$ ______ .

Solution:
To calculate the area I_A of the region between the graphs of the functions f and g, the difference $f - g$ with $f(x) - g(x) = 3 - (x-2)^2 - 2 \cdot (x-2)^4$ on the interval $[1; 3]$ is considered.



From the drawing of the graphs of the functions we see that the difference $f(x) - g(x)$ is greater or equal to zero for $x \in [1; 3]$. This can also be seen by calculation: According to the assumption, we have $1 \leq x \leq 3$ and thus $-1 \leq x - 2 \leq 1$. Hence, $(x-2)^2 \leq 1$ and thus $-(x-2)^2 \geq -1$ such that

$$f(x) - g(x) \geq 3 - 1 - 2 \cdot 1 = 0$$

for all $1 \leq x \leq 3$.

Thus, for the calculation of the area, we have to evaluate the integral $\int_1^3 (f(x) - g(x))\, dx$. For this purpose, the terms can be multiplied out and integrated according to the sum rule. Another way is to consider the terms of the functions in more detail: in this situation we have two terms, namely $(x-2)^2$ and $(x-2)^4$, that result from shifting the known terms z^2 and z^4 according to $z = x - 2$. An antiderivative of h with $h(z) = z^2$ is H with $H(z) = \frac{1}{3} \cdot z^3$. If we now consider F with $F(x) = 3x - \frac{1}{3} \cdot (x-2)^3$ accordingly, then we have (from the chain rule) $F'(x) = 3 - \frac{1}{3} \cdot 3 \cdot (x-2)^2 \cdot 1 = f(x)$. Here, the last factor results from the derivative of the inner function u with $u(x) = x - 2$. Therefore, F is an antiderivative of f. Likewise, it can be checked that G with $G(x) = \frac{2}{5} \cdot (x-2)^5$ is a antiderivative of g.

Thus, for the area I_A of the region between the graphs of the functions, we have

$$\begin{aligned} I_A &= \left| \int_1^3 \left(3 - (x-2)^2 - 2 \cdot (x-2)^4\right) dx \right| \\ &= \left| \left[3x - \frac{1}{3}(x-2)^3 - \frac{2}{5}(x-2)^5 \right]_1^3 \right| \\ &= \left| 27 - \frac{1}{3} - \frac{2}{5} - \left(3 + \frac{1}{3} + \frac{2}{5}\right) \right| \\ &= 22 + \frac{8}{15} \,. \end{aligned}$$

In the next exercise, a physical problem will be formulated in mathematical terms, where the description involves a simplification. This shall exemplify that the mathematical notation can, in principle, also be used in applications. In practise, shorter or simpler formulations may occur. For example, domain and range are not given explicitly if they can be deduced from the context.

Exercise 8.3.3
Calculate the work W done by a force on a small spherical homogeneous body k with mass m in lifting it *against* the gravitational force $F : [r_1; \infty[\to \mathbb{R}, r \mapsto F(r) := -\gamma \cdot \frac{m \cdot M}{r^2}$ from the surface of spherical homogeneous body K with radius $r_1 = 1$ and mass $M = 2$ to a height of $r_2 = 4$ (all lengths are measured with respect to the centre of the body K). Here, the mass m and the gravitational constant γ are assumed to be given, and the smaller body k is assumed to be point-like in comparison to the body K.

Answer: $W =$ [ ] .

Solution:
The force F_s that acts along the path on the small body k with mass m to lift it from the surface of the body K is directed *against* the gravitational force F. Thus, we have $F_s = -F$.

Hence, the work W done by the force in lifting the small body k from $r_1 = 1$ to $r_2 = 4$ is

$$\begin{aligned} W = \int_1^4 F_s(r)\,dr = -\int_1^4 F(r)\,dr &= -\int_1^4 -\gamma \cdot \frac{2 \cdot m}{r^2}\,dr \\ &= \int_1^4 \gamma \cdot \frac{2 \cdot m}{r^2}\,dr \\ &= \left[-\gamma \cdot \frac{2 \cdot m}{r}\right]_1^4 \\ &= -\gamma \cdot 2 \cdot m \left(\frac{1}{4} - \frac{1}{1}\right) \\ &= \gamma \cdot \frac{3m}{2}\,. \end{aligned}$$

8.4 Final Test

8.4.1 Final Test Module 8

Exercise 8.4.1
Find an antiderivative for each of the following functions:

a. $\int 3\sqrt{x}\,dx = \square$

b. $\int \left(2x - e^{x+\pi}\right) dx = \square$

Exercise 8.4.2
Calculate the integrals

$$\int_1^e \frac{1}{2x}\,dx = \square \quad \text{and} \quad \int_5^8 \frac{6}{x-4}\,dx = \square$$

Exercise 8.4.3
Calculate the integrals

$$\int_0^3 x \cdot \sqrt{x+1}\,dx = \square \quad \text{and} \quad \int_\pi^{\frac{3\pi}{4}} 5\sin(4x - 3\pi)\,dx = \square$$

Exercise 8.4.4
Fill in the boxes.

$$2\int_{\square}^4 |x^3|\,dx = \int_{-4}^4 |x^3|\,dx \;\square\; \left|\int_{-4}^4 x^3\,dx\right|$$

Exercise 8.4.5
Calculate the area I_A of the region A that is bounded by the graphs of the two functions f and g on $[-3;2]$ with $f(x) = x^2$ and $g(x) = 6 - x$.

Answer: $I_A = \square$

Exercise 8.4.6
Let an antiderivative F of the function f and an antiderivative G of the function g be given. Moreover, the function id with $\mathrm{id}(x) = x$ is given.

Which of the following statements are always true (provided the corresponding combinations/compositions are possible)?

True:	False:	Statement:
☐	☐	$\mathrm{id} \cdot F$ is an antiderivative of $\mathrm{id} \cdot f$
☐	☐	$F \circ G$ is an antiderivative of $f \circ g$
☐	☐	$F - G$ is an antiderivative of $f - g$
☐	☐	F/G is an antiderivative of f/g
☐	☐	$F \cdot G$ is an antiderivative of $f \cdot g$
☐	☐	$-20 \cdot F$ is an antiderivative of $-20 \cdot f$

9 Objects in the Two-Dimensional Coordinate System

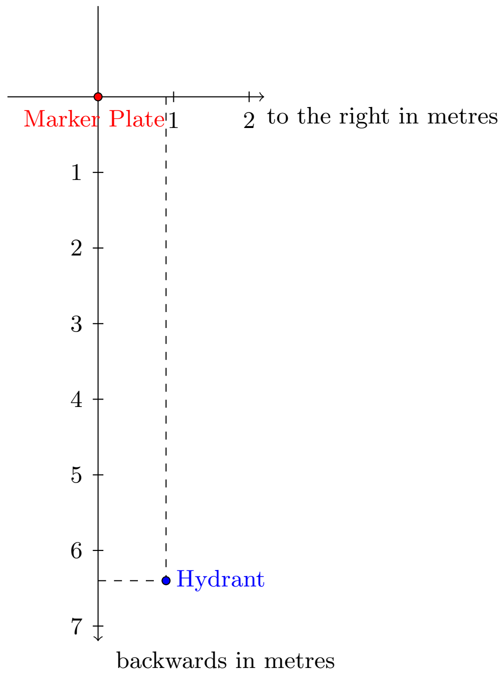
Module Overview

9.1 Cartesian Coordinate System in the Plane

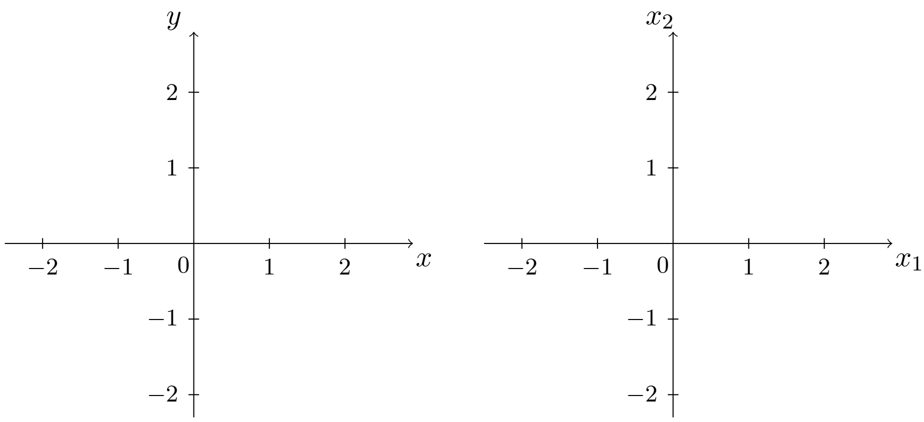
9.1.1 Introduction

In Module 5, we considered objects in the plane - lines and circles - in a purely geometric way. If we want to analyse such algebraically as well (by means of equations), then we can introduce **coordinate systems**, in which we can specify points uniquely. The basic idea of a coordinate system is very easy: to specify the position of a point exactly, we use a reference point (called the **origin**) and a certain unit length (e.g. kilometre). In this way, the position of a point can be given by two numbers. In mathematics, these two numbers are called the **coordinates** of the point. In the real world coordinates can be found, for example, on marker plates that specify the position of hydrants in the ground (see figure below).

In this case, the origin is the position at which the plate is mounted, the unit length is metre, and the numbers 0.9 and 6.4 specify that the hydrant can be found 0.9 metres to the right and 6.4 metres backwards. Hence, the two numbers 0.9 and 6.4 are the coordinates of the hydrant in the coordinate system defined by the position of the marker plate and the unit length of a metre (see figure below).

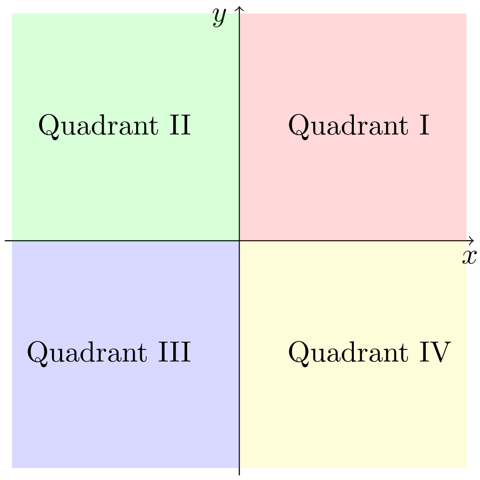


In mathematics, directions in a coordinate system are rarely called "right" and "backwards". In drawing a coordinate system, the horizontal direction is often called x-direction or x_1-direction, and the vertical direction is called y-direction or x_2-direction. Moreover, by mathematical convention the x-direction is to the right and the y-direction is up.



The unit length used can also be specified, but from a purely mathematical point of view this is not necessary. Note that negative coordinates are to the left and below the origin. Finally, the drawn **axes** are often called x-axis or x_1-axis and y-axis and x_2-axis according to the coordinates. Furthermore, the names **axis of abscissa** for the horizontal axis and **axis of ordinates** for the vertical axis are in common use. (Correspondingly, the coordinate values are then called **abscissas** and **ordinates**.)

From the figures above, we can see that such a coordinate system divides the plane into four regions. These regions have special names: **quadrants I** to **IV** (see figure below).



The coordinate systems described above are called **Cartesian coordinate systems**. This is means that the axes are perpendicular to each other, i.e. they intersect in the origin at an angle of 90°. Non-Cartesian coordinate systems do also exist, but they will not be considered in this course. Here, we will assume that every coordinate system is Cartesian.

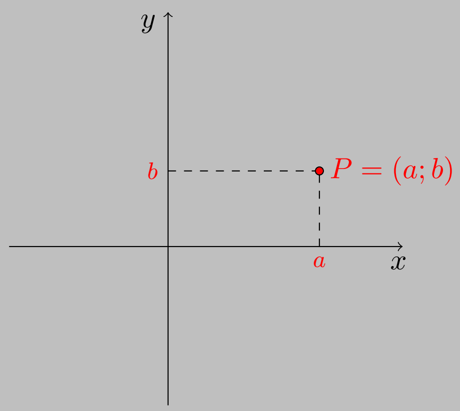
Once the concept of a coordinate system in the plane is clear, it is obvious that the position of points in (three-dimensional) space can be described in a similar way. For example, to describe the position of an aircraft precisely we need only its position with respect to the control tower but also its altitude. Thus, a third coordinate and hence a coordinate system with three axes are required. We will deal with these three-dimensional coordinate systems in Module 10.

9.1.2 Points in Cartesian Coordinate Systems

Now, if we want to describe points in the plane by coordinates, we use variables. Typically, points are denoted by upper-case Latin letters $A, B, C, \ldots$ or $P, Q, R, \ldots$, and their coordinates are denoted by lower-case Latin letters $a, b, c, \ldots$ or $x, y, \ldots$. First, we will define what is meant by a point in the plane in which a coordinate system is given, and we will fix the notation that will be used in the rest of this course.

Info9.1.1

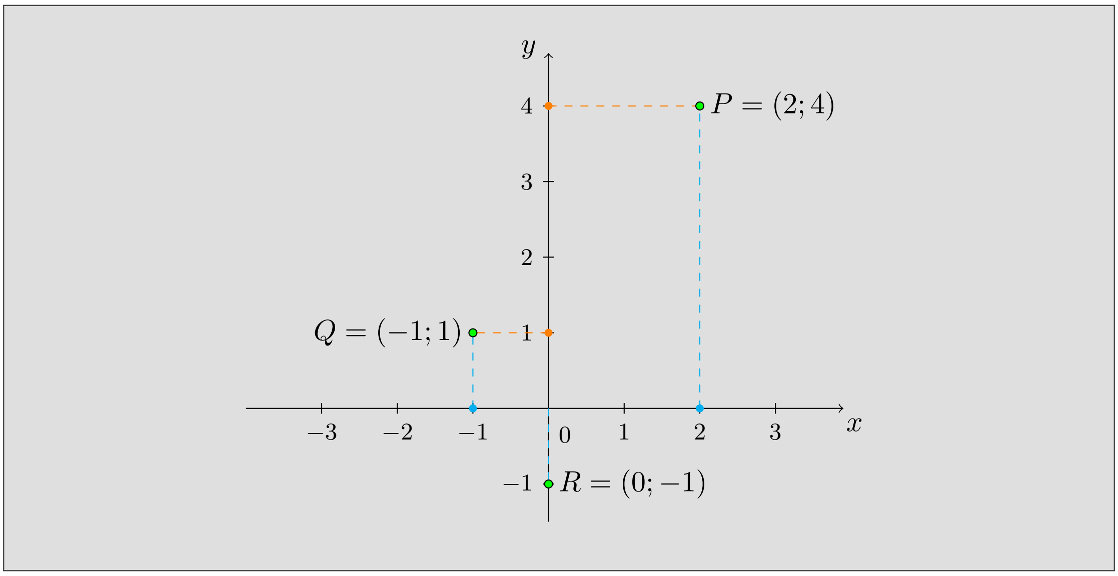
With respect to a given coordinate system, a **point** in the plane is described by $P = (a; b)$, where P is the variable denoting the point and a and b are its coordinates. Its abscissa or x-coordinate is a, and its ordinate or y-coordinate is b as shown in the figure below.



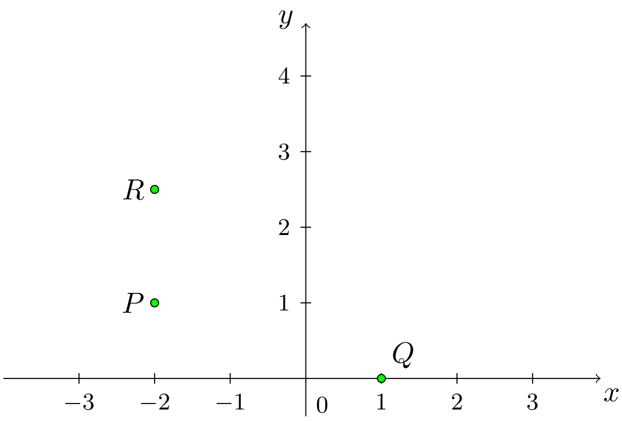
For points, there is some variation in notation. In schools, $P(a|b)$ or $P(a, b)$ is often written instead of $P = (a; b)$. Throughout this course, the notation $P = (a; b)$ will be used. Since points are uniquely determined by their coordinates, we will not distinguish between a point P and its coordinates $(a; b)$ in the following but we will consider both as the same object. For every coordinate system, the origin (the point with the coordinates $(0; 0)$) is a special point. Generally, it is denoted by the variable O: $O = (0; 0)$.

Example 9.1.2

The figure below shows the three points $P = (2; 4)$, $Q = (-1; 1)$, and $R = (0; -1)$. The point Q, for example, has the x-coordinate -1 (one unit length to the left on the axis of abscissas) and the y-coordinate 1 (one unit length upwards on the axis of ordinates).



Exercise 9.1.1
Specify the coordinates of the points drawn in the following coordinate system.



a. P = ☐ .

b. Q = ☐ .

c. R = ☐ .

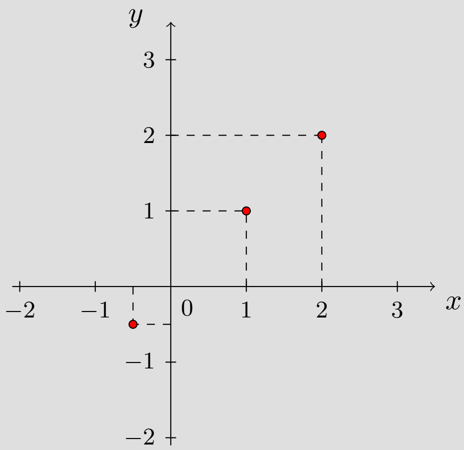
Solution:
The coordinates of the given points are:

$$P = (-2;1)\ ,\quad Q = (1;0)\ ,\quad R = \left(-2;\frac{5}{2}\right) = (-2;2.5)\ .$$

In the following sections we will describe further geometrical objects, such as lines and circles, by coordinates. For this purpose, we first have to understand that points in the plane (described by their coordinates with respect to a given coordinate system) can be collected into so-called **sets of points**. This is illustrated by the example below.

Example 9.1.3

In the figure below, three points are plotted.



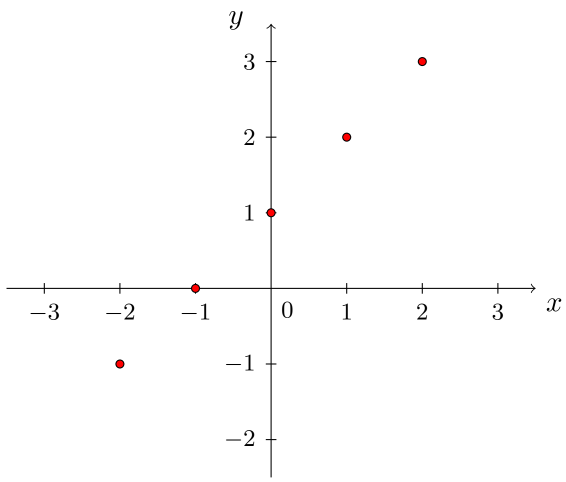
This set of points can be described by the following specification:

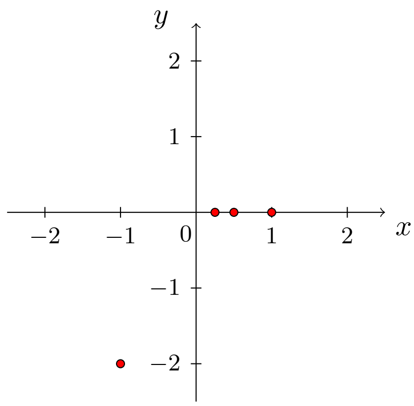
$$\{(-0.5;-0.5);(1;1);(2;2)\} = \{(a;a)\ :\ a \in \{-0.5;1;2\}\}$$

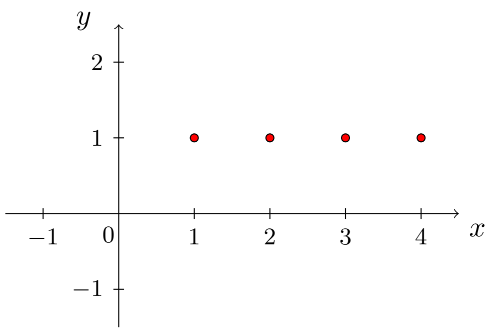
Exercise 9.1.2
Draw the following sets of points in a Cartesian coordinate system.

1. $\{(i;i+1)\ :\ i \in \{-2,-1,0,1,2\}\}$
2. $\{\left(\frac{1}{n};0\right)\ :\ n=1 \vee n=2 \vee n=4\} \cup \{(-1;-2)\}$
3. The set of all points in the first quadrant with integer abscissa smaller than 5 and ordinate 1

Solution:

1. 

2. 

3. 

As we know from Module 5, lines and circles are sets of an infinite number of points. It will be the subject of the following sections to describe their coordinates by means of sets of points and appropriate equations. A special infinite set of points is the collection of *all* points in a coordinate system in the plane. For this set, a specific notion exists.

Info9.1.4

The set of all points in the plane described by the pairs of coordinates in a given Cartesian coordinate system is denoted by

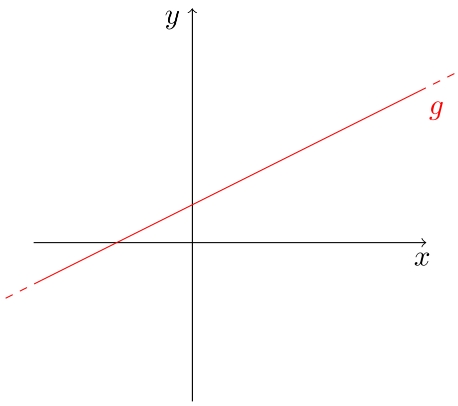
$$\mathbb{R}^2 := \{(x;y) \,:\, x \in \mathbb{R} \wedge y \in \mathbb{R}\} \,.$$

The symbol $\mathbb{R}^2$ reads as “$\mathbb{R}$ two”, “$\mathbb{R}$ to the power of two”, or “$\mathbb{R}$ squared”. This indicates that every point can be described by a pair of coordinates (also denoted as ordered pair) that consists of two real numbers.

9.2 Lines in the Plane

9.2.1 Introduction

In Module 5, lines in the plane were defined as line segments that are continued on both ends indefinitely. In the context of the previous section, these lines can now be considered as infinite sets of points in the plane with respect to a Cartesian coordinate system. The elements of these sets of points then have to satisfy certain (linear) equations. Usually, lines are denoted by lowercase Latin letters $g, h, \ldots$. If a line is drawn in a coordinate system, only a *section* or *segment* of the line can be drawn. However, the line itself extends indefinitely (see figure below).



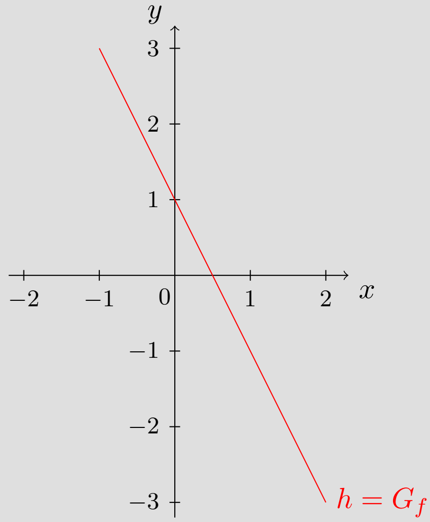
From Module 6 we already know special cases of lines that are described as infinite sets in $\mathbb{R}^2$. Namely, these are the graphs of linear affine functions that we discussed in Section Linear Functions and Polynomials in Module 6. The example below revises the relevant terms of this section before they are used again further on.

Example 9.2.1

The graph (G_f) of the linear affine function

$$f : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & -2x+1 \end{array} \right.$$

is a line h with the y-intercept 1 and the slope -2 (see figure below).



9.2.2 Coordinate Form of Equations of a Line

Let us first introduce the most general form of a coordinate equation of a line. Using this equation, every line in the plane can be specified as an infinite set of points with respect to a given coordinate system.

Info9.2.2

A **line** g in $\mathbb{R}^2$ is a set of points

$$g = \{(x; y) \in \mathbb{R}^2 \,:\, px + qy = c\} \,.$$

Here, p, q, c are real numbers that define the line. At least *one of the numbers* p and q *must be non-zero*. The linear equation

$$px + qy = c$$

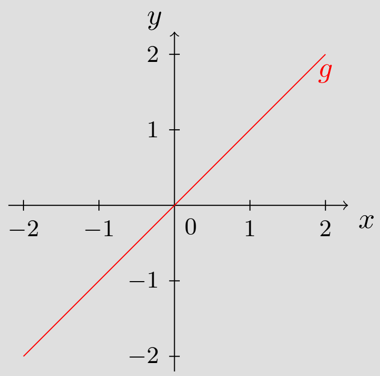
is called the **equation of a line** or, more specifically, to distinguish it from other forms of equations of a line, as **coordinate form of the equation of a line**. A common abbreviation for the explicit set notation given above is to specify only the variable of the line and the equation of the line:

$$g\colon\; px + qy = c \,.$$

The example below shows a few lines and their set notations or equations.

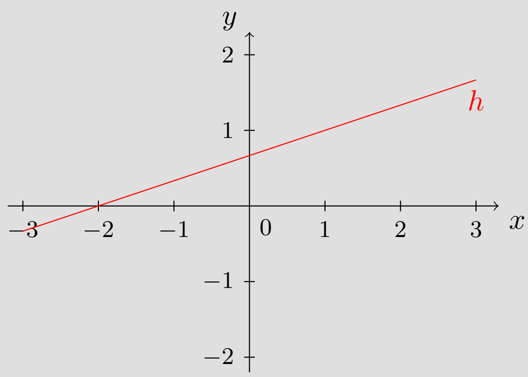
Example 9.2.3

a) $g = \{(x;y) \in \mathbb{R}^2 \,:\, x - y = 0\}$



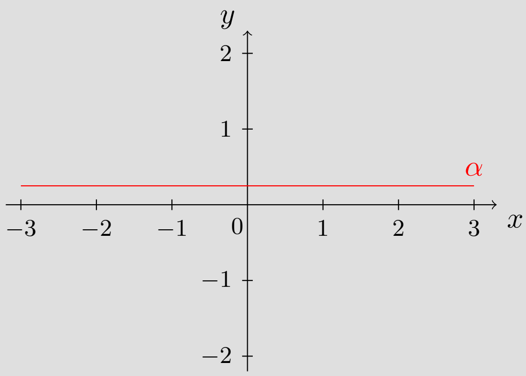
Here, we have $p = 1$, $q = -1$, and $c = 0$.

b) $h\colon \; -x - 2 = -3y \;\Leftrightarrow\; h\colon \; -x + 3y = 2$



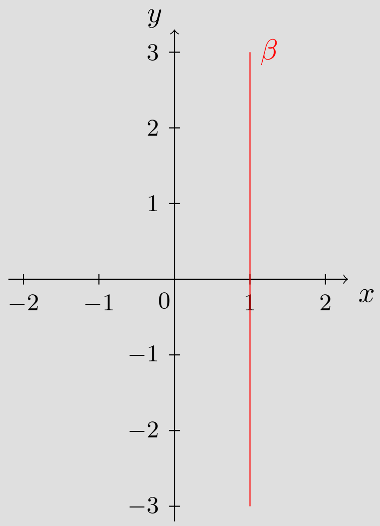
Here, we have $p = -1$, $q = 3$, and $c = 2$.

c) $\alpha\colon\; 4y = 1$



Here, we have $p = 0$, $q = 4$, and $c = 1$.

d) $\beta = \{(x; y) \in \mathbb{R}^2 \,:\, x - 1 = 0\}$



Here, we have $p = 1$, $q = 0$, and $c = 1$.

Now we want to be able to draw a line correctly in a coordinate system in $\mathbb{R}^2$. The line can be uniquely defined by an equation of a line, or by other data. To do this, we have to establish a relation to the graphs of linear affine functions. We also need to know by what kind of data a line in the plane is uniquely defined. This information will be given below.

Info9.2.4

A line g given by an equation of a line in coordinate form

$$g:\ px + qy = c$$

can be converted into **normal form** if $q \neq 0$. In this case, the equation of a line $px + qy = c$ can be solved for y, and the normal form of g is then

$$g:\ y = -\frac{p}{q}x + \frac{c}{q}\ .$$

In this form, the line describes a graph of a linear affine function f with the slope $-\frac{p}{q}$ and the y-intercept $\frac{c}{q}$:

$$f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & y = f(x) = -\frac{p}{q}x + \frac{c}{q}\ . \end{cases}$$

Since the slope and the y-intercept can be read off from the equation of a line in normal form, lines can be drawn in the same way as the graphs of linear affine functions.

Example 9.2.5

The line

$$g = \{(x; y) \in \mathbb{R}^2 \ :\ -x - 2y = 2\}$$

has the equation $-x - 2y = 2$ in coordinate form. This equation can be converted into the form $y = -\frac{1}{2}x - 1$ by equivalent transformations of linear equations. Thus, the line g has the normal form

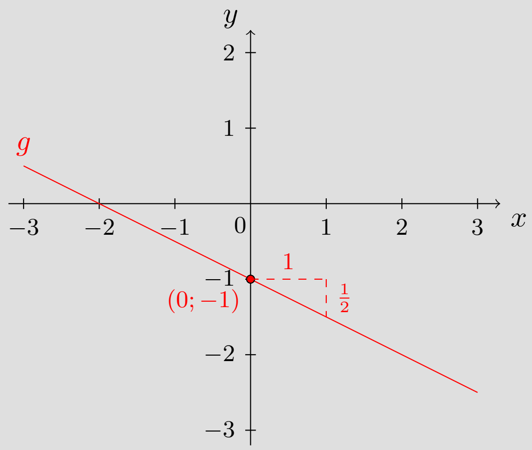
$$g:\ y = -\frac{1}{2}x - 1\ ,$$

and it describes the graph of the linear affine function

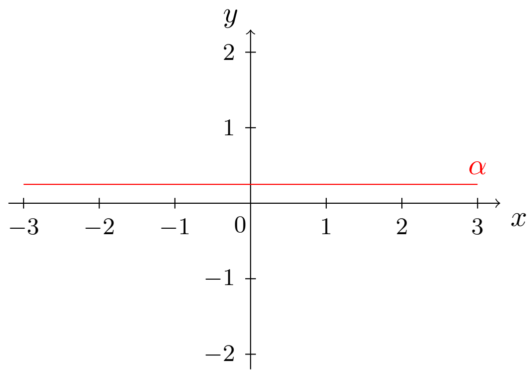
$$f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & y = f(x) = -\frac{1}{2}x - 1 \end{cases}$$

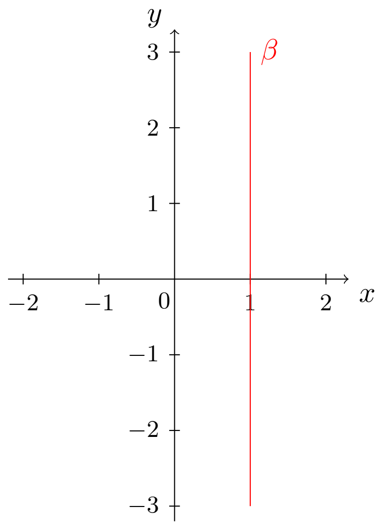
with the slope $-\frac{1}{2}$ and the y-intercept -1.

To draw g one has to understand the following: the y-intercept -1 implies that the point $(0; -1)$ lies on g. Starting from this point, the line g can be drawn correctly by constructing a slope triangle of slope $-\frac{1}{2}$ (by $x = 1$ units to the right and by $y = \frac{1}{2}$ units downwards) in the correct direction:



There two special cases: they can be demonstrated by the two lines $\alpha\colon 4y = 1$ and $\beta\colon x - 1 = 0$ in Example 9.2.3 (see figure below).





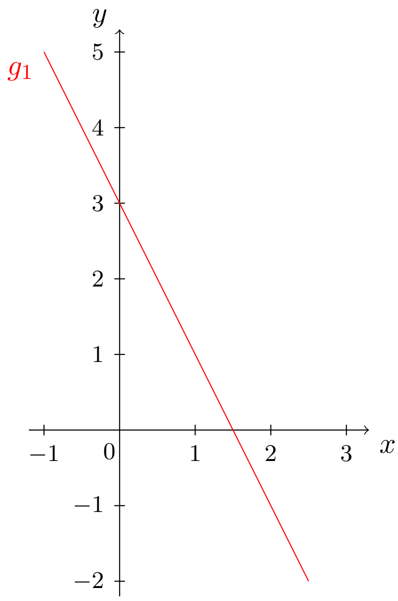
The line α is parallel to the x-axis. Thus, its normal form $\alpha\colon\, y = \frac{1}{4}$ describes the graph of a constant function as a special case of the linear affine functions. The line β is parallel to the y-axis. Its equation of a line cannot be converted into normal form since $q = 0$. This is true for all lines that are parallel to the y-axis.For such lines a normal form does not exist, since these lines cannot be a graph of a function (as discussed in Section 6.1.4). Lines that are parallel to the y-axis have neither have a y-intercept (since they do not intersect the y-axis) nor a slope. However, for the sake of consistency, they can be assigned a slope of ∞.

Exercise 9.2.1
Draw the following lines in a coordinate system. Convert the corresponding equation of a line (if required and possible) into normal form first.

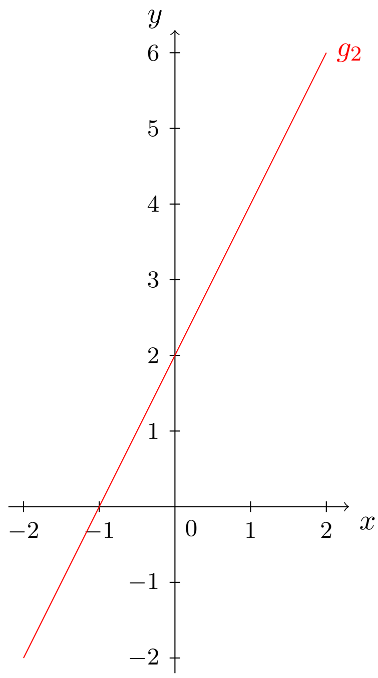
1. $g_1\colon\; y = -2x + 3$
2. $g_2\colon\; -2x + y - 2 = 0$
3. $g_3\colon\; x + 3 = 0$

Solution:

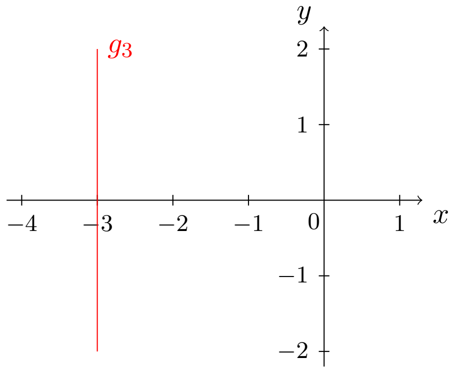
1. The equation of the line is given in normal form, so no transformation is required.



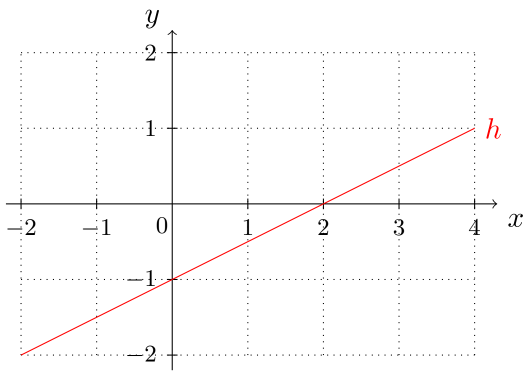
2. The normal form of the line is $g_2\colon\; y = 2x + 2$.



3. The equation of the line cannot be converted into normal form. The line is parallel to the y-axis.



Exercise 9.2.2
Let a line h be given by the figure below.



Specify the equation of the line h in normal form.
$h\colon\ y =$ [ ]

Solution:
The equation of a line of h in normal form is $y = \frac{1}{2}x - 1 = 0.5 \cdot x - 1$.

As well as by an equations, a line in the plane can also be uniquely defined by other data. From these data, the corresponding equation of a line can be derived, and the line can be drawn in a coordinate system.

Info9.2.6

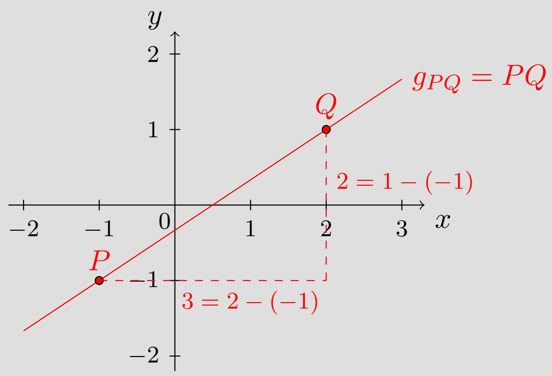
There are two alternative ways to uniquely define a line in the plane:

- **"A line is uniquely defined by two points"** If two points P and Q in $\mathbb{R}^2$ are given, there exists exactly one line g that passes through the points P and Q. The line passing through P and Q is then also denoted by $g = g_{PQ} = g_{QP}$ or simply by $g = PQ$.
- **"A line is uniquely defined by a point and a slope"** If a point P in $\mathbb{R}^2$ and a slope m are given, there exists exactly one line g that passes through P and has the slope m.

The two following examples illustrate how the equation of a line can be derived from the data that uniquely define this line, and how this line can be drawn.

Example 9.2.7

Let the points $P = (-1; -1)$ and $Q = (2; 1)$ be given. The line $g_{PQ} = PQ$ passing through these two points can be drawn immediately. To determine the line of equation it is useful to construct a slope triangle from the two given points:



From the x-coordinates -1 and 2 of P and Q we obtain the width 3 of the slope triangle. From the corresponding y-coordinates -1 and 1 we obtain its height 2. Thus, the slope in the equation of the line is $m = \frac{2}{3}$. For the normal form of the equation of the line g_{PQ} we thus obtain:

$$g_{PQ}\colon\ y = mx + b = \frac{2}{3}x + b\,.$$

Now, only the y-intercept b has to be determined. We know that the line g_{PQ} passes through the two points P and Q. Therefore, we can substitute the x- and y-coordinates of one of these points into the equation of the line, and calculate b. Substituting, for example, the coordinates of the point $Q = (2;1)$ results in

$$1 = \frac{2}{3} \cdot 2 + b \quad \Leftrightarrow \quad b = 1 - \frac{4}{3} = -\frac{1}{3} \ .$$

Using the point $P = (-1;-1)$ would result in the same equation. Thus, the required equation of the line in normal form is

$$g_{PQ}\colon\ y = \frac{2}{3}x - \frac{1}{3} \ .$$

Example 9.2.8

Let the point $R = (2;-1)$ and the slope $m = \frac{1}{2}$ be given. Find the line g that passes through the point R and has the slope $m = \frac{1}{2}$. As in Example 9.2.7, the equation of the line g in normal form can be specified immediately while the y-intercept is still unknown:

$$g\colon\ y = mx + b = \frac{1}{2}x + b \ .$$

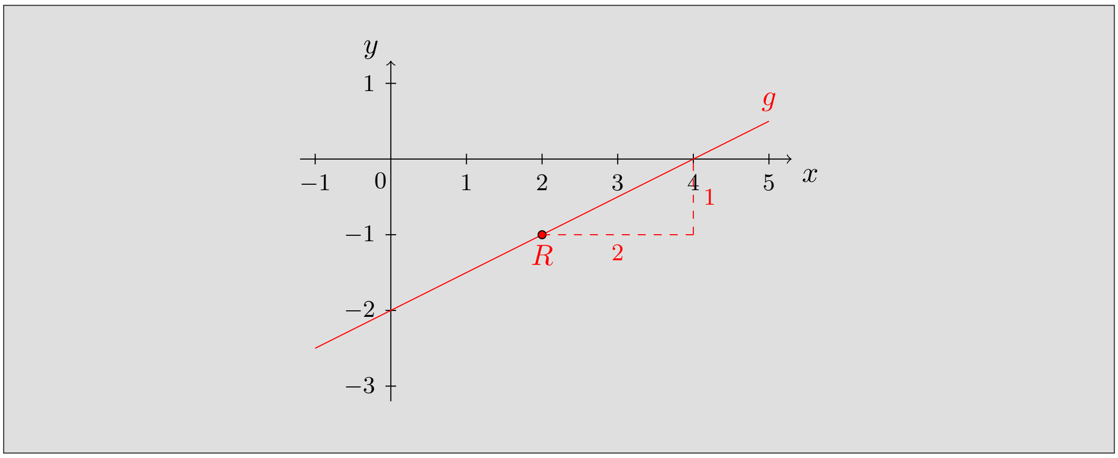
Moreover, both the coordinates x and y of the point R are given here from which the y-intercept can be calculated, as in Example 9.2.7:

$$-1 = \frac{1}{2} \cdot 2 + b \quad \Leftrightarrow \quad b = -2 \ .$$

Thus, the required equation of the line is

$$g\colon\ y = mx + b = \frac{1}{2}x - 2 \ .$$

Using the point $R = (2;-1)$ and the slope $m = \frac{1}{2}$, the line g can also be drawn immediately as illustrated by the figure below.



Exercise 9.2.3
Every set of data given here defines a unique line. For each one, give the equation of the line and then sketch it.

a. The points $A = (1; 5)$ and $B = (3; 1)$ are on the line.
$g_{AB}\colon\ y =$ ☐

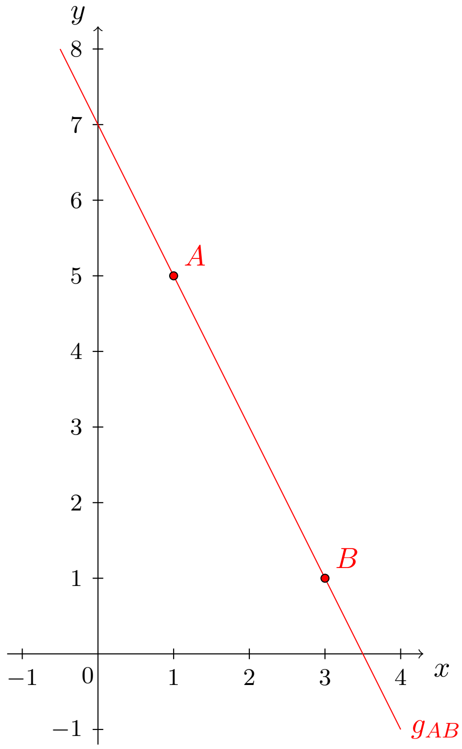
b. The points $S = (1.5; -0.5)$ and $T = \left(\frac{3}{2}; 2\right)$ are on the line.
$g_{ST}\colon\ x =$ ☐

c. The line g passes trough the point $(-4; 3)$ with a slope of -1.
$g\colon\ y =$ ☐

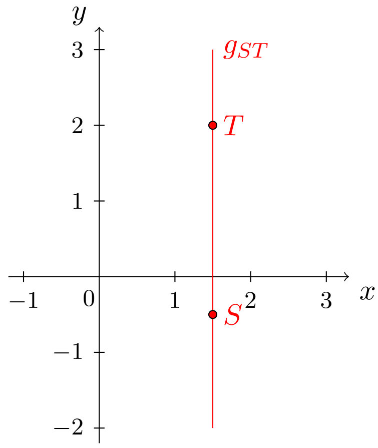
d. The line h passes trough the point $(42; 2)$ with a slope of 0.
$h\colon\ y =$ ☐

Solution:

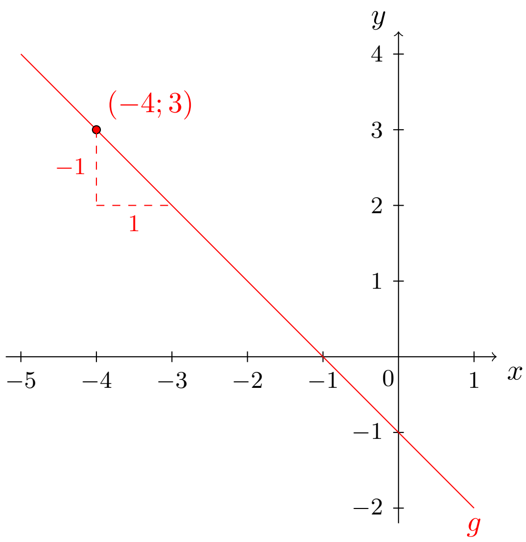
1. $g_{AB}\colon\ y = -2x + 7$



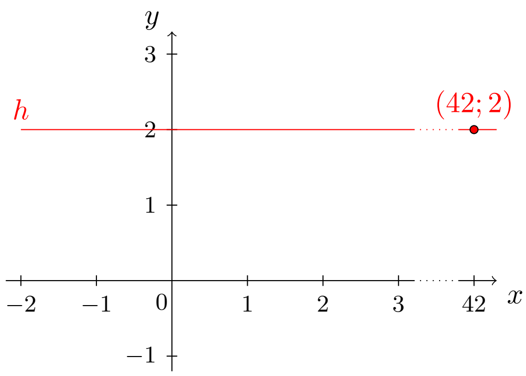
2. $g_{ST}\colon\ x = \frac{3}{2}$



3. $g\colon\ y = -x - 1$



4. $h\colon\ y = 2$



9.2.3 Relative Positions of Lines

In the previous Section 9.2.2 we discussed how to describe lines by means of equations in coordinate form and how to find the equations of a line from given data. This section investigates the relative positions of lines (given by equations) with respect to each other and with respect to other given points. The latter question can be answered very easily since a point either lies on a line or it does not.

Info9.2.9

Let a line

$$g = \{(x; y) \,:\, px + qy = c\}$$

and a point $P = (a; b)$ in $\mathbb{R}^2$ be given. The point P lies on the line (i.e. $P \in g$) if and only if its abscissa and its ordinate satisfy the equation of the line, i.e. if we have

$$pa + qb = c\ .$$

Thus, using an equation of the line, we can check whether points lie on the line or do not.

Example 9.2.10

Let us consider the line

$$h\colon\ x + 2y = -1\ .$$

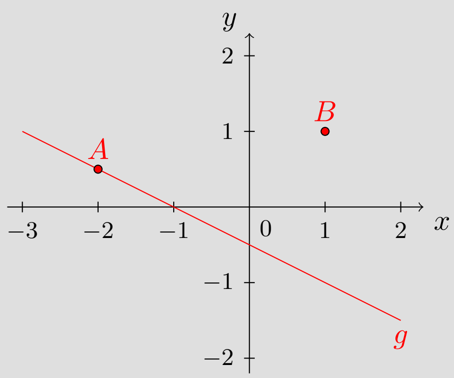
We see that the point $A = \left(-2; \frac{1}{2}\right)$ lies on h since its coordinates $x = -2$ and $y = \frac{1}{2}$ satisfy the equation of a line, i.e.

$$-2 + 2 \cdot \frac{1}{2} = -2 + 1 = -1\ .$$

The point $B = (1; 1)$, however, does not lie on h since its coordinates $x = 1$ and $y = 1$ do not satisfy the equation of a line, i.e.

$$1 + 2 \cdot 1 = 3 \neq -1\ .$$

This is illustrated by the figure below.



Exercise 9.2.4
Decide whether the given points lie on the given lines by inserting and calculating. Tick the points that lie on the line.

