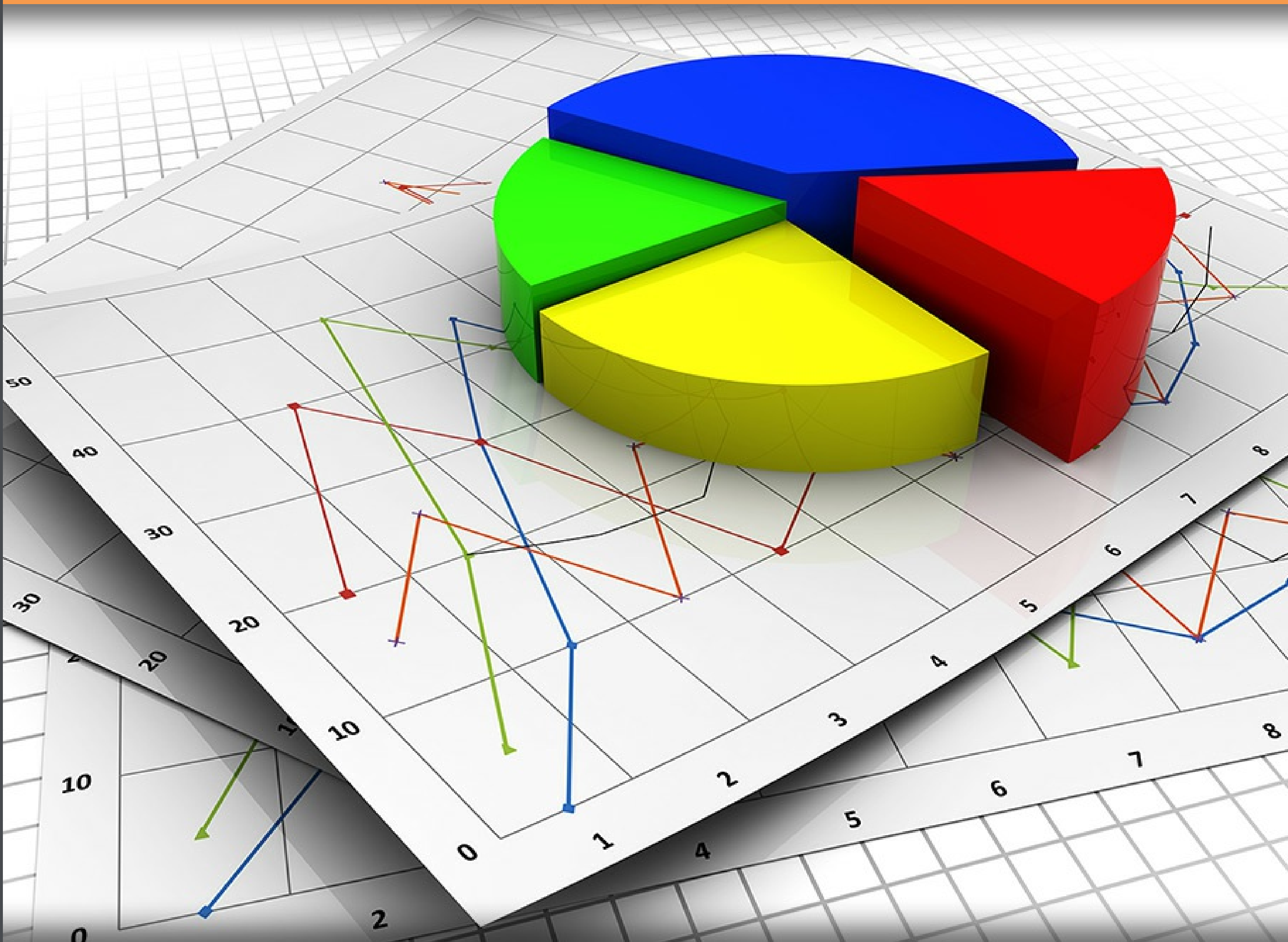
bookboon.com

A Refresher Course in Mathematics

Frank Werner



FRANK WERNER

A REFRESHER COURSE IN MATHEMATICS

A Refresher Course in Mathematics
1st edition

ISBN 978-87-403-1319-2
Peer reviewed by Dr. Larysa Burtseva, The Engineering Institute of the Autonomous University of Baja California, Mexicali, Mexico

Contents

Preface

Since more than 30 years, I taught several courses of basic mathematics for beginners at the university. This concerns students of engineering sciences, students of natural sciences and students of economical sciences. In particular, since almost 20 years I taught a course 'Mathematics of Economics and Management' at the Otto-von-Guericke University Magdeburg. The latter class included all relevant subjects from calculus and algebra. Roughly 10 years ago, I jointly with my co-author Yuri Sotskov from Minsk wrote a book 'Mathematics of Economics and Business' following exactly the structure of this lecture. This book appeared 2006 at Routledge, and I used it as the first item on my reading list for this class. However, this book explicitly includes only some refreshments from school in a short form in Chapter 4, namely how to work with real numbers. The foundations of calculus discussed in this book are, of course, also already a subject of school education in the upper classes so that there is a larger overlapping with mathematical subjects from secondary school education.