$g\colon\ 2x - 4(\frac{y}{2} + x) + 2y = -3$:

- ☐ $P = (1.5; 2)$
- ☐ $Q = \left(-\frac{3}{2}; -4\right)$
- ☐ $R = (0.5; 0)$
- ☐ $S = \left(\frac{9}{6}; 0\right)$
- ☐ $T = (0; -\pi)$

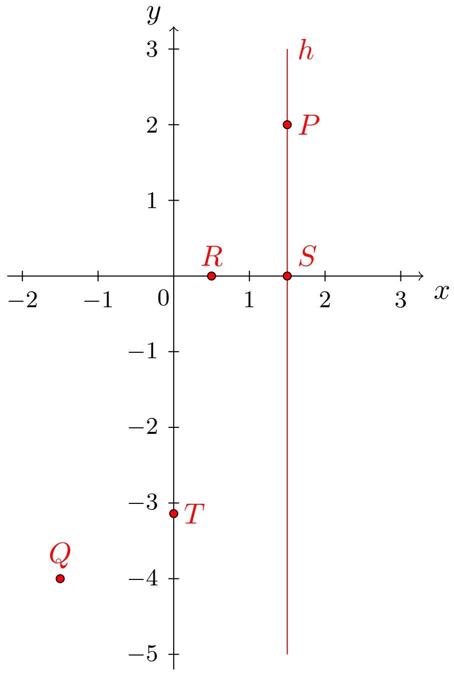
<u>Solution:</u>
The equation of the line can be simplified:

$$2x - 4(\frac{y}{2} + x) + 2y = -3 \quad \Leftrightarrow \quad 2x - 2y - 4x + 2y = -3 \quad \Leftrightarrow \quad -2x = -3 \quad \Leftrightarrow \quad x = \frac{3}{2}\ .$$

Thus, the line h is parallel to the axis of ordinates, and all points ith abscissa 1.5 lie on h. From

$$\frac{9}{6} = \frac{3}{2} = 1.5$$

we see that, on our list, these are only the points P and S. All our other points do not lie on the line h. This is illustrated by the figure below.



Two lines on a plane can have three different relative positions:

Info9.2.11

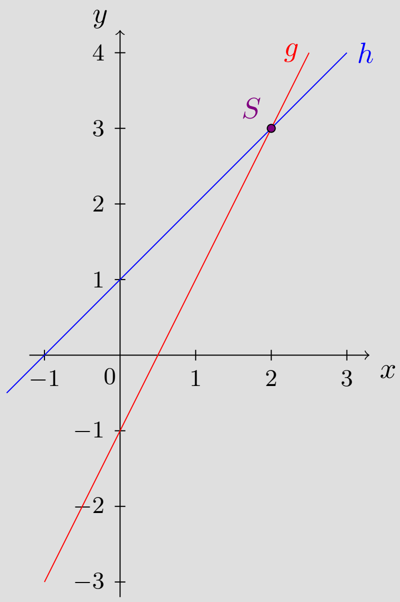
Let g and h be two lines in the plane that are described by equations of a line with respect to a coordinate system. Then the lines have exactly one of the following relative positions with respect to each other:

1. The lines g and h have exactly one point in common, i.e. they intersect. The common point is called the **intersection point**.
2. The lines g and h do not have any points in common, i.e. they do not intersect at all. In this case, the lines are **parallel**.
3. The lines g and h have all their point in common, i.e. they are identical. In this case, the lines are coincident.

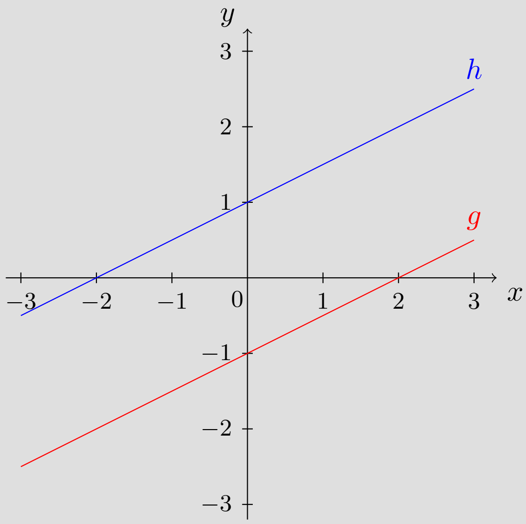
The last case seems to be a little strange at first glance. You may wonder why two names (g and h) exist for the same object. Different equations can in fact describe exactly one and the same line if the equations arise from each other by equivalent transformations. Sometimes this is not obvious but can be seen from a detailed investigation. This idea is illustrated by the example below.

Example 9.2.12

1. The lines $g\colon\ y = 2x - 1$ and $h\colon\ y = x + 1$ intersect. The only point they have in common is the intersection point $S = (2;3)$:

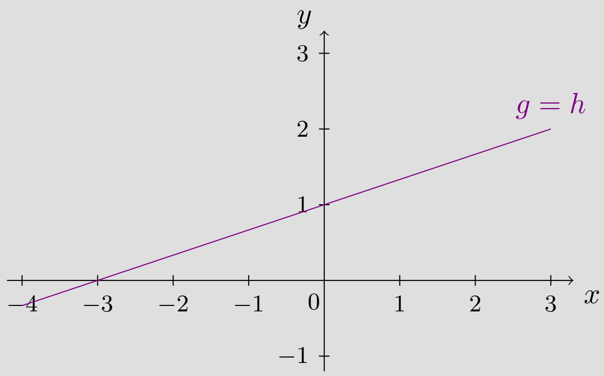


2. The lines $g\colon\ y = \frac{1}{2}x - 1$ and $h\colon\ x - 2y = -2$ do not intersect. They are parallel to each other:



3. The lines $g\colon\ y = \frac{1}{3}x + 1$ and $h\colon\ 2x - 6y = -6$ are coincident since they arise from each other by equivalent transformations:

$$y = \frac{1}{3}x + 1 \quad \Leftrightarrow \quad y - \frac{1}{3}x = 1 \ \mid \cdot (-6) \quad \Leftrightarrow \quad 2x - 6y = -6$$



The methods for calculating the intersection point of lines are the methods for solving a system of two linear equations in two variables (in this case these are the equations of a line) that were discussed in detail in Module 4. In particular, the geometrical aspect of the intersection point of lines was discussed in detail in Section 4.2 in this context. Hence, for the methods of finding an intersection point, we refer to Section 4.2 and highly recommend a brief repetition of the material presented there.

However, if two equations of a line are given in normal form, i.e. if their slopes and

y-intercepts are known, then it can be seen immediately (without a calculation) which of the three relative positions defined in Info Box 9.2.11 applies to the two lines:

Info9.2.13

Let two lines g and h in the plane be given by equations in normal form.

1. If the slopes of g and h are *different*, the two lines intersect.
2. If the slopes of g and h are *the same*, but their y-intercepts are *different*, the two lines are parallel.
3. If the slopes and y-intercepts of g and h are *the same*, the two lines are coincident.

Exercise 9.2.5
Decide by calculation whether the given lines intersect. Tick the corresponding boxes and enter the intersection points for the lines that do intersect. Sketch the pairs of lines.

a. $f\colon\ y = x - 2$ and $g\colon\ y = 2 - x$:
   - ☐ do not intersect (are parallel),
   - ☐ are coincident,
   - ☐ have an intersection point.

b. $f\colon\ y = 1 - x$ and $g\colon\ y = 4 \cdot (3x + 1) - x - 3$:
   - ☐ do not intersect (are parallel),
   - ☐ are coincident,
   - ☐ have an intersection point.

c. $f\colon\ y = 4(x + 1) - x - 1$ and $g\colon\ y = 3x - 3$:
   - ☐ do not intersect (are parallel),
   - ☐ are coincident,
   - ☐ have an intersection point.

d. $f\colon\ y = 5x - 2$ and $g\colon\ y = (2x + 1) + (3x - 3)$:
   - ☐ do not intersect (are parallel),
   - ☐ are coincident,
   - ☐ have an intersection point.

The first intersection point is ______ , the second intersection point is ______ .

Solution:
Equating both functions results in

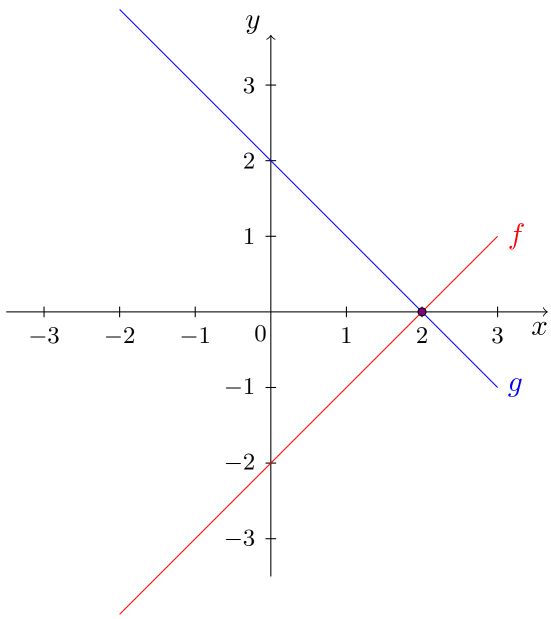
a.
$$x - 2 \;=\; 2 - x \;\Leftrightarrow\; 2x \;=\; 4 \;\Leftrightarrow\; x \;=\; 2\,.$$
Thus, the two lines have the intersection point $P = (2; 0)$.

b.
$$1 - x \;=\; 4(3x + 1) - x - 3 \;\Leftrightarrow\; 12x \;=\; 0 \;\Leftrightarrow\; x \;=\; 0\,.$$
Thus, the two lines have the intersection point $P = (0; 1)$.

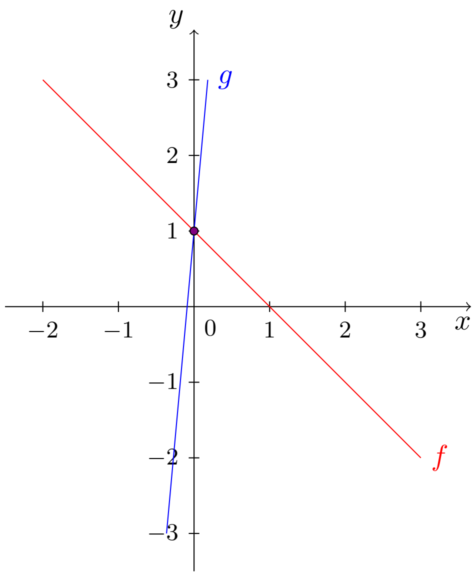
c.
$$4(x + 1) - x - 1 \;=\; 3x - 3 \;\Leftrightarrow\; 0 \;=\; -6\,.$$
These two lines do not intersect since the equation cannot be solved.

d. $(2x + 1) + (3x - 3) = 5x - 2$. These two lines are coincident.

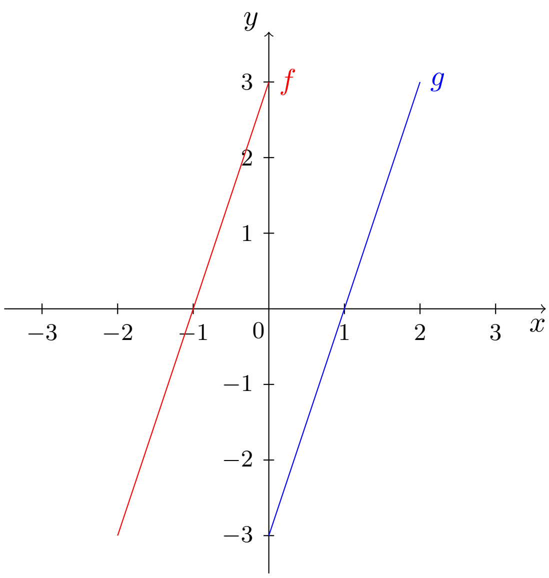
Sketch 1:



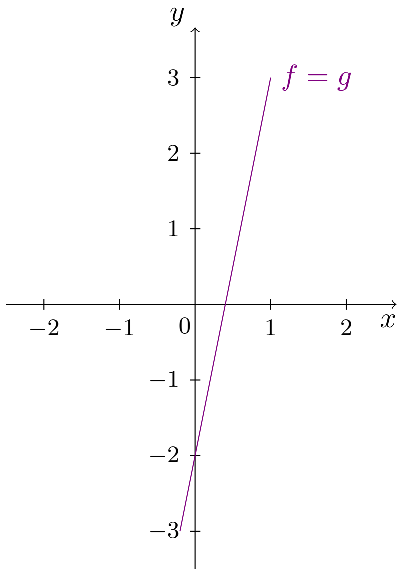
Sketch 2:



Sketch 3:



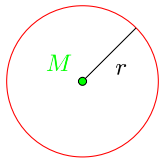
Sketch 4:


y
3
2
1
−2
−1
0
1
2
x
−1
−2
−3
f = g

9.3 Circles in the Plane

9.3.1 Introduction

Everybody has an intuitive understanding of what a **circle** is:

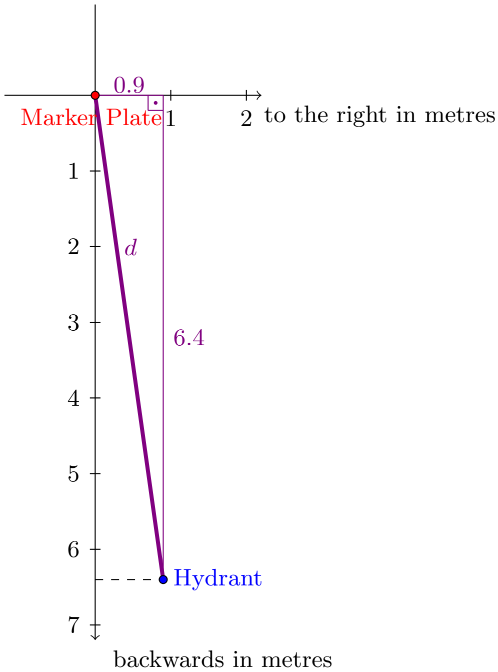


All points on the (red) circle have the same **distance** from exactly one point, namely the **centre** M. This distance r is the so-called **radius** of the circle (see Module 5). If we now want to describe circles in a given coordinate system by equations, as we did in the previous section for lines, and use these descriptions for calculations, we will have to examine the meaning of the terms circle, centre, distance, and radius in more detail. This will enable us to specify an **equation of a circle**. This is the subject of this section.

9.3.2 Distance and Length of a Line Segment

If we recall the first example of a hydrant in Section 9.1.1, we see that we are now able to specify the position of the hydrant in a coordinate system by means of the data given on the hydrant's plate. However, if we are interested in the distance of the hydrant from the plate, then we have to calculate this distance from the coordinates.

For this purpose, Pythagoras' theorem is useful:



For the distance d between the plate and the hydrant, we have

$$d^2 = 0.9^2 + 6.4^2 \ .$$

Thus, the distance d can be calculated (approximately):

$$d = \sqrt{0.81 + 40.96} \approx 6.46 \ .$$

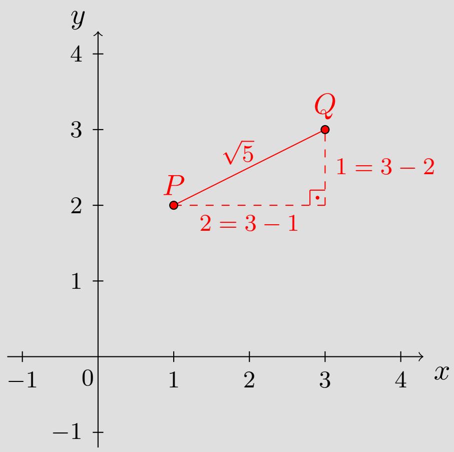
The distance between the plate and the hydrant is (measured in the unit lengths of metres) about 6 metres and 46 centimetres. For purely mathematical purposes, the unit length is not relevant, and that we will omit it again from here onwards.

The example of the plate and the hydrant above can easily be generalised. The distance between two points in $\mathbb{R}^2$ can always be determined using an appropriate right triangle and Pythagoras' theorem.

Example 9.3.1

The points $P = (1; 2)$ and $Q = (3; 3)$ have the distance

$$\sqrt{(3-1)^2 + (3-2)^2} = \sqrt{2^2 + 1^2} = \sqrt{5} \ .$$



Thus, the distance between two points in the plane can be calculated by determining the side lengths of a right triangle from their abscissas and ordinates and then applying Pythagoras' theorem. Furthermore, it is obvious from Example 9.3.1 above that the distance between the points P and Q equals the length of a finite segment of the line PQ, namely the segment between P and Q. This finite segment of the line PQ is called **line segment** between P and Q and is denoted by the symbol $\overline{PQ}$. The **length of the line segment** is the distance between P and Q and is denoted by the symbol $[\overline{PQ}]$.

Info9.3.2

The **distance** of two points $P = (x_0; y_0)$ and $Q = (x_1; y_1)$ in $\mathbb{R}^2$ is given by

$$[\overline{PQ}] = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2} \, .$$

Two points have the distance zero if they coincide.

Exercise 9.3.1 a. Calculate the distance between the two points $A = (-1; -5)$ and $B = (4; 7)$.
$[\overline{AB}] =$ ☐

b. Calculate *the square* of the distance between the two points $P = (3; 0)$ and $Q = (1; \psi)$ depending on ψ.
$[\overline{PQ}]^2 =$ ☐

c. Calculate the coordinates of the point V in the third quadrant that has the distance $3\sqrt{5}$ from the point $U = (2; 1)$ and lies on the line with the slope 2 that passes through the point U.
$V =$ ☐

In the second part of the exercise, ψ is an unknown constant that can be entered as `psi`.

Solution:

a.
$$[\overline{AB}] = \sqrt{(-1-4)^2 + (-5-7)^2} = \sqrt{5^2 + 12^2} = \sqrt{169} = 13$$

b.
$$[\overline{PQ}]^2 = (3-1)^2 + (0-\psi)^2 = 4 + \psi^2$$

c. According to Section 9.2.2, the line with the slope 2 passing through the point $U = (2; 1)$ has the equation $y = 2x + b$ with the y-intercept b that can be found by substituting the coordinates of U into the equation:

$$1 = 2 \cdot 2 + b \quad \Leftrightarrow \quad b = -3 \, .$$

Thus, for the line with the slope 2 passing through the point U, we have the equation $y = 2x - 3$. The coordinates of the points lying on this line are $(x; 2x - 3)$.

We now have to find the point that has the distance $3\sqrt{5}$ from U and lies in the third quadrant. Depending on x, we have for the distance between $U = (2; 1)$ and a point $(x; 2x-3)$:

$$[\overline{U(x;2x-3)}] = \sqrt{(2-x)^2 + (1-(2x-3))^2} = \sqrt{(x-2)^2 + (2x-4)^2} =$$
$$= \sqrt{(x-2)^2 + 4(x-2)^2} = \sqrt{5(x-2)^2} = \sqrt{5}|x-2| \; .$$

Solving the equation

$$\sqrt{5}|x-2| = 3\sqrt{5} \quad \Leftrightarrow \quad |x-2| = 3$$

using the methods described in Section 2.2 results in two values of x, namely -1 and 5. Only $x = -1$ corresponds to an abscissa of a point in the third quadrant. The corresponding ordinate results from substituting this value of x into the equation of the line:

$$y = 2 \cdot (-1) - 3 = -5 \; .$$

Hence, the required point is $V = (-1; -5)$.

9.3.3 Coordinate Equations of Circles

With a coordinate system in the plane at hand, we are now able to describe the points on a circle using the term of distance introduced in the previous Section 9.3.2 by an equation, the so-called equation of a circle. In practice, the compulsory root in the equation for the distance is avoided by using the square of the distance. This is allowed since distances are always non-negative. Thus, for two points $P_1 = (x_1; y_1)$ and $P_2 = (x_2; y_2)$ we have:

$$[\overline{P_1P_2}] = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2} \quad \Leftrightarrow \quad [\overline{P_1P_2}]^2 = (x_2-x_1)^2 + (y_2-y_1)^2 \; .$$

This is summarised in the Info Box below.

Info9.3.3

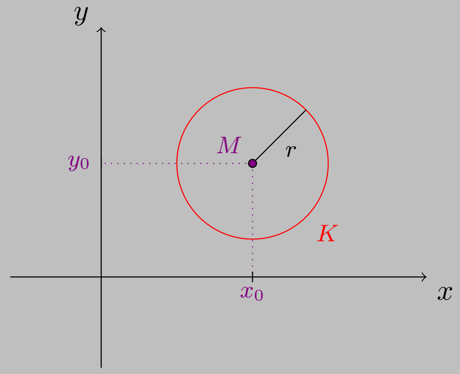
A **circle** K in the plane with a given coordinate system is the set of all points that have a fixed distance $r > 0$, the **radius**, from a common **centre** $M = (x_0; y_0)$. By specifying the radius and the centre a circle is uniquely defined. Thus, we have:

$$K = \{(x; y) \in \mathbb{R}^2 \, : \, (x-x_0)^2 + (y-y_0)^2 = r^2\} \; .$$

As for lines, often just the **equation of a circle** is given:

$$K\colon \; (x-x_0)^2 + (y-y_0)^2 = r^2 \; .$$

All points with coordinates that satisfy the equation of a given circle belong to the circle. This is illustrated by the figure below.



Using the equation of a circle we are now able to describe arbitrary circles in the plane as well as points that lie on that circle and points that do not.

Example 9.3.4

The circle with centre $P = (2;1)$ and radius $r = 2$ is described by the following equation:

$$(x-2)^2 + (y-1)^2 = 2^2 = 4.$$

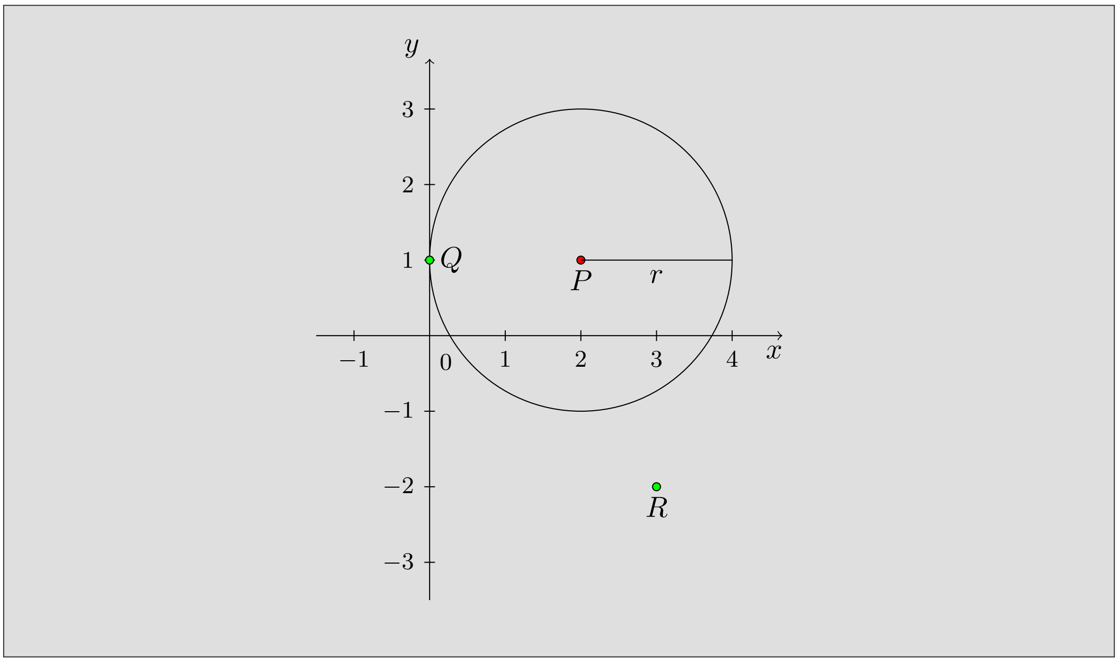
Thus, all points that have the distance of 2 from the point P lie on the circle. For example, the point $Q = (0;1)$ lies on the circle since we have

$$(0-2)^2 + (1-1)^2 = (-2)^2 + 0^2 = 4.$$

In contrast, the point $R = (3;-2)$ does not lie on the circle since its distance from the point P is

$$[\overline{PR}] = \sqrt{(2-3)^2 + (1-(-2))^2} = \sqrt{10} \neq 2.$$

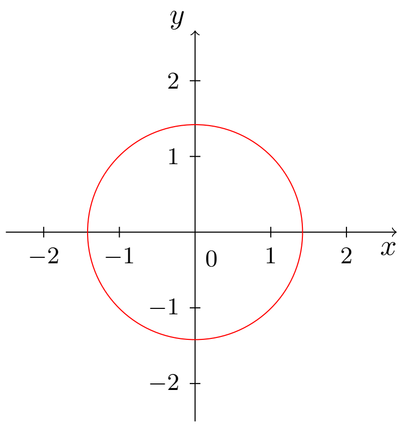
The coordinates of the point R do not satisfy the given equation of a circle.



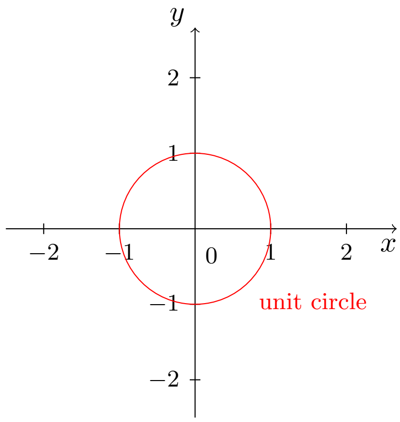
An important special case is any circle that has the origin of the coordinate system as its centre. For example, the equation

$$x^2 + y^2 = 2$$

describes a circle with a radius of $\sqrt{2}$ centred at the origin $(0;0)$ (see figure below).



The most special case of this is the circle with a radius of 1 at the origin $(0;0)$.



This circle is called the **unit circle**, and it is important in trigonometry (see Section 5.6 auf Seite 201 and Section 6.5 auf Seite 275).

Exercise 9.3.2 a. Let a circle Ξ be given by the equation

$$\Xi\colon x^2 + (y+2)^2 = 8.$$

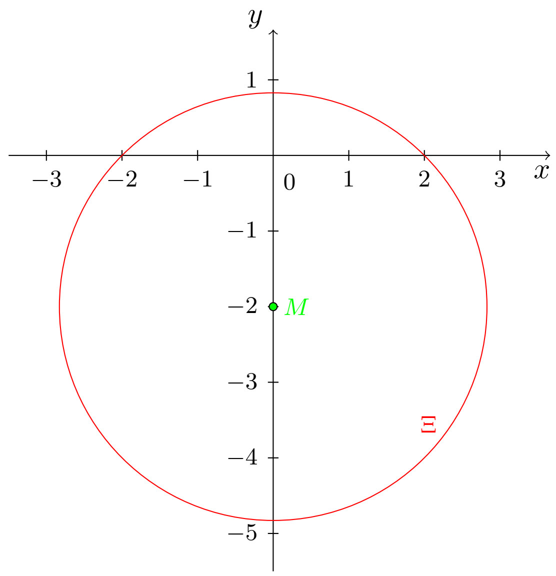
Its centre is at $M =$ [ ] and its radius is $r =$ [ ] . Draw the circle.

b. The equation of a circle with a radius of 1 at $(-2;-1)$ is [ ] $= 1$.
Decide whether the given points lie on the circle. Tick those points that lie on the circle.
- [ ] The origin
- [ ] $(1;1)$
- [ ] $(-2;0)$
- [ ] $\left(-\frac{3}{2};\frac{\sqrt{3}-2}{2}\right)$

Solution:

a.

$$M = (0;-2)$$
$$r = 2\sqrt{2}$$



b.

$$(x+2)^2+(y+1)^2=1$$

The origin and the point $(1;1)$ do not lie on the circle but the points $(-2;0)$ and $\left(-\frac{3}{2};\frac{\sqrt{3}-2}{2}\right)$ do, as can be seen by substituting the coordinates of the points into the given equation of a circle:

$$(0+2)^2+(0+1)^2=4+1=5\neq 1,$$

$$(1+2)^2+(1+1)^2=9+4=13\neq 1,$$

$$(-2+2)^2+(0+1)^2=0+1=1,$$

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

As the examples and exercises above show, it is rather easy to read off the centre and the radius of the circle from its equation of a circle if it is given in the form

$$(x-x_0)^2+(y-y_0)^2=r^2$$

described in Info Box 9.3.3. Therefore, this form is also called **normal form**. Unfortunately, the equation of a circle is often not given in this simple form but has to be transformed by a few calculation steps to enable us to read off the the centre and the radius. This approach is illustrated in the example below.

Example 9.3.5

Let a circle K be given by the equation

$$K\colon x^2 + y^2 - 6x + y + \frac{21}{4} = 0.$$

You can neither immediately see that this is an equation of a circle nor read off the centre and the radius. This equation can be transformed into normal form using the method of completing the square. We will apply this method separately to the terms containing x and the terms containing y in the equation of a circle given above.

For the terms containing x, we have

$$x^2 - 6x = x^2 - 2 \cdot 3x = x^2 - 2 \cdot 3x + 3^2 - 3^2 = x^2 - 6x + 9 - 9 = (x-3)^2 - 9,$$

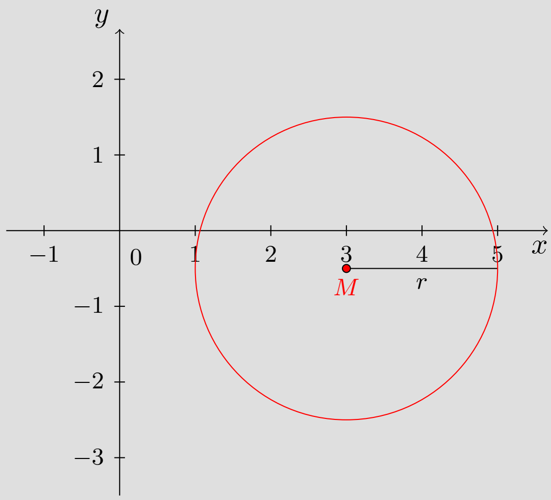
and for the terms containing y, we have

$$y^2 + y = y^2 + 2 \cdot \frac{1}{2} y = y^2 + 2 \cdot \frac{1}{2} y + \left(\frac{1}{2}\right)^2 - \left(\frac{1}{2}\right)^2 = y^2 + y + \frac{1}{4} - \frac{1}{4} = \left(y + \frac{1}{2}\right)^2 - \frac{1}{4}\,.$$

For the equation of the circle, this implies:

$$x^2 + y^2 - 6x + y + \frac{21}{4} = 0 \quad \Leftrightarrow \quad (x-3)^2 - 9 + \left(y + \frac{1}{2}\right)^2 - \frac{1}{4} + \frac{21}{4} = 0 \quad \Leftrightarrow \quad (x-3)^2 + \left(y + \frac{1}{2}\right)^2 = 4$$

This is the normal form of the equation of this circle, and we can read off the centre $M = \left(3; -\frac{1}{2}\right)$ and the radius $r = 2$ easily (see figure below).



Exercise 9.3.3
Find the centre P and the radius ρ of the circle

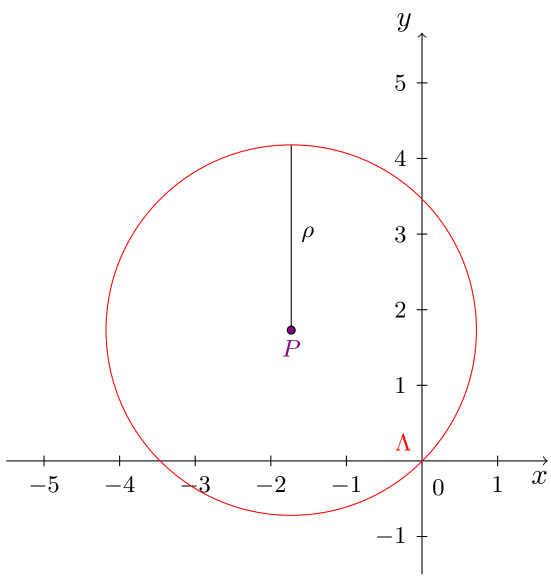
$$\Lambda = \{(x;y) \,:\, x^2 + 2\sqrt{3}x = 2\sqrt{3}y - y^2\}.$$

Transform the equation of the circle into normal form using the method of completing the square. In addition, sketch the circle.
$P =$ ☐
$\rho =$ ☐

<u>Solution:</u>

$$x^2+2\sqrt{3}x = 2\sqrt{3}y-y^2 \quad\Leftrightarrow\quad x^2+2\sqrt{3}x+y^2-2\sqrt{3}y = 0 \quad\Leftrightarrow\quad x^2+2\sqrt{3}x+3+y^2-2\sqrt{3}y+3 = 6$$
$$\Leftrightarrow\quad (x+\sqrt{3})^2 + (y-\sqrt{3})^2 = 6$$

Thus, we have $P = (-\sqrt{3}; \sqrt{3})$ and $\rho = \sqrt{6}$.



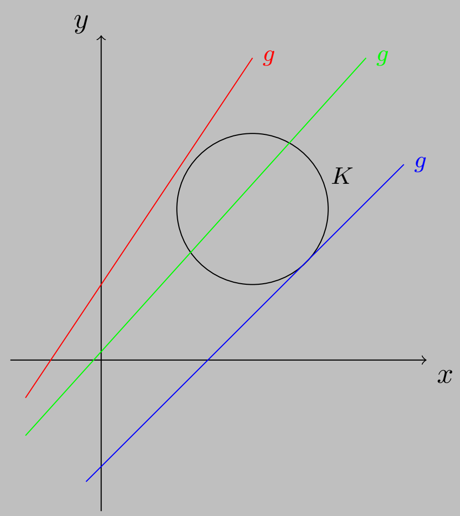
9.3.4 Relative Positions of Circles

We may also raise the question of the relative position in the coordinate system for a circle and a line or for two circles, as we did for two lines. That is, we have to determine whether the two objects intersect, osculate (touch in a single point), or do not have any points in common. In the case of a circle and a line, the intersection point or the osculation point can be calculated easily. For two circles, this is more difficult and goes beyond the scope of this course. For two circles we will only discuss whether or not they intersect or osculate, not at which points this happens.

Info9.3.6

Let a circle K and a line g in the plane be given (by equations in coordinate form with respect to a fixed coordinate system). Then the circle and the line have exactly one of the following three relative positions with respect to each other.

1. The circle K and the line g do not have any points in common. This is true for the red line in the figure below. Such a line is called an **exterior line** to the circle.
2. The circle K and the line g have exactly one point in common, i.e. the line osculates the circle. This is true for the blue line in the figure below. Such a line is called a **tangent line** to the circle.
3. The circle K and the line g have two points in common, i.e. the line intersects the circle. This is true for the green line in the figure below. Such a line is called a **secant line** to the circle.



If a line is a tangent or a secant, the circle and the line have one or two points in common, respectively. How can the intersection point or points be calculated? Since the points lie on both the circle and on the line, they have to satisfy both the equation of the circle and the equation of the line in coordinate form. Thus, we have two equations for the two unknown coordinates of the intersection points, and we are able to calculate the coordinates. However, since in the equations of a circle in coordinate form the unknown coordinates are squared, we do *not* have two linear equations. Hence, the methods for solving systems of linear equations described in Module 5 unfortunately *cannot* be applied. The Info Box below outlines the method for calculating the intersection points.

Info9.3.7

Let a circle K in the plane with centre $(x_0; y_0)$ and radius r be given by an equation of a circle in coordinate form

$$K\colon (x - x_0)^2 + (y - y_0)^2 = r^2 \ ,$$

and a line with slope m and y-intercept b by an equation in normal form

$$g\colon y = mx + b \ .$$

To calculate the potential intersection points, the equation of g can be substituted into the equation of K resulting in a quadratic equation in the variable x:

$$(x - x_0)^2 + (\underbrace{mx + b}_{=y} - y_0)^2 = r^2 \ .$$

For a quadratic equation three cases can occur (see Section 2.1.5):

1. The quadratic equation has no solution. In this case, g is an exterior line to K. An x-coordinate of a common point cannot be found.
2. The quadratic equation has exactly one solution. In this case, g is a tangent line to K. For the x-coordinate (the solution of the quadratic equation), a corresponding y-coordinate can be calculated from the given equation of a line. The two coordinates define the osculation point of the tangent line g and the circle K.
3. The quadratic equation has two solutions. In this case, g is a secant line to K. For both x-coordinates (the two solutions of the quadratic equation), corresponding y-coordinates can be calculated from the given equation of a line. The two pairs of coordinates define the intersection points of the secant line and the circle K.

Example 9.3.8

Let a circle K with the centre $(2;2)$ and a radius of 1 be given by

$$K\colon (x-2)^2 + (y-2)^2 = 1 \ ,$$

and the lines g_1, g_2, and g_3 by

$$g_1\colon y = x - \sqrt{2}$$

$$g_2\colon y = x + 1$$

$$g_3\colon y = 2x + 2 \ .$$

- Line g_1:
  Substituting the equation of the line g_1 into the given equation of the circle results in

  $$(x-2)^2 + (x - \sqrt{2} - 2)^2 = 1 \quad \Leftrightarrow \quad (x-2)^2 + [(x-2) - \sqrt{2}]^2 = 1$$

  $$\Leftrightarrow \quad (x-2)^2 + (x-2)^2 - 2\sqrt{2}(x-2) + (\sqrt{2})^2 = 1 \quad \Leftrightarrow \quad 2(x-2)^2 - 2\sqrt{2}(x-2) + 1 = 0$$

  $$\Leftrightarrow \quad (\sqrt{2}(x-2)-1)^2 = 0 \quad \Leftrightarrow \quad \sqrt{2}(x-2)-1 = 0 \quad \Leftrightarrow \quad x-2 = \frac{1}{\sqrt{2}} \quad \Leftrightarrow \quad x = 2 + \frac{1}{\sqrt{2}} \ .$$

  The resulting quadratic equation has only the solution $x = 2 + \frac{1}{\sqrt{2}}$. Thus, g_1 is a tangent line to K that osculates K in a point with the x-coordinate $2 + \frac{1}{\sqrt{2}}$. The corresponding y-coordinate is calculated from the given equation of the line:

  $$y = 2 + \frac{1}{\sqrt{2}} - \sqrt{2} = 2 + \frac{\sqrt{2}}{2} - \sqrt{2} = 2 - \frac{1}{\sqrt{2}} \ .$$

  Thus, the tangent line g_1 osculates the circle K at the point $P = \left(2 + \frac{1}{\sqrt{2}}; 2 - \frac{1}{\sqrt{2}}\right)$.

- Line g_2:
  Substituting the equation of the line g_2 into the given equation of the circle results in

  $$(x-2)^2 + (x+1-2)^2 = 1 \quad \Leftrightarrow \quad (x-2)^2 + (x-1)^2 = 1$$

  $$\Leftrightarrow \quad x^2 - 4x + 4 + x^2 - 2x + 1 = 1 \quad \Leftrightarrow \quad 2x^2 - 6x + 4 = 0 \quad \Leftrightarrow \quad x^2 - 3x + 2 = 0$$

  $$\Leftrightarrow \quad x_{1,2} = \frac{3 \pm \sqrt{9-8}}{2} = \frac{3 \pm 1}{2} \quad \Leftrightarrow \quad x_1 = 2 \ \wedge \ x_2 = 1 \ .$$

  The resulting quadratic equation has two solutions. Thus, g_2 is a secant line to the circle K that intersects K in two points with the x-coordinates $x_1 = 2$

and $x_2 = 1$. The corresponding y-coordinates are calculated from the given equation of the line:

$$y_1 = x_1 + 1 = 2 + 1 = 3 \text{ and } y_2 = x_2 + 1 = 1 + 1 = 2 .$$

Thus, the two points $Q_1 = (2; 3)$ and $Q_2 = (1; 2)$ are the intersection points of the secant line g_2 with the circle K.

- Line g_3:
  Substituting the equation of the line g_3 into the given equation of the circle results in

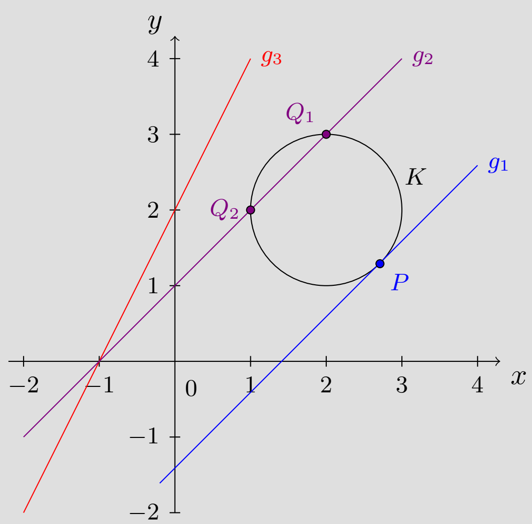
  $$(x-2)^2+(2x+2-2)^2 = 1 \Leftrightarrow (x-2)^2+(2x)^2 = 1 \Leftrightarrow x^2-4x+4+4x^2 = 1$$
  $$\Leftrightarrow 5x^2 - 4x + 3 = 0 .$$

  The discriminant (see Section 2.1.5) of this quadratic equation is

  $$(-4)^2 - 4 \cdot 5 \cdot 3 = 16 - 60 < 0 .$$

  Thus, the equation does not have a solution, and the line g_3 is an exterior line to the circle K.

The circle, all three lines, and the intersection points are shown in the figure below.



Exercise 9.3.4

Let a circle K be given by the equation

$$K\colon x^2 + (y-1)^2 = 2\ ,$$

and two lines by

$$g\colon y = \sqrt{7}x + 5$$
$$h\colon y = 2x + 2\ .$$

a. Show that g is a tangent line to K, and calculate the x-coordinate and y-coordinate of the osculation point $(x; y)$ of g and K.
   $x =$ ☐
   $y =$ ☐

b. Show that h is a secant line to K, and calculate the x-coordinates and y-coordinates of the intersection points $(x_1; y_1)$ and $(x_2; y_2)$ of h and K.
   $x_1 =$ ☐
   $y_1 =$ ☐
   $x_2 =$ ☐
   $y_2 =$ ☐

Solution:

a. Substituting g into K results in

$$x^2 + (\sqrt{7}x + 5 - 1)^2 = 2 \;\Leftrightarrow\; x^2 + (\sqrt{7}x + 4)^2 = 2 \;\Leftrightarrow\; x^2 + 7x^2 + 8\sqrt{7}x + 16 = 2$$
$$\Leftrightarrow\; 8x^2 + 8\sqrt{7}x + 14 = 0 \;\Leftrightarrow\; 4x^2 + 4\sqrt{7}x + 7 = 0 \;\Leftrightarrow\; (2x + \sqrt{7})^2 = 0\ .$$

Thus, the quadratic equation has exactly one solution, and g is a tangent line to K. For the x-coordinate of the osculation point we have:

$$(2x + \sqrt{7})^2 = 0 \;\Leftrightarrow\; 2x + \sqrt{7} = 0 \;\Leftrightarrow\; x = -\frac{\sqrt{7}}{2}\ .$$

Substituting x into the given equation of a line results in:

$$y = \sqrt{7}\left(-\frac{\sqrt{7}}{2}\right) + 5 = -\frac{7}{2} + 5 = \frac{3}{2}\ .$$

b. Substituting h in K results in

$$x^2 + (2x + 2 - 1)^2 = 2 \;\Leftrightarrow\; x^2 + (2x + 1)^2 - 2 = 0 \;\Leftrightarrow\; x^2 + 4x^2 + 4x + 1 - 2 = 0$$
$$\Leftrightarrow\; 5x^2 + 4x - 1 = 0 \;\Leftrightarrow\; x_{1,2} = \frac{-4 \pm \sqrt{16 + 20}}{10} = \frac{-4 \pm 6}{10} \;\Leftrightarrow\; x_1 = \frac{1}{5} \wedge x_2 = -1\ .$$

Since the quadratic equation has two solutions, h is a secant line to K. The corresponding y-coordinates result again from substituting the x-coordinates into the given equation of a line:

$$y_1 = 2\left(\frac{1}{5}\right) + 2 = \frac{2}{5} + 2 = \frac{12}{5}$$

and

$$y_2 = 2\,(-1) + 2 = -2 + 2 = 0\ .$$

Exercise 9.3.5
Let a circle K be given by the equation

$$K\colon (x-3)^2 + y^2 = 2\ ,$$

and the line g with the slope 2 and the y-intercept b by the equation

$$g\colon y = 2x + b$$

in normal form. Find the interval in which b must lie such that g is a secant line to K.
$b \in]$ [ ] ; [ ] $[$.

Solution:
Substituting the equation of the line (with the unknown b) into the equation of a circle results in

$$(x-3)^2+(2x+b)^2 = 2 \;\Leftrightarrow\; x^2-6x+9+4x^2+4bx+b^2 = 2 \;\Leftrightarrow\; 5x^2+(4b-6)x+b^2+7 = 0\ .$$

For g to be a secant line to K, this quadratic equation (in the variable x) must have two solutions. This is the case if its discriminant is positive, i.e. if we have:

$$(4b-6)^2-4\cdot5(b^2+7) > 0 \;\Leftrightarrow\; 16b^2-48b+36-20b^2-140 > 0 \;\Leftrightarrow\; -4b^2-48b-104 > 0$$

$$\Leftrightarrow\; b^2+12b+26 < 0\ .$$

The solutions of the quadratic equation $b^2+12b+26=0$ (in the variable b) are

$$b_{1,2} = \frac{-12 \pm \sqrt{144-104}}{2} = \frac{-12 \pm \sqrt{40}}{2} = \frac{-12 \pm 2\sqrt{10}}{2} = -6 \pm \sqrt{10}\ .$$

Thus, the required interval is $]-6-\sqrt{10}; -6+\sqrt{10}[$.

Of course, lines defined by equations that cannot be transformed into normal form (i.e. lines that are parallel to the y-axis) can intersect or osculate circles as well. The method described above cannot be applied directly to such lines . The example below illustrates the method which is applied in this case.

Example 9.3.9

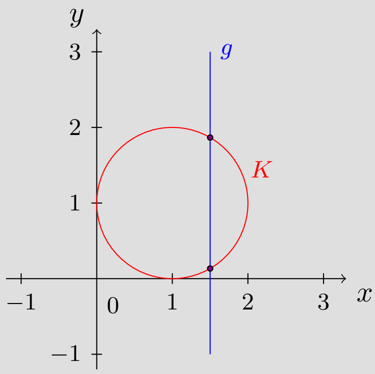
Let a circle K be given by the equation

$$K\colon (x-1)^2+(y-1)^2=1\ ,$$

and line g by the equation

$$g\colon x=\frac{3}{2}\ .$$

The two objects are shown in the figure below.



Obviously, g is a secant line to K. In this case, the intersection point cannot be calculated by solving the equation of a line for y. Instead, the equation of a line $x=\frac{3}{2}$ is simply substituted into the given equation of a circle. This results in two values of y, i.e. the y-coordinates of the intersection points:

$$(\frac{3}{2}-1)^2+(y-1)^2=1 \quad\Leftrightarrow\quad \frac{1}{4}+y^2-2y+1=1 \quad\Leftrightarrow\quad y^2-2y+\frac{1}{4}=0$$

$$\Leftrightarrow\quad y_{1,2}=\frac{2\pm\sqrt{4-1}}{2}=1\pm\frac{\sqrt{3}}{2} \quad\Leftrightarrow\quad y_1=1+\frac{\sqrt{3}}{2}\ \wedge\ y_2=1-\frac{\sqrt{3}}{2}\ .$$

Obviously, both corresponding x-coordinates are equal to $\frac{3}{2}$ since the intersection points lie on the line g. Thus, the two intersection points of the line g and the circle K are $\left(\frac{3}{2};1+\frac{\sqrt{3}}{2}\right)$ and $\left(\frac{3}{2};1-\frac{\sqrt{3}}{2}\right)$.

The Info Box below lists the different cases for the relative positions of two circles together with general criteria that allow to decide which of the cases for two given circles applies.

Info9.3.10

Let two (different) circles, K_1 with centre M_1 and radius r_1 and K_2 with centre M_2 and radius r_2, be given (by equations of a circle in the plane with respect to a fixed coordinate system). Then the circles have exactly one of the following three relative positions with respect to each other:

1. The circles K_1 and K_2 do not have any points in common, i.e. they do not intersect. This is the case if and only if for the radii r_1 and r_2 and the distance $[\overline{M_1M_2}]$ of the two centres, we have the inequalities

$$[\overline{M_1M_2}] > r_1 + r_2 \ \text{ or } \ [\overline{M_1M_2}] < |r_1 - r_2| \,.$$

2. The circles K_1 and K_2 have one point in common, i.e. they osculate. This is the case if and only if for the radii r_1 and r_2 and the distance $[\overline{M_1M_2}]$ of the two centres, we have the inequalities

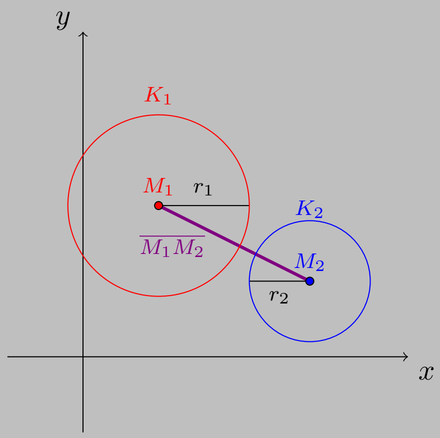
$$[\overline{M_1M_2}] = r_1 + r_2 \ \text{ or } \ [\overline{M_1M_2}] = |r_1 - r_2| \,.$$

3. The circles K_1 and K_2 have two points in common, i.e. they intersect. This is the case if and only if for the radii r_1 and r_2 and the distance $[\overline{M_1M_2}]$ of the two centres, we have the inequality

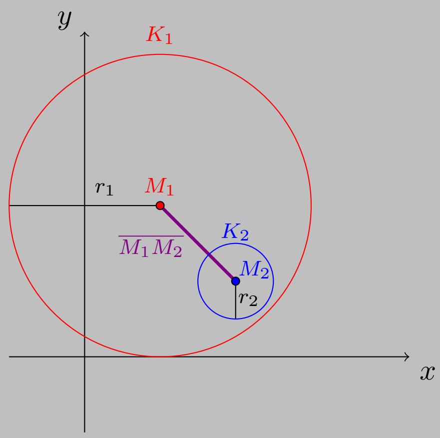
$$|r_1 - r_2| < [\overline{M_1M_2}] < r_1 + r_2 \,.$$

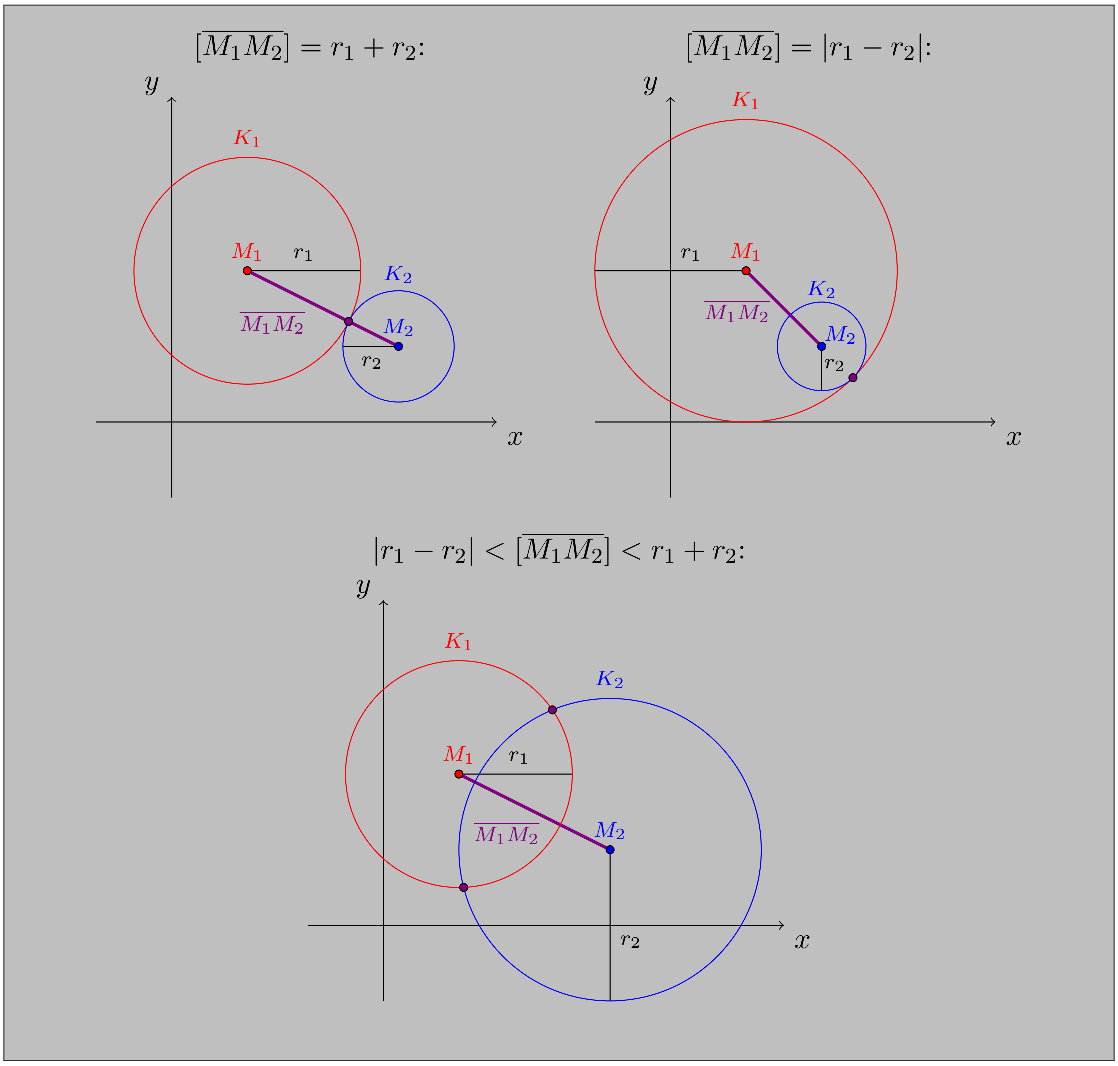
The three cases are illustrated in the figure below.

$[\overline{M_1M_2}] > r_1 + r_2$:



$[\overline{M_1M_2}] < |r_1 - r_2|$:





Thus, the relative position of two circles can be determined from their centres and radii. The calculation of possible intersection points is more difficult and goes beyond the scope of this course.

Example 9.3.11

Let a circle K with centre M_K and radius r_K be given by the equation

$$K: x^2 + y^2 - 2y = 1 ,$$

and a circle L with centre M_L and unknown radius $r > 0$ by the equation

$$L\colon (x+1)^2 + (y-1)^2 = r^2 \ .$$

Find the values of r for which the circles K and L intersect, osculate, or do not intersect.

From the given equation of a circle L the centre M_L can be read off directly: $M_L = (-1;1)$. However, the equation for K is not in normal form. By completing the square this equation can easily be transformed into normal form such that the centre M_K and the radius r_K can be read off:

$$x^2 + y^2 - 2y = 1 \quad \Leftrightarrow \quad x^2 + y^2 - 2y + 1 - 1 = 1 \quad \Leftrightarrow \quad x^2 + (y-1)^2 = 2 \ .$$

Thus, we have $M_K = (0;1)$ and $r_K = \sqrt{2}$. Hence, the distance between the two centres is

$$[\overline{M_K M_L}] = \sqrt{(0+1)^2 + (1-1)^2} = 1 \ ,$$

and the sum of the two radii is

$$r + r_K = r + \sqrt{2} \ .$$

Since $r > 0$, we have

$$r + \sqrt{2} > \sqrt{2} > 1 = [\overline{M_K M_L}] \ .$$

Hence $[\overline{M_K M_L}] \geq r + \sqrt{2}$ cannot occur in this case. The geometric reason for this is that the centre of L lies within K, as can be seen from the figure below. According to the criteria listed in Info Box 9.3.10, the two circles osculate if

$$|r - \sqrt{2}| = 1 = [\overline{M_K M_L}] \ .$$

According to the section on absolute value inequalities, this equation is satisfied if

$$r - \sqrt{2} = 1 \ \text{ or } \ r - \sqrt{2} = -1 \ .$$

Hence, the two radii r for which the circle L osculates the circle K are:

$$r = \sqrt{2} + 1 \ \text{ and } \ r = \sqrt{2} - 1 \ .$$

Now we have still to investigate the remaining possible values of r. For $0 < r < \sqrt{2} - 1$, we have

$$|r - \sqrt{2}| = -(r - \sqrt{2}) = \sqrt{2} - r > 1 = [\overline{M_K M_L}] \ ,$$

i.e. K and L do not have any points in common. For $\sqrt{2} - 1 < r < \sqrt{2} + 1$, we have to distinguish the cases $\sqrt{2} - 1 < r \leq \sqrt{2}$ and $\sqrt{2} < r < \sqrt{2} + 1$ if we investigate the absolute value. In the case $\sqrt{2} - 1 < r \leq \sqrt{2}$, we have

$$|r - \sqrt{2}| = -(r - \sqrt{2}) = \sqrt{2} - r < 1 = [\overline{M_K M_L}] \ ,$$

and in the case $\sqrt{2} < r < \sqrt{2}+1$, we have

$$|r - \sqrt{2}| = r - \sqrt{2} < 1 = [\overline{M_K M_L}] \ .$$

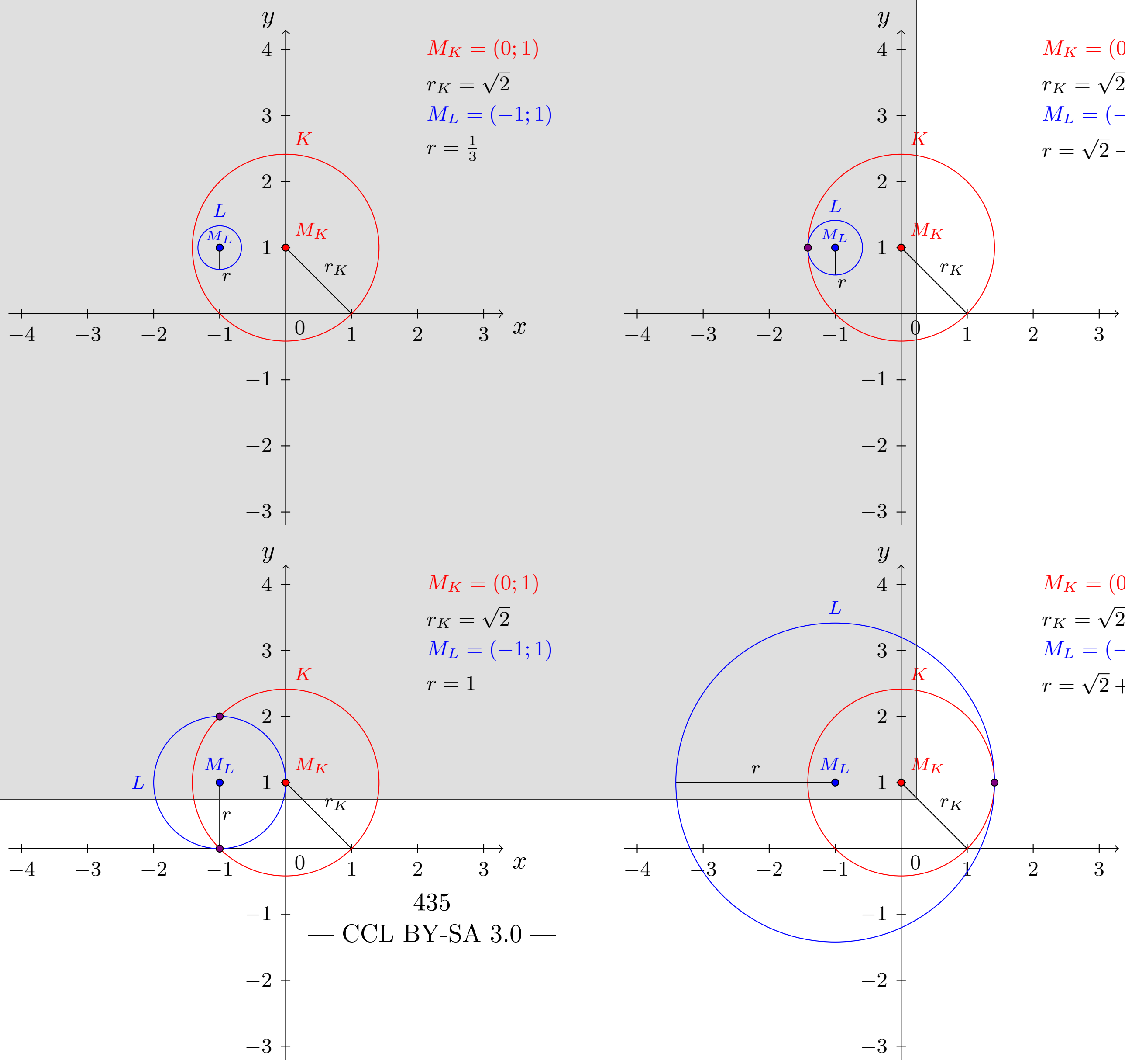
Thus, the circles K and L have two points in common for $\sqrt{2}-1 < r < \sqrt{2}+1$, i.e. the circles intersect. Finally, we have in the case $r > \sqrt{2}+1$:

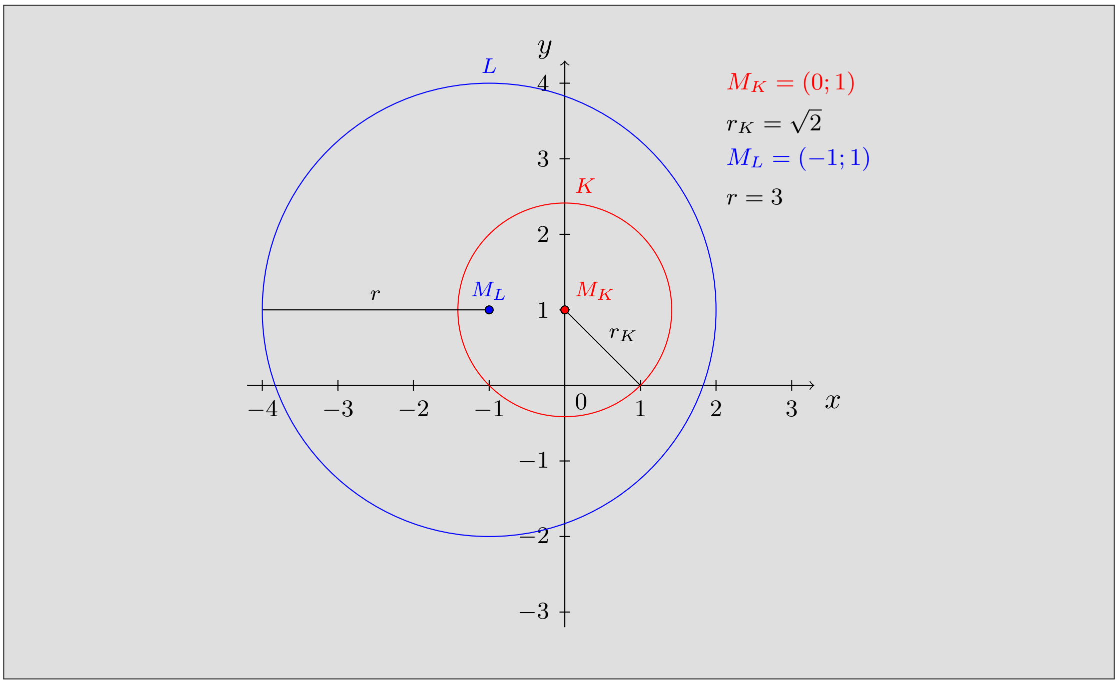
$$|r - \sqrt{2}| = r - \sqrt{2} > 1 = [\overline{M_K M_L}] \ .$$

The circles K and L also do not have any points in common in this case. In summary, we have the following conditions for the relative position of the two circles K and L:

- If $r \in \{\sqrt{2}-1; \sqrt{2}+1\}$, the two circles K and L osculate at one point.
- If $r \in \,]\sqrt{2}-1; \sqrt{2}+1[$, the two circles K and L intersect at two points.
- If $r \in \,]0; \sqrt{2}-1[\,\cup\,]\sqrt{2}+1; \infty[$, the two circles K and L do not have any points in common.

The figure below shows some of the cases for the circles K and L.





Exercise 9.3.6
Let two circles K_1 and K_2 be given by the equations

$$K_1\colon (x+6)^2 + (y+4)^2 = 64$$

$$K_2\colon x^2 + 2x + y^2 - 16y + 40 = 0\ .$$

The two circles K_1 and K_2

- ☐ osculate at one point,
- ☐ intersect at two points,
- ☐ do not have any points in common.

Solution:
The centre M_1 and the radius r_1 of the circle K_1 can immediately be read off: $M_1 = (-6;-4)$ and $r_1 = 8$. To transform the equation of the circle K_2 into normal form, we have to complete the square:

$$x^2 + 2x + y^2 - 16y + 40 = 0 \quad\Leftrightarrow\quad x^2 + 2x + 1 + y - 2\cdot 8y + 64 + 40 - 64 - 1 = 0$$

$$\Leftrightarrow\quad (x+1)^2 + (y-8)^2 = 25\ .$$

Now we can also read off the centre M_2 and the radius r_2 of K_2: $M_2 = (-1;8)$ and $r_2 = 5$. From these values, we calculate

$$[\overline{M_1 M_2}] = \sqrt{(-1+6)^2 + (8+4)^2} = \sqrt{25+144} = \sqrt{169} = 13$$

and

$$r_1 + r_2 = 13 \; .$$

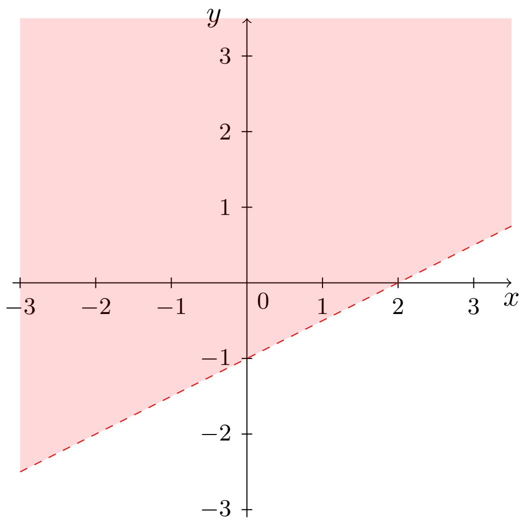
In this case we have $[\overline{M_1 M_2}] = 13 = r_1 + r_2$, and according to the criteria listed in Info Box 9.3.10, the circles K_1 and K_2 osculate at one point.

9.4 Regions in the plane

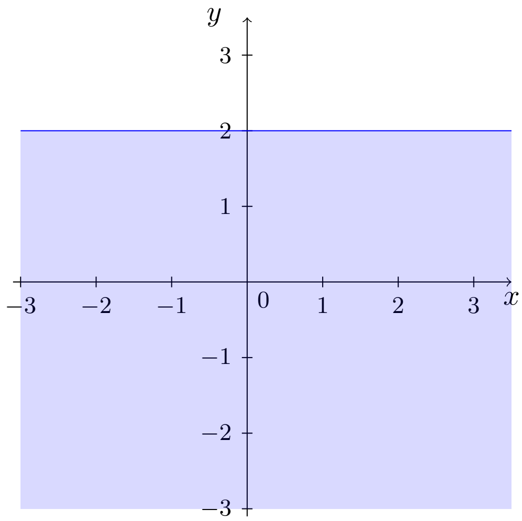
9.4.1 Introduction

While in the previous sections curves in the plane (lines or circles) were investigated by means of coordinate equations, in this section we will replace the coordinate *equations* by coordinate *inequalities*. Thereby not curves but *regions* in the plane are described, which are bounded by the corresponding curves. Depending on whether the inequality is strict ($<$ or $>$) or not ($\leq$ or $\geq$), the bounding curve is a part of the region or not. Regions can be, for example, areas above and below lines, areas within or outside circles, or even intersections of these. A few examples are shown in the figures below.

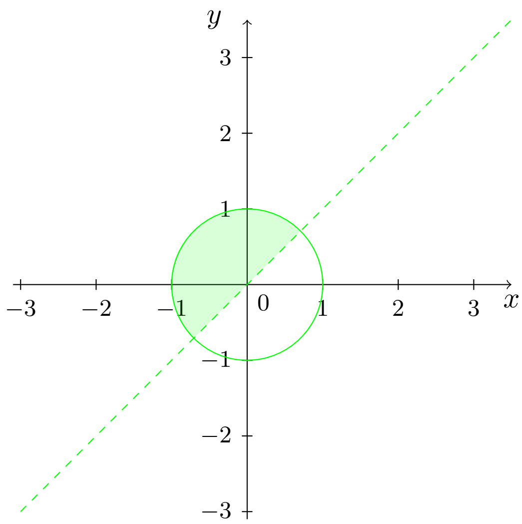
- Region above the line $y = \frac{1}{2}x - 1$ excluding the line itself:



- Region below the line $y = 2$ including the line itself:



- Region above the line $y = x$ and within the unit circle $x^2 + y^2 = 1$ including the points on the circle but excluding the points on the line:



Here and in the following we will use the general convention that curves included in the region are drawn as solid lines and curves *excluded* from the region are drawn as dashed lines.

9.4.2 Regions bounded by Lines and Circles

The Info Box below lists the regions that can occur if the equals sign in the equation of a line is replaced by an inequality sign.

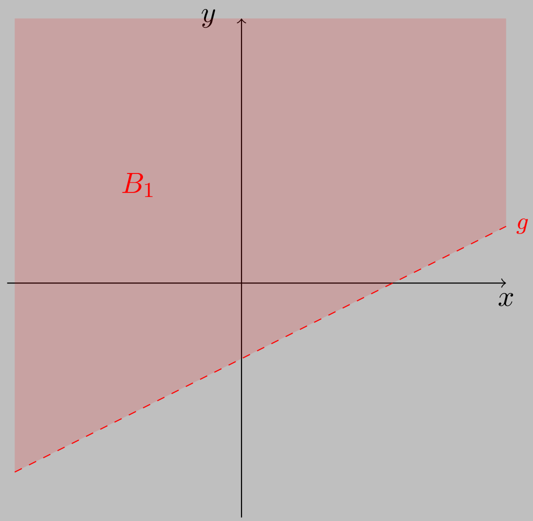
Info9.4.1

Let a line g in the plane (with slope m and y-intercept b) be given by

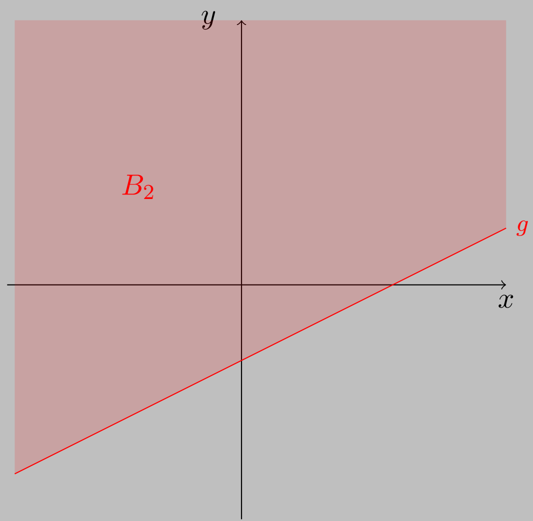
$$g\colon y = mx + b$$

in normal form with respect to a fixed coordinate system. Substituting an inequality sign for the equals sign results in the following sets that describe regions in the plane:

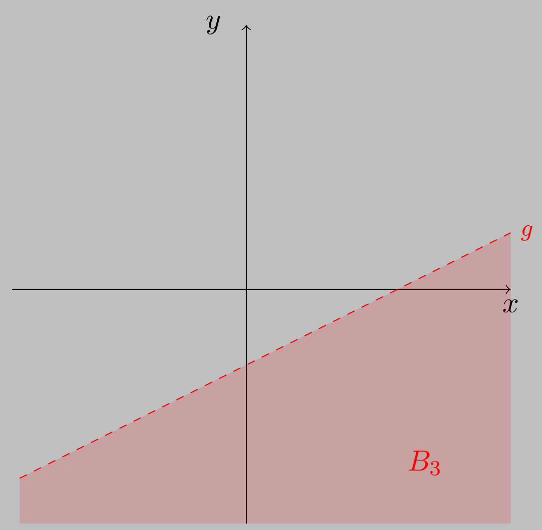
- $B_1 := \{(x;y) \in \mathbb{R}^2 : y > mx + b\}$ = "region above the line excluding the points on the line itself"



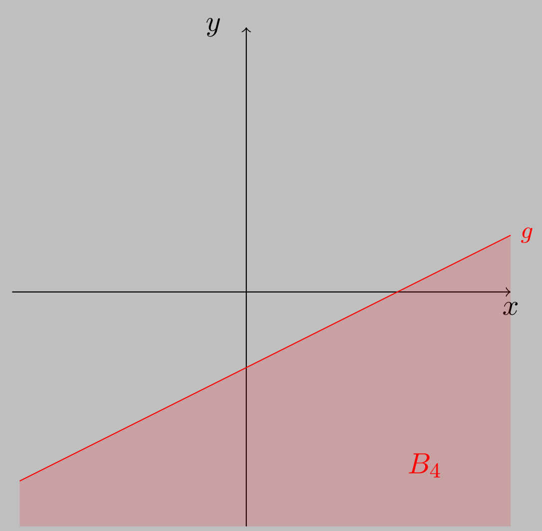
- $B_2 := \{(x;y) \in \mathbb{R}^2 : y \geq mx + b\}$ = "region above the line including the points on the line itself"



- $B_3 := \{(x;y) \in \mathbb{R}^2 : y < mx + b\}$ = "region below the line excluding the points on the line itself"



- $B_4 := \{(x;y) \in \mathbb{R}^2 : y \leq mx + b\}$ = "region below the line including the points on the line itself"



For equations of a line that cannot be transformed into normal form, the line of thought on the resulting regions is analogous. The example below shows two simple cases.

Example 9.4.2

Let two lines be given by the equations

$$g\colon y = -x + 1$$

$$h\colon x = -1 \ .$$

Find and sketch the following sets:

$$A = \text{“region above } g \text{ excluding the points on the line } g \text{ itself”},$$

$$B = \text{“region to the right of line } h \text{ including the points on the line } h \text{ itself”},$$

and $A \cap B$.

From the Info Box above, we have

$$A = \{(x; y) \in \mathbb{R}^2 \ : \ y < -x + 1\}$$

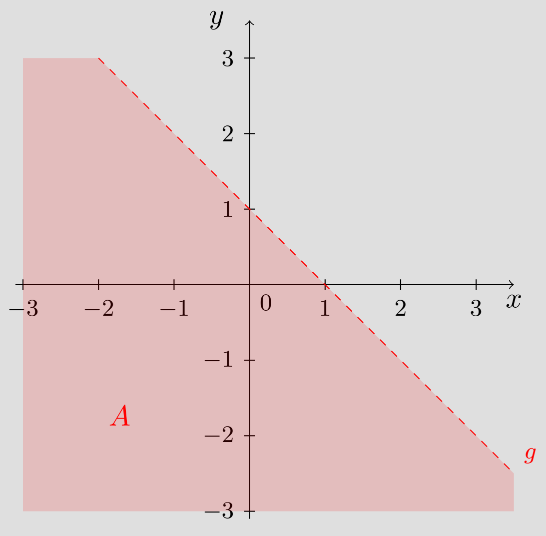
and

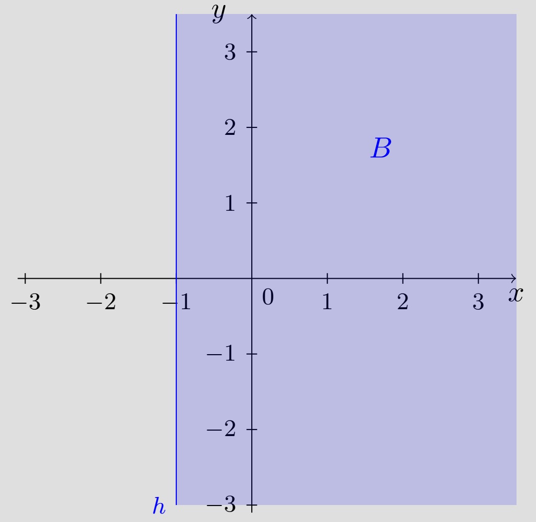
$$B = \{(x; y) \in \mathbb{R}^2 \ : \ x \geq -1\}$$

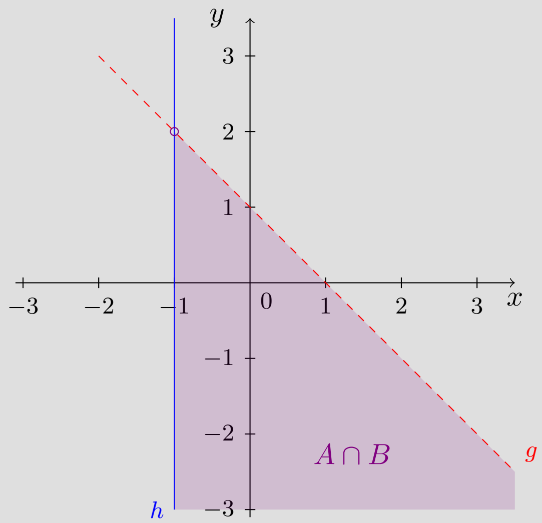
since the points to the right of line h have x-coordinates that are greater than -1. Thus, the intersection $A \cap B$ is the set of points with coordinates that satisfy both the conditions:

$$A \cap B = \{(x; y) \in \mathbb{R}^2 \ : \ y < -x + 1 \wedge x \geq -1\} \ .$$

This is illustrated by the figures below.







Here, the intersection point $(-1; 2)$ of the two lines is not an element of $A \cap B$ since points on the line g are generally excluded from the region $A \cap B$. In the figure above this is indicated by a small empty circle at this point.

The example above illustrates the following: regions given by coordinate inequalities derived from equations of lines are easy to specify. It becomes more difficult if intersections of such regions are considered. The following, more difficult example shows that even absolute values can be involved.

Example 9.4.3

Describe the set defined by

$$M = \{(x; y) \,:\, |x - y| < 1\}$$

in words and sketch it.

For absolute values (see Section 2.2) a case analysis is required as usual:

1. $x - y \geq 0 \;\Leftrightarrow\; x \geq y$
   In this case the inequality $|x - y| < 1$ can be solved for

   $$|x - y| < 1 \;\Leftrightarrow\; x - y < 1 \;\Leftrightarrow\; y > x - 1\,.$$

   Thus, the set M contains all points $(x; y)$ with coordinates that satisfy the inequalities $x \geq y$ and $y > x - 1$, i.e. those points that lie above the line $y = x - 1$ but below the angle bisector $y = x$.

2. $x - y < 0 \;\Leftrightarrow\; x < y$
   In this case the inequality $|x - y| < 1$ can be solved for

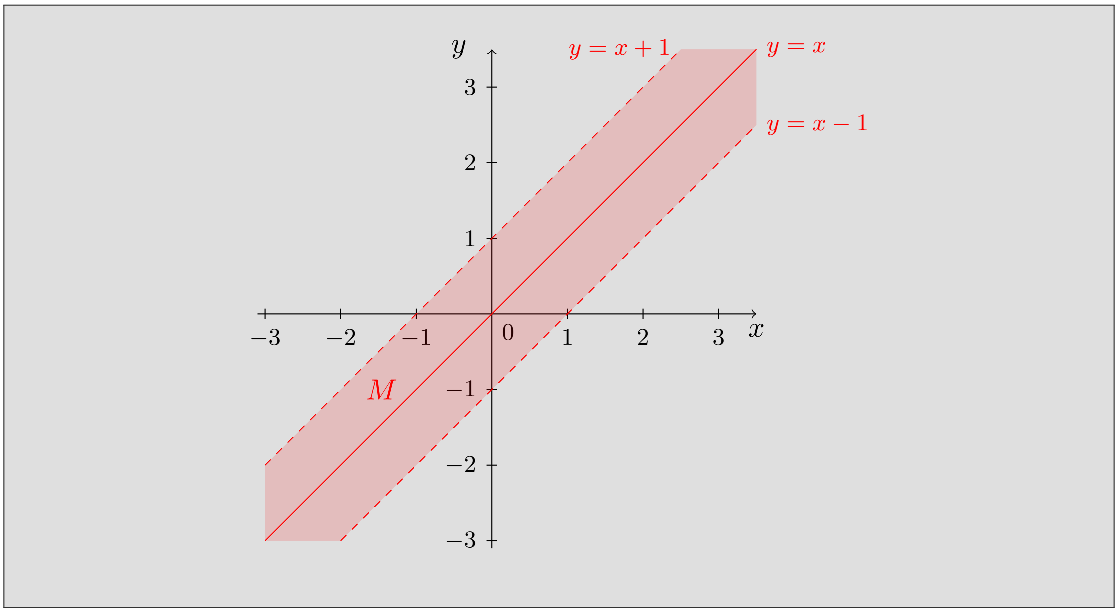
   $$|x - y| < 1 \;\Leftrightarrow\; -(x - y) < 1 \;\Leftrightarrow\; y < x + 1\,.$$

   Thus, the set M contains all points $(x; y)$ with coordinates that satisfy the inequalities $x < y$ and $y < x + 1$, i.e. those points that lie below the line $y = x + 1$ but above the angle bisector $y = x$.

From these two cases, we obtain the following description of the set M:

$M =$ “all points between the lines $y = x-1$ and $y = x+1$ that do not lie on those lines”

The figure below shows the corresponding sketch.



Exercise 9.4.1
Describe and sketch the following sets:

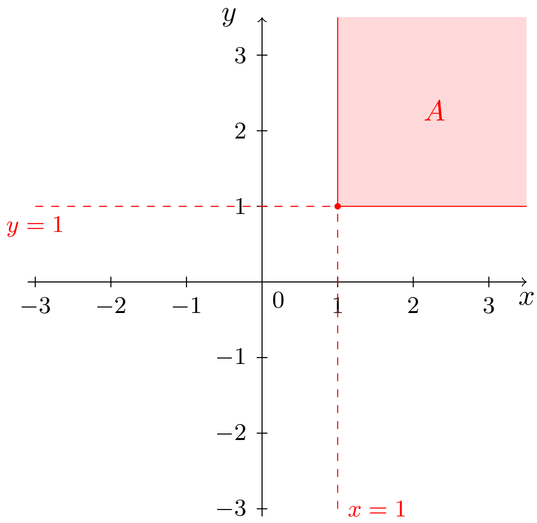
a. $A = \{(x; y) \in \mathbb{R}^2 : y \geq 1\} \cap \{(x; y) \in \mathbb{R}^2 : x \geq 1\}$

b. $B = \{(x; y) \in \mathbb{R}^2 : |2x - y| \geq 1\}$

c. $C = \{(x; y) \in \mathbb{R}^2 : |y| > x + 1\}$

Solution:

a. All points with coordinates that satisfy the inequality $y \geq 1$ lie on or above the line $y = 1$, and all points with coordinates that satisfy the inequality $x \geq 1$ lie on or to the right of the line $x = 1$. The set A contains the points that satisfy both conditions (intersection of two sets). Thus, we have:



b. Case analysis:

a) $2x - y \geq 0 \;\Leftrightarrow\; y \leq 2x$:

$$|2x - y| \geq 1 \;\Leftrightarrow\; 2x - y \geq 1 \;\Leftrightarrow\; y \leq 2x - 1$$

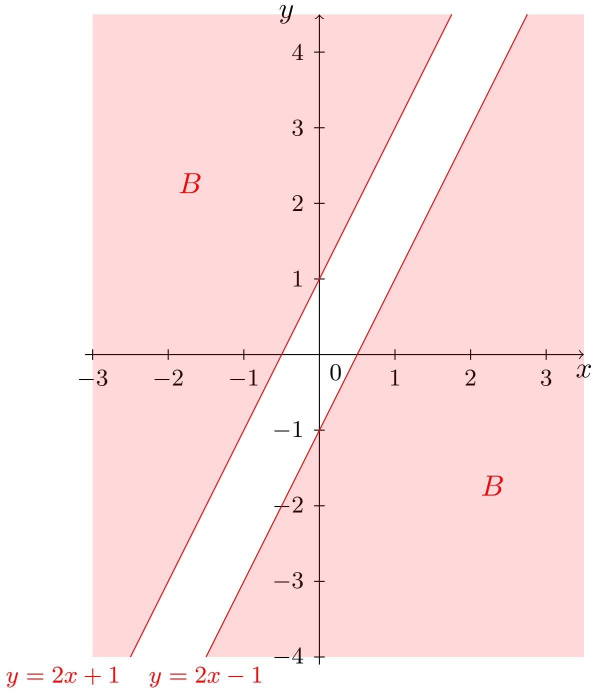
This case includes all points that lie on or below the line $y = 2x - 1$.

b) $2x - y < 0 \;\Leftrightarrow\; y > 2x$:

$$|2x - y| \geq 1 \;\Leftrightarrow\; -(2x - y) \geq 1 \;\Leftrightarrow\; y \geq 2x + 1$$

This case includes all points that lie on or above the line $y = 2x + 1$.

All in all the following regions are included:



c. Case analysis:

a) $y \geq 0$:

$$|y| > x + 1 \;\Leftrightarrow\; y > x + 1$$

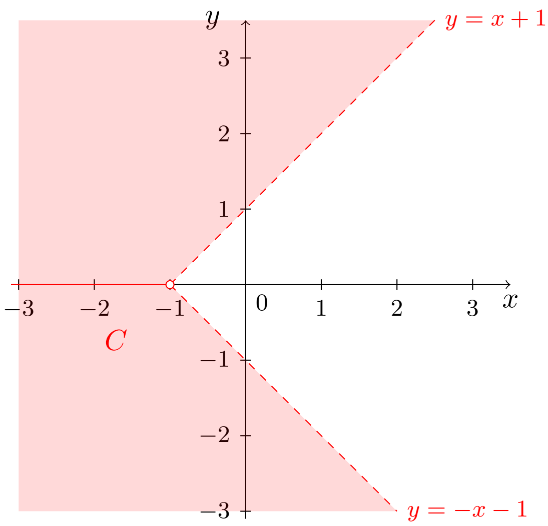
This case includes all points with a non-negative y-coordinate that lie above the line $y = x + 1$.

b) $y < 0$:

$$|y| > x + 1 \;\Leftrightarrow\; -y > x + 1 \;\Leftrightarrow\; y < -x - 1$$

This case includes all points with a negative y-coordinate that lie below the line $y = -x - 1$.

All in all the following region results:



The point $(-1; 0)$ does not belong to the set C.

The Info Box below lists the regions in the plane that can be bounded by a circle.

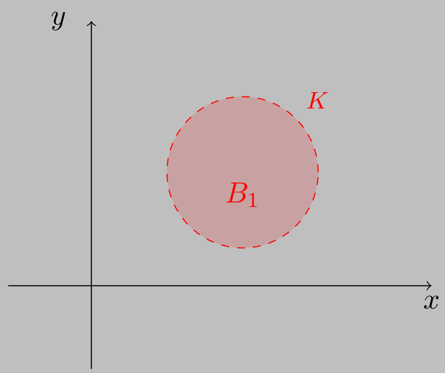
Info9.4.4

Let a circle K in the plane (with the centre $M = (x_0; y_0)$ and the radius r) be given by the equation

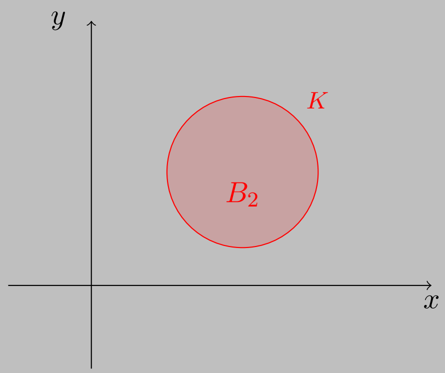
$$K\colon (x - x_0)^2 + (y - y_0)^2 = r^2$$

in normal form with respect to a fixed coordinate system. Then replacing the equals sign with an inequality results in the following sets that describe regions an a plane:

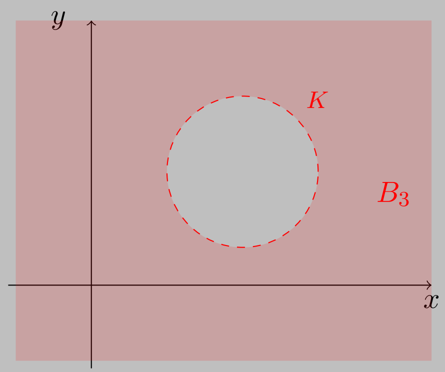
- $B_1 := \{(x; y) \in \mathbb{R}^2 \,:\, (x - x_0)^2 + (y - y_0)^2 < r^2\}$ = "region within the circle excluding the points on the circle itself"



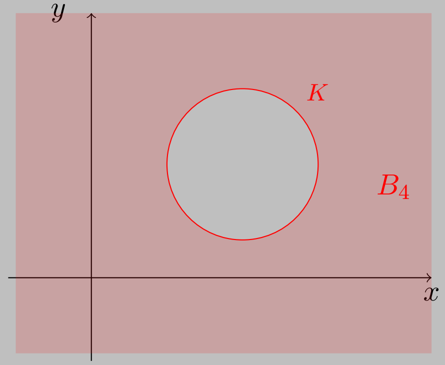
- $B_2 := \{(x; y) \in \mathbb{R}^2 : (x - x_0)^2 + (y - y_0)^2 \leq r^2\}$ = "region within the circle including the points on the circle itself"



- $B_3 := \{(x; y) \in \mathbb{R}^2 : (x - x_0)^2 + (y - y_0)^2 > r^2\}$ = "region outside the circle excluding the points on the circle itself"



- $B_4 := \{(x; y) \in \mathbb{R}^2 : (x - x_0)^2 + (y - y_0)^2 \geq r^2\}$ = "region outside the circle including the points on the circle itself"



The example below shows a few special cases of regions that are bounded by circles as well as a few more complex cases that arise by combining several regions bounded by circles or lines.

Example 9.4.5

Let two circles K_1 and K_2 be given by the equations

$$K_1\colon x^2 + y^2 = 4$$

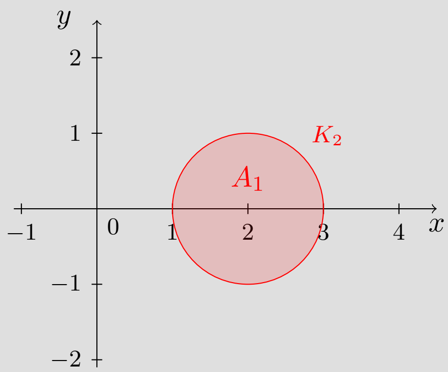
and

$$K_2\colon (x-2)^2 + y^2 = 1\ ,$$

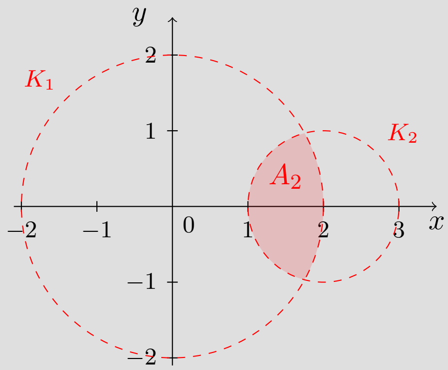
and let the line g be given by the equation

$$g\colon y = -x + 1\ .$$

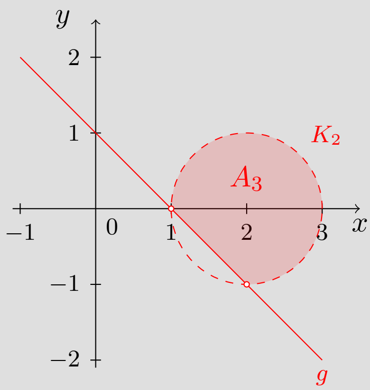
- The set $A_1 := \{(x;y) \in \mathbb{R}^2 : (x-2)^2 + y^2 \leq 1\}$ consists of all points within and on the circle K_2:



- The set $A_2 := \{(x;y) \in \mathbb{R}^2 : (x-2)^2 + y^2 < 1\} \cap \{(x;y) \in \mathbb{R}^2 : x^2 + y^2 < 4\}$ consists of all points that lie both within the circle K_1 and within the circle K_2, i.e. within the intersection of the two discs, excluding the points on the circles themselves:



- The set $A_3 := \{(x;y) \in \mathbb{R}^2 : (x-2)^2 + y^2 < 1 \wedge y \geq -x+1\}$ consists of all points that lie both within the circle K_2 – excluding the points on the circle – and above the line g:



The intersection points between the circle and the line do not belong to the set A_3.

Exercise 9.4.2
Sketch the given sets:

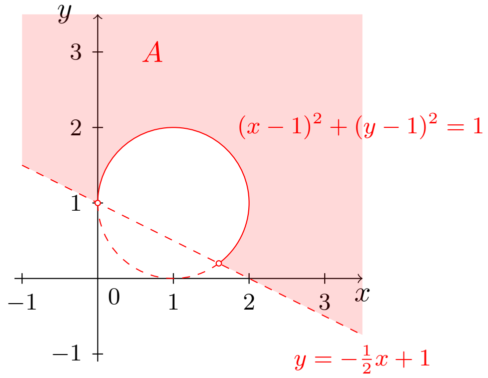
a. $A = \{(x;y) \in \mathbb{R}^2 : (x-1)^2 + (y-1)^2 \geq 1\} \cap \{(x;y) \in \mathbb{R}^2 : y > -\frac{1}{2}x + 1\}$

b. $B = \{(x;y) \in \mathbb{R}^2 : |x| < 1\} \cup \{(x;y) \in \mathbb{R}^2 : (x+3)^2 + (y-1)^2 \leq 4\}$

c. $C = \{(x;y) \in \mathbb{R}^2 \,:\, x^2 + y^2 < 4 \wedge x^2 + (y+1)^2 \geq 1\}$

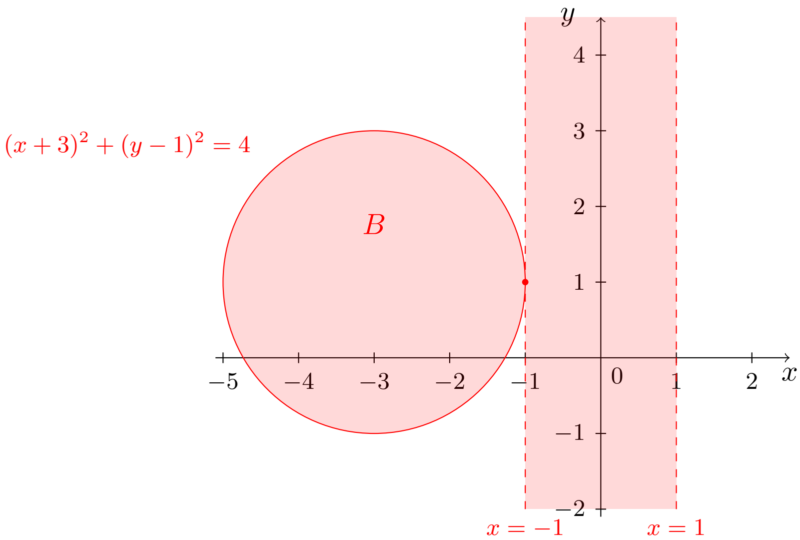
Solution:

a. The set A consists of all points that lie both outside or on the circle at $(1;1)$ with radius 1 and above the line $y = -\frac{1}{2}x + 1$:



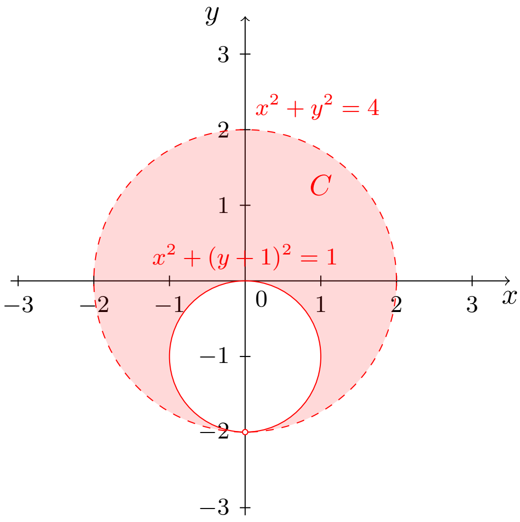
The intersection points of the circle and the line do not belong to A.

b. The set B consists of the points in the plane with coordinates that satisfy the inequality $|x| < 1$, i.e. that satisfy $-1 < x < 1$. Additionally, the (union!) set B contains the points within and on the circle $(x+3)^2 + (y-1)^2 = 4$:



The point $(-1;1)$ belongs to the set B.

c. The set C consists of all points that lie both within the circle $x^2 + y^2 = 4$ and outside the circle $x^2 + (y+1)^2 = 1$:



The point $(0; -2)$ does not belong to C.

9.5 Final Test

9.5.1 Final Test Module 5

Exercise 9.5.1
Specify the normal form of the equation of the line PQ that passes through the two points $P = (1; 3)$ and $Q = (-1; 7)$.

Answer: $y =$ ☐ .

Exercise 9.5.2
Let a line be given by the equation $6x + 2y = 4$.

a. The normal form of this equation is $y =$ ☐ .

b. What is the relative position of this line with respect to the line described by the equation $y = 3x - 2$?
   - ☐ There is no intersection point at all.
   - ☐ There is exactly one intersection point.
   - ☐ The lines coincide.

Exercise 9.5.3
Find the intersection point between the line described by the equation $y = 2x + 2$ and the line described by the equation $2x = 6$.

Answer: The intersection point is ☐ .

Tick the possible reasons why the equation of a line $2x = 6$ of the second line cannot be transformed into normal form. (Several statements can be true.)
- ☐ The equation of the line cannot be solved for x.
- ☐ The equation of the line cannot be solved for y.
- ☐ The equation of the line is not cancelled completely.
- ☐ The line is parallel to the x-axis.
- ☐ The line is parallel to the y-axis.
- ☐ The slope of the line is not finite.
- ☐ The line does not intersect the x-axis.

Exercise 9.5.4
Decide which of the following points lie on the circle with the centre $P = (3; -1)$ and a radius of $r = \sqrt{10}$:

- [ ] The origin
- [ ] $(2;3)$
- [ ] $(4;2)$
- [ ] $(3;2)$
- [ ] $(0;\sqrt{10})$

Exercise 9.5.5
Let a circle be defined by the following equation:

$$(x-2)^2+(y+3)^2 \;=\; 9\;.$$

What are the properties of this circle?

a. Its radius is $r \;=$ [ ] .

b. Its centre is $M \;=$ [ ] .

c. It intersects the line that passes though M and a second unknown point P
- [ ] at one point,
- [ ] at two points,
- [ ] at three points,
- [ ] not at all.

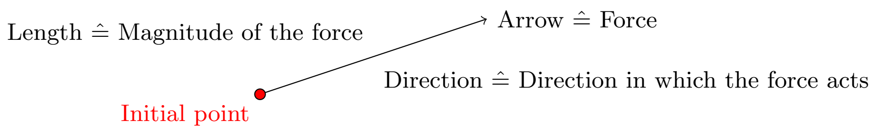
10 Basic Concepts of Descriptive Vector Geometry

Module Overview

10.1 From Arrows to Vectors

10.1.1 Introduction

The basic idea underlying the mathematical concept of a vector is rooted in physics. In science there are quantities that are described by one single number, their magnitude. Such quantities include voltage, work, or power. In mathematical terms, these quantities are simply described by elements of the set of real numbers. There are other quantities which have not only a certain magnitude but also a certain *direction*, such as force or velocity. For example, a force that acts on a body at a certain point can be visualised as an arrow of a corresponding length starting at that point. The direction of the arrow then corresponds to the direction in which the force acts. This is illustrated in the figure below.



In mathematics, such quantities are described by vectors. In the sciences, the magnitudes of vectors have certain units of measure (e.g. forces are measured in Newton). From a purely mathematical point of view, however, these units of measure are not relevant, so they are omitted here. The concept of a vector is the main topic of this section and will be discussed in detail in Subsection 10.1.3. Since vectors are considered not only in two-dimensional space (i.e. in the plane) but also in three-dimensional space (i.e. in the space), the concept of coordinate systems and points introduced in Module 9 will be extended to three dimensions. This is done in Subsection 10.1.2. Finally, we see that certain operations can be applied to vectors. In Subsection 10.1.4 we will study how these vector operations are carried out.

10.1.2 Coordinate Systems in Three-Dimensional Space

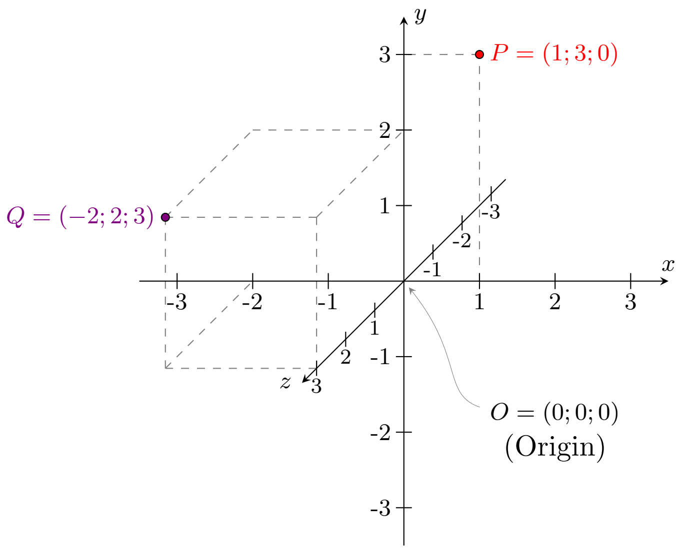
In Section 9.1 of the previous Module 9 we introduced Cartesian coordinate systems and points in the plane described by coordinates with respect to these coordinate systems. A solid understanding of these concepts is now presumed in this Module. To describe a **point in three-dimensional space**, three **coordinates** are required. Thus, a **coordinate system** in three-dimensional space needs three **axes**, the x-axis, y-axis and z-axis (sometimes called the x_1-axis, x_2-axis, and x_3-axis). Usually, points will be denoted by upper-case Latin letters $P, Q, R, \ldots$, and their coordinates will be denoted by lower-case Latin letters $a, b, c, x, y, z, \ldots$ as variables. Extending the notation from Module 9, the coordinates of a point are written, for example, as follows:

$$P = (1; 3; 0)$$

or

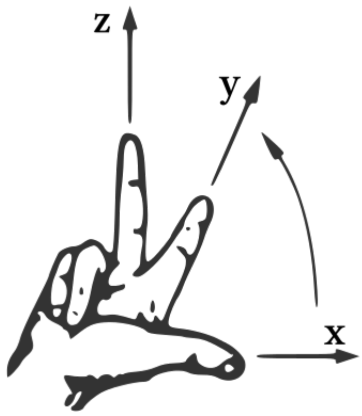
$$Q = (-2;2;3) .$$

Here, the x-coordinate of the point Q is -2, its y-coordinate is 2 and its z-coordinate is 3. The point with the coordinates $(0;0;0)$ is called the **origin**, and it is denoted by the symbol O. All these points are drawn in the figure below.



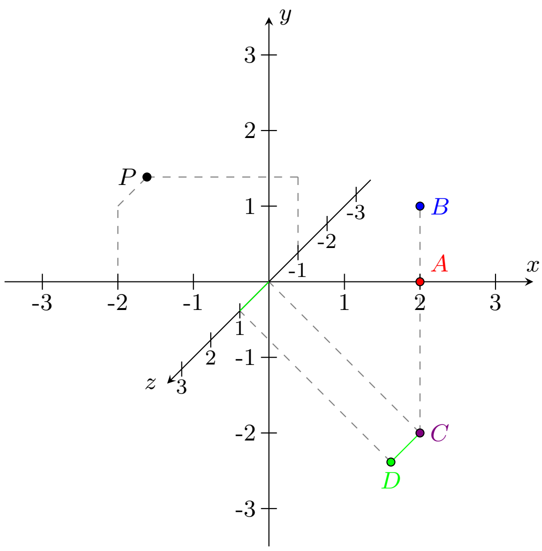
The dashed lines in this figure indicate how the coordinates of points in such a three-dimensional representation can be drawn and read off. Note that these lines are all parallel to the coordinate axes.

We will only consider coordinate systems in three-dimensional space with perpendicular coordinate axes - these are **Cartesian coordinate systems**. Furthermore, we will use the common mathematical convention that coordinate systems in three-dimensional space are **right-handed**. Sometimes these are also called **positively oriented**. This means that the positive directions of the x, y, and z-axis can be determined by means of the *right-hand* rule as illustrated in the figure below.



However, there are various possible representations. In the figure above showing the points P and Q, the x-axis points to the right, the y-axis points up, and the z-axis points perpendicularly outwards from the drawing plane. In the figure which illustrates the right-hand rule, the x-axis points to the right, the y-axis points backwards into the drawing plane, and the z-axis points up. However, both coordinate systems are right-handed.

Exercise 10.1.1
Specify the coordinates of the points indicated in the figure below. Consider how all indicated points can be collected into one mathematical object.



a. A = [ ] .

b. B = [ ] .

c. C = [ ] .

d. D = [ ] .

e. P = [ ] .

Solution:
The coordinate triples of the indicated points are:

$$A = (2;0;0) ,$$
$$B = (2;1;0) ,$$
$$C = (2;-2;0) ,$$
$$D = (2;-2;1) ,$$
$$P = (-2;1;-1) .$$

The set of all points indicated in the figure above is

$$\{A;B;C;D;P\} = \{(2;0;0);(2;1;0);(2;-2;0);(2;-2;1);(-2;1;-1)\} .$$

As in the two-dimensional case discussed in Section 9.1.2, an arbitrary number of points in three-dimensional space can be collected into a **set of points**. The following notation is used:

Info10.1.1

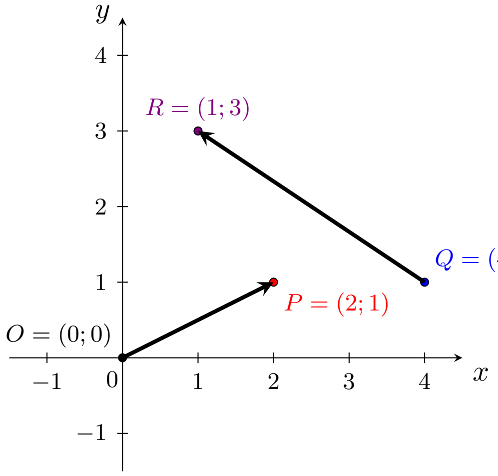
The set of all points (in space) specified as coordinate triples with respect to a given Cartesian coordinate system is denoted by

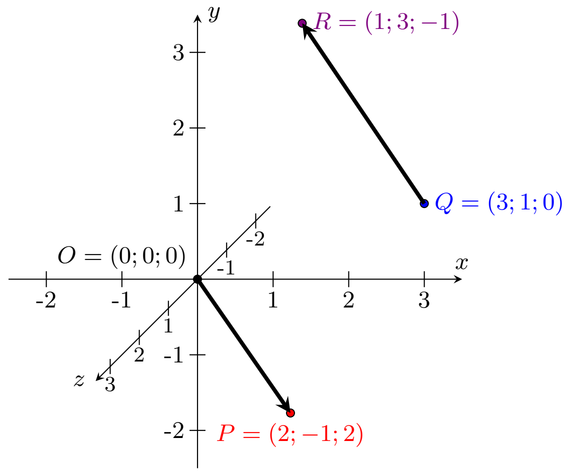
$$\mathbb{R}^3 := \{(x;y;z) \,:\, x \in \mathbb{R} \wedge y \in \mathbb{R} \wedge z \in \mathbb{R}\} .$$

The symbol $\mathbb{R}^3$ reads as "$\mathbb{R}$ three" or "$\mathbb{R}$ to the power of three". This indicates that a point can be uniquely described by a coordinate triple (also known as an ordered triple) consisting of three real numbers.

10.1.3 Vectors in the Plane and in the Space

Points on the plane or in space that are defined as ordered pairs or triples with respect to a given coordinate system can be connected by line segments. Assigning a direction to these line segments (one of the end points of the segment is specified as the initial point and the other one is specified as the end point) results in arrows that point from one point to the other (see left and right figure below for the two-dimensional and the three-dimensional cases).





According to these figures, an arrow provides the following information: it specifies how to get from the initial point (at the foot of the arrow) to the end point (at the tip of the arrow). For example, the arrow that connects the point Q to the point R in the left figure above specifies that starting from the initial point Q one has to move 3 units to the left and 2 units upwards to get to the terminal point R. In more mathematical terms: starting from Q, shift by -3 in the x-direction and by 2 in the y-direction. It is even simpler for the arrow that connects the origin to the point P in the left figure above: to get from O to P move 2 units in the x-direction and 1 unit in the y-direction. Of course, these values are exactly the coordinates of the point P.

Exercise 10.1.2
For the arrows in the right-hand side figure above (three-dimensional case), specify the (signed) movements in the three coordinate directions that are required to get from the initial point to the terminal point of the corresponding arrow. Proceed in the same way as explained above for two dimensions.

For the arrow from O to P, we have:

a. in the x-direction: [ ] ,

b. in the y-direction: [ ] ,

c. in the z-direction: [ ] .

For the arrow from Q to R, we have:

a. in the x-direction: [ ] ,

b. in the y-direction: [ ] ,

c. in the z-direction: [ ] .

Solution:

For the arrow from O to P, we have:

a. in the x-direction: 2,

b. in the y-direction: -1,

c. in the z-direction: 2.

For the arrow from Q to R, we have:

a. in the x-direction: -2,

b. in the y-direction: 2,

c. in the z-direction: -1.

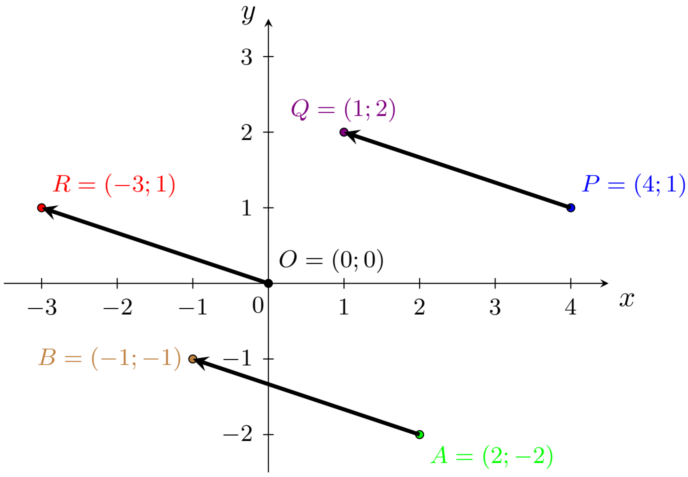
For the arrows connecting Q to R, movements in the corresponding coordinate directions are determined by the coordinates of the points at the foot and at the tip of the arrows. Thus, in the two-dimensional case, we have:

$$\left.\begin{array}{r} R=(1;3) \\ Q=(4;1) \end{array}\right\} \Rightarrow \left\{\begin{array}{l} \text{in } x\text{-direction: } -3=1-4 \\ \text{in } y\text{-direction: } 2=3-1\ , \end{array}\right.$$

and in the three-dimensional case:

$$\left.\begin{array}{r} R=(1;3;-1) \\ Q=(3;1;0) \end{array}\right\} \Rightarrow \left\{\begin{array}{l} \text{in } x\text{-direction: } -2=1-3 \\ \text{in } y\text{-direction: } 2=3-1 \\ \text{in } z\text{-direction: } -1=-1-0\ . \end{array}\right.$$

The movements in the different coordinate directions are the differences of the coordinates of the terminal point and the initial point of the arrow. This means that all arrows connecting pairs of points with equal coordinate differences only differ from each other by a parallel translation, i.e. they retain their direction. The pairs of points P and Q, A and B, O and R indicated in the figure below are each connected by arrows that can be made to coincide by parallel translations.



Here, an infinite number of pairs of points can be found that are connected by such an arrow. This idea works analogously in the three-dimensional case.

Each arrow in the figure above provides the same information, namely a shift by -3 in the x-direction and 1 in the y-direction. So what could be more natural than regarding each of these arrows only as a representation (a so-called **representative**) of a more basic object? This basic mathematical object is called **vector**, and in this case it has the two **components**: -3 (x-component) and 1 (y-component) that are written as a so-called 2-**tuple** one above the other:

$$\text{vector represented in the figure above } = \begin{pmatrix} -3 \\ 1 \end{pmatrix} .$$

The Info Box below outlines these conclusions and a few more notations and dictions concerning vectors.

Info10.1.2

A two- or three-dimensional **vector** is a 2- or 3-**tuple** with 2 or 3 **components** called the x-, y- (and z-)components. In general, vectors are denoted by lowercase italic letters accented by a right arrow or by upright boldface lowercase letters. The components of a vector are often denoted by the same lowercase italic letter as the vector, with the corresponding coordinate direction as its index:

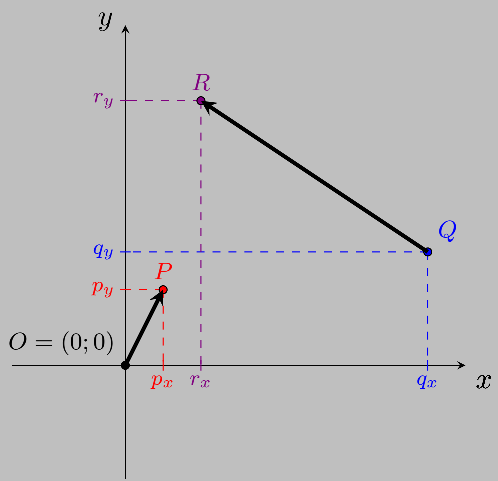
$$\vec{a} = \begin{pmatrix} a_x \\ a_y \end{pmatrix} , \ \vec{b} = \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} .$$

An arrow in the plane or in space is called a **representative** of the vector if the arrow connects two points in the plane or in space such that the differences between the coordinates at the initial and end points of the arrow give the components of the vector.

Often, a point P or two points Q and R in the plane or in space are given and one wants to specify the vector that has the arrow from the origin O to the given point P or the arrow connecting Q to R as its representatives. The Info Box below outlines the notation and diction as well as the required vector operation:

Info10.1.3

- **Two-dimensional case:**
  Let $P = (p_x; p_y)$, $Q = (q_x; q_y)$, and $R = (r_x; r_y)$ be points in the plane. Then the vector
  $$\overrightarrow{QR} := \begin{pmatrix} r_x - q_x \\ r_y - q_y \end{pmatrix}$$
  is called the **connecting vector from the initial point Q to the terminal point R**, and
  $$\overrightarrow{P} := \overrightarrow{OP} = \begin{pmatrix} p_x \\ p_y \end{pmatrix}$$
  is called the **position vector of the point P**. These are exactly those vectors whose representatives include the connecting arrows of the points (see figure below).



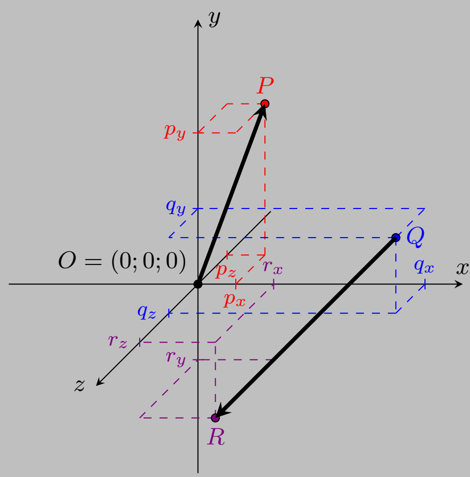
- **Three-dimensional case:**
  Let $P = (p_x; p_y; p_z)$, $Q = (q_x; q_y; q_z)$, and $R = (r_x; r_y; r_z)$ be points in space. Then the vector

$$\overrightarrow{QR} := \begin{pmatrix} r_x - q_x \\ r_y - q_y \\ r_z - q_z \end{pmatrix}$$

  is called the **connecting vector from the initial point Q to the point R**, and

$$\overrightarrow{P} := \overrightarrow{OP} = \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix}$$

  is called the **position vector of the point P**. These are exactly those vectors whose representatives include the connecting arrows of the points (see figure below).



Example 10.1.4

- **Two-dimensional case:**
  The point $P = (-1; -2)$ has the position vector

$$\overrightarrow{P} = \begin{pmatrix} -1 \\ -2 \end{pmatrix} .$$

  The vector

$$\vec{v} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

is the connecting vector from the point $A = (1;1)$ to the point $B = (3;1)$. Thus, we have

$$\vec{v} = \overrightarrow{AB} \,.$$

However, $\vec{v} \neq \overrightarrow{BA}$ since

$$\overrightarrow{BA} = \begin{pmatrix} 1-3 \\ 1-1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \end{pmatrix} \neq \begin{pmatrix} 2 \\ 0 \end{pmatrix} \,.$$

- **Three-dimensional case:**
  Consider the two points $Q = (1;1;1)$ and $R = (-2;0;2)$. The connecting vector from Q to R is:

$$\overrightarrow{QR} = \begin{pmatrix} -2-1 \\ 0-1 \\ 2-1 \end{pmatrix} = \begin{pmatrix} -3 \\ -1 \\ 1 \end{pmatrix} \,.$$

However, the connecting vector from R to Q is

$$\overrightarrow{RQ} = \begin{pmatrix} 1-(-2) \\ 1-0 \\ 1-2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix} \,.$$

Obviously, the vector

$$\begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}$$

is also the position vector of the point $(3;1;-1)$.

The example above reveals an interesting fact: reversing the orientation of a vector (and thus the orientation of all its representative arrows) results in a vector in which all components have the opposite sign. This vector is also called the **opposite vector**. This suggests that vector calculations can be carried out component-wise. This will be discussed in detail in Subsection 10.1.4.

Obviously, the also exists a vector with all its components equal to 0 in the two- and three-dimensional cases:

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \,.$$

This vector is called the (two-dimensional or three-dimensional) **zero vector**. One can imagine the zero vector as having "arrows of zero length" as its representatives, i.e. arrows that connect a point with itself. In other words, the zero vector is the position vector of the origin.

Exercise 10.1.3
Let the points

$$A = \left(-1; \frac{3}{2}\right) \text{ and } B = (\pi; -2)$$

be given in the plane, and the points

$$P = (0.5; 1; -1) \text{ and } Q = \left(\frac{1}{2}; -1; 1\right)$$

in space as well, as the (two- and three-dimensional) vectors

$$\vec{a} = \begin{pmatrix} \pi \\ -1 \end{pmatrix} \text{ and } \vec{v} = \begin{pmatrix} 0 \\ 3 \\ -3 \end{pmatrix} .$$

- Find the following vectors:

  a. $\overrightarrow{AB} =$ ☐

  b. $\overrightarrow{BA} =$ ☐

  c. $\overrightarrow{PQ} =$ ☐

  d. $\overrightarrow{QP} =$ ☐

- Find the points C in the plane and R in space such that the following statements are true:

  a. $\vec{a} = \overrightarrow{CB} \Leftrightarrow C =$ ☐

  b. $\vec{v} = \overrightarrow{QR} \Leftrightarrow R =$ ☐

- Draw at least three representatives of the vector a.

Solution:

- The required vectors are:

  a. $\overrightarrow{AB} = \begin{pmatrix} \pi + 1 \\ -\frac{7}{2} \end{pmatrix}$

  b. $\overrightarrow{BA} = \begin{pmatrix} -1 - \pi \\ \frac{7}{2} \end{pmatrix}$

c. $\overrightarrow{PQ} = \begin{pmatrix} 0 \\ -2 \\ 2 \end{pmatrix}$

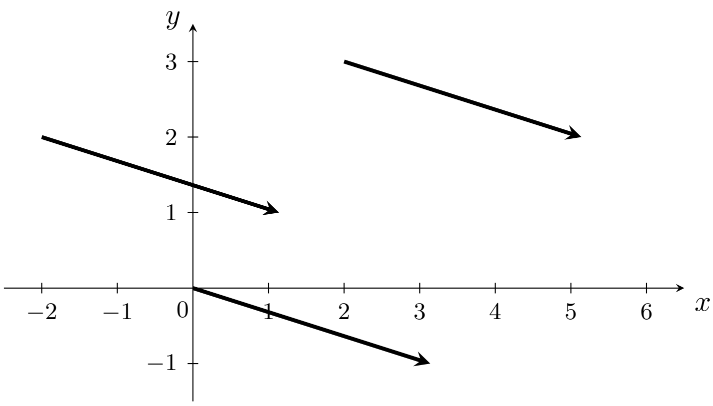
d. $\overrightarrow{QP} = \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix}$

- The required points are:

a. $C = (0; -1)$

b. $R = \left(\frac{1}{2}; 2; -2\right)$

- The figure below shows three possible representatives of $\vec{a} = \begin{pmatrix} \pi \\ -1 \end{pmatrix}$:



The previous introduction of the concept of a vector reveals the close relationship between vectors and points. Indeed, there is a one-to-one correspondence between points and position vectors: for every point there exists exactly one vector that is the position vector of this point. Conversely, for every vector there exists exactly one point that has this vector as its position vector. This is true in both the two-dimensional and the three-dimensional cases. This justifies the convention we will follow below: describing points by their position vectors. A point $P = (2; 1)$, for example, is often described by its corresponding position vector $\overrightarrow{P} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ instead of its coordinates $(2; 1)$. If geometric objects such as lines and planes are investigated (see e.g. Subsection 10.2.2), this convention also involves certain advantages in the description of these objects (in particular in the three-dimensional case).

Moreover, this one-to-one correspondence between points and position vectors also justifies to use the abbreviations $\mathbb{R}^2$ and $\mathbb{R}^3$ not only for the set of all points in the plane or in space but also for the set of all two-dimensional or three-dimensional vectors. These abbreviations will also be used throughout the following sections.

10.1.4 Simple Vector Operations

In this section we discuss what kinds of vector operations we can carry out on the vectors introduced in the previous Section 10.1.3. We can view vector operations in two different ways. First you can carry out the vector operations of addition, subtraction and – with a certain restriction – multiplication on vectors specified as 2- or 3-tuples. On the other hand, these operations can be interpreted graphically with respect to the arrows representing the vectors. That is, the vector operations on vectors can be interpreted as geometric operations on their representatives. This geometric interpretation of vector operations leads to a deeper understanding of position vectors and connecting vectors of points.

Vector operations in two and three dimensions are carried out in essentially the same way. For all component-wise vector operations presented below both cases (two-dimensional and three-dimensional) are considered. The corresponding figures illustrate the geometric interpretation of the operations on the representatives of the vectors and visualise the position and connecting vectors. They will mostly show only arrows and points without a coordinate system. In this way, the figures apply to both the two-dimensional and the three-dimensional case.

Since vectors will be transformed below by equivalent transformations into each other, we need to specify under which conditions two vectors are considered to be equal.

Info10.1.5

Two vectors $\vec{a}, \vec{b} \in \mathbb{R}^2$ or $\mathbb{R}^3$ are **equal** (written as $\vec{a} = \vec{b}$) if and only if they satisfy one (and hence all) of the following equivalent conditions:

- $\vec{a}$ and $\vec{b}$ have the same components:

$$\vec{a} = \vec{b} \Leftrightarrow \begin{pmatrix} a_x \\ a_y \end{pmatrix} = \begin{pmatrix} b_x \\ b_y \end{pmatrix} \Leftrightarrow a_x = b_x \text{ and } a_y = b_y$$

  in the two-dimensional case or

$$\vec{a} = \vec{b} \Leftrightarrow \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} = \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} \Leftrightarrow a_x = b_x \text{ and } a_y = b_y \text{ and } a_z = b_z$$

in the three-dimensional case. This is also known as **comparison of coordinates** or **components**.

- $\vec{a}$ and $\vec{b}$ have the same representatives.
- $\vec{a}$ and $\vec{b}$ are both the position vector of the same point.
- $\vec{a}$ and $\vec{b}$ are both the connecting vector of the same two points.

From the statements in the Info Box above we see that two vectors of different dimensions (i.e. $\vec{a} \in \mathbb{R}^2$ and $\vec{b} \in \mathbb{R}^3$) can never be equal. Since these vectors have a different number of components, they are not even comparable. Thus, vector operations are only carried out on vectors with an equal number (two or three) of components , and these operations will always result in a vector with this fixed number of components.

Info10.1.6

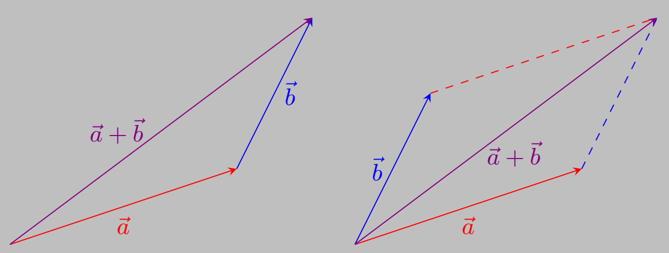
The **addition of two vectors** involves the addition of all their components, i.e.

$$\vec{a} + \vec{b} = \begin{pmatrix} a_x \\ a_y \end{pmatrix} + \begin{pmatrix} b_x \\ b_y \end{pmatrix} = \begin{pmatrix} a_x + b_x \\ a_y + b_y \end{pmatrix}$$

in the two-dimensional case and

$$\vec{a} + \vec{b} = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} + \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} = \begin{pmatrix} a_x + b_x \\ a_y + b_y \\ a_z + b_z \end{pmatrix}$$

in the three-dimensional case. Geometrically, vector addition can be interpreted as either a“linking” of two arrows or a completion of two arrows to a parallelogram, depending on which representatives of the vectors are used:



The laws of associativity and commutativity apply to the addition of two vectors as to the addition of real numbers (see Section 1.1.3):

$$\vec{a}+\vec{b}=\vec{b}+\vec{a}$$

and

$$\vec{a}+\vec{b}+\vec{c}=(\vec{a}+\vec{b})+\vec{c}=\vec{a}+(\vec{b}+\vec{c})\ .$$

The **zero vector** $\overrightarrow{O}=\begin{pmatrix}0\\0\end{pmatrix}$ or $\overrightarrow{O}=\begin{pmatrix}0\\0\\0\end{pmatrix}$ plays the same role for vectors as the number 0 for the real numbers:

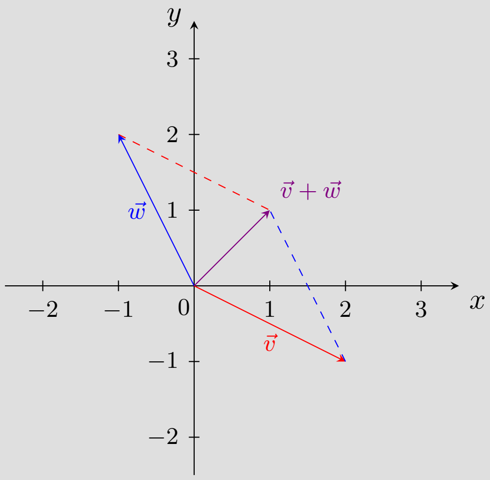
$$\vec{a}+\overrightarrow{O}=\overrightarrow{O}+\vec{a}=\vec{a}\ .$$

Example 10.1.7

Consider two vectors $\vec{v}=\begin{pmatrix}2\\-1\end{pmatrix}$ and $\vec{w}=\begin{pmatrix}-1\\2\end{pmatrix}$. Then, we have

$$\vec{v}+\vec{w}=\begin{pmatrix}2\\-1\end{pmatrix}+\begin{pmatrix}-1\\2\end{pmatrix}=\begin{pmatrix}2-1\\-1+2\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}\ .$$

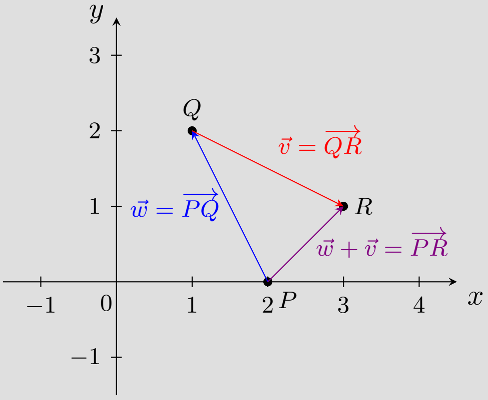
This is illustrated by the figure below.



Furthermore, $\vec{w}$ is the connecting vector from the point $P=(2;0)$ to the point $Q=(1;2)$, i.e. $\vec{w}=\overrightarrow{PQ}$, and $\vec{v}$ is the connecting vector from the point $Q=(1;2)$ to the point $R=(3;1)$, i.e. $\vec{v}=\overrightarrow{QR}$.

$$\vec{w}+\vec{v}=\overrightarrow{PQ}+\overrightarrow{QR}=\overrightarrow{PR}\ .$$

This is illustrated by the figure below.



The example above shows that expressions for connecting vectors of points can also be simplified by means of vector operations. This will be discussed for the subtraction of vectors in more detail below.

Exercise 10.1.4 a. Let the vectors $\vec{u} = \begin{pmatrix} 1 \\ 0 \\ -8 \end{pmatrix}$ and $\vec{v} = \begin{pmatrix} -3 \\ -4 \\ 3 \end{pmatrix}$ be given. Calculate $\vec{v} + \vec{u}$.

$\vec{v} + \vec{u} =$ ☐

b. Let the points P, Q, and R be given. Which of the following expressions are equivalent to the expression $(\overrightarrow{P} + \overrightarrow{PQ}) + \overrightarrow{QR}$?

- ☐ $\overrightarrow{P} + (\overrightarrow{PQ} + \overrightarrow{QR})$
- ☐ $\overrightarrow{PR}$
- ☐ $\overrightarrow{QR}$
- ☐ $\overrightarrow{OR}$
- ☐ $\overrightarrow{R}$
- ☐ $\overrightarrow{Q} + \overrightarrow{QR}$

Solution:

a.

$$\vec{v} + \vec{u} = \begin{pmatrix} -3 \\ -4 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ -8 \end{pmatrix} = \begin{pmatrix} -3+1 \\ -4+0 \\ 3-8 \end{pmatrix} = \begin{pmatrix} -2 \\ -4 \\ -5 \end{pmatrix} .$$

b.

$$(\overrightarrow{P}+\overrightarrow{PQ})+\overrightarrow{QR}=\overrightarrow{P}+(\overrightarrow{PQ}+\overrightarrow{QR}) \text{ (associativity)}$$
$$=(\overrightarrow{OP}+\overrightarrow{PQ})+\overrightarrow{QR}=\overrightarrow{OQ}+\overrightarrow{QR}=\overrightarrow{Q}+\overrightarrow{QR}=\overrightarrow{OR}=\overrightarrow{R}\ .$$

When we study possible operations on vectors further, we will see that the component-wise multiplication or division of vectors is not a meaningful operation. To understand this, however, would go far beyond the scope of this course. At this point you simply have to accept that vectors cannot be multiplied that simply, let alone divided. What we certainly can do is multiply vectors by real numbers and – based on that – subtract vectors. Please note: if we speak of the length of a vector in the following section, we mean the geometric length of the arrows representing this vector. The concept of the length (a.k.a. absolute value or norm) of a vector will be discussed in more detail later on.

Info10.1.8

The **multiplication of a vector by a real number** involves the multiplication of each component by this real number. If $\vec{a}$ is a vector and $s \in \mathbb{R}$, then we have

$$s\cdot\vec{a}=s\vec{a}=s\begin{pmatrix}a_x\\a_y\end{pmatrix}=\begin{pmatrix}sa_x\\sa_y\end{pmatrix}$$

in the two-dimensional case and

$$s\cdot\vec{a}=s\vec{a}=s\begin{pmatrix}a_x\\a_y\\a_z\end{pmatrix}=\begin{pmatrix}sa_x\\sa_y\\sa_z\end{pmatrix}$$

in the three-dimensional case. The division of a vector by a number $0\neq s\in\mathbb{R}$ is then simply defined as its multiplication by the reciprocal $\frac{1}{s}$:

$$\frac{\vec{a}}{s}=\frac{1}{s}\vec{a}\ .$$

Thus, multiplying a vector by a real number $s\in\mathbb{R}$, $s>0$ results in an equally oriented vector that is stretched by a factor of s . For $s\in\mathbb{R}$, $s<0$, the resulting vector is also stretched by a factor of s but it is flipped around by an angle of π. In the special case of $s=0$, we obviously have

$$0\vec{a}=\overrightarrow{O}$$

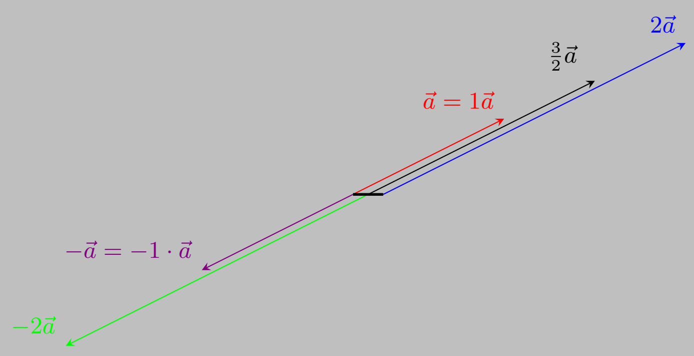
for every vector $\vec{a}$. Two other relevant cases are the multiplication by a factor of $s=1$

$$1\cdot\vec{a}=\vec{a}$$

– which obviously leaves the vector unchanged – and the multiplication by $s = -1$

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

– which results in the so-called **opposite vector**. This is a vector of equal length and opposite orientation. This is illustrated in the figure below.



Since the multiplication by real numbers changes the length (*scales* the vectors), real numbers with respect to vectors are often called **scalars**, and the multiplication of a vector by a real number is called **scalar multiplication**.

Example 10.1.9

Let the vector $\vec{v} = \begin{pmatrix} 3 \\ \frac{3}{2} \end{pmatrix}$ be given. Then, for example, we have

$$2\vec{v} = 2\begin{pmatrix} 3 \\ \frac{3}{2} \end{pmatrix} = \begin{pmatrix} 2 \cdot 3 \\ 2 \cdot \frac{3}{2} \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix}$$

and

$$-\frac{\vec{v}}{3} = -\frac{1}{3}\vec{v} = -\frac{1}{3}\begin{pmatrix} 3 \\ \frac{3}{2} \end{pmatrix} = \begin{pmatrix} -1 \\ -\frac{1}{2} \end{pmatrix} .$$

Moreover, $\vec{v} = \overrightarrow{PQ}$ for $P = \left(3; \frac{1}{2}\right)$ and $Q = (6; 2)$ since

$$\vec{v} = \begin{pmatrix} 3 \\ \frac{3}{2} \end{pmatrix} = \begin{pmatrix} 6-3 \\ 2-\frac{1}{2} \end{pmatrix} .$$

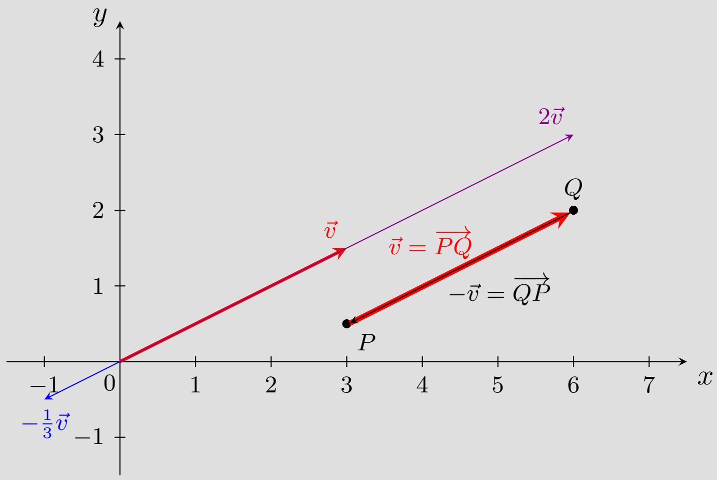
Then we have

$$-\vec{v} = -\overrightarrow{PQ} = \overrightarrow{QP}$$

since

$$-\vec{v} = \begin{pmatrix} -3 \\ -\frac{3}{2} \end{pmatrix} = \begin{pmatrix} 3-6 \\ \frac{1}{2}-2 \end{pmatrix} .$$

This is illustrated in the figure below.



The following calculation rules apply to scalar multiplication:

Info10.1.10

Let two real numbers r and s and two vectors $\vec{a}$ and $\vec{b}$ be given. Then the following calculation rules apply:

1. $r\vec{a} = \vec{a}r$
2. $rs\vec{a} = (rs)\vec{a} = r(s\vec{a})$
3. $(r+s)\vec{a} = r\vec{a} + s\vec{a}$
4. $r(\vec{a}+\vec{b}) = r\vec{a} + r\vec{b}$
5. $r(-\vec{a}) = (-r)\vec{a} = -(r\vec{a})$
6. $r\vec{a} = 0 \Leftrightarrow r = 0$ or $\vec{a} = \vec{O}$

The first law is called the commutativity law of scalar multiplication, the second is the associativity law of scalar multiplication, and the third and fourth are the distributive laws of scalar multiplication.

Using the concept of an opposite vector we are now able to specify what the subtraction of vectors means.

Info10.1.11

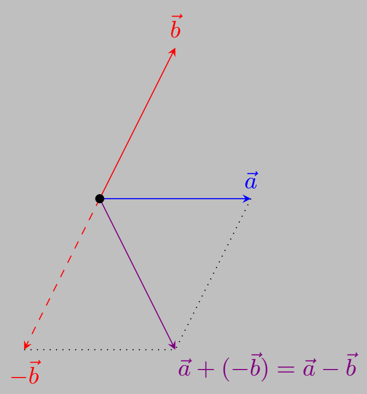
Let two vectors $\vec{a}$ and $\vec{b}$ be given. Then, their **difference** $\vec{a} - \vec{b}$ is the sum of $\vec{a}$ and the opposite vector of $\vec{b}$. Thus, we have

$$\vec{a} - \vec{b} = \vec{a} + (-\vec{b}) = \begin{pmatrix} a_x \\ a_y \end{pmatrix} + \begin{pmatrix} -b_x \\ -b_y \end{pmatrix} = \begin{pmatrix} a_x - b_x \\ a_y - b_y \end{pmatrix}$$

in the two-dimensional case and

$$\vec{a} - \vec{b} = \vec{a} + (-\vec{b}) = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} + \begin{pmatrix} -b_x \\ -b_y \\ -b_z \end{pmatrix} = \begin{pmatrix} a_x - b_x \\ a_y - b_y \\ a_z - b_z \end{pmatrix}$$

in the three-dimensional case. The difference of vectors can also be interpreted geometrically by means of the representatives as illustrated in the figure below.



If we only consider the difference of vectors by means of their components, the question arises what the concept of an opposite vector here is required for. Indeed, the difference of vectors in componentwise notation could also be written analogously to the sum without using the concept of an opposite vector. However, if we think of the geometrical interpretation of the difference by means of representatives (see figure in the Info Box above), we see that a geometrical interpretation is only possible using the concept of an opposite vector.

Example 10.1.12

This example shows typical problems involving the calculation rules for vectors presented so far.

1. Simplify the following vector expressions:

   (i) $\begin{pmatrix} 1 \\ -2 \\ 0.5 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}$,

   (ii) $2(\vec{v} - \vec{w}) + 3r\vec{w} - r \cdot (-2\vec{v})$ für $r \in \mathbb{R}$.

   Applying the calculation rules results in

   (i)
   $$\begin{pmatrix} 1 \\ -2 \\ 0.5 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 0.5 \end{pmatrix} - \begin{pmatrix} 1 \\ -\frac{3}{2} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1-1 \\ -2-(-\frac{3}{2}) \\ 0.5-\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{1}{2} \\ 0 \end{pmatrix} .$$

   (ii)
   $$2(\vec{v} - \vec{w}) + 3r\vec{w} - r \cdot (-2\vec{v}) = 2\vec{v} - 2\vec{w} + 3r\vec{w} + 2r\vec{v} = 2(r+1)\vec{v} + (3r-2)\vec{w} .$$

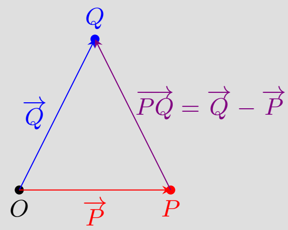
2. Let two vectors $\vec{a} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\vec{b} = \begin{pmatrix} -8 \\ 3 \end{pmatrix}$ be given. Find the unknown vector $\vec{x}$ in the equation
   $$\vec{a} - 2\vec{b} - (3\vec{a} + \vec{x}) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} .$$

   Solving for $\vec{x}$ and substituting $\vec{a}$ and $\vec{b}$ results in
   $$-(3\vec{a}+\vec{x}) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} - \vec{a} + 2\vec{b} \;\Leftrightarrow\; -\vec{x} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} - \vec{a} + 2\vec{b} + 3\vec{a} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} + 2(\vec{a}+\vec{b})$$
   $$\Leftrightarrow\; \vec{x} = -\begin{pmatrix} 0 \\ -1 \end{pmatrix} - 2\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} -8 \\ 3 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - 2\begin{pmatrix} -7 \\ 5 \end{pmatrix} \;\Leftrightarrow\; \vec{x} = \begin{pmatrix} 14 \\ -9 \end{pmatrix} .$$

3. Using the difference of vectors, specify the connecting vector $\overrightarrow{PQ}$ of the two points P and Q by means of the position vectors $\overrightarrow{P}$ and $\overrightarrow{Q}$. We have:

$$\overrightarrow{Q} - \overrightarrow{P} = \overrightarrow{OQ} - \overrightarrow{OP} = -\overrightarrow{QO} - \overrightarrow{OP} = -(\overrightarrow{QO} + \overrightarrow{OP}) = -\overrightarrow{QP} = \overrightarrow{PQ}\,.$$

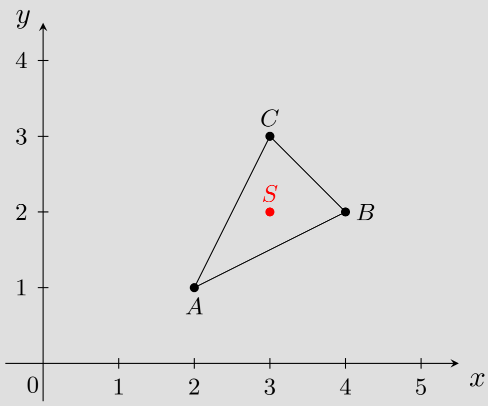
The connecting vector $\overrightarrow{PQ}$ from a point P to a point Q is always the difference of the position vector $\overrightarrow{Q}$ (to the terminal point of the connecting vector) and the position vector $\overrightarrow{P}$ (to the initial point of the connecting vector). This is illustrated in the figure below and can also be seen from the calculation rule for connecting vectors outlined in Info Box 10.1.3.



4. The points $A = (2;1)$, $B = (4;2)$, and $C = (3;3)$ are the vertices of a triangle. The (geometric) centroid S of this triangle can be calculated by means of the corresponding position vectors:

$$\overrightarrow{S} = \frac{1}{3}(\overrightarrow{A} + \overrightarrow{B} + \overrightarrow{C}) = \frac{1}{3}\left(\begin{pmatrix}2\\1\end{pmatrix} + \begin{pmatrix}4\\2\end{pmatrix} + \begin{pmatrix}3\\3\end{pmatrix}\right) = \frac{1}{3}\begin{pmatrix}9\\6\end{pmatrix} = \begin{pmatrix}3\\2\end{pmatrix}\,.$$

Thus, we have $S = (3;2)$. This is illustrated in the figure below.



Exercise 10.1.5 a. Let P, Q, R, and S be points in space. Simplify the expression

$$\overrightarrow{PQ} - (\overrightarrow{PQ} - \overrightarrow{QR}) + \overrightarrow{RS}$$

as far as possible.

b. Show that the points $A = (1;2)$, $B = (4;3)$, and $C = (3;1)$ together with the origin form the vertices of a parallelogram.

Solution:

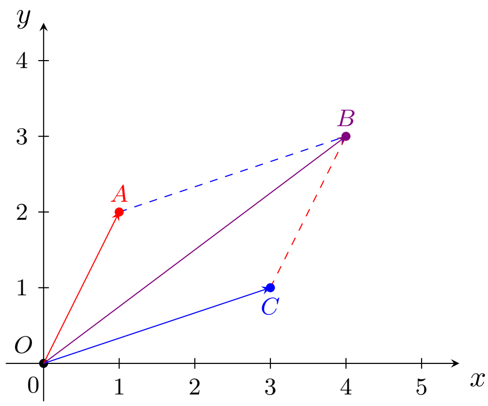
a.

$$\overrightarrow{PQ} - (\overrightarrow{PQ} - \overrightarrow{QR}) + \overrightarrow{RS} = \overrightarrow{PQ} - \overrightarrow{PQ} + \overrightarrow{QR} + \overrightarrow{RS} = \vec{O} + \overrightarrow{QR} + \overrightarrow{RS} = \overrightarrow{QS}\,.$$

b. According to the geometrical interpretation of vector addition, the four points form a parallelogram if one of the position vectors $\overrightarrow{A}$, $\overrightarrow{B}$, or $\overrightarrow{C}$ is the sum of the two other position vectors. Since we have

$$\overrightarrow{A} + \overrightarrow{C} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix} = \overrightarrow{B}\ ,$$

the three points together with the origin form the vertices of a parallelogram. This is illustrated in the figure below.



Exercise 10.1.6
Simplify the following expressions as far as possible:

a. $2\begin{pmatrix} -1 \\ 4 \\ 2 \end{pmatrix} - 3\begin{pmatrix} 1 \\ 6 \\ -2 \end{pmatrix} =$ [ ] .

b. $-2\begin{pmatrix} -t \\ 3 \end{pmatrix} - \left(\begin{pmatrix} -1 \\ 0 \end{pmatrix} + \frac{t}{2}\begin{pmatrix} 4 \\ -42 \end{pmatrix}\right) =$ [ ] .

Solution:

a.

$$2\begin{pmatrix}-1\\4\\2\end{pmatrix}-3\begin{pmatrix}1\\6\\-2\end{pmatrix}=\begin{pmatrix}-2\\8\\4\end{pmatrix}-\begin{pmatrix}3\\18\\-6\end{pmatrix}=\begin{pmatrix}-5\\-10\\10\end{pmatrix}.$$

b.

$$-2\begin{pmatrix}-t\\3\end{pmatrix}-\left(\begin{pmatrix}-1\\0\end{pmatrix}+\frac{t}{2}\begin{pmatrix}4\\-42\end{pmatrix}\right)=\begin{pmatrix}2t\\-6\end{pmatrix}+\begin{pmatrix}1\\0\end{pmatrix}-\begin{pmatrix}2t\\-21t\end{pmatrix}=\begin{pmatrix}1\\21t-6\end{pmatrix}.$$

Exercise 10.1.7
Find the vector $\vec{y}$ in the equation

$$3\left(\begin{pmatrix}1\\-1\\1\end{pmatrix}-\vec{y}\right)=-8\begin{pmatrix}0.25\\0.25\\-0.25\end{pmatrix}+\vec{y}\,.$$

$\vec{y}=$ [ ]

Solution:

$$3\left(\begin{pmatrix}1\\-1\\1\end{pmatrix}-\vec{y}\right)=-8\begin{pmatrix}0.25\\0.25\\-0.25\end{pmatrix}+\vec{y}\;\Leftrightarrow\;\begin{pmatrix}3\\-3\\3\end{pmatrix}+2\begin{pmatrix}1\\1\\-1\end{pmatrix}=4\vec{y}$$

$$\Leftrightarrow\;4\vec{y}=\begin{pmatrix}5\\-1\\1\end{pmatrix}\;\Leftrightarrow\;\vec{y}=\begin{pmatrix}\frac{5}{4}\\-\frac{1}{4}\\\frac{1}{4}\end{pmatrix}.$$

Since any vector has an arbitrary number of arrows as its representatives (which arise from each other by parallel translation), all these representatives have the same geometrical length (which is always the distance between the two points they connect). Therefore, it is reasonable to speak of the length of a vector. In mathematics, the length of a vector is called the **absolute value** or **norm**.

Info10.1.13

The **absolute value** or **norm** of a vector $\vec{a}$ is denoted by $|\vec{a}|$ and equals the distance between the origin O and the point P that has the vector $\vec{a}$ as its corresponding position vector (i.e. $\overrightarrow{P} = \vec{a}$). Thus, we have generally

$$|\vec{a}| = [\overline{OP}]$$

and hence in the two-dimensional case

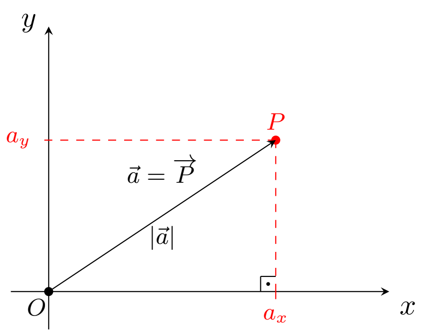
$$|\vec{a}| = \left|\begin{pmatrix} a_x \\ a_y \end{pmatrix}\right| = \sqrt{a_x^2 + a_y^2}$$

or in the three-dimensional case

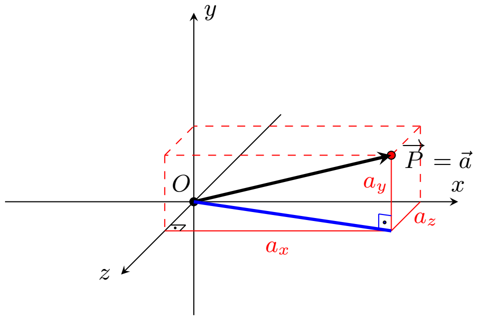
$$|\vec{a}| = \left|\begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}\right| = \sqrt{a_x^2 + a_y^2 + a_z^2}\ .$$

A vector with an absolute value of 1 is also called the **unit vector**.

Applying the formula given in the Info Box 9.3.2 in Module 9 for the distance of two points in a two-dimensional coordinate system immediately results in the formula for the two-dimensional case:



This is another simple application of Pythagoras' theorem. The three-dimensional case is not much more complex: Pythagoras' theorem can still be applied:



In the figure above two right triangles can be identified. From the right triangle lying in the xz-plane we obtain $\sqrt{a_x^2 + a_z^2}$ for the length of the blue line. From the second right triangle we then obtain for the length of the segment from O to P, i.e. for $|\vec{a}|$:

$$\sqrt{a_y^2 + \left(\sqrt{a_x^2 + a_z^2}\right)^2} = \sqrt{a_x^2 + a_y^2 + a_z^2}\ .$$

For norms of vectors the following calculation rules apply:

Info10.1.14

Let two vectors $\vec{a}$ and $\vec{b}$ (both in $\mathbb{R}^2$ or in $\mathbb{R}^3$) and a real number $r \in \mathbb{R}$ be given. Then, we have:

1. $|\vec{a}| \geq 0$ and $|\vec{a}| = 0 \ \Leftrightarrow\ \vec{a} = \overrightarrow{O}$,
2. $|r\vec{a}| = |r| \cdot |\vec{a}|$, and
3. $|\vec{a} + \vec{b}| \leq |\vec{a}| + |\vec{b}|$.

The first rule states that norms are always non-negative and that only the norm of the zero vector is 0. The second rule is especially useful in calculating the norms of constant multiples of vectors. The third rule is called the **triangle inequality**.

Example 10.1.15

- The absolute value of the vector $\frac{1}{4}\begin{pmatrix}3\\-1\\\sqrt{6}\end{pmatrix}$ is

$$\left|\frac{1}{4}\begin{pmatrix}3\\-1\\\sqrt{6}\end{pmatrix}\right| = \left|\frac{1}{4}\right|\left|\begin{pmatrix}3\\-1\\\sqrt{6}\end{pmatrix}\right| = \frac{1}{4}\sqrt{3^2+(-1)^2+6} = \frac{\sqrt{16}}{4} = 1\ .$$

  Hence, this is a unit vector.

- Find a number $q \in \mathbb{R}$ such that $\left|\begin{pmatrix}q^2-2\\4\end{pmatrix} - 2\begin{pmatrix}q-1\\q\end{pmatrix}\right| = 0$.

$$\left|\begin{pmatrix}q^2-2\\4\end{pmatrix} - 2\begin{pmatrix}q-1\\q\end{pmatrix}\right| = \left|\begin{pmatrix}q^2-2q\\4-2q\end{pmatrix}\right| = 0 \quad\Leftrightarrow\quad \begin{pmatrix}q^2-2q\\4-2q\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

$$\Leftrightarrow\quad q^2-2q=0 \text{ and } 4-2q=0 \quad\Leftrightarrow\quad q(q-2)=0 \text{ and } 2(2-q)=0 \quad\Leftrightarrow\quad q=2\ .$$

Exercise 10.1.8
Calculate
$\left|-\frac{1}{3}\begin{pmatrix}2\\-14\\5\end{pmatrix}\right| =$ [ ] .

Solution:

$$\left|-\frac{1}{3}\begin{pmatrix}2\\-14\\5\end{pmatrix}\right| = \left|-\frac{1}{3}\right|\cdot\left|\begin{pmatrix}2\\-14\\5\end{pmatrix}\right| = \frac{1}{3}\sqrt{2^2+(-14)^2+5^2} = \frac{1}{3}\sqrt{225} = \frac{15}{3} = 5\ .$$

Exercise 10.1.9
Find the number $\chi > 3$ such that $\left|\begin{pmatrix}3\\\chi\end{pmatrix} - \begin{pmatrix}\chi\\3\end{pmatrix}\right| = 2\sqrt{2}$.

$\chi =$ [ ] .

Solution:

$$\left|\begin{pmatrix}3\\\chi\end{pmatrix} - \begin{pmatrix}\chi\\3\end{pmatrix}\right| = \left|\begin{pmatrix}3-\chi\\\chi-3\end{pmatrix}\right| = \sqrt{(3-\chi)^2+(\chi-3)^2} = 2\sqrt{2} \quad\Leftrightarrow\quad \sqrt{2(3-\chi)^2} = 2\sqrt{2}$$

$$\Leftrightarrow \quad |3-\chi| = 2 \quad \Leftrightarrow \quad \chi = 1 \text{ or } \chi = 5\ .$$

From $\chi > 3$, it follows that $\chi = 5$.

Exercise 10.1.10
Show that the points $A = (4;2;7)$, $B = (3;1;9)$, and $C = (2;3;8)$ are the vertices of an equilateral triangle.

Solution:
The triangle is equilateral if

$$|\overrightarrow{AB}| = |\overrightarrow{AC}| = |\overrightarrow{BC}|\ .$$

$$|\overrightarrow{AB}| = |\vec{B} - \vec{A}| = \left| \begin{pmatrix} 3 \\ 1 \\ 9 \end{pmatrix} - \begin{pmatrix} 4 \\ 2 \\ 7 \end{pmatrix} \right| = \left| \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \right| = \sqrt{(-1)^2 + (-1)^2 + 2^2} = \sqrt{6}\ ,$$

$$|\overrightarrow{AC}| = |\vec{C} - \vec{A}| = \left| \begin{pmatrix} 2 \\ 3 \\ 8 \end{pmatrix} - \begin{pmatrix} 4 \\ 2 \\ 7 \end{pmatrix} \right| = \left| \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \right| = \sqrt{(-2)^2 + 1^2 + 1^2} = \sqrt{6}\ ,$$

$$|\overrightarrow{BC}| = |\vec{C} - \vec{B}| = \left| \begin{pmatrix} 2 \\ 3 \\ 8 \end{pmatrix} - \begin{pmatrix} 3 \\ 1 \\ 9 \end{pmatrix} \right| = \left| \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} \right| = \sqrt{(-1)^2 + 2^2 + (-1)^2} = \sqrt{6}\ .$$

Thus, the triangle is equilateral.

10.2 Lines and Planes

10.2.1 Introduction

In this section vectors are used (first of all) to describe lines in the plane. Then we will see that this description of lines can be extended to the three-dimensional case. Besides lines, there are other mathematical objects in space which can easily be described by means of vectors, namely planes. Finally, we will discuss the possible relative positions of points, lines, and planes with respect to each other.

For this purpose, the concepts outlined in the Info Box below are important.

Info10.2.1

- Two vectors $\vec{a}$ and $\vec{b}$ in $\mathbb{R}^2$ or $\mathbb{R}^3$ ($\vec{a}, \vec{b} \neq \overrightarrow{O}$) are called **collinear** if there exists a number $s \in \mathbb{R}$ such that
$$\vec{a} = s\vec{b}\,.$$
- Three vectors $\vec{a}$, $\vec{b}$, and $\vec{c}$ in $\mathbb{R}^3$ ($\vec{a}, \vec{b}, \vec{c} \neq \overrightarrow{O}$) are called **coplanar** if there exist two numbers $s, t \in \mathbb{R}$ such that
$$\vec{a} = s\vec{b} + t\vec{c}\,.$$

Further contents:
In this course, the zero vector is excluded from the definition of collinearity and coplanarity since it will be not needed for our simple description of lines and planes. If the conditions for collinearity and coplanarity are made slightly more complex, the definition also extends to the zero vector. The required considerations will naturally involve the (important) terms of **linear independence** and **linear span** that, however, go far beyond the scope of this course.

The following considerations and figures illustrate why the concepts of collinearity and coplanarity are relevant for the investigation of lines and planes.

Collinear vectors are multiples of each other. The representatives of collinear vectors with the same initial point lie on the same line. For example, the vectors

$$\vec{x} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

and

$$\vec{y} = \begin{pmatrix} -1 \\ \frac{1}{2} \end{pmatrix}$$

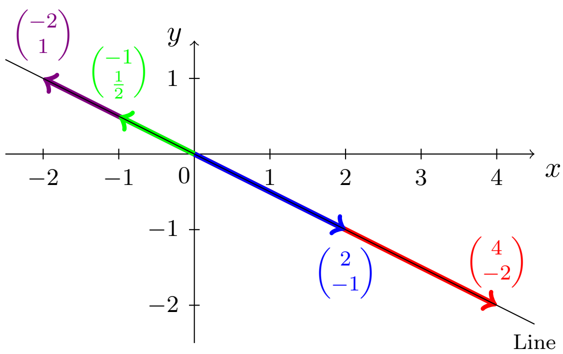
are collinear since

$$\vec{y} = \begin{pmatrix} -1 \\ \frac{1}{2} \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = -\frac{1}{2}\vec{x}$$

(or also $\vec{x} = -2\vec{y}$). Further vectors that are collinear to both $\vec{x}$ and $\vec{y}$ are, for example, $\begin{pmatrix} 4 \\ -2 \end{pmatrix}$ or $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$. In contrast, the vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is not collinear to $\vec{x}$ (and hence not collinear to $\vec{y}$) since there *cannot* be any number $s \in \mathbb{R}$ that satisfies the equation

$$\vec{x} = \begin{pmatrix} 2 \\ -1 \end{pmatrix} = s \begin{pmatrix} 1 \\ 1 \end{pmatrix} .$$

Representatives of collinear vectors with the same initial point, such as the arrows of the corresponding position vectors (see figure below), all lie on the same line.



For coplanar vectors in space, their representatives lie in the same plane if they have the same initial point. For example, the vectors

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} , \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}$$

are coplanar since

$$\begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} = 2\vec{e}_1 + 3\vec{e}_2 = 2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} .$$

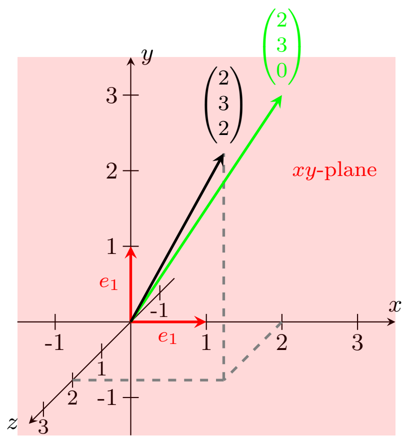
The arrows of their corresponding position vectors all lie in the xy-plane of a coordinate system in space. In contrast, the vectors

$$\vec{e}_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}, \ \vec{e}_2 = \begin{pmatrix}0\\1\\0\end{pmatrix} \text{, and } \begin{pmatrix}2\\3\\2\end{pmatrix}$$

are *not* coplanar since the vector $\begin{pmatrix}2\\3\\2\end{pmatrix}$ has a non-zero z-component, i.e. all its representatives are perpendicular to the xy-plane. It can easily be seen that there *cannot* be any numbers $s, t \in \mathbb{R}$ such that the equation

$$\begin{pmatrix}2\\3\\2\end{pmatrix} = s\begin{pmatrix}1\\0\\0\end{pmatrix} + t\begin{pmatrix}0\\1\\0\end{pmatrix}$$

is satisfied. This is illustrated in the figure below.



10.2.2 Lines in the Plane and in Space

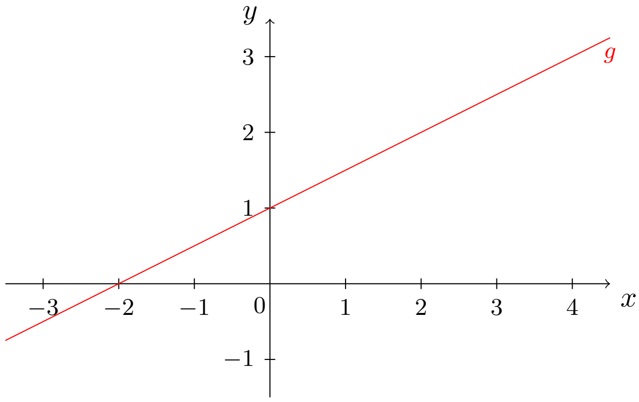
In Module 9 we described lines in the plane using coordinate equations for the points lying on the lines with respect to a fixed coordinate system. In this way, for example, a line g with slope $\frac{1}{2}$ and y-intercept 1 is given as the set of points

$$g = \{(x; y) \in \mathbb{R}^2 \,:\, y = \frac{1}{2}x + 1\} \,,$$

which is often abbreviated by specifying only the coordinate equation of a line (here in normal form):

$$g\colon y = \frac{1}{2}x + 1 \,.$$

The figure below shows this line.



In the following sections we want to describe the points on the line by only their corresponding position vectors. The considerations outlined below result in the following description: the coordinates of the points $(x; y)$ lying on the line g satisfy the equation

$$y = \frac{1}{2}x + 1 \, .$$

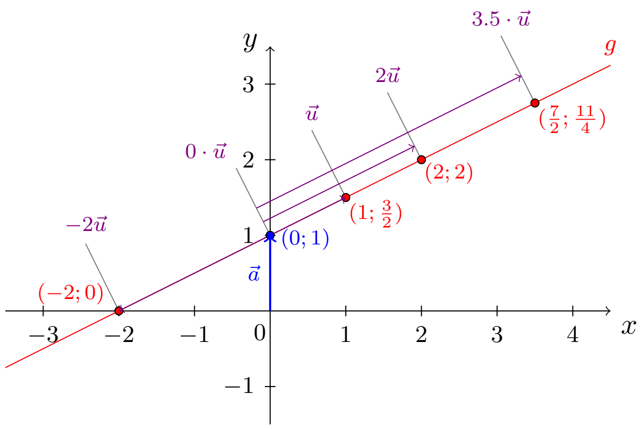
This equation can be substituted into the description of the point. Then, we see that the line g consists of points of the form $\left(x; \frac{1}{2}x + 1\right)$ with $x \in \mathbb{R}$. Let the corresponding position vectors of these points be denoted by $\vec{r}$. Then, we have

$$\vec{r} = \begin{pmatrix} x \\ \frac{1}{2}x + 1 \end{pmatrix} = x \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

with $x \in \mathbb{R}$. Using position vectors, the line g can also be described by

$$g\colon \vec{r} = x \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ , \quad x \in \mathbb{R} \, .$$

In other words: The points of g are defined by the sum of the vector $\vec{a} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and all possible multiples of the vector $\vec{u} = \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}$, i.e. all vectors collinear to the vector $\begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}$. The figure below illustrates this approach.



The Info Box below outlines the most relevant terms, methods and concepts for this so-called **vector form** or **parametric form** of an equation of a line.

Info10.2.2

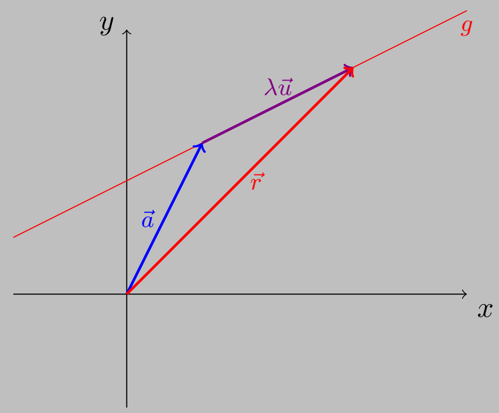
- A line g in the plane is given in **vector form** or in **parametric form** as the set of position vectors

$$g = \left\{ \vec{r} = \lambda\vec{u} + \vec{a} \in \mathbb{R}^2 \,:\, \lambda \in \mathbb{R} \right\} \;,$$

  often written in short as

$$g\colon \vec{r} = \lambda\vec{u} + \vec{a} \;, \quad \lambda \in \mathbb{R} \;.$$

  Here, λ is called a **parameter**, $\vec{a}$ is called the **reference vector**, and $\vec{u} \neq \overrightarrow{O}$ is called he **direction vector** of the line. The position vectors $\vec{r}$ point to the individual points on the line. The reference vector $\vec{a}$ is the position vector of a fixed point on the line that is called a **reference point**. The multiples $\lambda\vec{u}$ of the direction vector $\vec{u}$ are all vectors collinear to $\vec{u}$ (see figure below).



- For a line g given in normal form by the equation

$$g\colon y = mx + b$$

the vector form of the equation of a line can be found by generating the position vectors $\begin{pmatrix} x \\ mx+b \end{pmatrix} = x \begin{pmatrix} 1 \\ m \end{pmatrix} + \begin{pmatrix} 0 \\ b \end{pmatrix}$. Then, the vector form of the equation of a line is

$$g\colon \vec{r} = x \begin{pmatrix} 1 \\ m \end{pmatrix} + \begin{pmatrix} 0 \\ b \end{pmatrix} \;, \quad x \in \mathbb{R}$$

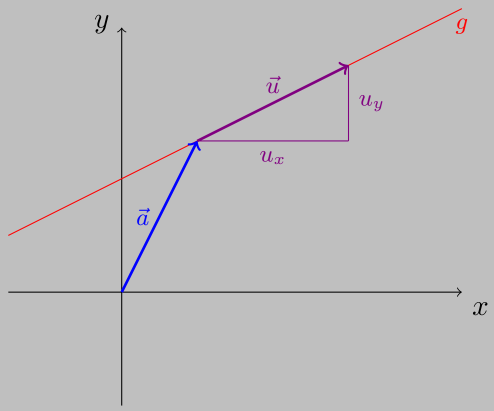
with the direction vector $\vec{u} = \begin{pmatrix} 1 \\ m \end{pmatrix}$ and the reference vector $\vec{a} = \begin{pmatrix} 0 \\ b \end{pmatrix}$.

- For a line g given by the equation in parametric form:

$$g\colon \vec{r} = \lambda\vec{u} + \vec{a} \;, \quad \lambda \in \mathbb{R}$$

the corresponding equation of the line can be found as follows: the direction vector $\vec{u} = \begin{pmatrix} u_x \\ u_y \end{pmatrix}$ immediately provides the slope of the line via a slope triangle. We have

$$m = \frac{u_y}{u_x} \;.$$



(Here, u_x must be non-zero, i.e. $u_x \neq 0$. The special case $u_x = 0$ is discussed in the example below.) From Section 9.2.2 we know that we only need another point on the line to determine the y-intercept b and specify the equation of the line in normal form. It is easiest to use the reference point $\vec{a}$ for this.

One can immediately see that the parametric form of the equation of a line is not unique. Every point on the line can be used as reference point, and the direction vector can be chosen from an arbitrary number of collinear vectors. For example, the line g with the equation of a line

$$g\colon \frac{1}{2}x + 1$$

in coordinate form discussed in the first example of the subsection is not only described by the vector equation

$$g\colon \vec{r} = x\begin{pmatrix}1\\ \frac{1}{2}\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix} \quad , \quad x \in \mathbb{R}$$

in parametric form, but also by the equations

$$g\colon \vec{r} = \lambda\begin{pmatrix}2\\1\end{pmatrix} + \begin{pmatrix}2\\2\end{pmatrix} \quad , \quad \lambda \in \mathbb{R}$$

or

$$g\colon \vec{r} = \nu\begin{pmatrix}-2\\-1\end{pmatrix} + \begin{pmatrix}-2\\0\end{pmatrix} \quad , \quad \nu \in \mathbb{R}\ .$$

Often representations are chosen such that the direction vector is as simple as possible. However, for representations of the same line by different direction vectors or reference vectors, different parameter values have to be used since in different representations the same parameter value defines different points on the line. For example, the parameter value $\lambda = 1$ defines, in the corresponding parametric equation of g, the point

$$1 \cdot \begin{pmatrix}2\\1\end{pmatrix} + \begin{pmatrix}2\\2\end{pmatrix} = \begin{pmatrix}4\\3\end{pmatrix},$$

but the parameter value $\nu = 1$ defines, in the corresponding parametric equation of g, the point

$$1 \cdot \begin{pmatrix} -2 \\ -1 \end{pmatrix} + \begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} -4 \\ -1 \end{pmatrix} .$$

The example below lists a few applications of equations of a line in vector form.

Example 10.2.3

- Let the line g in the plane be given by the equation

  $$g\colon 2y - 3x = 6 .$$

  Find two different equations of a line of g in vector form.

  First, we transform the equation of a line into normal form:

  $$2y - 3x = 6 \quad \Leftrightarrow \quad y = \frac{3}{2}x + 3 .$$

  Thus, points on the line g have the form $\left(x; \frac{3}{2}x + 3\right)$, $x \in \mathbb{R}$ described by the position vector $\begin{pmatrix} x \\ \frac{3}{2}x + 3 \end{pmatrix}$, $x \in \mathbb{R}$. Hence, one possible parametric form is given by

  $$g\colon \vec{r} = x \begin{pmatrix} 1 \\ \frac{3}{2} \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \end{pmatrix} \quad , \quad x \in \mathbb{R} .$$

  Choosing another direction vector collinear to $\begin{pmatrix} 1 \\ \frac{3}{2} \end{pmatrix}$ and another reference point on g results in another equation of a line in parametric form. For example, $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ is collinear to $\begin{pmatrix} 1 \\ \frac{3}{2} \end{pmatrix}$ since $\begin{pmatrix} 2 \\ 3 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ \frac{3}{2} \end{pmatrix}$. $\begin{pmatrix} 2 \\ 6 \end{pmatrix}$ is another appropriate reference vector since the coordinates of the point $(2; 6)$ obviously satisfy the equation of a line. Hence,

  $$g\colon \vec{r} = \sigma \begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 2 \\ 6 \end{pmatrix} \quad , \quad \sigma \in \mathbb{R}$$

  is another possible parametric form of an equation of the line g.

- In the case of a line with an equation that cannot be transformed into normal form, such as

  $$h\colon x = 2 ,$$

an equation in parametric form can still be given.

All points on the line h have the form $(2; y)$ for $y \in \mathbb{R}$, with the corresponding position vector $\begin{pmatrix} 2 \\ y \end{pmatrix}$ for $y \in \mathbb{R}$. Since $\begin{pmatrix} 2 \\ y \end{pmatrix} = y \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, one possible vector form of h is given by

$$h\colon \vec{r} = y \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 0 \end{pmatrix} \;,\;\; y \in \mathbb{R} \;.$$

- Let the line α be given in parametric form by

$$\alpha\colon \vec{r} = \mu \begin{pmatrix} -3 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \;,\;\; \mu \in \mathbb{R} \;.$$

Find the corresponding equation of a line in normal form.

The direction vector $\begin{pmatrix} -3 \\ 2 \end{pmatrix}$ gives the slope $m = \frac{2}{-3} = -\frac{2}{3}$. Thus, the equation of the line in normal form is

$$\alpha\colon y = -\frac{2}{3}x + b \;.$$

The reference vector of α is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. The reference point $(1; 1)$ can be substituted into the equation of the line to determine the y-intercept:

$$1 = -\frac{2}{3} \cdot 1 + b \;\Leftrightarrow\; b = \frac{5}{3} \;.$$

Thus, we have

$$\alpha\colon y = -\frac{2}{3}x + \frac{5}{3} \;.$$

- If, for example, a line is given by the equation

$$\beta\colon \vec{r} = \lambda \begin{pmatrix} 0 \\ -2 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \end{pmatrix} \;,\;\; \lambda \in \mathbb{R}$$

in parametric form, where the x-component of the direction vector is 0, the corresponding equation of the line in component form can be found.

The direction vector with an x-component of 0 implies that the line is parallel to the y-axis. Hence, the equation of the line has the form

$$\beta\colon x = c \;.$$

The constant c can be determined by substituting the reference point $(-1; 1)$ into this equation resulting in $-1 = c$, and we get

$$\beta\colon x = -1 \;.$$

- Given the two points $P = (-1; -1)$ and $Q = (3; 2)$, find the equation of a line of PQ in parametric form.

  For the direction vector we the connection vector

  $$\vec{u} = \overrightarrow{PQ} = \overrightarrow{Q} - \overrightarrow{P} = \begin{pmatrix} 3 \\ 2 \end{pmatrix} - \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

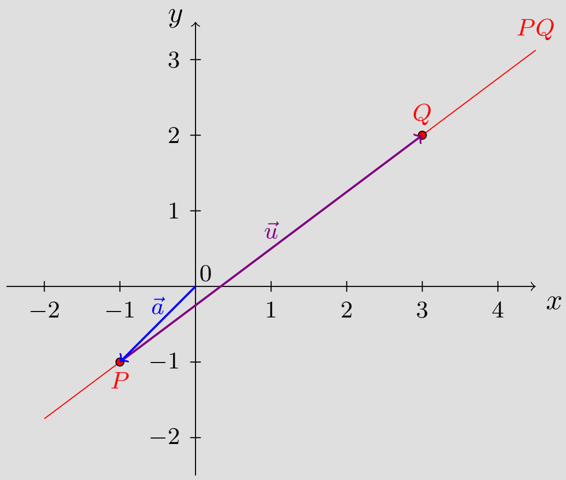
  and for the reference vector we use the position vector of a given point, for example,

  $$\vec{a} = \overrightarrow{P} = \begin{pmatrix} -1 \\ -1 \end{pmatrix} .$$

  Thus, we have

  $$PQ \colon \vec{r} = \lambda \begin{pmatrix} 4 \\ 3 \end{pmatrix} + \begin{pmatrix} -1 \\ -1 \end{pmatrix} , \quad \lambda \in \mathbb{R} .$$

  The line is shown in the figure below.



Exercise 10.2.1 a. Let the line g be given by the equation

$$g \colon \vec{r} = t \begin{pmatrix} -1 \\ 5 \end{pmatrix} + \begin{pmatrix} 2 \\ 5 \end{pmatrix} , \quad t \in \mathbb{R}$$

in parametric form. Find the equation g in normal form:
$g \colon y =$ [ ] .

b. The line h with the equation of a line in coordinate form

$$h \colon \frac{1}{2} y + x + 2 = 0$$

has the parametric form

$$h\colon \vec{r} = s\begin{pmatrix} a \\ 2 \end{pmatrix} + \begin{pmatrix} b \\ -5 \end{pmatrix}\ , \quad s \in \mathbb{R}\ .$$

Find the missing values of a and b.
$a =$ ☐
$b =$ ☐

c. Consider the two points $A = (-2; -1)$ and $B = (3; -\frac{3}{2})$. Which of the following parametric equations are correct representations of the line AB?

☐ (i) $AB : \vec{r} = s\begin{pmatrix} 5 \\ -\frac{1}{2} \end{pmatrix} + \begin{pmatrix} 0 \\ -\frac{6}{5} \end{pmatrix},\ s \in \mathbb{R}$

☐ (ii) $AB : \vec{r} = t\begin{pmatrix} 5 \\ \frac{1}{2} \end{pmatrix} + \begin{pmatrix} -2 \\ -1 \end{pmatrix},\ t \in \mathbb{R}$

☐ (iii) $AB : \vec{r} = u\begin{pmatrix} -5 \\ \frac{1}{2} \end{pmatrix} + \begin{pmatrix} -12 \\ 0 \end{pmatrix},\ u \in \mathbb{R}$

☐ (iv) $AB : \vec{r} = v\begin{pmatrix} 10 \\ -1 \end{pmatrix} + \begin{pmatrix} 4 \\ -\frac{8}{5} \end{pmatrix},\ v \in \mathbb{R}$

☐ (v) $AB : \vec{r} = w\begin{pmatrix} -1 \\ 10 \end{pmatrix} + \begin{pmatrix} -22 \\ 1 \end{pmatrix},\ w \in \mathbb{R}$

☐ (vi) $AB : \vec{r} = z\begin{pmatrix} \frac{5}{2} \\ -\frac{1}{4} \end{pmatrix} + \begin{pmatrix} \frac{6}{5} \\ -1 \end{pmatrix},\ z \in \mathbb{R}$

d. Find the value of ψ such that the point P with the position vector

$$\vec{P} = \begin{pmatrix} -2 \\ \psi \end{pmatrix}$$

lies on the line

$$i\colon \vec{r} = \tau\begin{pmatrix} 1 \\ -3 \end{pmatrix} + \begin{pmatrix} -1 \\ 2 \end{pmatrix}\ , \quad \tau \in \mathbb{R}\ ,$$

and find the value of the parameter τ such that $\vec{r} = \vec{P}$.
$\psi =$ ☐
$\tau =$ ☐

Solution:

a. The slope of the line can be found from the direction vector $\begin{pmatrix} -1 \\ 5 \end{pmatrix}$: $m = \frac{5}{-1} = -5$.
Thus, we have

$$g\colon y = -5x + b\ .$$

Substituting the reference point $(2; 5)$ into the equation results in:

$$5 = -5 \cdot 2 + b \;\Leftrightarrow\; b = 15 \,.$$

Thus, we have:

$$g\colon y = -5x + 15 \,.$$

b. Transforming the equation of a line into normal form results in

$$\frac{1}{2}y + x + 2 = 0 \;\Leftrightarrow\; y = -2x - 4 \,.$$

The reference point $(b; -5)$ corresponding to the reference vector $\begin{pmatrix} b \\ -5 \end{pmatrix}$ must lie on the line, i.e. its coordinates must satisfy the equation of the line:

$$-5 = -2b - 4 \;\Leftrightarrow\; b = \frac{1}{2} \,.$$

The position vectors of the points on h have the form $\begin{pmatrix} x \\ -2x - 4 \end{pmatrix} = x \begin{pmatrix} 1 \\ -2 \end{pmatrix} + \begin{pmatrix} 0 \\ -4 \end{pmatrix}$, so $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ is the direction vector of h. Other direction vectors of h are collinear to this vector. Since

$$-1 \cdot \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

we have $a = -1$.

c. From the given points A and B, we have for the direction vector

$$\overrightarrow{AB} = \vec{B} - \vec{A} = \begin{pmatrix} 3 \\ -\frac{3}{2} \end{pmatrix} - \begin{pmatrix} -2 \\ -1 \end{pmatrix} = \begin{pmatrix} 5 \\ -\frac{1}{2} \end{pmatrix} \,.$$

The direction vectors in the cases (ii) and (v) are not collinear to this vector. Thus, (ii) and (v) do not represent the line AB correctly. From the direction vector $\overrightarrow{AB} = \begin{pmatrix} 5 \\ -\frac{1}{2} \end{pmatrix}$, we know the slope $m = \frac{-\frac{1}{2}}{5} = -\frac{1}{10}$ of the line AB. Thus, the equation of the line is

$$AB\colon y = -\frac{1}{10}x + b \,,$$

and substituting the coordinates of A into the equation results in the following value of the y-intercept b:

$$-1 = -\frac{1}{10} \cdot (-2) + b \;\Leftrightarrow\; b = -\frac{6}{5} \,.$$

Hence, we have

$$AB\colon y = -\frac{1}{10}x - \frac{6}{5} \,.$$

The equation of the line is satisfied by the coordinates of the reference points in cases (i) to (v) but not by the coordinates of the reference point in the case (vi). Thus, the parametric equations in the cases (i), (iii), and (iv) represent the line correctly, the equations in the cases (ii), (v), and (vi), however, do not.

d. The condition

$$\begin{pmatrix} -2 \\ \psi \end{pmatrix} = \tau \begin{pmatrix} 1 \\ -3 \end{pmatrix} + \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

results in the two equations

$$-2 = \tau - 1 \quad \text{und} \quad \psi = -3\tau + 2 \,.$$

Thus, we first have $\tau = -1$ and hence, $\psi = -3 \cdot (-1) + 2 = 5$.

In contrast to lines in the plane, lines in space *cannot* be described by an equation of a line in coordinate form. However, the description by a parametric equation can be easily extended from two to three dimensions (see Info Box below).

Info10.2.4

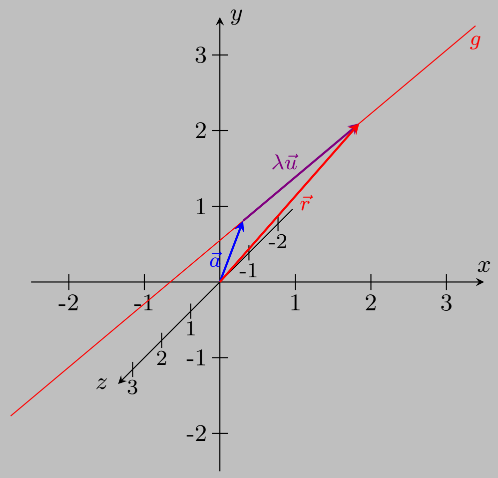
A line g in space is given in **vector form** or **parametric form** as the set of position vectors

$$g = \left\{ \vec{r} = \lambda\vec{u} + \vec{a} \in \mathbb{R}^3 \,:\, \lambda \in \mathbb{R} \right\} \,,$$

often written in short as

$$g\colon \vec{r} = \lambda\vec{u} + \vec{a} \,, \quad \lambda \in \mathbb{R} \,.$$

As in the two-dimensional case: λ is called a **parameter**, $\vec{a}$ is called the **reference vector**, and $\vec{u} \neq \overrightarrow{O}$ is called the **direction vector** of the line g (see figure below).



In this three-dimensional case, as in the plane, the parametric form of the equation of a line is not unique. The example below lists a few applications of parametric equations of a line in space.

Example 10.2.5

Let the two points $P = (-1; -2; \frac{1}{2})$ and $Q = (2; 0; 8)$ in space be given. Find two different representations of the line PQ in parametric form.

We use the connecting vector $\overrightarrow{PQ}$ as direction vector:

$$\overrightarrow{PQ} = \overrightarrow{Q} - \overrightarrow{P} = \begin{pmatrix} 2 \\ 0 \\ 8 \end{pmatrix} - \begin{pmatrix} -1 \\ -2 \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ \frac{15}{2} \end{pmatrix} .$$

The point Q can be used as reference point. For the parametric form we get

$$PQ\colon \vec{r} = t \begin{pmatrix} 3 \\ 2 \\ \frac{15}{2} \end{pmatrix} + \begin{pmatrix} 2 \\ 0 \\ 8 \end{pmatrix} , \quad t \in \mathbb{R} .$$

Further possible direction vectors must be collinear to the vector $\overrightarrow{PQ}$. For example:

$$\begin{pmatrix} -6 \\ -4 \\ -15 \end{pmatrix} = -2 \begin{pmatrix} 3 \\ 2 \\ \frac{15}{2} \end{pmatrix} .$$

We can also use the point P as reference point which results in

$$PQ\colon \vec{r} = s \begin{pmatrix} -6 \\ -4 \\ -15 \end{pmatrix} + \begin{pmatrix} -1 \\ -2 \\ \frac{1}{2} \end{pmatrix} , \quad s \in \mathbb{R}$$

as another correct equation of the line PQ in vector form.

Exercise 10.2.2 a. The line $h = AB$ through the points $A = (-1; -1; 0)$ and $B =$

$(-3; 0; 1)$ has the parametric equation

$$h\colon \vec{r} = \lambda \begin{pmatrix} 4 \\ a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \\ -4 \end{pmatrix} \;, \quad \lambda \in \mathbb{R}\,.$$

Find the missing values of a, b, c, and d.

$a =$ [ ]

$b =$ [ ]

$c =$ [ ]

$d =$ [ ]

b. Find the value of χ such that the point P with the position vector

$$\vec{P} = \begin{pmatrix} -2 \\ \chi \\ -8 \end{pmatrix}$$

lies on the line

$$g\colon \vec{r} = \nu \begin{pmatrix} 1 \\ -3 \\ 8 \end{pmatrix} + \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \;, \quad \nu \in \mathbb{R}\,,$$

and find the value of the parameter ν such that $\vec{r} = \vec{P}$.

$\chi =$ [ ]

$\nu =$ [ ]

Solution:

a. From the given points A and B, we have for the direction vector

$$\overrightarrow{AB} = \vec{B} - \vec{A} = \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}\,.$$

Thus, $\begin{pmatrix} 4 \\ a \\ b \end{pmatrix}$ is collinear to $\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$, and hence it is also a possible direction vector of h for $a = b = -2$ since in this case, we have

$$\begin{pmatrix} 4 \\ -2 \\ -2 \end{pmatrix} = -2 \cdot \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}\,.$$

The reference vector $\begin{pmatrix} c \\ d \\ -4 \end{pmatrix}$ must correspond to a point on h. Using $\overrightarrow{AB}$ as direction vector and $\overrightarrow{A}$ as reference vector, one possible equation of the line h in parametric form is

$$h\colon \vec{r} = t \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} \;, \quad t \in \mathbb{R} \,.$$

This results in the equation

$$\begin{pmatrix} c \\ d \\ -4 \end{pmatrix} = t \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2t-1 \\ t-1 \\ t \end{pmatrix} \;,$$

and from the third component we immediately read off $t = -4$. Thus, we have

$$\begin{pmatrix} c \\ d \\ -4 \end{pmatrix} = -4 \cdot \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 7 \\ -5 \\ -4 \end{pmatrix}$$

and hence $c = 7$ and $d = -5$.

b. The condition

$$\begin{pmatrix} -2 \\ \chi \\ -8 \end{pmatrix} = \nu \begin{pmatrix} 1 \\ -3 \\ 8 \end{pmatrix} + \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} \nu - 1 \\ -3\nu + 2 \\ 8\nu \end{pmatrix}$$

results for the first and the third component in $\nu = -1$, and from the second component we have $\chi = -3 \cdot (-1) + 2 = 5$.

10.2.3 Planes in Space

Starting from a vector $\vec{u}$ in space one obtains all vectors that are collinear to $\vec{u}$ (see Info Box 10.2.1) by taking all multiples $\lambda\vec{u}$, $\lambda \in \mathbb{R}$ of this vector. Interpreted as position vectors, all these collinear vectors in combination with an arbitrary reference vector constitute the parametric equation of a line as discussed in the previous Subsection 10.2.2. With this in mind, one may ask which object we obtain starting from two fixed (but *non-collinear*) vectors $\vec{u}$ and $\vec{v}$ and considering all their coplanar vectors (all vectors that result from $\lambda\vec{u} + \mu\vec{v}$; $\lambda, \mu \in \mathbb{R}$ - see Info Box 10.2.1). This, in combination with an arbitrary reference vector, generalises the concept of the parametric equation of a line resulting in the parametric equation of a plane in space which is outlined in the Info Box below.

Planes are usually denoted by uppercase Latin letters (E, F, G, ...). Of course, the concept of a plane is only meaningful in $\mathbb{R}^3$.

Info10.2.6

A plane E in space is given in **vector form** or **parametric form** as the set of position vectors

$$E = \{\vec{r} = \vec{a} + \lambda\vec{u} + \mu\vec{v} \, : \, \lambda, \mu \in \mathbb{R}\} \, ,$$

often written

$$E \colon \vec{r} = \vec{a} + \lambda\vec{u} + \mu\vec{v} \, ; \quad \lambda, \mu \in \mathbb{R} \, .$$

Here, λ and μ are called **parameters**, $\vec{a}$ is called the **reference vector**, and $\vec{u}, \vec{v} \neq \vec{O}$ is called the **direction vector** of the plane. Here, the direction vectors $\vec{u}$ and $\vec{v}$ are *non-collinear*. The position vectors point to individual points in the plane. The reference vector $\vec{a}$ is the position vector of a fixed point in the plane, called the **reference point**.

(This figure will be released shortly.)

Just as two points in space uniquely define a line (see Section 10.2.2), three given points in space uniquely define a plane. From these three given points, the parametric form of the equation of the corresponding plane can be determined rather easily. The vector form of the equation of a given plane plane is, as for a line, not unique. An infinite number of equivalent equations in vector form exists to represent a given plane. The example below lists a few typical applications.

Example 10.2.7

- The reference vector $\vec{a} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and the direction vectors $\vec{u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ define an equation in parametric form

  $$E \colon \vec{r} = \vec{a} + \lambda\vec{u} + \mu\vec{v} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \; ; \; \lambda, \mu \in \mathbb{R}$$

  of a plane that lies at an altitude of 1 parallel to the xz-plane in the coordinate system (see figure below).

  (This figure will be released shortly.)

The parametric equation of the plane E given above is not the only possible one. Each point in the plane E can be used as a reference point. For example, the point defined by the position vector $\vec{a}' = \begin{pmatrix}1\\1\\1\end{pmatrix}$ lies in E since for $\lambda = \mu = 1$ we have:

$$\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}0\\1\\0\end{pmatrix} + 1\cdot\begin{pmatrix}1\\0\\0\end{pmatrix} + 1\cdot\begin{pmatrix}0\\0\\1\end{pmatrix} .$$

Thus, this point can be used as a reference vector. All vectors that are coplanar to $\vec{u}$ and $\vec{v}$ but not collinear to each other can be used as alternative direction vectors. Examples are the vectors $\vec{u}' = \begin{pmatrix}1\\0\\1\end{pmatrix} = 1\cdot\begin{pmatrix}1\\0\\0\end{pmatrix} + 1\cdot\begin{pmatrix}0\\0\\1\end{pmatrix}$ and $\vec{v}' = \begin{pmatrix}1\\0\\-1\end{pmatrix} = 1\cdot\begin{pmatrix}1\\0\\0\end{pmatrix} - 1\cdot\begin{pmatrix}0\\0\\1\end{pmatrix}$. Then, another representation of E in parametric form is given by the equation

$$E\colon \vec{r} = \vec{a}' + s\vec{u}' + t\vec{v}' = \begin{pmatrix}1\\1\\1\end{pmatrix} + s\begin{pmatrix}1\\0\\1\end{pmatrix} + t\begin{pmatrix}1\\0\\-1\end{pmatrix} \quad ; \quad s,t \in \mathbb{R} .$$

- Consider three points $A = (1;0;-2)$, $B = (4;1;2)$, and $C = (0;2;1)$. Find the equation of the plane F that is specified by these three points, in parametric form.

  One of these three points, for example the point A, is used as the reference point. $\overrightarrow{A} = \begin{pmatrix}1\\0\\-2\end{pmatrix}$ is the corresponding reference vector. The connecting vectors from the reference point to the two other points are used as the direction vectors:

$$\overrightarrow{AB} = \overrightarrow{B} - \overrightarrow{A} = \begin{pmatrix}4\\1\\2\end{pmatrix} - \begin{pmatrix}1\\0\\-2\end{pmatrix} = \begin{pmatrix}3\\1\\4\end{pmatrix} ,$$

$$\overrightarrow{AC} = \overrightarrow{C} - \overrightarrow{A} = \begin{pmatrix}0\\2\\1\end{pmatrix} - \begin{pmatrix}1\\0\\-2\end{pmatrix} = \begin{pmatrix}-1\\2\\3\end{pmatrix} .$$

  Hence, the equation

$$F\colon \vec{r} = \begin{pmatrix}1\\0\\-2\end{pmatrix} + \rho\begin{pmatrix}3\\1\\4\end{pmatrix} + \sigma\begin{pmatrix}-1\\2\\3\end{pmatrix} \quad ; \quad \rho,\sigma \in \mathbb{R}$$

is a correct representation of the plane F in parametric form.

(This figure will be released shortly.)

- Consider the two points $P = (1; 2; 3)$ and $Q = (2; 6; 6)$. Verify whether they lie in the plane G given by the equation

$$G\colon \vec{r} = \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \nu \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \quad ; \quad \mu, \nu \in \mathbb{R}$$

in parametric form.

The points P or Q lie in the plane G if their position vectors arise for specific parameter values of μ and ν as position vectors from the equation of G, i.e. $\overrightarrow{P} = \vec{r}$ or $\overrightarrow{Q} = \vec{r}$ for appropriate values of μ and ν. For the point P, we have:

$$\overrightarrow{P} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \nu \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \mu \\ 3 + 2\mu + \nu \\ 2 + 3\mu + 2\nu \end{pmatrix} .$$

From the first component of this vector equation we get $\mu = 1$. Substituting this parameter value into the second and third component provides two contradicting equations in the parameter ν:

$$2 = 3 + 2 \cdot 1 + \nu \;\Leftrightarrow\; \nu = -3$$

and

$$3 = 2 + 3 \cdot 1 + 2\nu \;\Leftrightarrow\; \nu = -1 .$$

There are no parameter values μ and ν providing in the parametric equation of the plane G the position vector $\overrightarrow{P}$, so the point P does not lie in the plane G. For Q, however, we have:

$$\overrightarrow{Q} = \begin{pmatrix} 2 \\ 6 \\ 6 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \nu \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \mu \\ 3 + 2\mu + \nu \\ 2 + 3\mu + 2\nu \end{pmatrix} .$$

From the first component we get $\mu = 2$. Substituting this parameter value into the second and third component results in

$$6 = 3 + 2 \cdot 2 + \nu \;\Leftrightarrow\; \nu = -1$$

and

$$6 = 2 + 3 \cdot 2 + 2\nu \;\Leftrightarrow\; \nu = -1 .$$

This is not a contradiction. We see that the parameter values $\mu = 2$ and $\nu = -1$ provide the position vector $\overrightarrow{Q}$. Hence, the point Q lies in the plane G.

(This figure will be released shortly.)

As well as by three points, a plane can also be defined by a line and a point that does not lie on the line. The example below shows how this can be reduced to the case of three given points.

Example 10.2.8

Let a point $P = (2; 1; -3)$ be given. In addition, let a line g be given in parametric form by the equation

$$g\colon \vec{r} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix} \ , \quad t \in \mathbb{R} \ .$$

The point P does not lie on the line g since there is no value of the parameter $t \in \mathbb{R}$ such that

$$\overrightarrow{P} = \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 2t \\ -1 \\ -t \end{pmatrix} \ .$$

The second component of this vector equation results in the contradiction $1 = -1$. The point P and the line g uniquely define a plane E that contains both P and g. A parametric equation of this plane can by found by choosing two additional points on g besides the given point P that can be used as a reference point and then proceeding as in the example above for three given points. Hence, the reference vector is in this case

$$\overrightarrow{P} = \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix} \ ,$$

and the two additional points Q_1 and Q_2 on g result from the equation of the line for two different values of the parameter t, for example, $t = 0$ and $t = 1$. Choosing $t = 0$ results in the reference point of the line as position vector:

$$\overrightarrow{Q}_1 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} \ .$$

Choosing $t = 1$ results in

$$\overrightarrow{Q}_2 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} \ .$$

Thus, the direction vectors are

$$\overrightarrow{PQ_1} = \overrightarrow{Q}_1 - \overrightarrow{P} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix} = \begin{pmatrix} -2 \\ -2 \\ 3 \end{pmatrix}$$

and

$$\overrightarrow{PQ}_2 = \overrightarrow{Q}_2 - \overrightarrow{P} = \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ 2 \end{pmatrix} .$$

Hence, the plane E is given by the vector equation

$$E\colon \vec{r} = \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix} + v \begin{pmatrix} -2 \\ -2 \\ 3 \end{pmatrix} + w \begin{pmatrix} 0 \\ -2 \\ 2 \end{pmatrix} \quad ; \quad v, w \in \mathbb{R} .$$

(This figure will be released shortly.)

In the following Section 10.2.4 we will further discuss the relative positions of planes and lines, as well as other data that can be used to define a plane uniquely.

Exercise 10.2.3
The plane E uniquely defined by the three points $A = (0; 0; 8)$, $B = (3; -1; 10)$, and $C = (-1; -2; 11)$ has the parametric equation

$$E\colon \vec{r} = \begin{pmatrix} 2 \\ -3 \\ x \end{pmatrix} + s \begin{pmatrix} y \\ 1 \\ -1 \end{pmatrix} + t \begin{pmatrix} 5 \\ z \\ -4 \end{pmatrix} \quad ; \quad s, t \in \mathbb{R} .$$

Find the missing components x, y, and z.
$x =$ [ ]
$y =$ [ ]
$z =$ [ ]

Solution:

The reference vector $\overrightarrow{A} = \begin{pmatrix} 0 \\ 0 \\ 8 \end{pmatrix}$ and the direction vectors

$$\overrightarrow{AB} = \overrightarrow{B} - \overrightarrow{A} = \begin{pmatrix} 3 \\ -1 \\ 10 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 8 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} ,$$

$$\overrightarrow{AC} = \overrightarrow{C} - \overrightarrow{A} = \begin{pmatrix} -1 \\ -2 \\ 11 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 8 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \\ 3 \end{pmatrix}$$

define the following parametric equation:

$$E\colon \vec{r} = \begin{pmatrix}0\\0\\8\end{pmatrix} + \mu \begin{pmatrix}3\\-1\\2\end{pmatrix} + \nu \begin{pmatrix}-1\\-2\\3\end{pmatrix} \quad ; \quad \mu, \nu \in \mathbb{R} \; .$$

The reference point $\begin{pmatrix}2\\-3\\x\end{pmatrix}$ lies in the plane E if

$$\begin{pmatrix}2\\-3\\x\end{pmatrix} = \begin{pmatrix}0\\0\\8\end{pmatrix} + \mu \begin{pmatrix}3\\-1\\2\end{pmatrix} + \nu \begin{pmatrix}-1\\-2\\3\end{pmatrix} = \begin{pmatrix}3\mu - \nu\\-\mu - 2\nu\\8 + 2\mu + 3\nu\end{pmatrix} \; .$$

This is a system of linear equations in the three variables μ, ν, and x that can be solved using the methods described in Section 4.3. Considering the first and second components results in the two equations

$$2 = 3\mu - \nu \;\text{ and }\; -3 = -\mu - 2\nu$$

with the solution $\mu = \nu = 1$. Substituting these values into the third component results in

$$x = 8 + 2 + 3 = 13 \; .$$

The two vectors $\begin{pmatrix}y\\1\\-1\end{pmatrix}$ and $\begin{pmatrix}5\\z\\-4\end{pmatrix}$ are direction vectors of the plane E if they are coplanar to both $\overrightarrow{AB}$ and $\overrightarrow{AC}$. For the first vector, we have

$$\begin{pmatrix}y\\1\\-1\end{pmatrix} = a \begin{pmatrix}3\\-1\\2\end{pmatrix} + b \begin{pmatrix}-1\\-2\\3\end{pmatrix} = \begin{pmatrix}3a - b\\-a - 2b\\2a + 3b\end{pmatrix} \; ,$$

which is again a system of linear equations in three variables. Considering the second and the third component results in the two equations

$$1 = -a - 2b \;\text{ and }\; -1 = 2a + 3b$$

with the solution $a = 1$, $b = -1$. Substituting this values into the first equation results in

$$y = 3 + (-1) \cdot (-1) = 4 \; .$$

For the second vector we have

$$\begin{pmatrix}5\\z\\-4\end{pmatrix} = a \begin{pmatrix}3\\-1\\2\end{pmatrix} + b \begin{pmatrix}-1\\-2\\3\end{pmatrix} = \begin{pmatrix}3a - b\\-a - 2b\\2a + 3b\end{pmatrix} \; .$$

Considering the first and the third component results in the two equations

$$5 = 3a - b \text{ and } -4 = 2a + 3b$$

with the solution $a = 1$, $b = -2$ which finally results in

$$z = -1 - 2 \cdot (-2) = 3 \, .$$

Exercise 10.2.4
Consider the points $P = (h; 2; -2)$, $Q = (1; i; 6)$, $R = (-3; 2; j)$ and the plane E be given by an equation

$$E\colon \vec{r} = \begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix} + s \begin{pmatrix} 2 \\ 1 \\ 7 \end{pmatrix} + t \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix} \quad ; \quad s, t \in \mathbb{R}$$

in parametric form. Find the missing components h, i, and j such that the points P, Q, and R lie in the plane E.
$h =$ ☐
$i =$ ☐
$j =$ ☐

Solution:
For the plane E, we have the equation

$$\vec{r} = \begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix} + s \begin{pmatrix} 2 \\ 1 \\ 7 \end{pmatrix} + t \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 3 + 2s + 3t \\ s + 2t \\ 2 + 7s + 5t \end{pmatrix} \, .$$

The conditions

$$\vec{P} = \vec{r} \Leftrightarrow \begin{pmatrix} h \\ 2 \\ -2 \end{pmatrix} = \begin{pmatrix} 3 + 2s + 3t \\ s + 2t \\ 2 + 7s + 5t \end{pmatrix} \, ,$$

$$\vec{Q} = \vec{r} \Leftrightarrow \begin{pmatrix} 1 \\ i \\ 6 \end{pmatrix} = \begin{pmatrix} 3 + 2s + 3t \\ s + 2t \\ 2 + 7s + 5t \end{pmatrix} \, ,$$

and

$$\vec{R} = \vec{r} \Leftrightarrow \begin{pmatrix} -3 \\ 2 \\ j \end{pmatrix} = \begin{pmatrix} 3 + 2s + 3t \\ s + 2t \\ 2 + 7s + 5t \end{pmatrix}$$

each result in a system of linear equations in the variables s, t, and h or i or j that can be solved using the methods described in Section 4.3.

For the point P, considering the second and the third component results in a system of two linear equations in the variables s and t

$$2 = s + 2t \text{ and } -4 = 7s + 5t$$

with the solution $s = -2$, $t = 2$. Substituting the solution into the first component results in

$$h = 3 + 2 \cdot (-2) + 3 \cdot 2 = 5 \, .$$

For the point Q, considering the first and the third component results in a system of two linear equations in the two variables s and t

$$-2 = 2s + 3t \text{ and } 4 = 7s + 5t$$

with the solution $s = 2$, $t = -2$. Substituting the solution into the second component results in

$$i = 2 + 2 \cdot (-2) = -2 \, .$$

For the point R, considering the first and the second component results in a system of two linear equations in the two variables s and t

$$-6 = 2s + 3t \text{ and } 2 = s + 2t$$

with the solution $s = -18$, $t = 10$. Substituting the solution into the third component results in

$$j = 2 + 7 \cdot (-18) + 5 \cdot 10 = -74 \, .$$

10.2.4 Relative Positions of Lines and Planes in Space

While two lines in the plane can only have three different relative positions with respect to each other (lines are parallel, coincide, or intersect, see Section 9.2.3), two lines in space can have four different relative positions with respect to each other. These will be outlined in the Info Box below.

Info10.2.9

Let two lines in space be given by vector equations. The line g has the reference vector $\vec{a}$ and the direction vector $\vec{u}$, and the line h has the reference vector $\vec{b}$ and the direction vector $\vec{v}$:

$$g\colon \vec{r} = \vec{a} + s\vec{u} \; ; \quad s \in \mathbb{R} \, ,$$

$$h\colon \vec{r} = \vec{b} + t\vec{v}\ ;\quad t \in \mathbb{R}\ .$$

The two lines g and h can have four different relative positions:

1. The lines are **identical**. In this case, the lines g and h have all their points in common, they coincide. This is the case if and only if the two direction vectors $\vec{u}$ and $\vec{v}$ are collinear and the lines have any one point in common.
2. The lines are **parallel**. This is the case if and only if the two direction vectors $\vec{u}$ and $\vec{v}$ are collinear and the two lines do *not* have any points in common.
3. The lines intersect. In this case, the lines g and h have exactly one point in common. This point is called the **intersection point**. This is the case if and only if the two direction vectors $\vec{u}$ and $\vec{v}$ are *not* collinear and the two lines have exactly one point in common.
4. Lines that are neither identical nor parallel and do not intersect are called **skew**. This is the case if and only if the two direction vectors are *not* collinear, and the two lines do *not* have any points in common.

(This figure will be released shortly.)

In practice, the relative position of two lines in space is investigated according as follows: first, we examine two direction vectors for collinearity, then we check whether the two lines have points in common. This uniquely identifies one of the four cases. The example below illustrates this approach for all four cases.

Example 10.2.10

Let the four lines g, h, i, and j be given in parametric form by

$$g\colon \vec{r} = \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix} + s \begin{pmatrix} -2 \\ 2 \\ -4 \end{pmatrix}\ ;\quad s \in \mathbb{R}$$

$$h\colon \vec{r} = \begin{pmatrix} 1 \\ -2 \\ 7 \end{pmatrix} + t \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}\ ;\quad t \in \mathbb{R}$$

$$i\colon \vec{r} = \begin{pmatrix} 4 \\ 0 \\ 8 \end{pmatrix} + u \begin{pmatrix} -3 \\ 3 \\ -6 \end{pmatrix}\ ;\quad u \in \mathbb{R}$$

$$j\colon \vec{r} = \begin{pmatrix}1\\3\\2\end{pmatrix} + v \begin{pmatrix}1\\3\\-3\end{pmatrix} \quad ; \quad v \in \mathbb{R}\ .$$

- The lines g and h are identical. The two direction vectors $\begin{pmatrix}-2\\2\\-4\end{pmatrix}$ of g and $\begin{pmatrix}1\\-1\\2\end{pmatrix}$ of h are collinear. We have

$$\begin{pmatrix}-2\\2\\-4\end{pmatrix} = -2 \cdot \begin{pmatrix}1\\-1\\2\end{pmatrix} \ .$$

The point described by the position vector $\begin{pmatrix}1\\-2\\7\end{pmatrix}$ lies both on the line h (as reference point) and on the line g since we have for the line g:

$$\begin{pmatrix}1\\-2\\7\end{pmatrix} = \begin{pmatrix}-1\\0\\3\end{pmatrix} + s \begin{pmatrix}-2\\2\\-4\end{pmatrix} = \begin{pmatrix}-1-2s\\2s\\3-4s\end{pmatrix} \ \Leftrightarrow\ s = -1\ .$$

Thus, the vector $\begin{pmatrix}1\\-2\\7\end{pmatrix}$ results from the equation of g for the parameter value $s = -1$.

- The lines h and i (and hence the lines g and i) are parallel. The two direction vectors $\begin{pmatrix}1\\-1\\2\end{pmatrix}$ of h and $\begin{pmatrix}-3\\3\\-6\end{pmatrix}$ of i are collinear.

$$\begin{pmatrix}-3\\3\\-6\end{pmatrix} = -3 \cdot \begin{pmatrix}1\\-1\\2\end{pmatrix} \ .$$

However, the lines h and i do not have any points in common: the reference point of one of the two lines is not a point on the other line. Here, we can check whether the reference vector $\begin{pmatrix}4\\0\\8\end{pmatrix}$ of the line i can result as a position vector of the line h:

$$\begin{pmatrix}4\\0\\8\end{pmatrix} = \begin{pmatrix}1\\-2\\7\end{pmatrix} + t \begin{pmatrix}1\\-1\\2\end{pmatrix} = \begin{pmatrix}1+t\\-2-t\\7+2t\end{pmatrix} \ .$$

In this vector equation, $t = 3$ results from the first component and $t = -2$ from the second, which is a contradiction. Hence, the two lines do not have any points in common.

- The lines i and j intersect. First we see that for these two lines the two direction vectors $\begin{pmatrix}-3\\3\\-6\end{pmatrix}$ and $\begin{pmatrix}1\\3\\-3\end{pmatrix}$ are not collinear. There is no number $a \in \mathbb{R}$ such that
$$\begin{pmatrix}-3\\3\\-6\end{pmatrix} = a \begin{pmatrix}1\\3\\-3\end{pmatrix}$$
since the equality of the first component results in $a = -3$ and equality of the second component results in $a = 1$, which is a contradiction. However, these two lines have a point in common that can be found by equating the position vectors for i and j:
$$\begin{pmatrix}4\\0\\8\end{pmatrix} + u \begin{pmatrix}-3\\3\\-6\end{pmatrix} = \begin{pmatrix}4-3u\\3u\\8-6u\end{pmatrix} = \begin{pmatrix}1+v\\3+3v\\2-3v\end{pmatrix} = \begin{pmatrix}1\\3\\2\end{pmatrix} + v \begin{pmatrix}1\\3\\-3\end{pmatrix} .$$
Equating the first two components results in the two equations
$$3 - 3u = v \text{ and } u = 1 + v$$
in the variables u and v with the solution $v = 0$, $u = 1$. Substituting these values into the equation for the third component results in
$$8 - 6 \cdot 1 = 2 - 3 \cdot 0 \Leftrightarrow 2 = 2 .$$
Thus, the vector equation for the position vectors is satisfied for the parameter values $u = 1$ and $v = 0$. Hence, the position vector of the intersection point results from substituting the parameter value $u = 1$ into the equation of the line i or from substituting the parameter value $v = 0$ into the equation of the line j. For the intersection point of the lines we have $(1; 3; 2)$.

- The lines g and j (and hence the lines h and j) are skew. As in the previous case of the intersecting lines it can be easily seen that the two direction vectors $\begin{pmatrix}-2\\2\\-4\end{pmatrix}$ of g and $\begin{pmatrix}1\\3\\-3\end{pmatrix}$ of j are not collinear. However, in this case the lines do not have any point in common which again can be found by equating the position vectors:
$$\begin{pmatrix}-1\\0\\3\end{pmatrix} + s \begin{pmatrix}-2\\2\\-4\end{pmatrix} = \begin{pmatrix}-1-2s\\2s\\3-4s\end{pmatrix} = \begin{pmatrix}1+v\\3+3v\\2-3v\end{pmatrix} = \begin{pmatrix}1\\3\\2\end{pmatrix} + v \begin{pmatrix}1\\3\\-3\end{pmatrix} .$$

This vector equation involves a contradiction; there are no pairs of parameter values of s and v such that the equation is satisfied, and hence g and j do not have any points in common. Considering the first and the second components results in the two equations

$$-2s = 2 + v \text{ and } 2s = 3 + 3v$$

with the solution $v = -\frac{5}{4}$, $s = -\frac{3}{8}$. However, substituting this into the equation for the third component results in the contradiction

$$3 - 4(-\frac{3}{8}) = 2 - 3(-\frac{5}{4}) \Leftrightarrow \frac{9}{2} = \frac{23}{4} .$$

Exercise 10.2.5
Tick the true statements:
The two lines given by the equations

$$g\colon \vec{r} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} + x \begin{pmatrix} -5 \\ 10 \\ -15 \end{pmatrix} \; ; \; x \in \mathbb{R}$$

and

$$h\colon \vec{r} = \begin{pmatrix} 4 \\ 0 \\ 7 \end{pmatrix} + y \begin{pmatrix} 3 \\ -2 \\ 3 \end{pmatrix} \; ; \; y \in \mathbb{R}$$

intersect since

- ☐ the two direction vectors are collinear,
- ☐ the two direction vectors are not collinear,
- ☐ the two direction vectors are collinear and the lines have a point in common,
- ☐ the two direction vectors are not collinear and the lines have a point in common,
- ☐ the two direction vectors are not collinear and the lines do not have any points in common.

Find the intersection point S of the two lines g and h.
$S =$ ☐

The position vector of the intersection point $\overrightarrow{S}$ results from the lines g and h for the parameter values
$x =$ ☐ and
$y =$ ☐ .

Solution:

The two direction vectors $\begin{pmatrix}-5\\10\\-15\end{pmatrix}$ and $\begin{pmatrix}3\\-2\\3\end{pmatrix}$ are not collinear since there is no number $a \in \mathbb{R}$ such that the equation

$$\begin{pmatrix}-5\\10\\-15\end{pmatrix} = a\begin{pmatrix}3\\-2\\3\end{pmatrix}$$

is satisfied. Considering the second component of this vector equation results in $a = -5$, considering the first component of this vector equation results in $a = -\frac{5}{3}$; this is a contradiction. The two lines have a point in common, namely the intersection point S that will be calculated below. According to Info Box 10.2.9 these conditions and only these conditions are sufficient that the two lines intersect.

The intersection point results from equating the two position vectors:

$$\begin{pmatrix}1\\2\\4\end{pmatrix} + x\begin{pmatrix}-5\\10\\-15\end{pmatrix} = \begin{pmatrix}1-5x\\2+10x\\4-15x\end{pmatrix} = \begin{pmatrix}4+3y\\-2y\\7+3y\end{pmatrix} = \begin{pmatrix}4\\0\\7\end{pmatrix} + y\begin{pmatrix}3\\-2\\3\end{pmatrix} .$$

Considering the first two components of this vector equation results in the system of two linear equations

$$-5x = 3+3y \text{ and } 2+10x = -2y ,$$

with the solution $x = 0$, $y = -1$. Substituting these values into the equation for the third component results in

$$4 - 15\cdot 0 = 7 + 3\cdot(-1) \Leftrightarrow 4 = 4 .$$

Thus, the vector equation is satisfied for these parameter values, and the two lines g and h have a point in common. Its position vector results from substituting, for example, the parameter value $x = 0$ into the equation of g:

$$\overrightarrow{S} = \begin{pmatrix}1\\2\\4\end{pmatrix} + 0\cdot\begin{pmatrix}-5\\10\\-15\end{pmatrix} = \begin{pmatrix}1\\2\\4\end{pmatrix} .$$

Exercise 10.2.6

The two lines given by the equations

$$\gamma\colon \vec{r} = \begin{pmatrix}-4\\6\\0\end{pmatrix} + s\begin{pmatrix}3\\-2\\-2\end{pmatrix} \; ; \; s\in\mathbb{R}$$

and

$$\kappa\colon \vec{r} = \begin{pmatrix}a\\b\\4\end{pmatrix} + t\begin{pmatrix}-\frac{3}{2}\\c\\1\end{pmatrix} \; ; \; t\in\mathbb{R}$$

are parallel. Find the value of c and specify which values the parameters a and b must **not** take simultaneously to ensure that the two lines are parallel.

$a \neq$ [ ]
$b \neq$ [ ]
$c =$ [ ]

Solution:
The two lines are parallel if the two direction vectors are collinear. From this condition

$$\begin{pmatrix} -\frac{3}{2} \\ c \\ 1 \end{pmatrix} = s \begin{pmatrix} 3 \\ -2 \\ -2 \end{pmatrix} \quad ,$$

we find $s = -\frac{1}{2}$ and hence $c = 1$. To ensure that the lines are really parallel and not identical the reference point of κ must not lie on γ. Thus, the parameters a and b must have values such that the vector equation

$$\begin{pmatrix} a \\ b \\ 4 \end{pmatrix} = \begin{pmatrix} -4 \\ 6 \\ 0 \end{pmatrix} + s \begin{pmatrix} 3 \\ -2 \\ -2 \end{pmatrix} = \begin{pmatrix} -4+3s \\ 6-2s \\ -2s \end{pmatrix}$$

cannot be satisfied for any value of the parameter s. Equating the third component immediately results in $s = -2$, i.e. the equation is satisfied for this value. Substituting this value into the equation for the first and the second component results in $a = -10$ and $b = 10$. Hence, the lines are really parallel if $a \neq -10$ or $b \neq 10$.

Consider two lines in space be given that are truly parallel or intersecting. Then these two lines uniquely define a plane (see figure below).

(This figure will be released shortly.)

This is always the plane that contains both lines.

The example below shows how to derive the parametric equation of the plane from two truly parallel or intersecting lines.

Example 10.2.11

- The two lines given by the equations

$$g\colon \vec{r} = \begin{pmatrix}1\\2\\4\end{pmatrix} + s\begin{pmatrix}-5\\10\\-15\end{pmatrix} \quad ; \quad s\in\mathbb{R}$$

and

$$h\colon \vec{r} = \begin{pmatrix}4\\0\\7\end{pmatrix} + t\begin{pmatrix}3\\-2\\3\end{pmatrix} \quad ; \quad t\in\mathbb{R}$$

intersect at the point $S = (1;2;4)$ (see Exercise 10.2.5). Thus, they uniquely define a plane E that contains both g and h. For the equation of the plane E in parametric form, the position vectors of three appropriate points are used resulting from the given equations of the lines g and h. The intersection point is a suitable reference point of E, with the reference vector

$$\overrightarrow{S} = \begin{pmatrix}1\\2\\4\end{pmatrix}.$$

Then, the position vectors of two points P and Q result, for example, from substituting the parameter values $s = 1$ and $t = 1$ into the equations of the lines in parametric form:

$$\overrightarrow{P} = \begin{pmatrix}1\\2\\4\end{pmatrix} + 1\cdot\begin{pmatrix}-5\\10\\-15\end{pmatrix} = \begin{pmatrix}-4\\12\\-11\end{pmatrix},$$

$$\overrightarrow{Q} = \begin{pmatrix}4\\0\\7\end{pmatrix} + 1\cdot\begin{pmatrix}3\\-2\\3\end{pmatrix} = \begin{pmatrix}7\\-2\\10\end{pmatrix}.$$

Thus, we have the direction vectors

$$\overrightarrow{SP} = \overrightarrow{P} - \overrightarrow{S} = \begin{pmatrix}-4\\12\\-11\end{pmatrix} - \begin{pmatrix}1\\2\\4\end{pmatrix} = \begin{pmatrix}-5\\10\\-15\end{pmatrix}$$

and

$$\overrightarrow{SQ} = \overrightarrow{Q} - \overrightarrow{S} = \begin{pmatrix}7\\-2\\10\end{pmatrix} - \begin{pmatrix}1\\2\\4\end{pmatrix} = \begin{pmatrix}6\\-4\\11\end{pmatrix}.$$

Hence, a possible equation of the plane E in parametric form is given by

$$E\colon \vec{r} = \begin{pmatrix}1\\2\\4\end{pmatrix} + \mu\begin{pmatrix}-5\\10\\-15\end{pmatrix} + \nu\begin{pmatrix}6\\-4\\11\end{pmatrix} \quad ; \quad \mu,\nu\in\mathbb{R}.$$

- The two lines g and h given by the equations

$$g\colon \vec{r} = \begin{pmatrix} -4 \\ 6 \\ 0 \end{pmatrix} + s \begin{pmatrix} 3 \\ -2 \\ -2 \end{pmatrix} \quad ; \quad s \in \mathbb{R}$$

and

$$h\colon \vec{r} = \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix} + t \begin{pmatrix} -\frac{3}{2} \\ 1 \\ 1 \end{pmatrix} \quad ; \quad t \in \mathbb{R}$$

are parallel (see Exercise 10.2.6). This uniquely defines a plane F that contains both g and h. The parametric equation of the plane F is derived from the position vectors of three appropriate points resulting from the equations of the lines g and h. The position vectors of points on g or h that result, for example, from substituting the parameter values $s = 0$, $s = 1$, and $t = 0$ are suitable. The first vector

$$\vec{A} = \begin{pmatrix} -4 \\ 6 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 3 \\ -2 \\ -2 \end{pmatrix} = \begin{pmatrix} -4 \\ 6 \\ 0 \end{pmatrix}$$

can be used as a reference vector. From

$$\vec{B} = \begin{pmatrix} -4 \\ 6 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 3 \\ -2 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \\ -2 \end{pmatrix}$$

it immediately results that as the first direction vector $\overrightarrow{AB}$ of the plane F the direction vector $\begin{pmatrix} 3 \\ -2 \\ -2 \end{pmatrix}$ of g can be used. The second direction vector of the plane results from the position vector of a point B on h for the parameter value $t = 0$, i.e. the reference point of h:

$$\vec{B} = \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix} .$$

Hence,

$$\overrightarrow{AB} = \vec{B} - \vec{A} = \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix} - \begin{pmatrix} -4 \\ 6 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ -5 \\ 4 \end{pmatrix}$$

is the second direction vector of F. Thus, a parameter form of the equation of F is

$$F\colon \vec{r} = \begin{pmatrix} -4 \\ 6 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ -2 \\ -2 \end{pmatrix} + \mu \begin{pmatrix} 5 \\ -5 \\ 4 \end{pmatrix} \quad ; \quad \lambda, \mu \in \mathbb{R} .$$

A line and a plane in space can only have three different relative positions with respect to each other, as outlined in the Info Box below.

Info10.2.12

Let a line g with reference point $\vec{a}$ and direction vector $\vec{u}$ and a plane E with the reference vector $\vec{b}$ and the direction vectors $\vec{v}$ and $\vec{w}$ in space be given in parametric form by the equations

$$g\colon \vec{r} = \vec{a} + \lambda\vec{u}\ ;\quad \lambda \in \mathbb{R}$$

and

$$E\colon \vec{r} = \vec{b} + \mu\vec{v} + \nu\vec{w}\ ;\quad \mu, \nu \in \mathbb{R}\ .$$

Then the line g and the plane E can have three different relative positions:

1. The line g lies in the plane E. This is the case if and only if the three direction vectors $\vec{u}$, $\vec{v}$, and $\vec{w}$ are coplanar, and the reference point of the line lies in the plane.
2. The line g is parallel to the plane E. This is the case if and only if the three direction vectors $\vec{u}$, $\vec{v}$, and $\vec{w}$ are coplanar and the reference point of the line does *not* lie in the plane.
3. The lines g and the plane E intersect. This is the case if and only if the three direction vectors $\vec{u}$, $\vec{v}$, and $\vec{w}$ are not coplanar.

(This figure will be released shortly.)

To investigate the relative position of a given line and plane, we first examine the three direction vectors for collinearity, then we check whether the reference point of the line lies in the plane. This uniquely identifies one of the three possible cases. If the line and the plane intersect, we can calculate the intersection point. The example below illustrates a few approaches.

Example 10.2.13

Let the plane E be given by the parametric equation

$$E\colon \vec{r} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} + s \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}\ ;\quad s, t \in \mathbb{R}\ .$$

- A line with the vector $\begin{pmatrix} 3 \\ -1 \\ -4 \end{pmatrix}$ as its direction vector either lies in the plane or is parallel to the plane E, since $\begin{pmatrix} 3 \\ -1 \\ -4 \end{pmatrix}$ is coplanar to the two direction vectors of E. From the condition

$$\begin{pmatrix} 3 \\ -1 \\ -4 \end{pmatrix} = s \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} ,$$

we have $s = 1$, $t = -2$. Hence, the line

$$g\colon \vec{r} = \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix} + x \begin{pmatrix} 3 \\ -1 \\ -4 \end{pmatrix} \quad ; \quad x \in \mathbb{R}$$

lies in the plane E, since the reference point $(-1; 3; 0)$ lies in E.

$$\begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} + s \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 2+3s \\ 2-s \\ 2+2t \end{pmatrix} \Leftrightarrow s = t = -1 \, .$$

Hence, the position vector $\begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix}$ of the line results from the parametric equation of the plane for the parameter values $s = t = -1$. In contrast, the line

$$h\colon \vec{r} = y \begin{pmatrix} 3 \\ -1 \\ -4 \end{pmatrix} \quad ; \quad y \in \mathbb{R}$$

is parallel to the plane E since h has the origin $(0; 0; 0)$ as its reference point. The origin does not lie in the plane E since there are no parameter values s and t such that the vector equation

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} + s \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 2+3s \\ 2-s \\ 2+2t \end{pmatrix}$$

is satisfied. Considering the first component implies $s = -\frac{2}{3}$, and $s = 2$ results from the second component; this is a contradiction.

- Every line with a direction vector that is not coplanar to the two direction vectors $\begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$ of E intersect the plane E at exactly one point. An example of such a line is

$$k\colon \vec{r} = \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad ; \quad \mu \in \mathbb{R} \, .$$

The direction vector $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ is not coplanar to $\begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$ since the condition

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = a \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 3a \\ -a \\ 2b \end{pmatrix}$$

cannot be satisfied by any choice of $a, b \in \mathbb{R}$. Considering the first component would imply $a = \frac{1}{3}$ and the second would imply $a = -1$; this is a contradiction. From equating the position vectors of the line k and the plane E, the intersection point can be calculated:

$$\begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -3+\mu \\ 1+\mu \\ \mu \end{pmatrix} = \begin{pmatrix} 2+3s \\ 2-s \\ 2+2t \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} + s \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} \, .$$

If we are only interested in the intersection point, it is sufficient to determine the parameter value of the line for which this vector equation is satisfied. The position vector of the intersection point then results from substituting the determined parameter value into the equation of the line. Considering the first two components of this vector equation results in a system of two linear equations in the variables μ and s:

$$\mu = 5 + 3s \ \text{ and } \ \mu = 1 - s \, ,$$

with the solution $\mu = 2$. Thus, the intersection point has the position vector

$$\begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ 2 \end{pmatrix} \, .$$

Exercise 10.2.7
Let the plane E be given by the equation

$$E\colon \vec{r} = \begin{pmatrix} 8 \\ -2 \\ 0 \end{pmatrix} + s \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} + t \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} \quad ; \quad s, t \in \mathbb{R}$$

and the line g by the equation

$$g\colon \vec{r} = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} + u \begin{pmatrix} 0 \\ 4 \\ c \end{pmatrix} \quad ; \quad u \in \mathbb{R}$$

whose reference point does not lie in the plane E.

Find the missing component c such that the line g is parallel to the line E.
$c =$ [ ]

For all other values of c, calculate the intersection point $S = (x; y; z)$ depending on c. Specify the three components of S separately.
$x =$ [ ]
$y =$ [ ]
$z =$ [ ]

Solution:

The line g is parallel to the plane E if the direction vector $\begin{pmatrix} 0 \\ 4 \\ c \end{pmatrix}$ of the line g is coplanar to the two direction vectors $\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}$ of the plane E. The condition

$$\begin{pmatrix} 0 \\ 4 \\ c \end{pmatrix} = s \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} + t \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}$$

is only satisfied for $c = 1$ since considering the first and the second component separately results in the system of two linear equations

$$s - t = 0 \,, \quad 3s + t = 4$$

in the two variables s and t with the solution $s = t = 1$. Then, for the third component, we must have

$$c = 2 \cdot 1 - 1 = 1 \,.$$

By equating the position vectors of the plane and the line, the intersection point can be calculated:

$$\begin{pmatrix}0\\2\\1\end{pmatrix} + u\begin{pmatrix}0\\4\\c\end{pmatrix} = \begin{pmatrix}0\\2+4u\\1+cu\end{pmatrix} = \begin{pmatrix}8+s-t\\-2+3s+t\\2s-t\end{pmatrix} = \begin{pmatrix}8\\-2\\0\end{pmatrix} + s\begin{pmatrix}1\\3\\2\end{pmatrix} + t\begin{pmatrix}-1\\1\\-1\end{pmatrix} .$$

This vector equation corresponds to a system of three linear equations in the variables s, t, and u with the parameter c. It can be solved using the methods described in Section 4.4. Solving the system, we get

$$u = \frac{10}{1-c} .$$

Thus, substituting this value of u into the equation of the line g results in the position vector of the intersection point S:

$$\overrightarrow{S} = \begin{pmatrix}0\\2\\1\end{pmatrix} + \frac{10}{1-c}\begin{pmatrix}0\\4\\c\end{pmatrix} = \begin{pmatrix}0\\2+\frac{40}{1-c}\\1+\frac{10c}{1-c}\end{pmatrix} .$$

Exercise 10.2.8
Let the line h be given by the equation

$$h\colon \vec{r} = \begin{pmatrix}3\\2\\1\end{pmatrix} + \rho\begin{pmatrix}-8\\9\\1\end{pmatrix} \quad ; \quad \rho \in \mathbb{R} .$$

Find the following:

a. Value of the parameter ρ for which the line h intersects the xy-plane: $\rho =$ [ ]

b. Value of the parameter ρ for which the line h intersects the yz-plane: $\rho =$ [ ]

c. Value of the parameter ρ for which the line h intersects the xz-plane: $\rho =$ [ ]

Solution:

a. In the xy-plane, we have $z = 0$, hence the third component of the position vector of the line h must be equal to zero:

$$1 + \rho = 0 \;\Leftrightarrow\; \rho = -1 .$$

b. In the yz-plane, we have $x = 0$, hence the first component of the position vector of the line h must be equal to zero:

$$3 - 8\rho = 0 \;\Leftrightarrow\; \rho = \frac{3}{8} \; .$$

c. In the xz-plane, we have $y = 0$, hence the second component of the position vector of the line h must be equal to zero:

$$2 + 9\rho = 0 \;\Leftrightarrow\; \rho = -\frac{2}{9} \; .$$

If we consider two planes in space, we find that they can have three different relative positions with respect to each other which correspond to the three different relative positions of two lines described in Section 9.2.3.

These three cases are outlined in the Info Box below.

Info10.2.14

Let the plane E_1 with the reference vector $\vec{a}_1$ and the two direction vectors $\vec{u}_1$ and $\vec{v}_1$ and the plane E_2 with the reference vector $\vec{a}_2$ and the two direction vectors $\vec{u}_2$ and $\vec{v}_2$ be given by the equations

$$E_1 : \vec{r} = \vec{a}_1 + \mu\vec{u}_1 + \nu\vec{v}_1 \; ; \quad \mu, \nu \in \mathbb{R}$$
$$E_2 : \vec{r} = \vec{a}_2 + \rho\vec{u}_2 + \sigma\vec{v}_2 \; ; \quad \rho, \sigma \in \mathbb{R} \; .$$

The planes E_1 and E_2 can have three different possible relative positions with respect to each other:

1. The planes E_1 and E_2 are identical if they have all points in common. This is the case if and only if the three direction vectors $\vec{u}_1$, $\vec{v}_1$, $\vec{u}_2$ and the three direction vectors $\vec{u}_1$, $\vec{v}_1$, $\vec{v}_2$ are coplanar and the reference point of E_1 lies in E_2.

2. The planes E_1 and E_2 are parallel if they do not have any points in common. This is the case if and only if the three direction vectors $\vec{u}_1$, $\vec{v}_1$, $\vec{u}_2$ and the three direction vectors $\vec{u}_1$, $\vec{v}_1$, $\vec{v}_2$ are coplanar and the reference point of E_1 does *not* lie in E_2.

3. The planes E_1 and E_2 intersect if the points they have in common form a line. This is the case if and only if the three direction vectors $\vec{u}_1$, $\vec{v}_1$, $\vec{u}_2$ *or* the three direction vectors $\vec{u}_1$, $\vec{v}_1$, $\vec{v}_2$ are *not* coplanar.

(This figure will be released shortly.)

Of course, in the conditions of the three cases outlined in the Info Box above, the planes can be exchanged; it can also be checked whether the reference vector of E_2 lies in E_1; this makes no difference. If the planes intersect, the set of intersection points can be determined. Sets of intersection points were already discussed in Section 4.3, where the solvability of systems of linear equations in three variables was interpreted geometrically. A sound understanding of this interpretation is now presumed in this Module and a brief repetition of the material presented in Section 4.3 is highly recommended. The example below illustrates how the relative position of two planes is determined.

Example 10.2.15

Let the three planes E, F, and G be given by the equations

$$E\colon \vec{r} = \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix} + a \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + b \begin{pmatrix} 4 \\ 0 \\ -2 \end{pmatrix} \quad ; \quad a, b \in \mathbb{R} \; ,$$

$$F\colon \vec{r} = \begin{pmatrix} 5 \\ 1 \\ 0 \end{pmatrix} + c \begin{pmatrix} 5 \\ -2 \\ -1 \end{pmatrix} + d \begin{pmatrix} -3 \\ -2 \\ 3 \end{pmatrix} \quad ; \quad c, d \in \mathbb{R} \; ,$$

and

$$G\colon \vec{r} = \begin{pmatrix} 5 \\ 0 \\ 1 \end{pmatrix} + x \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} \quad ; \quad x, y \in \mathbb{R} \; .$$

- The planes E and F are parallel. The directions vectors $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 4 \\ 0 \\ -2 \end{pmatrix}$ of E and the first direction vector $\begin{pmatrix} 5 \\ -2 \\ -1 \end{pmatrix}$ of F are coplanar since the condition

  $$a \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + b \begin{pmatrix} 4 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \\ -1 \end{pmatrix}$$

  is satisfied for $a = b = 1$. Likewise, the direction vectors $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 4 \\ 0 \\ -2 \end{pmatrix}$ of E and the second direction vector $\begin{pmatrix} -3 \\ -2 \\ 3 \end{pmatrix}$ of F are coplanar since the condition

  $$a \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + b \begin{pmatrix} 4 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} -3 \\ -2 \\ 3 \end{pmatrix}$$

is satisfied for $a = 1$ and $b = -1$. Moreover, the reference point of F does not lie in E since the condition

$$\begin{pmatrix} 5 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix} + a \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + b \begin{pmatrix} 4 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} a + 4b \\ 2 - 2a \\ -2 + a - 2b \end{pmatrix}$$

cannot be satisfied for any value of a and b. Considering the second component results in $a = \frac{1}{2}$. Substituting this value if a into the equation for the first component results in $b = \frac{9}{8}$. Substituting these two values into the equation for the third component results in the contradiction $0 = -2 + \frac{1}{2} - \frac{9}{4}$. However, choosing another reference point for F, for example the same as for E, would result in an equation that describes a plane identical to the plane E, i.e. another equivalent parametric representation of one and the same plane. For example, the equation

$$F' : \vec{r} = \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix} + \alpha \begin{pmatrix} 5 \\ -2 \\ -1 \end{pmatrix} + \beta \begin{pmatrix} -3 \\ -2 \\ 3 \end{pmatrix} \quad ; \quad \alpha, \beta \in \mathbb{R}$$

represents such a plane.

- The planes E and G intersect. Both direction vectors $\begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix}$ of G are not coplanar to the direction vectors $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 4 \\ 0 \\ -2 \end{pmatrix}$ of E. For the second direction vector of G, we find that the condition

$$\begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} = a \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + b \begin{pmatrix} 4 \\ 0 \\ -2 \end{pmatrix}$$

cannot be satisfied for any value of a and b. Considering the first two components would result in $a = b = 0$. Substituting these values into the equation for the third equation would result in a contradiction. The intersection line of the two planes is calculated by equating the position vectors of the two planes. Here, we have

$$\begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix} + a \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + b \begin{pmatrix} 4 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} a + 4b \\ 2 - 2a \\ -2 + a - 2b \end{pmatrix} = \begin{pmatrix} 5 + x \\ -2x \\ 1 + 3y \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \\ 1 \end{pmatrix} + x \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} .$$

This vector equation corresponds to a system of three linear equations in four variables x, y, a, and b. According to the methods described in Section 4.4, it is solved by taking one variable as a parameter and solving the system for the other variables as functions of this parameter. This left over parameter will be the parameter in the equation of the intersection line in vector form. Which of the variables is taken for the parameter doesn't matter. Here, we use x as parameter. Considering the first two components of the vector equation result in the system of two linear equations

$$a + 4b = 5 + x \text{ and } 2 - 2a = -2x \ ,$$

with the solution $a = 1 + x$, $b = 1$. Substituting this solution into the third component results in

$$-2 + (1 + x) - 2 \cdot 1 = 1 + 3y \ \Leftrightarrow \ y = \frac{1}{3}x - \frac{4}{3} \ .$$

Now, substituting $y = \frac{1}{3}x - \frac{4}{3}$ or $a = 1 + x$ and $b = 1$ into the equation of the plane G or the plane E results – for the same parameter value – in the same parametric representation of the line h, namely the intersection line of the two planes. Specifically, substituting the parameters into the equation of G results in:

$$h\colon \vec{r} = \begin{pmatrix} 5 \\ 0 \\ 1 \end{pmatrix} + x \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} + (\tfrac{1}{3}x - \tfrac{4}{3}) \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \\ -3 \end{pmatrix} + x \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \quad ; \quad x \in \mathbb{R} \ .$$

Exercise 10.2.9
Let the two planes E and F be given by the equations

$$E\colon \vec{r} = \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix} + a \begin{pmatrix} 2 \\ -5 \\ 8 \end{pmatrix} + b \begin{pmatrix} 0 \\ -1 \\ 4 \end{pmatrix} \quad ; \quad a, b \in \mathbb{R}$$

and

$$F\colon \vec{r} = \begin{pmatrix} 5 \\ 0 \\ 0 \end{pmatrix} + c \begin{pmatrix} -2 \\ 3 \\ x \end{pmatrix} + d \begin{pmatrix} 2 \\ y \\ 12 \end{pmatrix} \quad ; \quad c, d \in \mathbb{R} \ ,$$

where the reference point of F does not lie in the plane E.

Find the values of the missing components x and y of F such that the planes F and E are parallel.

$x =$ ____

$y =$ ____

Solution:
The two planes are parallel if the two direction vectors of F are each coplanar to the two direction vectors of E. For the first direction vector of F this results in the condition

$$\begin{pmatrix} -2 \\ 3 \\ x \end{pmatrix} = a \begin{pmatrix} 2 \\ -5 \\ 8 \end{pmatrix} + b \begin{pmatrix} 0 \\ -1 \\ 4 \end{pmatrix} .$$

Considering the first and the second component results in a system of two linear equations with the solution $a = -1$, $b = 2$. Substituting these values into the equation for the third component implies

$$x = -8 + 2 \cdot 4 = 0 .$$

For the second direction vector of F this results in the condition

$$\begin{pmatrix} 2 \\ y \\ 12 \end{pmatrix} = a \begin{pmatrix} 2 \\ -5 \\ 8 \end{pmatrix} + b \begin{pmatrix} 0 \\ -1 \\ 4 \end{pmatrix} .$$

Considering the first and the third components results in a system of two linear equations with the solution $a = b = 1$. Substituting these values into the equation for the second component implies

$$y = -5 - 1 = -6 .$$

Exercise 10.2.10
Let the two planes E and F be given by the equations

$$E\colon \vec{r} = a \begin{pmatrix} 2 \\ -5 \\ 8 \end{pmatrix} + b \begin{pmatrix} 0 \\ -1 \\ 4 \end{pmatrix} \quad ; \quad a, b \in \mathbb{R}$$

and

$$F\colon \vec{r} = c \begin{pmatrix} 0 \\ 3 \\ 4 \end{pmatrix} + d \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \quad ; \quad c, d \in \mathbb{R} .$$

These planes intersect, and the intersection line is given by the equation

$$g\colon \vec{r} = \xi \begin{pmatrix} 4 \\ x \\ y \end{pmatrix} \quad ; \quad \xi \in \mathbb{R} .$$

Find the values of the missing components x and y of the direction vector of the intersection line.

$x =$ [ ]

$y =$ [ ]

Solution:
Equating the position vectors of the two planes results in

$$a \begin{pmatrix} 2 \\ -5 \\ 8 \end{pmatrix} + b \begin{pmatrix} 0 \\ -1 \\ 4 \end{pmatrix} = \begin{pmatrix} 2a \\ -5a - b \\ 8a + 4b \end{pmatrix} = \begin{pmatrix} 2d \\ 3c - d \\ 4c \end{pmatrix} = c \begin{pmatrix} 0 \\ 3 \\ 4 \end{pmatrix} + d \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} .$$

The solution of the corresponding system of linear equations with the parameter d is $a = d$, $b = -\frac{5}{2}d$, $c = -\frac{1}{2}d$. Substituting the condition $c = -\frac{1}{2}d$ into the equation of the plane F results in

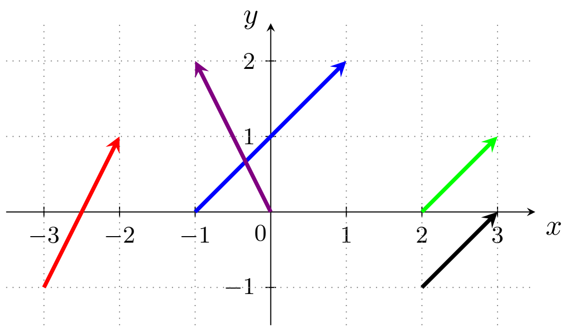
$$-\tfrac{1}{2}d \begin{pmatrix} 0 \\ 3 \\ 4 \end{pmatrix} + d \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} = d \begin{pmatrix} 2 \\ -\frac{5}{2} \\ -2 \end{pmatrix} .$$

Thus, an appropriate direction vector for the intersection lines is $\begin{pmatrix} 2 \\ -\frac{5}{2} \\ -2 \end{pmatrix}$. However, the direction vector in the given parametric equation of g has 4 as its first component, it is coplanar to this direction vector. Hence, the direction vector of the intersection line is $\begin{pmatrix} 4 \\ -5 \\ -4 \end{pmatrix}$, i.e. $x = -5$ and $y = -4$.

10.3 Final Test

10.3.1 Final Test Module 3

Exercise 10.3.1
Specify the vectors that have the arrows shown in the figure below as their representatives.



a. Red vector: ☐ .

b. Purple vector: ☐ .

c. Blue vector: ☐ .

d. Green vector: ☐ .

e. Black vector: ☐ .

Exercise 10.3.2
In still air conditions, a sports aircraft can fly with a velocity of 150 kilometres per hour due south. However, a crosswind blowing from the west with a velocity of 30 kilometres per hour causes the plane to drift. Represent the velocity of the aircraft as the sum of two vectors in the plane, where the second component corresponds to the north-south-direction (positive values for north) and the first component corresponds to the east-west-direction (positive values for east). Drop the unit of measure (kilometres per hour) in your calculation:

a. In still air conditions, the velocity is ☐ .

b. The wind causes an additional velocity of ☐ .

c. The drifting aircraft has in total the velocity vector ☐ .

d. The length of this vector (absolute value of the velocity) is ☐ .

Exercise 10.3.3
Let three points $P = (3;4)$, $Q = (1;0)$, and $R = (-2;1)$ be given in the plane. Calculate the following vectors:

a. $\overrightarrow{PQ}$ = ☐ .

b. $\overrightarrow{QR}$ = ☐ .

c. $\overrightarrow{RR}$ = ☐ .

d. $\overrightarrow{QP}$ = ☐ .

e. $\overrightarrow{RP}$ = ☐ .

Exercise 10.3.4
Let three points $P = (1;2;3)$, $Q = (3;0;0)$, and $R = (-1;2;2)$ be given in space. Calculate the following vectors:

a. $\overrightarrow{PQ}$ = ☐ .

b. $\overrightarrow{RQ}$ = ☐ .

Find the position vector $\overrightarrow{M}$ of the midpoint M of the line segment $\overline{PR}$: $\overrightarrow{M}$ = ☐ .

Exercise 10.3.5
Find the intersection point S of the two lines given by the equations in vector form

$$\vec{r} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + \alpha \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \;;\; \alpha \in \mathbb{R} \text{ and } \vec{r} = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} + \beta \cdot \begin{pmatrix} -2 \\ -4 \\ 4 \end{pmatrix} \;;\; \beta \in \mathbb{R}\,.$$

a. The position vector of the intersection point is $\overrightarrow{S}$ = ☐ .

b. It results from the first equation, for the parameter value α = ☐ .

c. It results from the second equation, for the parameter value β = ☐ .

11 Language of Descriptive Statistics

Module Overview

In this Module we discuss the most important basics of descriptive statistics. In particular, we will discuss rounding and percentage calculations (which actually are not subjects of descriptive statistics, but will be needed there). The competent handling of percentage calculation is essential in economics. Experience tells that these elementary subjects are already taught at secondary school but often not thoroughly enuogh. For example, a test has shown that half of first year students are not able to calculate the amount of VAT in a gross invoice. This module consists of the following sections:

- Terminology: we introduce the fundamental concepts of statistics and explain different rounding methods for numbers.
- Frequency Distributions and Percentage Calculation: we introduce frequency distributions and explain the percentage calculation involved as well as the visualisation of the results using typical types of diagrams.
- Statistical Measures: we explain the fundamental statistical measures of descriptive statistics such as arithmetic mean and sample variance.
- Final Test.

11.1 Terminology and Language

11.1.1 Introduction

For statistical observations (surveys) of appropriately chosen units of observation (a.k.a. units of investigation or experimental units), the values or attributes of a property or properties are determined. Here, a property is a characteristic of the observation unit to be investigated. The terminology of descriptive statistics is as follows:

- The **unit of investigation** (also: **unit of observation**) is the smallest unit on which the observations are made.
- The **characteristic** or **property** is the statistical variable of the unit to be investigated. Characteristics are often denoted by upper-case Latin letters $(X, Y, Z, \ldots)$.
- **Characteristic attributes** or **property values** are values that properties can take. They are often denoted by lower-case Latin letters $(a, b, \ldots, x, y, z, a_1, a_2, \ldots)$.
- The set of units of observation that is investigated with respect to a property of interest is called **universe** or also **population**. It is the set of all possible observation units.
- A **sample** is a "random finite subset" of a certain population of interest. If this set consists of n elements, then this set is called a "sample of size n".
- **Data** are the observed values (attributes) of one or more characteristics or properties of a sample unit of observation of a certain population.
- The **original list** is the protocol that lists the sampled data in chronological order. Thus, the original list is a n-tuple (or vector, written here mostly in coordinate form):
  $$x = (x_1, \ldots, x_n)\,.$$
  This n-tuple is often called a "sample of size n".

Example 11.1.1

From a daily production of components in a factory, $n = 20$ samples of 15 parts each are taken and the number of defective parts in each sample is determined. Here, x_i is the number of defective parts in the ith sample, $i = 1, \ldots, 20$. The original list (sample of size $n = 20$) contains the following data:

$$x = (0, 4, 2, 1, 1, 0, 0, 2, 3, 1, 0, 5, 3, 1, 1, 2, 0, 0, 1, 0)\,.$$

In the second sample, $x_2 = 4$ defective parts were found. The population in this example is the set of all 15-element subsets of the daily production. The property of interest is in this case

$$X = \text{Number of defective workpieces in a sample of 15 elements}\,.$$

Info11.1.2

The variables in a statistical observation are called **characteristics** or **properties**. Values that the properties can take are called **property values** or **characteristic attributes**.

Properties are roughly classified into qualitative properties (that can be ascertained in a descriptive way) and quantitative properties (that can naturally be ascertained numerically):

- **Qualitative properties**:
  - Nominal properties: attributes classified according to purely qualitative aspects. Examples: skin colour, nationality, blood type.
  - Ordinal properties: attributes with a natural hierarchy, i.e. they can be ordered or sorted. Examples: grades, ranks, surnames.
- **Quantitative properties**:
  - Discrete properties: property values are isolated values (e.g. integers). Examples: numbers, years, age in years.
  - Continuous properties: property values can (at least in principle) take any value. Examples: body size, weight, length.

The transition between continuous and discrete properties is partly fluid, once we consider the possibility of rounding.

11.1.2 Rounding

The **rounding** of measurement values is an everyday process.

Info11.1.3

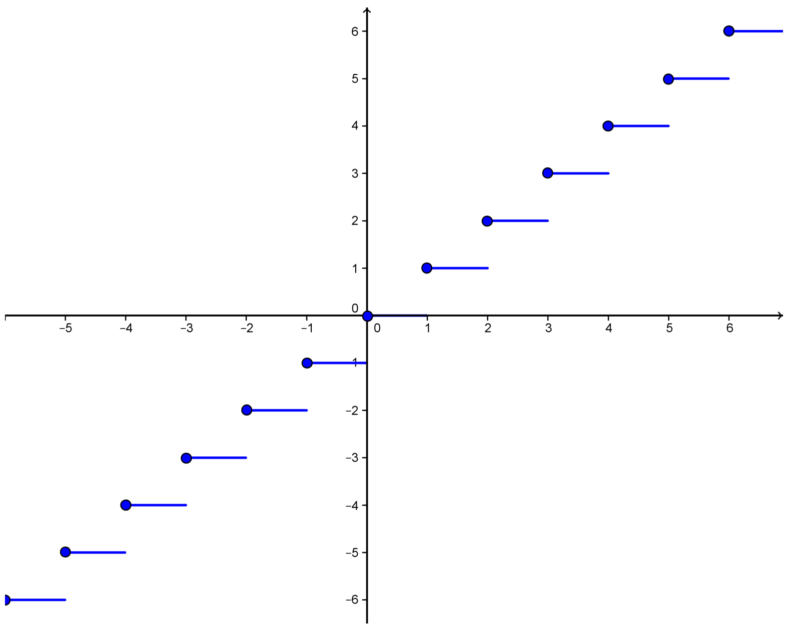
In principle, there are three ways of rounding:

- Rounding (off) using the floor function $\lfloor x \rfloor$.
- Rounding (up) using the ceil function $\lceil x \rceil$.
- Rounding using the round function (sometimes also called rnd function).

The **floor function** is defined as

$$\mathsf{floor}: \mathbb{R} \longrightarrow \mathbb{R} \ , \ x \longmapsto \mathsf{floor}(x) = \lfloor x \rfloor = \max\{k \in \mathbb{Z} : k \leq x\} .$$

If $x \in \mathbb{R}$ is a real number, then $\mathsf{floor}(x) = \lfloor x \rfloor$ is the largest integer that is smaller than or equal to x. It results from rounding off the value of x. If a positive real number x is written as a decimal, then $\lfloor x \rfloor$ equals the integer on the left of the decimal point: rounding (off) cuts off the digits on the right of the decimal point. For example $\lfloor 3.142 \rfloor = 3$ but $\lfloor -2.124 \rfloor = -3$. The floor function is a step function with jumps (in more mathematical terms, jump discontinuities) of height 1 at all points $x \in \mathbb{Z}$. The function values at the jumps always lie a step up. They are indicated by the small circles in the figure below, which shows the graph of the floor function.



Graph of the floor function

Let a real number $a \geq 0$ be given, written as a decimal number

$$a = g_n g_{n-1} \ldots g_1 g_0 . a_1 a_2 a_3 \ldots$$

This number a can be rounded to r fractional digits ($r \in \mathbb{N}_0$) using the floor function by

$$\tilde{a} = \frac{1}{10^r} \cdot \lfloor 10^r \cdot a \rfloor .$$

This process of rounding cuts off the decimal after the rth fractional digit. Thus, rounding using the floor function is in general a rounding off.

Example 11.1.4

Rounding the number $a_1 = 2.3727$ to 2 fractional digits using the floor function results in

$$\tilde{a}_1 = \frac{1}{10^2} \cdot \left\lfloor 10^2 \cdot 2.3727 \right\rfloor = \frac{1}{10^2} \cdot \lfloor 237.27 \rfloor = \frac{1}{10^2} \cdot 237 == 2.37 .$$

Alternatively, it can be rounded by cutting off the decimal after the second fractional digit (however, this is only possible if the number is given as a decimal which is rarely the case in a computer program).

Rounding the number $a_2 = \sqrt{2} = 1.414213562\ldots$ to 4 fractional digits using the floor function results in

$$\tilde{a}_2 = \frac{1}{10^4} \cdot \left\lfloor 10^4 \cdot \sqrt{2} \right\rfloor = \frac{1}{10^4} \cdot \lfloor 14142.1\ldots \rfloor = \frac{1}{10^4} \cdot 14142 = 1.4142 .$$

Rounding the number

$$a_3 = \pi = 3,141592654\ldots$$

to 2 fractional digits using the floor function results in

$$\tilde{a}_3 = \frac{1}{10^2} \cdot \left\lfloor 10^2 \cdot \pi \right\rfloor = \frac{1}{10^2} \cdot \lfloor 314.159\ldots \rfloor = \frac{1}{10^2} \cdot 314 = 3.14 .$$

The rounding method using the floor function is often applied for calculating final grades in certificates ("academic rounding"). If a mathematics student has the individual grades

Subject	Grade
Mathematics 1	1.3
Mathematics 2	2.3
Mathematics 3	2.0

then the arithmetic mean of these grades is calculated by

$$\frac{1.3 + 2.3 + 2.0}{3} = \frac{5.6}{3} = 1.8\overline{6} \,.$$

Rounding to the first fractional digit using the floor function would result in the final grade of $\tilde{a} = 1.8$. The rounding methods for calculating final grades always have to be described exactly in the examination regulations.

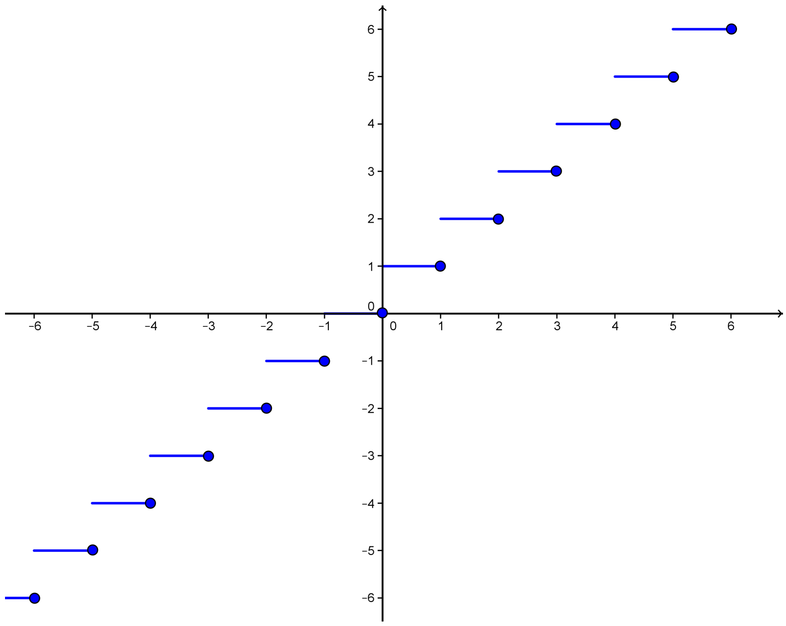
The counterpart to the floor function is the ceil (a.k.a. ceiling) function:

Info11.1.5

The **ceil function** is defined as

$$\mathsf{ceil} : \mathbb{R} \longrightarrow \mathbb{R} \ , \ x \longmapsto \mathsf{ceil}(x) = \lceil x \rceil = \min\{k \in \mathbb{Z} \,:\, k \geq x\} \,.$$

If $x \in \mathbb{R}$ is a real number, then $\mathsf{ceil}(x) = \lceil x \rceil$ is the smallest integer that is greater than or equal to x. The ceil function is a step function with jumps (jump discontinuities) of height 1 at all points $x \in \mathbb{Z}$. The function values at the jumps always lie at the bottom. They are indicated by the small circles in the figure below showing the graph of the ceil function.



Graph of the ceil function

Let a real number $a \geq 0$ be given as a decimal number

$$a = g_n\, g_{n-1} \, \ldots \, g_1\, g_0 \, . \, a_1\, a_2\, a_3 \, \ldots$$

This number a can be rounded to r fractional digits ($r \in \mathbb{N}_0$) using the ceil function by

$$\hat{a} = \frac{1}{10^r} \cdot \lceil 10^r \cdot a \rceil \, .$$

Rounding using the ceil function is in general a rounding up to the next decimal digit.

Example 11.1.6

Rounding the number $a_1 = 2.3727$ to 2 fractional digits using the ceil function results in

$$\hat{a}_1 = \frac{1}{10^2} \cdot \left\lceil 10^2 \cdot 2.3727 \right\rceil = \frac{1}{10^2} \cdot \lceil 237.27 \rceil = \frac{1}{10^2} \cdot 238 = 2.38 \, .$$

Analogously, rounding the number $a_2 = \sqrt{2} = 1.414213562\ldots$ to 4 fractional digits using the ceil function results in

$$\hat{a}_2 = \frac{1}{10^4} \cdot \left\lceil 10^4 \cdot \sqrt{2} \right\rceil = \frac{1}{10^4} \cdot \lceil 14142.1\ldots \rceil = \frac{1}{10^4} \cdot 14143 = 1.4143 \, .$$

Rounding the number $a_3 = \pi = 3.141592654\ldots$ to 2 fractional digits using the ceil function results in

$$\hat{a}_3 = \frac{1}{10^2} \cdot \left\lceil 10^2 \cdot \pi \right\rceil = \frac{1}{10^2} \cdot \lceil 314.15\ldots \rceil = \frac{1}{10^2} \cdot 315 = 3.15 \, .$$

The rounding method using the ceil function is often applied, for example, in craftsmen's invoices. A craftsman is mostly paid by the hour. If a repair takes 50 minutes (i.e. $0.8\overline{3}$ hours as a decimal), then a craftsmen will round up and invoice a full working hour. Colloquially, rounding mostly means mathematical rounding:

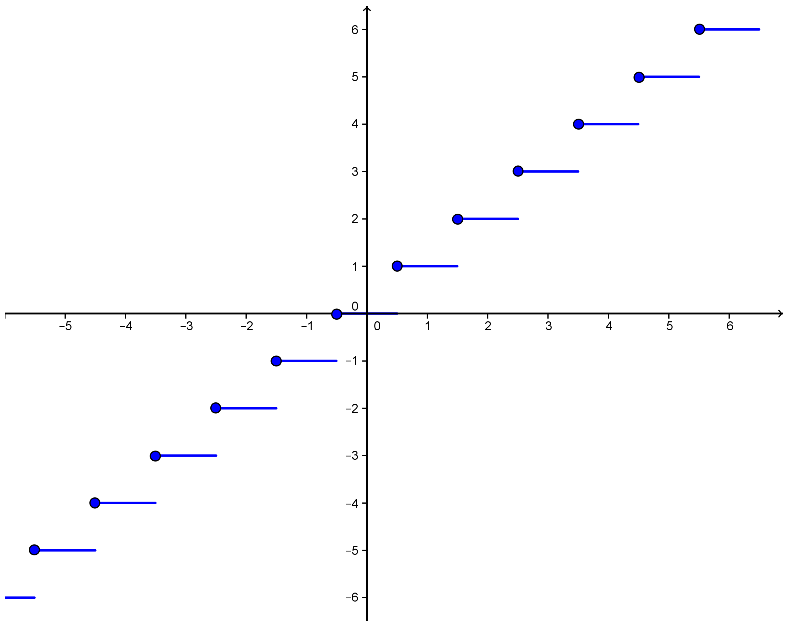
Info11.1.7

The **round function** (or mathematical rounding) is defined as

$$\mathsf{round} : \mathbb{R} \longrightarrow \mathbb{R} \ , \ x \longmapsto \mathsf{round}(x) = \mathsf{floor}\left(x + \frac{1}{2}\right) = \left\lfloor x + \frac{1}{2} \right\rfloor \ .$$

In contrast to rounding up or rounding off, the maximum change to the number by this rounding is 0.5.

The round function is a step function with jumps (jump discontinuities) of height 1 at all points $x + \frac{1}{2},\ x \in \mathbb{Z}$. The function values at the jumps always lie a step up. They are indicated by the small circles in the figure below showing the graph of the round function.



Graph of the round function

Let a real number $a \geq 0$ be given as a decimal number

$$a = g_n\, g_{n-1}\, \ldots\, g_1\, g_0\, .\, a_1\, a_2\, a_3\, \ldots$$

This number a can be rounded to r fractional digits ($r \in \mathbb{N}_0$) using the round function:

$$\overline{a} = \frac{1}{10^r} \cdot \mathsf{round}(10^r \cdot a) = \frac{1}{10^r} \cdot \left\lfloor 10^r \cdot a + \frac{1}{2} \right\rfloor .$$

This rounding method is called mathematical rounding and corresponds to the “normal” rounding process.

Example 11.1.8

The number $a_1 = 1.49$ is rounded to one fractional digit using the round function to

$$\begin{aligned} \overline{a}_1 &= \frac{1}{10} \cdot \mathsf{round}(10 \cdot 1.49) = \frac{1}{10} \cdot \lfloor 10 \cdot 1.49 + 0.5 \rfloor \\ &= \frac{1}{10} \cdot \lfloor 14.9 + 0.5 \rfloor = \frac{1}{10} \cdot \lfloor 15.4 \rfloor = \frac{1}{10} \cdot 15 = 1.5 . \end{aligned}$$

The number $a_2 = 1.52$ is rounded to one fractional digit using the round function to

$$\begin{aligned} \overline{a}_2 &= \frac{1}{10} \cdot \mathsf{round}(10 \cdot 1.52) = \frac{1}{10} \cdot \lfloor 10 \cdot 1.52 + 0.5 \rfloor \\ &= \frac{1}{10} \cdot \lfloor 15.2 + 0.5 \rfloor = \frac{1}{10} \cdot \lfloor 15.7 \rfloor = \frac{1}{10} \cdot 15 = 1.5 . \end{aligned}$$

The number $a_3 = 2.3727$ is rounded to two fractional digits using the round function to

$$\begin{aligned}\overline{a}_3 &= \frac{1}{10^2}\cdot \mathsf{round}(10^2\cdot 2.3727) = \frac{1}{100}\cdot\lfloor 100\cdot 2.3727+0.5\rfloor\\ &= \frac{1}{100}\cdot\lfloor 237.27+0.5\rfloor = \frac{1}{100}\cdot\lfloor 237.77\rfloor = \frac{1}{100}\cdot 237 = 2.37\,.\end{aligned}$$

The number $a_4 = \sqrt{2} = 1.414213562\ldots$ is rounded to seven fractional digits using the round function to

$$\begin{aligned}\overline{a}_3 &= \frac{1}{10^7}\cdot \mathsf{round}(10^7\cdot \sqrt{2}) = \frac{1}{10^7}\cdot\left\lfloor 10^7\cdot 1.414213562\ldots+0.5\right\rfloor\\ &= \frac{1}{10^7}\cdot\lfloor 14142135.62\ldots+0.5\rfloor = \frac{1}{10^7}\cdot\lfloor 14142136.12\ldots\rfloor\\ &= \frac{1}{10^7}\cdot 14142136 = 1.4142136\,.\end{aligned}$$

Exercise 11.1.1
Using the round function, round the number $\pi = 3.141592654\ldots$ to four fractional digits:
$\overline{\pi} =$ [ ] .

Solution:

$$\begin{aligned}\overline{\pi} &= \frac{1}{10^4}\cdot \mathsf{round}(10^4\cdot \pi) = \frac{1}{10^4}\cdot\left\lfloor 10^4\cdot 3.141592654\ldots+0.5\right\rfloor\\ &= \frac{1}{10^4}\cdot\lfloor 31415.92654\ldots+0.5\rfloor = \frac{1}{10^4}\cdot\lfloor 31416.42654\ldots\rfloor\\ &= \frac{1}{10^4}\cdot 31416 = 3.1416\,.\end{aligned}$$

Exercise 11.1.2
Let the numbers

$$a = \frac{47}{17} \quad\text{and}\quad b = 3.7861$$

be given.

a. Round each of the numbers a and b to 2 fractional digits using the floor function. The roundings result in $\tilde{a} =$ [ ] and $\tilde{b} =$ [ ] .

b. Round each of the numbers a and b to 2 fractional digits using the ceil function.

The roundings result in $\hat{a}$ = [ ] and $\hat{b}$ = [ ] .

c. Round each of the numbers a and b to 2 fractional digits using the round function. The roundings result in $\overline{a}$ = [ ] and $\overline{b}$ = [ ] .

Solution:
First, we transform the fraction into an appropriate decimal fraction by dividing successively with remainder and substituting the results of the division as digits into the decimal:

$$\frac{47}{17} = 2 + \frac{13}{17}.$$

The digit left to the decimal point is 2. We have

$$2 + \frac{13}{17} = 2 + \frac{1}{10} \cdot \frac{130}{17} = 2 + \frac{1}{10} \cdot \left(7 + \frac{11}{17}\right).$$

The first fractional digit is 7. Proceeding further this way results in $a = 2.764705\ldots$, we only need three fractional digits for the required rounding. The complete calculation using the floor function results in

$$\tilde{a} = \frac{1}{10^2} \cdot \lfloor 10^2 \cdot 2.764705\ldots \rfloor = \frac{1}{10^2} \cdot \lfloor 276.4705\ldots \rfloor = 2.76\,.$$

This result can be obtained more quickly by simply cutting off the decimal after the second decimal digit:

$$\tilde{a} = 2.76 \text{ and } \tilde{b} = 3.78\,.$$

However, this is only allowed in this exercise since a and b are non-negative. A simple rounding up after the second decimal digit or a complete calculation using the ceil function results in

$$\hat{a} = 2.77 \text{ and } \hat{b} = 3.79\,.$$

The results for the mathematical rounding are obtained either by a complete calculation as in the examples above, or by rounding to the decimal with two fractional digits with the smallest difference from the original number:

$$\overline{a} = 2.76 \text{ and } \overline{b} = 3.79\,.$$

11.1.3 Remarks on the Rounding Processes

As the following considerations and examples will show, we have to be very careful when calculating with rounded results. Let us consider the set $M = \mathbb{R}_{\geq 0}$ of all non-negative real numbers. On this set, let us define the multiplication

$$M \times M \longrightarrow M \ , \ (a,b) \longmapsto a \odot b$$

by the calculation rule

$$a \odot b \;=\; \frac{1}{10^2} \cdot \mathsf{round}(10^2 \cdot a \cdot b) \;=\; \frac{1}{10^2} \cdot \left\lfloor 10^2 \cdot a \cdot b + \frac{1}{2} \right\rfloor \;,$$

i.e. the product $a \odot b$ is calculated by calculating the usual product $a \cdot b$ first and subsequently rounding the result mathematically to two fractional digits.

The law of associativity no longer applies to the rounded multiplication. For the numbers $a = 2.11$, $b = 3.35$, and $c = 2,61$, we have, for example,

$$a \odot b \;=\; 2.11 \odot 3.35 \;=\; 7.07 \text{ and } (a \odot b) \odot c \;=\; 7.07 \odot 2.61 \;=\; 18.45\,.$$

Changing the brackets, however, results in

$$b \odot c \;=\; 3.35 \odot 2.61 \;=\; 8.74 \;\text{ and }\; a \odot (b \odot c) \;=\; 2.11 \odot 8.74 \;=\; 18.44\,.$$

Info11.1.9

Since calculators (and computers) always calculate using rounded results, this means that the law of associativity does not apply unrestrictedly to multiplication on calculators.

Likewise, false results can be caused by careless rounding. Let us consider the numbers $a = 4.98$ and $b = 1.001$. Then, we have

$$a \cdot b \;=\; 4.98 \cdot 1.001 \;=\; 4.98498 \;\text{, i.e. } a \odot b \;=\; 4.98 \odot 1.001 \;=\; 4.98 \;=\; a\,.$$

Furthermore, we have

$$a \cdot b^{1000} \;=\; 4.98 \cdot 1.001^{1000} \;=\; a \cdot \underbrace{b \cdot \ldots \cdot b}_{1000 \text{ factors}} \approx 13.53028118\,.$$

Rounding after each multiplication to 2 fractional digits results (due to $a \odot b = a$) in the wrong result:

$$(\ldots((a \odot \underbrace{b) \odot b) \ldots \odot b}_{1000 \text{ factors}}) \;=\; a \;=\; 4.98\,.$$

Info11.1.10

In practice, this means that you must calculate at least with double precision (twice the required digits) and round the result only finally to the required digits.

11.2 Frequency Distributions and Percentage Calculation

11.2.1 Introduction

Let X be a given property. A sample of size n resulted in the original list (sample)

$$x = (x_1, x_2, \ldots, x_n) .$$

Info11.2.1

If a is a possible property value, then

$$H_x(a) = \text{number of } x_j \text{ within the original list } x \text{ with } x_j = a$$

is called the **absolute frequency** of the property a in the original list $x = (x_1, x_2, \ldots, x_n)$.

If $a_1, a_2, \ldots a_k$ are the possible property values in the original list $x = (x_1, x_2, \ldots, x_n)$, then we have

$$H_x(a_1) + H_x(a_2) + \ldots + H_x(a_k) = n$$

or in words: each of the n values is counted by exactly one of the frequencies.

Info11.2.2

The **relative frequency** of the property value a in the original list $x = (x_1, x_2, \ldots, x_n)$ is defined by

$$h_x(a) = \frac{1}{n} \cdot H_x(a) .$$

If $a_1, a_2, \ldots a_k$ are the possible property values in the original list $x = (x_1, x_2, \ldots, x_n)$, then we have

$$h_x(a_1) + h_x(a_2) + \ldots + h_x(a_k) = 1 .$$

Relative frequencies always lie in the interval $[0; 1]$ and are often specified in percentages,e.g. $h_x(a_1) = 34\%$ instead of $h_x(a_1) = 0.34$.

Info11.2.3

Collecting the absolute or relative frequencies of all occurring (or possible) property values in the original list (sample) $x = (x_1, x_2, \dots, x_n)$ in a table results in the **empirical frequency distribution**.

Example 11.2.4

In a data centre, the processing time (in seconds, rounded to one fractional digit) of 20 program jobs was determined. This resulted in the following original list of a sample of size $n = 20$:

3.9	3.3	4.6	4.0	3.8
3.8	3.6	4.6	4.0	3.9
3.9	3.9	4.1	3.7	3.6
4.6	4.0	4.0	3.8	4.1

The smallest value is 3.3 s, the largest value is 4.6 s, the increment is 0.1 s. Thus, we have the empirical frequency distribution listed (in tabular form) below. To keep the table short all values less than 3.3 and greater that 4.6 are not listed.

Result a	$H_x(a)$	$h_x(a)$	Percentage
3.3	1	$\frac{1}{20} = 0.05$	5%
3.4	0	0	0%
3.5	0	0	0%
3.6	2	$\frac{2}{20} = 0.1$	10%
3.7	1	$\frac{1}{20} = 0.05$	5%
3.8	3	$\frac{3}{20} = 0.15$	15%
3.9	4	$\frac{4}{20} = 0.2$	20%
4.0	4	$\frac{4}{20} = 0.2$	20%
4.1	2	$\frac{2}{20} = 0.1$	10%
4.2	0	0	0%
4.3	0	0	0%
4.4	0	0	0%
4.5	0	0	0%
4.6	3	$\frac{3}{20} = 0.15$	15%
Sum	20	1	100%

11.2.2 Percentage Calculation

In descriptive statistics, numerical values are often specified in percentages, so we will review the most relevant elements of percentage calculations in this section. Numbers given as percentages ("percent, hundredth") serve to illustrate ratios and to make them comparable by putting the numbers into relation to a unified base value (hundred).

Info11.2.5

Let $a \geq 0$ be a real number. Then, we have $a\,\% = \frac{a}{100}$, i.e. the symbol % can be interpreted as "divided by 100" (just as the symbol $\circ$ with respect to angles was interpreted in Module 5 as a multiplication by $\frac{\pi}{180}$).

For example:

- One percent is one hundredth: $1\,\% = \frac{1}{100} = 0.01$
- Ten percent is one tenth: $10\,\% = \frac{10}{100} = 0.1$
- 25 percent is one quarter: $25\,\% = \frac{25}{100} = 0.25$
- One hundred percent is a whole: $100\,\% = \frac{100}{100} = 1$
- 150 percent is a factor of 1.5: $150\,\% = \frac{150}{100} = 1.5$

In general, percentages describe ratios and relate to a certain base value. The base value is the initial value the percentage relates to. The percentage is expressed in percent and denotes a ratio with respect to the base value. The real value of this quantity is called the percent value. The percent value has the same unit as the base value.

Info11.2.6

The rule of three applies for the percent value, base value and percentage :

$$\text{percentage} \cdot \text{base value} = \text{percent value}\,.$$

11.2.3 Calculation of Interest

In the calculation of interest, we distinguish between simple interest and compound interest. For simple interest, the interest is paid at the end of the interest period. For compound interest, the interest is also paid on the previously accumulated interest.

Info11.2.7

When simple interest is applied, a quantity K that increases by $p\,\%$ every year will increase after t years ($t \in \mathbb{N}$) to

$$K_t \;=\; K \cdot \left(1 + t \cdot \frac{p}{100}\right).$$

Note that p itself can be a decimal, for example, for $p = 2.5$, the percent value is $2.5\% = 0.025$.

Exercise 11.2.1
What is the final capital for an initial capital of $K = 4{,}000$ EUR after an interest period of $t = 10$ years, when simple interest is applied at a rate of $p = 2.5\%$ p. a.?

Answer: K_{10} = ☐ EUR.

Solution:
Substituting the corresponding values into the interest formula for simple interest results in

$$K_{10} \;=\; K \cdot (1 + 10 \cdot 0.025) = 4{,}000 \text{ EUR} \cdot 1.25 \;=\; 5{,}000 \text{ EUR}\,.$$

Exercise 11.2.2
What is the initial capital K that had been deposited at the $1st$ of January, 2000 to get paid an end capital of $K_{12} = 10{,}000$ EUR at the $31st$ of December, 2011 when simple interest is applied at a rate of $p = 5\%$ p. a.?

Answer: K = ☐ EUR.

Solution:
Substituting the corresponding values into the interest formula for simple interest results in

$$K_{12} \;=\; 10{,}000 \text{ EUR} \;=\; K \cdot (1 + 12 \cdot 0.05) \;=\; K \cdot 1.6\,.$$

Solving this equation for K results in

$$K \;=\; \frac{10{,}000 \text{ EUR}}{1.6} \;=\; \frac{10{,}0000 \text{ EUR}}{16} \;=\; 6{,}250 \text{ EUR after cancelling the fraction}\,.$$

While simple**interest** is simply paid after an interest period, compound interest carries over into the next interest period, i.e. the interest will be added to the initial capital or will be **capitalised**:

Example 11.2.8

For a bank account at the end of an interest period, an initial capital of 1,000 EUR is deposited at an interest rate of 8%. After one year the deposit (in EUR) in the bank account is

- $1{,}000 + \frac{1{,}000 \cdot 8}{100} = 1{,}000 \cdot \left(1 + \frac{8}{100}\right) = 1{,}000 \cdot 1.08 = 1{,}080.$
- This deposit is invested for an additional year at the same interest rate of 8%. Then, the deposit (in EUR) after two years is $1{,}080 \cdot 1.08 = 1{,}000 \cdot 1.08^2 = 1{,}000 \cdot \left(1 + \frac{8}{100}\right)^2$.
- The deposit increases by a factor of 1.08 per year. Hence, the deposit (in EUR) after t years ($t \in \mathbb{N}_0$) is

$$1{,}000 \cdot 1.08^t \;=\; 1{,}000 \cdot \left(1 + \frac{8}{100}\right)^t\,.$$

Thus, the compound interest is based on the following formula:

Info11.2.9

A quantity K that increases every year by an amount of $p\,\%$ will have been increased after t years ($t \in \mathbb{N}_0$) to

$$K \cdot \left(1 + \frac{p}{100}\right)^t .$$

Here, $1 + \frac{p}{100}$ is called the growth factor for a growth of $p\,\%$.

In an advert offering deposit accounts or loans, the interest is usually given as a rate per year, even if the actual interest period differs. This interest period is the time between two successive dates at which the interest payments are due. On a deposit account the interest period is one year, though it is becoming more common to offer other interest periods. For example, for short-term loans the interest is paid daily or monthly.

If a bank offers a yearly interest rate of $9\,\%$ with monthly interest payments, then at the end of each month $\left(\frac{1}{12}\right) \cdot 9\% = 0,75\%$ of the capital will be credited.

Info11.2.10

The yearly rate is divided by the number of interest periods to obtain the periodic rate (the interest rate per period).

Suppose an investment of S_0 EUR yields $p\,\%$ interest per interest period. After t periods ($t \in \mathbb{N}_0$) the investment will have been increased to

$$S_t \;=\; S_0 \cdot (1+r)^t \;\text{ with }\; r \;=\; \frac{p}{100} .$$

In every period the investment increases by a factor of $1 + r$, and we say “the **interest rate** equals $p\,\%$” or “the **periodic rate** equals r”. Suppose interest is credited at a rate of $\frac{p}{n}\,\%$ to the capital at n different times, more or less evenly distributed over the year. Then, the capital is multiplied by a factor of

$$\left(1 + \frac{r}{n}\right)^n$$

every year. After t years the capital has increased to

$$S_0 \cdot \left(1 + \frac{r}{n}\right)^{n \cdot t} .$$

Example 11.2.11

A capital of 5,000 EUR is deposited for $t = 8$ years in a bank account at a yearly interest rate of 9 %, where the interest is paid quarterly. The periodic rate $\frac{r}{n}$ here is

$$\frac{r}{n} = \frac{0.09}{4} = 0.0225\,,$$

and for the number of periods $n \cdot t$ we have $n \cdot t = 4 \cdot 8 = 32$. Thus, after $t = 8$ years the deposit has increased to

$$5000 \cdot (1 + 0.0225)^{32} \approx 10190.52 \text{ EUR}\,.$$

Exercise 11.2.3
A capital of $K_0 = 8{,}750$ EUR is deposited for $t = 4$ years at an interest rate of $p = 3,5\%$ p. a., and the interest is capitalised.

a. After one year the amount of capital is $K_1 =$ [ ] .

b. After two years the amount of capital is $K_2 =$ [ ] .

c. After three years the amount of capital is $K_3 =$ [ ] .

d. The final capital is $K_4 =$ [ ] .

Specify all values rounded mathematically to the second fractional digit. Round only *after* carrying out the calculations. For these calculations, you are allowed to use a calculator.

Solution:
Substituting the values into the compound interest formula results in

$$\begin{aligned}
K_1 &= K_0 \cdot (1 + 0.035)^1 = 9056.25 \\
K_2 &= K_0 \cdot (1 + 0.035)^2 = 9373.22 \text{ (rounded)} \\
K_3 &= K_0 \cdot (1 + 0.035)^3 = 9701.28 \text{ (rounded)} \\
K_4 &= K_0 \cdot (1 + 0.035)^4 = 10040.83 \text{ (rounded)}.
\end{aligned}$$

The above mentioned error propagation for rounding requires calculating the power first and then rounding. For example, it is wrong to multiply the rounded value K_3 by 1.035 to obtain K_4.

A consumer who wants to take out a loan always faces several offers from competing banks. Thus, it is extremely useful to compare the different offers.

Example 11.2.12

Let us consider an offer providing an yearly interest rate of 9%, where the interest is charged at a monthly (12 times per year) rate of 0.75%. If no interest is paid off in the meantime, the initial debt will increase to

$$S_0 \cdot \left(1 + \frac{0.09}{12}\right)^{12} \approx S_0 \cdot 1.094$$

after one year. The interest to be paid off is approximately

$$1.094 \cdot S_0 - S_0 = 0.094 \cdot S_0 .$$

As long as no interest is paid off, the debt will increase at a constant rate which is approximately 9.4% per year. This is why we may speak of an "effective" annual interest rate. In the example above the effective annual interest rate is 9.4%.

Info11.2.13

If interest is paid n times per year at a periodic rate of $\frac{r}{n}$ per period, then the **effective annual interest rate** R is defined by

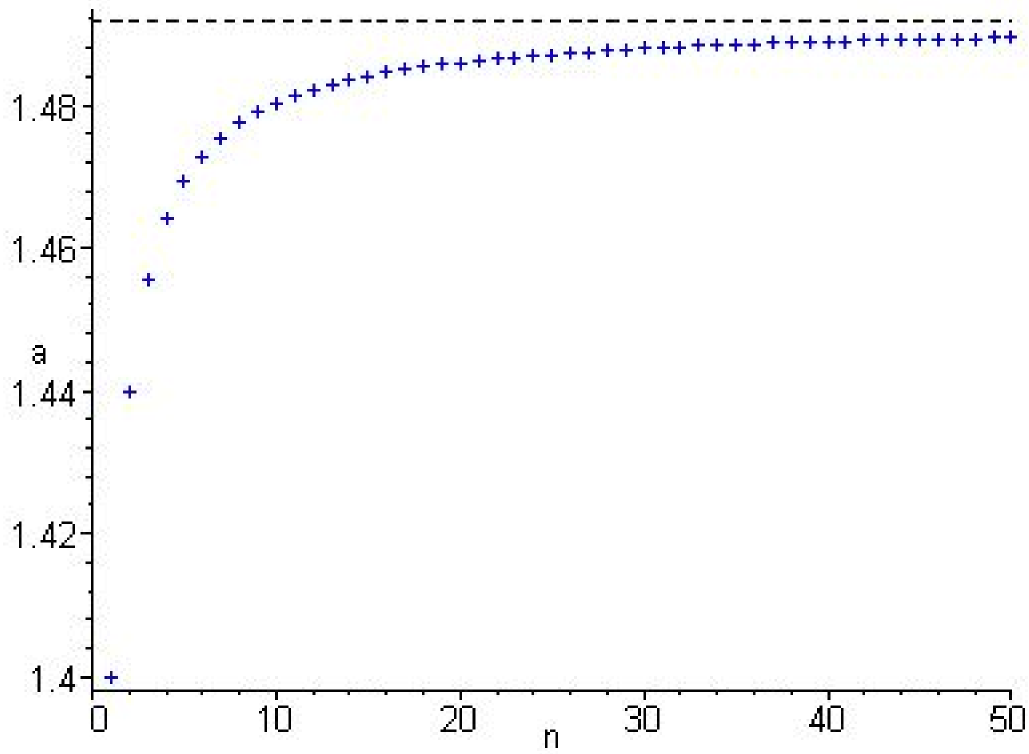
$$R = \left(1 + \frac{r}{n}\right)^n - 1 .$$

11.2.4 Continuous Compounding Interest

The expression $a_n = \left(1 + \frac{r}{n}\right)^n$ with $r \in \mathbb{R}$ can also be interpreted as a map depending on $n \in \mathbb{N}$

$$a : \mathbb{N} \longrightarrow \mathbb{R} \ , \ n \longmapsto a(n) = a_n = \left(1 + \frac{r}{n}\right)^n .$$

A map $\mathbb{N} \ni n \mapsto a_n \in \mathbb{R}$ is called a real **sequence**. The pairs (n, a_n) can be interpreted as points in the Euclidean plane. In this sense, the sequence $a_n = \left(1 + \frac{0.4}{n}\right)^n$ is shown in the figure below as a sequence of points in the Euclidean plane.



Two properties of this sequence can immediately be seen from the figure above:

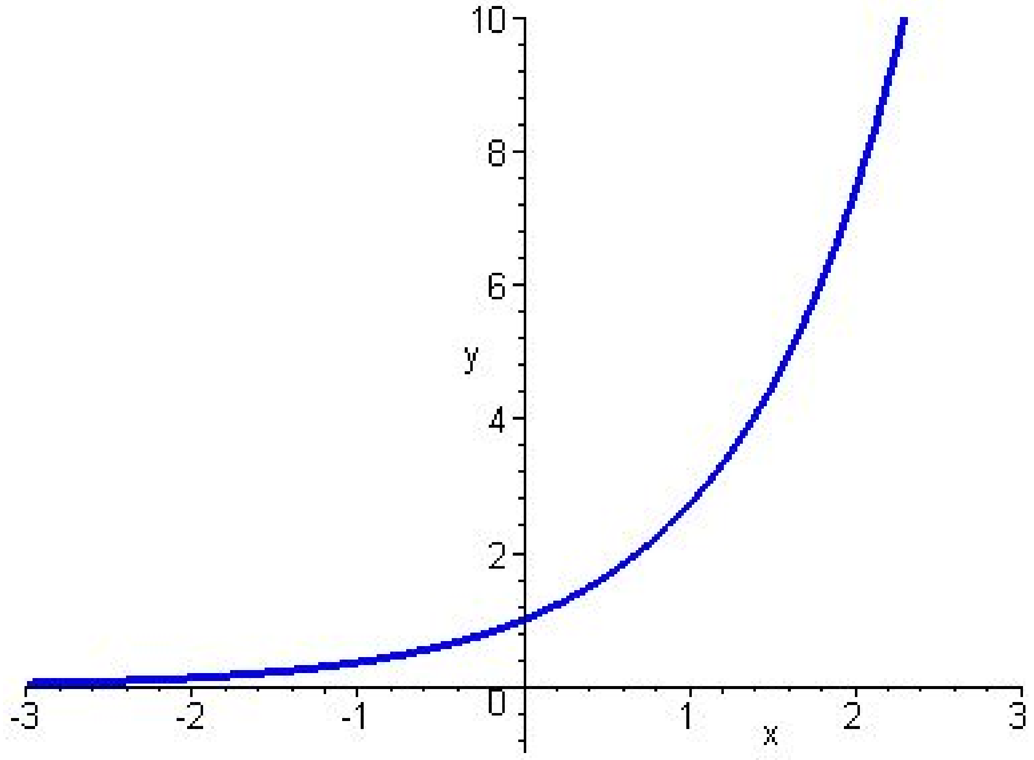
- The sequence a_n, $n \in \mathbb{N}$ is monotonically increasing, i.e. for $i \leq j$ $a_i \leq a_j$, for all $i, j \in \mathbb{N}$.
- The sequence approaches the value $a \in \mathbb{R}$ as $n \in \mathbb{N}$ increases. This number a is called limit of the sequence a_n, and is written

$$\lim_{n \to \infty} a_n = a \,.$$

In the lecture mathematics 1, the natural exponential function

$$\exp : \mathbb{R} \longrightarrow \mathbb{R} \quad , \quad x \longmapsto \exp(x) = e^x$$

will be studied in detail.



The natural exponential function

There, the following statement will be shown:

Info11.2.14

For an arbitrary number $x \in \mathbb{R}$, we have

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n \;=\; e^x\;.$$

For $x = 1$, the limit of this sequence is Euler's number (named after the Swiss mathematician Leonhard Euler, 1707–1783):

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \;=\; e \;\approx\; 2.7182\ldots\;.$$

It can be shown (with some difficulty) that Euler's number e is an irrational number, and hence it cannot be written as a fraction.

The exponent rules apply to the natural exponential function with arbitrary real numbers as its exponents:

- $\exp(x+y) = e^{x+y} = e^x \cdot e^y = \exp(x) \cdot \exp(y)$ for $x, y \in \mathbb{R}$.
- $\exp(x \cdot y) = e^{x \cdot y} = (e^x)^y = (e^y)^x$ for $x, y \in \mathbb{R}$.

Information on the compound interest process can be gained if the number of times n gets very large using the exponential function and the relation to the sequence $(1+\frac{x}{n})^n$: the capital is multiplied by a factor of $\left(1+\frac{r}{n}\right)^n$ every year if the interest at a rate of $\frac{r}{n}$ is credited to the initial capital S_0 at n different times in the year. After t years, $t \in \mathbb{N}$, the initial capital has increased to

$$S_0 \cdot \left(1+\frac{r}{n}\right)^{n \cdot t} .$$

If $n \to \infty$, the limit of this sequence is

$$\lim_{n\to\infty} \left(S_0 \cdot \left(1+\frac{r}{n}\right)^{n \cdot t}\right) = S_0 \cdot e^{r \cdot t} .$$

For increasing $n \in \mathbb{N}$ the interest is paid more and more frequently:

Info11.2.15

The limiting case is called the **continuous compounding interest**. For positive real numbers t, the formula

$$s(t) = S_0 \cdot e^{r \cdot t}$$

specifies to which amount an initial capital S_0 has increased after t years if continuous compounding interest is applied at a rate r per year.

Example 11.2.16

An investment of 5,000 EUR is deposited for $t = 8$ years in a bank account where continuous compounding interest is applied at a yearly interest rate of 9 %. After $t = 8$ years, this results in an investment of

$$5{,}000 \cdot e^{0.09 \cdot 8} = 5{,}000 \cdot e^{0.72} \approx 10{,}272.17 \text{ EUR} .$$

11.2.5 Types of Diagrams

Qualitative and quantitative discrete data gained from a sample are often presented graphically by **bar charts**.

Info11.2.17

The bar chart shows the absolute or relative frequencies as a function of a finite number of property values in the sample. The bar lengths are proportional to the values they represent.

This is now illustrated by an example. The species of 10 trees at the forest's edge was determined. The possible characteristic attributes are:

$$\begin{aligned} a_1 &= \text{Oak} \ , \\ a_2 &= \text{Beech} \ , \\ a_3 &= \text{Spruce} \ , \\ a_4 &= \text{Pine, etc.} \ . \end{aligned}$$

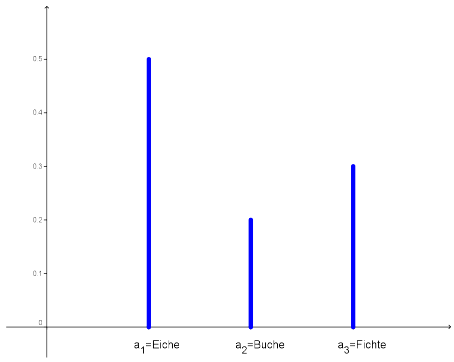
A sample resulted in the following original list:

i	1	2	3	4	5	6	7	8	9	10
x_i	a_2	a_1	a_1	a_3	a_1	a_2	a_1	a_1	a_3	a_3

This original list results in the following empirical frequency table:

Attribute	absolute	relative	in %
Oak	5	0.5	50
Beech	2	0.2	20
Spruce	3	0.3	30

The bar chart corresponding to this empirical frequency table is shown in the figure below.



Bar chart

Qualitative properties are often represented by **pie charts**:

Info11.2.18

A slice is assigned to each characteristic attribute according to its relative frequency, where

$$h_j = \frac{H_j}{n} = \frac{\alpha_j}{360°}.$$

Here, α_j is the angle (in degree measure) of the slice (circular sector) that corresponds to the attribute j within the original list $x = (x_1, x_2, \ldots, x_n)$.

This is again illustrated by an example.

A number $n = 1000$ of households were queried as to how satisfied they were with a new kind of barbecue. The possible answers were: very satisfied (1), satisfied (2), less satisfied (3) and not satisfied (4).

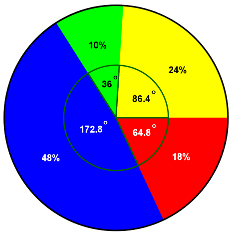
The survey resulted in the following empirical frequency table.

Attribute	Absolute frequencies	Relative frequencies	Percentage
Very satisfied	100	0.1	10%
Satisfied	240	0.24	24%
Less satisfied	480	0.48	48%
Not satisfied	180	0.18	18%
Sum	1000	1	100%

The corresponding angles are, according to the Info Box above,

- $\alpha_1 = 360° \cdot 0.1 = 36°$,
- $\alpha_2 = 360° \cdot 0.24 = 86.4°$,
- $\alpha_3 = 360° \cdot 0.48 = 172.8°$,
- $\alpha_4 = 360° \cdot 0.18 = 64.8°$.

This results in the following pie chart:



It is often pointless to present all possible attributes in a diagram. It is more convenient to classify them and draw only the frequencies of the classes into a diagram. This is the only way to visualise the frequencies of continuous characteristics in a bar or pie chart.

Let X be a quantitative (continuous) property, and $x = (x_1, x_2, \ldots, x_n)$ the original list for a sample of size n. An empirical frequency distribution is obtained according to the following approach:

- Find the minimum and the maximum sample value, i.e.

$$x_{(1)} = \min\{x_1, x_2, \ldots, x_n\} \text{ and } x_{(n)} = \max\{x_1, x_2, \ldots, x_n\}.$$

- List these and all values in between, rounded to the required fractional digit and sorted by size. This converts the (continuous) property X into a discrete property.
- Prepare a tally sheet and draw the corresponding empirical frequency distribution.

The empirical frequency distribution of a continuous property can be very broad. In particular zeros may appear, caused by measurement values that do not occur in the original list (sample). Due to this, the empirical frequency table gets very confusing and bulky. Hence, a **classification** is carried out to reduce the amount of data (data reduction). In fact, this corresponds to a reduction of measurement accuracy (rounding!).

Info11.2.19

Classes are half-open intervals of the form

$$(a;b] \;=\; \{x \in \mathbb{R} \,:\, a < x \leq b\} \;\text{ with }\; a, b \in \mathbb{R} \cup \{\pm\infty\}\,.$$

There is no general rule defining the number k of classes or the size of a class. However, the following guidelines are recommended:

- Uniform classification: Find $x_{(1)} = \min\{x_1, x_2, \ldots, x_n\}$ and $x_{(n)} = \max\{x_1, x_2, \ldots, x_n\}$. Then divide the interval $(x_{(1)} - \varepsilon; x_{(n)} + \varepsilon]$ with small $\varepsilon > 0$ into k uniform, non-overlapping, half-open subintervals.
- Avoid classes that are too small or too large.
- If possible, avoid classes with only a few observations.
- Find approximately $k \approx \sqrt{n}$ equally sized classes, where n is the number of samples.

Info11.2.20

The **histogram** is used for the graphical representation of quantitative data (variable). It represents the relative frequencies of the data in the class $(a, b]$ by a rectangle with base $(a, b]$ whose area represents the class.

A histogram is obtained through the following approach: let

$$x \;=\; (x_1, x_2, \ldots x_n)$$

be an original list for a sample of size n of a quantitative property X.

- Use a classification into k classes. Let the interval of the jth class $j = 1, 2, \ldots, k$ be $(t_j; t_{j+1}]$.
- Let H_j be the number of sample values in the interval $(t_j; t_{j+1}]$ for $j = 1, 2, \ldots, k$. The numbers H_j are also called absolute class frequencies.
- For each $j \in \{1, 2, \ldots, k\}$ draw a rectangle over the base $(t_j; t_{j+1}]$ of height d_j with the area $d_j \cdot (t_{j+1} - t_j) = h_j = \frac{H_j}{n}$. The areas h_j are the relative frequencies.

The total area of these rectangles equals 1.

This approach is now illustrated by a detailed example. In a data centre, the processing time (in s, rounded to one fractional digit) of 20 program jobs was determined. This resulted in the following original list of a sample of size $n = 20$:

3.9	3.3	4.6	4.0	3.8
3.8	3.6	4.6	4.0	3.9
3.9	3.9	4.1	3.7	3.6
4.6	4.0	4.0	3.8	4.1

The smallest value is 3.3 s, the largest value is 4.6 s, the increment is 0.1 s. According to the guidelines above, we should find approximately $k \approx \sqrt{n}$ equally sized classes. Here, we use the following classification into $k = 4$ classes.

Class	$(t_j; t_{j+1}],\ j = 1, 2, 3, 4$	Data
Class 1	$(3.25; 3.65]$	"From 3.3 to 3.6"
Class 2	$(3.65; 3.95]$	"From 3.7 to 3.9"
Class 3	$(3.95; 4.25]$	"From 4.0 to 4.2"
Class 4	$(4.25; 4.65]$	"From 4.3 to 4.6"

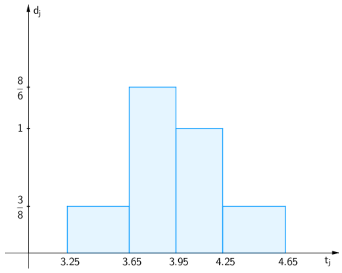
The table of the absolute and relative frequencies has the following form:

Class	abs. Class frequency H_j	rel. Class frequency h_j
Class 1	3	0.15
Class 2	8	0.4
Class 3	6	0.3
Class 4	3	0.15

The heights of the $k = 4$ rectangles are as follows:

- Class 1: $d_1 \cdot (t_2 - t_1) = d_1 \cdot 0.4 = h_1 = 0.15$, i.e. $d_1 = \frac{3}{8} = 0.375$.
- Class 2: $d_2 \cdot (t_3 - t_2) = d_2 \cdot 0.3 = h_2 = 0.4$, i.e. $d_2 = \frac{4}{3} = 1.\overline{3}$.
- Class 3: $d_3 \cdot (t_4 - t_3) = d_3 \cdot 0.3 = h_3 = 0.3$, i.e. $d_3 = 1$.
- Class 4: $d_4 \cdot (t_5 - t_4) = d_4 \cdot 0.4 = h_4 = 0.15$, i.e. $d_4 = \frac{3}{8} = 0.375$.

Thus, we have the following histogram:


d_j
$\frac{8}{6}$
1
$\frac{3}{8}$
3.25
3.65
3.95
4.25
4.65
t_j

11.3 Statistical Measures

11.3.1 Introduction

Suppose a sample of size n is given for some quantitative property X. Let the original list be given by

$$x = (x_1, x_2, \ldots, x_n)\,.$$

Info11.3.1

The **arithmetic mean** $\overline{x}$ (also called sample mean) of $x_1, x_2, \ldots, x_n$ is defined as

$$\overline{x} = \frac{1}{n} \cdot \sum_{k=1}^{n} x_k = \frac{x_1 + x_2 + \ldots + x_n}{n}\,.$$

In physical terms, $\overline{x}$ describes the centre of mass of a mass distribution given by equal masses at $x_1, x_2, \ldots, x_n$ on the massless number line.

Example 11.3.2

We have the following original list for a sample of size $n = 20$:

10	11	9	7	9
11	22	12	13	9
11	9	10	12	13
12	11	10	10	12

The investigated property could be, for example, the length of study (measured in semesters) of 20 mathematics students at the TU Berlin. Summing up the values results in

$$\sum_{k=1}^{20} x_k = 223\,,$$

so for the arithmetic mean we have

$$\overline{x} = \frac{1}{20} \cdot \sum_{k=1}^{20} x_k = \frac{223}{20} = 11.15\,.$$

The arithmetic mean is rather sensitive to so-called statistical outliers: measurement values that vary strongly form the other data can significantly affect the arithmetic mean.

Example 11.3.3

Let us again consider the original list for the sample of size $n = 20$ above. If we drop the value $x_7 = 22$, then for the arithmetic mean of the remaining 19 data values we have

$$\frac{1}{19} \cdot \sum_{k=1, k\neq 7}^{n} x_k \;=\; \frac{201}{19} \;\approx\; 10.58\,.$$

If a multiplicative or relative relation exists among the values in an original list (for example, for growth processes or continuous compounding interest), the arithmetic (additive) mean is not an appropriate measure. For such data values, the geometric mean is used:

Info11.3.4

Let data of the form $x_1 > 0,\ x_2 > 0,\ \ldots, x_n > 0$ be given. Then, the **geometric mean** $\overline{x}_G$ of $x_1, x_2, \ldots, x_n$ is given by

$$\overline{x}_G \;=\; \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n}\,.$$

Example 11.3.5

Let us consider a population that consists of 50 animals at time t_0. Every two years, the number of animals is counted again.

Year		Number of animals		Growth rate
t_0		50		
t_0+2		100		doubled ($x_1 = 2$)
t_0+4		400		quadrupled ($x_2 = 4$)
t_0+6		1200		tripled ($x_3 = 3$)

For the (geometric) mean growth rate, we have

$$\overline{x}_G \;=\; \sqrt[3]{2\cdot 4\cdot 3} \;=\; \sqrt[3]{24} \;\approx\; 2.8845\,.$$

This example illustrates that applying the arithmetic mean to growth processes gives misleading results. We would get

$$\overline{x} \;=\; \frac{1}{3}\cdot(2+4+3) \;=\; \frac{9}{3} =\; 3\,.$$

However, a theoretical tripling of the population size every two years would imply that the number of animals after six years would be 1,350 which is obviously not the case. From an average growth rate of 2.8845, we obtain the correct result: $50\cdot(2.8845)^3 \approx$ 1,200.

Exercise 11.3.1
The growth rates per year of an investment are as follows:

Year	2011	2012	2013	2014	2015
Growth rate	0.5%	1.1%	0.8%	1.2%	0.7%

Calculate the mean growth rate over five years in percent: $\overline{x}_G$ = [ ] %, rounded mathematically to two fractional digits.

In this exercise you are allowed to use a calculator.

Solution:
Taking the geometric mean results in

$$\overline{x}_G \;=\; \sqrt[5]{0.5\cdot 1.1\cdot 0.8\cdot 1.2\cdot 0.7} \;=\; \sqrt[5]{0.3696} \;=\; 0.819495159191\ldots$$

which is mathematically rounded to 0.82.

11.3.2 Robust Measures

The measures presented in this section are robust with respect to outliers: large deviations of single data values do not affect this measures (or only affect it slightly).

Consider an original list

$$x = (x_1, x_2, \ldots, x_n)$$

for a sample of size n. Let the data x_i be the property values of a quantitative property X.

Info11.3.6

The list $x_{(\,)} = (x_{(1)}, x_{(2)}, \ldots, x_{(n)})$ gained by ascending sorting

$$x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$$

of the original list is called an ordered list or ordered sample (of the original list x). The i*th* entry $x_{(i)}$ in the ordered list is the ith smallest value in the original list.

Example 11.3.7

Let us again consider the original list $x = (x_1, x_2, \ldots, x_{20})$ for the sample of size $n = 20$ from the examples above. Ascending sorting $x_{(\,)} = (x_{(1)}, x_{(2)}, \ldots, x_{(20)})$ results in the following ordered sample:

7	9	9	9	9	10	10	10	10	11	11	11	11	12	12	12	12	13	13	22

Info11.3.8

The (empirical) **median** $\tilde{x}$ of $x_1, x_2, \ldots, x_n$ is defined as

$$\tilde{x} = \begin{cases} x_{\left(\frac{n+1}{2}\right)} & \text{for } n \text{ odd} \\ \frac{1}{2} \cdot \left(x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)} \right) & \text{for } n \text{ even} \end{cases} .$$

In contrast to the arithmetic mean, the (empirical) mean is not sensitive to outliers. For example, the largest value in the ordered original list can be arbitrarily enlarged without changing the median.

Example 11.3.9

In the example above, the sample size $n = 20$ is even. Thus, we have for the median

$$\tilde{x} \;=\; \frac{1}{2} \cdot \left(x_{(10)} + x_{(11)}\right) \;=\; \frac{1}{2} \cdot (11 + 11) \;=\; 11 \,.$$

Approximately half of the values in the original list are less than or equal to the median, and half of the values are greater than or equal to the median $\tilde{x}$. This principle can be generalised to define quantiles. For this purpose, take an original list $x = (x_1, x_2, \ldots, x_n)$ for a sample of size n of a quantitative property X.

Info11.3.10

Let

$$x_{(\,)} \;=\; (x_{(1)}, x_{(2)}, \ldots, x_{(n)})$$

be the corresponding ordered sample and

$$\alpha \in (0,1) \;\text{ and }\; k = \mathsf{floor}(n \cdot \alpha) \;=\; \lfloor n \cdot \alpha \rfloor \,.$$

Then

$$\tilde{x}_\alpha \;=\; \begin{cases} x_{(k+1)} & \text{if } \; n \cdot \alpha \notin \mathbb{N} \\ \frac{1}{2} \cdot \left(x_{(k)} + x_{(k+1)}\right) & \text{if } \; n \cdot \alpha \in \mathbb{N} \end{cases}$$

is called a sample α-quantile or simply α-quantile of $x_1, x_2 \ldots, x_n$.

The 0.25-quantile is also called the lower **quartile**. It splits off approximately the lowest 25 % of data values from the highest 75 %. Accordingly, the 0.75-quantile is called the upper quartile. For $\alpha = 0.5$ we have the median, i.e. $\tilde{x} = \tilde{x}_{0.5}$. If $\alpha \in (0,1)$, the ordered list $x_1, x_2, \ldots, x_n$ is split so that approximately $\alpha \cdot 100\%$ of the data value are less or equal to $\tilde{x}_\alpha$ and approximately $(1 - \alpha) \cdot 100\%$ of the data values are greater or equal to $\tilde{x}_\alpha$.

Example 11.3.11

Consider again the original list $x = (x_1, x_2, \ldots, x_{20})$ for the sample of size $n = 20$ from the examples above together with the ordered sample $x_{(\,)} = (x_{(1)}, x_{(2)}, \ldots, x_{(20)})$

7 9 9 9 9 10 10 10 10 11 11 11 11 12 12 12 12 13 13 22

For $\alpha = 0.25$, the 25%-quantile is defined by $n \cdot \alpha = \frac{20}{4} = 5 \in \mathbb{N}$, i.e. for the lower quartile we have

$$\tilde{x}_{0.25} = \frac{1}{2} \cdot \left(x_{(5)} + x_{(6)}\right) = \frac{1}{2} \cdot (9 + 10) = \frac{19}{2} = 9.5\,.$$

For the upper quartile, we set $\alpha = 0.75$ and obtain $n \cdot \alpha = \frac{20 \cdot 3}{4} = 15 \in \mathbb{N}$, hence

$$\tilde{x}_{0,75} = \frac{1}{2} \cdot \left(x_{(15)} + x_{(16)}\right) = \frac{1}{2} \cdot (12 + 12) = 12\,.$$

again, let a sample of size n be given to a quantitative property X with the corresponding ordered sample

$$x_{(\,)} = (x_{(1)}, x_{(2)}, \ldots, x_{(n)})$$

and

$$\alpha \in [0,\ 0.5) \text{ and } k = \mathsf{floor}(n \cdot \alpha) = \lfloor n \cdot \alpha \rfloor\,.$$

Info11.3.12

The α-trimmed (or α-truncated) sample mean is defined as

$$\overline{x}_\alpha = \frac{1}{n - 2 \cdot k} \cdot \sum_{j=k+1}^{n-k} x_{(j)} = \frac{1}{n - 2 \cdot k} \cdot \left(x_{(k+1)} + \ldots + x_{(n-k)}\right)\,.$$

The α-trimmed mean is an arithmetic mean that discards the $\alpha \cdot 100\%$ largest and $\alpha \cdot 100\%$ smallest data points from the calculation. Thus, it is a flexible protection tool against outliers at the boundaries of the data range. However, we mustn't forget that we no longer take all data into account when we use this tool.

Example 11.3.13

In the already much considered data set, the ordered sample $x_{()} = (x_{(1)}, x_{(2)}, \ldots, x_{(20)})$ is given by

$$7 \quad 9 \quad 9 \quad 9 \quad 9 \quad 10 \quad 10 \quad 10 \quad 10 \quad 11 \quad 11 \quad 11 \quad 11 \quad 12 \quad 12 \quad 12 \quad 12 \quad 13 \quad 13 \quad 22\,,$$

and for $\alpha = 0.12$ and $k = \lfloor 20 \cdot 0.12 \rfloor = \lfloor 2.4 \rfloor = 2$ we obtain for the 12%-trimmed mean of the sample

$$\overline{x}_{0.12} \;=\; \frac{1}{16} \cdot \sum_{j=3}^{18} x_{(j)} \;=\; \frac{1}{16} \cdot 172 \;=\; 10.75\,.$$

It is less than the arithmetic mean $\overline{x} = 11.15$ since outliers, such as $x_{(20)} = 22$, were ignored.

11.3.3 Measures of Dispersion

Means and quantiles are measures of position, i.e. they give information on the absolute position of the qualitative values x_j. If we add a constant c to every value x_j, then the position measures also increase by c. In contrast, measures of dispersion are measures that give information on the dispersion or relative distribution of the data values independent of their absolute position. Consider a sample of size $n \geq 2$ of a quantitative property X. Let the original list be given by $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

Info11.3.14

The **sample variance** of the original list is defined as

$$s_x^2 \;=\; \frac{1}{n-1} \cdot \sum_{k=1}^{n} (x_k - \overline{x})^2 \;=\; \frac{(x_1 - \overline{x})^2 + \ldots + (x_n - \overline{x})^2}{n-1}\,.$$

The **sample standard deviation** is defined by $s_x = +\sqrt{s_x^2}$.

The sample variance is a measure of dispersion that describes the variability of the observation sample. The smaller the variance the "closer" the data values lie to each other. A variance $s_x^2 = 0$ is only possible if all data values are equal. Typically, it strongly increases with increasing n. The standard deviation is a more appropriate measure for the "broadness" of the distribution of data values. The two formulas given above have a few pitfalls:

- Before the variance can be calculated the mean $\overline{x}$ must already be known.
- The fact that in the definition of s_x^2 is divided by $n-1$ and not by n is for deeper mathematical reasons that can only be discussed in a statistics lecture.
- The notation $s_x = +\sqrt{s_x^2}$ is a little misleading. You must not cancel the square by the square root, since the *sum* s_x^2 must be calculated (and this value is not defined as a single square) to determine s_x.
- Be careful using a scientific calculator with statistical functions: the sample variance is available via the s^2 key. The σ^2 key, however, provides the sum with denominator n instead of $n-1$. This is not the sample standard deviation.

Example 11.3.15

The data sequence $x = (-1, 0, 1)$ has the mean $\overline{x} = 0$ and the sample standard deviation

$$s_x^2 = \frac{1}{n-1} \cdot \sum_{k=1}^{n} (x_k - \overline{x})^2 = \frac{1}{3-1} \cdot \left((-1-0)^2 + (0-0)^2 + (1-0)^2\right) = 1\,.$$

Adding further zeros to the data sequence does not change the position measure $\overline{x}$, but the measure of deviation s_x^2,does change since the data values here are more strongly concentrated at the mean. In contrast, shifting all data values by a constant does not change the variance. For example, the data sequence $(-5, -4, -3)$ has also variance 1.

Exercise 11.3.2

A data sequence (with an unknown number n of values) has the measures $\overline{x} = 4$, $s_x^2 = 10$, and the median $\tilde{x} = 3$. Suppose the values of a second data sequence satisfy the equation $y_k = (-2) \cdot x_k$ for every k. What are its measures?

Answer: the measures are $\overline{y}$ = [ ] , s_y^2 = [ ] , and $\tilde{y}$ = [ ] .

Hint: recall the definitions of the mean, the sample variance, and the median consider how multiplying all x-values by a factor of (-2) influences the entire expression.

Solution:
Substituting the new x-values results in

$$\begin{aligned}
\overline{y} &= \frac{1}{n}\sum_{k=1}^{n} y_k \;=\; \frac{1}{n}\sum_{k=1}^{n}(-2)\cdot x_k \;=\; (-2)\cdot\frac{1}{n}\sum_{k=1}^{n} y_k \;=\; (-2)\cdot\overline{x} \;=\; -8\,,\\
s_y^2 &= \frac{1}{n-1}\sum_{k=1}^{n}(y_k-\overline{y})^2 \;=\; \frac{1}{n-1}\sum_{k=1}^{n}((-2)x_k-(-2)\overline{x})^2\\
&= \frac{(-2)^2}{n-1}\sum_{k=1}^{n}(x_k-\overline{x})^2 \;=\; (-2)^2\cdot s_x^2 \;=\; 40\,,\\
\tilde{y} &= (-2)\tilde{x} \;=\; -6\,.
\end{aligned}$$

The conversion of the median uses the fact that a multiplication by a factor of (-2) reverses the ordering of the ordered original list, but the value at the mid position (for an odd number) or the two values at the mid positions (for an even number) stay at their positions and are multiplied by (-2) each.

11.4 Final Test

11.4.1 Final Test Module 4

Exercise 11.4.1
For bonds (e.g. government bonds),we distinguish between the nominal value and the issue price (quote). Bonds can be issued at nominal value, under nominal value or over nominal value. The issue price is closer to the nominal value the more the bond interest corresponds to the current market rate. A customer buys bonds with a nominal value of $K = 10{,}000$ EUR, an issue price of 100%, an interest rate of $p = 4,5\%$ p. a., and a period of $t = 10$ years.

a. How much interest is paid at the end of each interest period for an issue price of 100% when simple interest is applied. Answer: the annually paid interest is ☐ EUR.

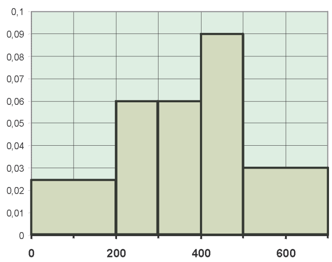
b. Specify the amount of the totally paid capital at the end of the period if simple interest is applied. Answer: the amount of capital paid at the end of the period is ☐ EUR.

Exercise 11.4.2
A capital of $K = 25{,}000$ EUR will be invested with an annually paid interest rate of $p = 3,5\%$ p. a. until the amount of the capital has doubled. How many years does the capital have to be invested, if continuous compounding interest is applied?

Answer: the required investment period is $t =$ ☐ years.

Round up your result to the text integer.

Exercise 11.4.3
Read off the properties of the described sample from the histogram shown in the figure below.



Histogram of the sample $x = (x_1, \ldots, x_n)$.

Specify the interval boundaries of the five classes and the corresponding relative frequencies. Fill in the frequency table. For this purpose, calculate the ratios of the areas of the single bars in the diagram to the total area.

Class	Interval	rel. Class frequencies h_j
Class 1	$[0; 200)$	$0,16$
Class 2		
Class 3		
Class 4		
Class 5		

Exercise 11.4.4
The measurement of the weight of $n = 11$ watermelons (in kilogram) resulted in the following values:

Number j	1	2	3	4	5	6	7	8	9	10	11
Weight x_j in kilogram	6.2	5.5	7.3	6.8	6.3	5.5	4.5	6.5	7.3	5.7	5.6

a. Find the arithmetic mean of the 11 sample values: $\overline{x}$ = ______ .

b. Find the median of the 11 sample values: $\tilde{x}$ = ______ .

Enter your result rounded to two fractional digits. Do not use a calculator but try to find the values by hand.

12 Entrance Test

Module Overview

Please note:

- The sample part of the test will introduce you to the use of the input boxes.
- Your answers in the sample part of the test will not be graded.
- During the test you are allowed to rely on literature or written notes. However, the test must be taken without any technical aids. In particular, you may not use a calculator.

For the sample test, please click here. To proceed immediately to the graded entrance test, please click here.

12.1 Test 1: Sample Part

12.1.1 Restart

Please read the instructions on the test procedure.

The sample part of the test is optional and will not be graded. It serves to introduce you to the use of the input boxes.

Important notice: please solve the exercises on a sheet of paper without using a calculator. Enter your solutions into the input boxes, and they will be checked. As soon as you are familiar with the use of the input boxes, you can click "next" in the upper right-hand corner and take the graded entrance test.

Exercise 12.1.1
Simplify the following complex fractions so that only a simple fraction remains:

a. $\dfrac{2+\frac{3}{2}}{1-\frac{1}{3}}$ is equivalent to $\boxed{}$.

b. $\dfrac{4x^2+y^2}{3x-y} - \dfrac{5x^2-2y^2}{y-3x}$ is equivalent to $\boxed{}$.

Exercise 12.1.2
Expand the brackets completely and collect like terms together:

$(x-2)(x+1)\cdot x = \boxed{}$.

Exercise 12.1.3
Apply one of the binomial formulas to transform each of the following terms:

a. $(-x-3)(-x+3) = \boxed{}$.

b. $(s+2r+t)^2 = \boxed{}$.

<u>Hint:</u>
First apply the binomial formula to $(s + (2r+t))^2$.

Exercise 12.1.4
Rewrite this power and radical expression as a simple power with a rational number as its exponent, without using the root sign:

$\sqrt{\sqrt{x} \cdot x} =$ [ ] .

Exercise 12.1.5
Transform the fractions such that the denominator disappears:

a. $\frac{2x}{\sqrt{2}-\sqrt{3}} =$ [ ] .

b. $\frac{1}{\sqrt{n+1}+\sqrt{n}} =$ [ ] .

Exercise 12.1.6
Solve the equation $\dfrac{t-2}{t+1} = 2$ for t.
Answer: $t =$ [ ] .

Exercise 12.1.7
Specify the solution sets of the following quadratic equations:

a. $x^2 + 3x - 10 = 0$ has the solution set [ ] .

b. $x^2 + 2x + 3 = 0$ has the solution set [ ] .

c. $(x-1)^2 - (x+1)^2 = 0$ has the solution set [ ] .

Exercise 12.1.8
Specify the solution sets of the following equations:

a. $\frac{1}{x} + 1 = x - 1$ has the solution set [ ] .

b. $\frac{1}{x} + \frac{2}{x} = \frac{4}{x}$ has the solution set [ ] .

c. $\sqrt{2x^2+1} = 3x$ has the solution set [ ] .

How many solutions does the equation $(b-a)^{100} + (b+a)^{100} = 0$ have if a and b are independent variables?

- [ ] No solution
- [ ] Exactly one solution for both a and b
- [ ] One solution for a and an infinite number of solutions for b

☐ An infinite number of solutions for both variables

Exercise 12.1.9
Rewrite the absolute value expression $|2x-1|-3x$ as a case analysis with two expressions that do not contain any absolute values.
Answer: $|2x-1|-3x =$ ☐ .

What to do here:
In a mathematical case analysis, absolute values are defined by different cases. For example, the absolute value expression $|x-1|$ can be written as:

$$|x-1| = \begin{cases} x-1 & \text{if } x \geq 1 \\ -x+1 & \text{if } x < 1 \end{cases}$$

Exercise 12.1.10
Find all solutions of the absolute value equation $|2x-7| = x-2$.

Answer: The solution set is $L =$ ☐ .

Exercise 12.1.11
Specify the solution set of the inequality $x^2 + 4x < 5$ as an interval.

Answer: $L =$ ☐ .

Exercise 12.1.12
Specify the domains and the solution sets of the following inequalities as intervals:

a. The inequality $\sqrt{x} > x$ has the domain ☐ and the solution set ☐ .

b. The inequality $\sqrt{\sqrt{x-1}+2} > \sqrt{x+1}$ has the domain ☐ and the solution set ☐ .

As soon as you are familiar with the use of the input boxes, you can click “next” in the upper right corner and take the graded entrance test.

12.2 Test 1: Graded Part To Be Submitted

12.2.1 Graded Test for the Online Course

Exercise 12.2.1
Check whether these mathematical expressions denote equations, inequalities, terms, or numbers (multiple checks are possible):

Mathematical expression	Equation	Inequality	Term	Number
$\sqrt{3+\frac{1}{2}}$	☐	☐	☐	☐
$5x-1$	☐	☐	☐	☐
$32x^2 = \frac{1}{y}$	☐	☐	☐	☐
$32 > 2^x$	☐	☐	☐	☐
$x \in \{1;2;3\}$	☐	☐	☐	☐
$b^2 - 4ac \geq 0$	☐	☐	☐	☐

Exercise 12.2.2
Simplify these complex fractions as far as possible:

a. $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4}$ is equivalent to [ ] .

b. $\dfrac{x - \frac{1}{x}}{1 - \frac{1}{x}}$ is equivalent to [ ] .

Exercise 12.2.3
Expand this term completely and collect like terms together: $(a-b)(2c+d) =$ [ ]
.

Exercise 12.2.4
Apply one of the binomial formulas to transform the following terms.The results should not contain any brackets or radical terms.

a. $(2\sqrt{3} + x\sqrt{3})^2 =$ [ ] .

b. $(3a - 4b)^2 =$ [ ] .

Exercise 12.2.5
Rewrite these power and radical expressions as a simple power with a rational number as its exponent, without using the root sign:

a. $\sqrt{ab} \cdot \frac{1}{a} \cdot \sqrt{b^3} =$ [ ] .

b. $\sqrt{\sqrt{\sqrt{a} \cdot b} \cdot \frac{1}{c}} =$ [ ] .

Exercise 12.2.6
Transform the following fractions such that the denominator disappears. The solution must not contain any fractions or powers.

a. $\frac{12}{\sqrt{3}} =$ [ ] .

b. $\frac{1+x-\sqrt{4x}}{\sqrt{x}-1} =$ [ ] .

c. $\frac{u^3}{\sqrt{u^2+1}+1} =$ [ ] .

Exercise 12.2.7
The single solution of the equation $\dfrac{x-2}{|x+1|} = \frac{1}{2}$ is $x =$ [ ] .

Exercise 12.2.8
Specify the solution sets of the following equations:

a. $\frac{1}{x} + 2x = 2 - x$ has the solution set [ ] .

b. $x^2 - 1 = (x-1)^3$ has the solution set [ ] .

c. $\sqrt[7]{r} = r^5$ has the solution set [ ] .

Exercise 12.2.9
Find all solutions of the absolute value equation $|5x-1| = x^2 - 1$.

The solution set is $L =$ [ ] .

Do not use a calculator! Your solution may contain radical terms and fractions.

Exercise 12.2.10
Specify the solution set of the inequality $x^2 + 6 < -5x$ as an interval.

The solution interval is [ ] .

Exercise 12.2.11
Specify the solution sets of the following inequalities as intervals. Be careful with the specification of the endpoints.

a. $(x-1)^2 < x$ has the solution set [ ] .

b. $\sqrt{x^2-1} < x$ has the solution set [ ] .

Exercise 12.2.12
Find the solution set of the following system of linear equations:

$$\begin{aligned} -x+2y &= -5 \\ 3x+y &= 1 \end{aligned}$$

The solution set [ ] is empty,
[ ] contains exactly one solution: $x =$ [ ] , $y =$ [ ] ,
[ ] contains an infinite number of solution pairs (x, y).

Exercise 12.2.13
Find the two-digit number such that its digit sum is 6, and exchanging the digit positions results in a number that is smaller by 18.

Answer: [ ]

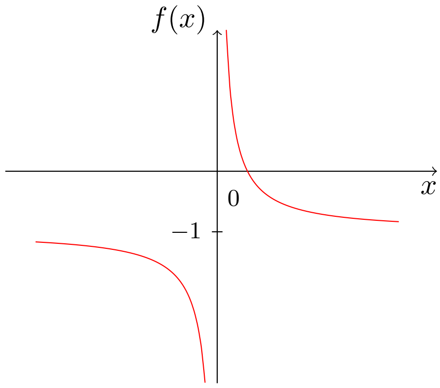
Exercise 12.2.14
Find the value of the real parameter α for which the following system of linear equations

$$\begin{aligned} 2x+y &= 3 \\ 4x+2y &= \alpha \end{aligned}$$

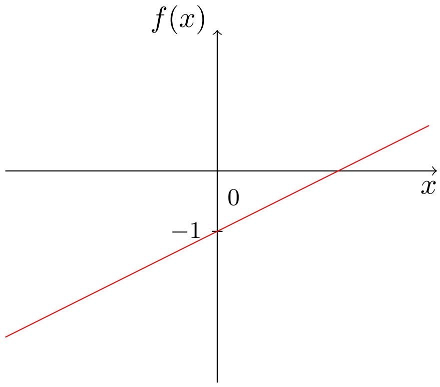
has an infinite number of solutions.

Answer: $\alpha =$ [ ]

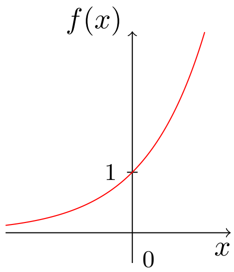
Exercise 12.2.15
Assign to each of the graphs in (a)–(e) the correct mapping rules of the corresponding functions.



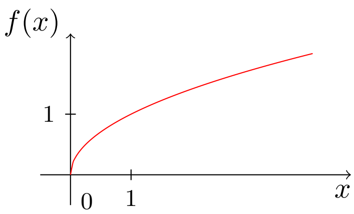
(a)



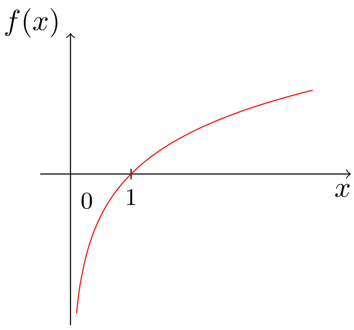
(b)



(c)



(d)



(e)

a. Graph (a) corresponds to the function $f(x) =$ [ ] .

b. Graph (b) corresponds to the function $f(x) =$ [ ] .

c. Graph (c) corresponds to the function $f(x) =$ [ ] .

d. Graph (d) corresponds to the function $f(x) =$ [ ] .

e. Graph (e) corresponds to the function $f(x) =$ [ ] .

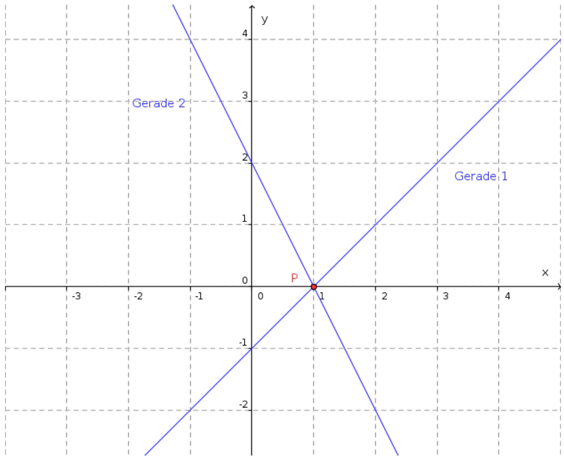
Select your answers from the mapping rules and input terms listed below (not all of them are used).

$f(x) = \sqrt{x}$
$f(x) = \frac{1}{2}x - 1$
$f(x) = \ln(1 - x)$
$f(x) = \ln(x)$
$f(x) = x^{1,5}$
$f(x) = \exp(x) = e^x$
$f(x) = (0,5)^x$
$f(x) = \frac{1}{x}$
$f(x) = \frac{1}{2x} - 1$
$f(x) = \frac{1}{2} - x$

Specify the asymptote of the function with the mapping rule (a).

It is $\lim_{x\to\infty} f(x) =$ [ ] .

Exercise 12.2.16
The figure below shows two lines in two-dimensional space.



Find the two equations describing the lines.

Line 1: $y =$ [ ]

Line 2: $y =$ [ ]

What is the correct number of solutions to the corresponding system of linear equations?

The system of linear equations has
- [ ] no solution,
- [ ] exactly one solution, or
- [ ] an infinite number of solutions.

Exercise 12.2.17
Find the solution set of the following system of linear equations consisting of three equations in three variables.

$$\begin{aligned} x + 2z &= 3 \\ -x + y + z &= 1 \\ 2y + 3z &= 5 \end{aligned}$$

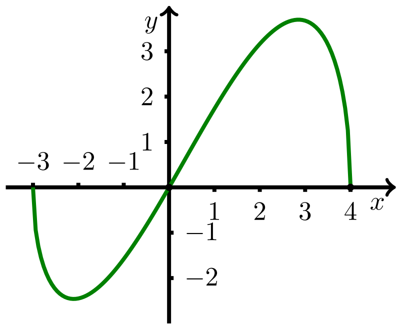
The solution set
- [ ] is empty,
- [ ] contains exactly one solution: $x =$ [ ] , $y =$ [ ] , $z =$ [ ] ,
- [ ] contains an infinite number of solutions (x, y, z).

Exercise 12.2.18
The mileage counter of a van displays 20 km when the vehicle starts its journey at 6 a.m. Four hours later it arrives at its destination. The odometer is now displaying 280 km. Calculate the average velocity v (the average rate of change of position) between the start point and destination. For this purpose, insert the missing numbers and mathematical

symbols (+, −, ·, /) into the following calculation.

$$v = \Big(280 \ \square \ \square\Big) \ \square \ \Big(\square - 6\Big) = \square$$

Exercise 12.2.19
Consider the function $f : [-3; 4] \to \mathbb{R}$, the graph of which is shown in the figure below.



a. The derivative at $x_1 = 4$ is ☐ equal to 0, ☐ not defined, ☐ infinite.

b. The derivative at $x_2 = 0$ is ☐ positive, ☐ equal to 0, ☐ negative.

Exercise 12.2.20
Calculate the first and the second derivative of the function

$f : \mathbb{R} \to \mathbb{R}, x \mapsto \frac{5x}{3+x^2}$, and give your results with all terms cancelled or collected together as appropriate.

a. First derivative $f'(x) =$ ☐

b. Second derivative $f''(x) =$ ☐

Exercise 12.2.21
Specify the regions on which the function f with $f(x) := \frac{\ln x}{x}$ for $x > 0$ is monotonically increasing and where it is decreasing. Specify the regions as maximum open intervals (r, s).

a. f is monotonically increasing on ☐ .

b. f is monotonically decreasing on ☐ .

Which of the points $x_1 = 1$, $x_2 = 2$, or $x_3 = 6$ belongs to an interval on which the function f is convex?

Answer: ☐

Exercise 12.2.22
Find an antiderivative for each of the following functions.

a. $\int \left(5x^4 + 8x\right)\ dx =$ [ ]

b. $\int 6\sin(2x)\ dx =$ [ ]

Exercise 12.2.23
Calculate the integrals

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left(x^3 + \cos(x)\right) dx = \boxed{} \quad \text{and} \quad \int_1^4 x \cdot \sqrt{x} dx = \boxed{}$$

Exercise 12.2.24
We have $\int_{-5}^{5} x \cdot \cos(4x) dx = 0$ since the interval of integration is [ ] with respect to 0 and the integrand is an [ ] function.

Exercise 12.2.25
The graph of the function $f : [-1; 3] \to \mathbb{R}$ with $f(x) := x^3 - 3x^2 - x + 3$ for $-1 \leq x \leq 3$ and the x-axis enclose a region A. Specify the intersection points of the graph of f with the x-axis, and calculate the area I_A of A. Answer: $I_A =$ [ ]

Exercise 12.2.26
Specify the intersection point of the following two lines:

- the line $y = 3x + 3$,
- the line with the general equation $2x - 2y = 6$.

The intersection point is [ ] .

Exercise 12.2.27
Let

$$(x-1)^2 + (y+1)^2 \ = \ d$$

be the general equation of a circle, where d is an unknown positive constant. Specify the properties of this circle.

a. Its radius is $r \ =$ [ ] .

b. Its centre is at $P =$ ☐ .

c. The circle intersects the line $\overline{PQ}$ passing trough the points $P = (-3; 3)$ and $Q = (3; -3)$

- ☐ at one point,
- ☐ at two points,
- ☐ at three points,
- ☐ not at all,
- ☐ the answer depends on the value of the constant d.

Exercise 12.2.28

Take the vectors

$$\vec{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} , \vec{y} = \begin{pmatrix} 0 \\ -4 \end{pmatrix} , \vec{z} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} .$$

From these, calculate the following vectors:

a. $\vec{x} + \vec{y} - \vec{z} =$ ☐ .

b. $2\vec{x} - \frac{1}{2}\vec{y} =$ ☐ .

c. $2(\vec{x} - \vec{y}) + 3\vec{z} =$ ☐ .

Index