I noticed that at the beginning of their study, the majority of students has some partial knowledge about the basic mathematical subjects from school, but not at the required extent. It seems so that this tendency is even increasing currently. In any case, one can observe that beginners of a university study enormously range in their mathematical skills and aptitudes. I know that for many beginners at the university, mathematical subjects appear to be rather difficult. However, if these gaps are not filled at the beginning of the study, this will definitively cause subsequent difficulties in other courses. Without any doubt, nowadays a solid mathematical knowledge is the base for most (almost all) study courses.

So, I felt that there is a need to present some necessary foundations from the mathematical education at school in more detail and also some additional supplementary material, where I wish that university beginners are familiar with. When writing this booklet, I also used the experience collected in several classes of extra-occupational study courses, among them also a bridge course, which refreshes the main subjects from mathematics in school. Typically, the latter students have even more difficulties with mathematical subjects because their school education finished already some years ago. Summarizing, I found that there is a need to write such a booklet from my personal point of view, using the experience collected over the past decades. The goal was roughly not to exceed 250 pages.

Sure, the content of the mathematical education in secondary school varies from country to country a bit. So, I tried to cover a broad range of subjects which might be useful for a university study from an overall point of view. The booklet consists of 12 chapters. Chapters 1 - 2 discuss some basics: some mathematical foundations as well as real numbers and arithmetic operations. Chapters 3 - 5 deal with equations and inequalities. Chapter 6 surveys some basics from analytic geometry in the plane. This is nowadays not so intensively taught as at the time when I attended school, but nevertheless it addresses some useful subjects. Chapters 7 - 10 treat classical subjects from calculus. Chapter 11 presents some aspects of vectors. Chapter 12 discusses some foundations from combinatorics, probability theory and statistics.

I tried to write the chapters as independent as possible. So, it is not necessary to read all chapters beginning from the first one. Instead, the student can go immediately to a particular subject. Sure, the chapters are not completely independent since in mathematics, there are often specific relationships between different subjects. Nevertheless, there is no necessity to study the chapters systematically one by one in the given sequence for the understanding of the book. Moreover, since it is an repetition and summary of elementary material of mathematics, I avoided the formal use of theorems and definitions. Instead of, the major notions are shaded in grey and in addition, important formulas and properties are given in boxes.

Each chapter gives the learning objectives at the beginning. Moreover, every chapter finishes with a number of exercises. The solutions to the exercises (i.e., the concrete results) are given on my homepage so that the reader can verify whether to be able or not to solve typical problems from a particular topic. They can be downloaded as a pdf file under:

http://www.math.uni-magdeburg.de/∼werner/solutions-refresher-course.pdf

The author is grateful to many people for suggestions and comments. In particular, I would like to thank Dr. Michael Höding and my Ph.D. student Ms. Julia Lange from the Institute of Mathematical Optimization of the Faculty of Mathematics at the Otto-von-Guericke-University Magdeburg for their many useful hints during the preparation of this booklet and the support in the preparation of the figures, respectively. I would also like to use this opportunity to thank both for their long-term support in the teaching process at the Otto-von-Guericke-University Magdeburg.

I hope that this small booklet will help the students to overcome their initial difficulties when studying the required mathematical foundations at the university. Typically, the first term at the university is the hardest one due to many changes compared with secondary school. I want to finish this preface with a hint: Long time ago, I was told that learning mathematics is somehow like learning swimming. Nobody learns it by looking how other people do it! One does not learn mathematics by exclusively listening to the lecturer or tutor and copying notes from the blackboard or slides. *Only own practice contributes to a significant progress.* The time necessary for getting a sufficient progress varies for the individual students significantly, so everybody has to find this out by solving a sufficient number of exercises. So, students in their first year at the university should not be afraid of mathematics but should take into account that some (or even a lot of) time is needed to get a sufficiently wide experience in applying mathematical tools.

The author is also grateful for all hints that improve further the content and the presentation of this edition. All suggestions should be addressed preferably to the email address given below. It is my pleasure to thank the publisher Bookboon for the delightful cooperation during the preparation of this booklet.

Magdeburg, March 2016

Frank Werner
Otto-von-Guericke-University Magdeburg
Faculty of Mathematics
PSF 4120
39016 Magdeburg / Germany
Internet: http://www.math.uni-magdeburg.de/∼werner
email: frank.werner@ovgu.de

Chapter 1

Some Mathematical Foundations

This chapter intends to refresh some basic mathematical foundations which are necessary to understand the elementary mathematics in first- or second-year classes on mathematically oriented subjects at universities. We deal with

- sets and operations on sets,
- the use of the sum and product notations and
- mathematical proofs by induction.

Proofs by induction can be used e.g. for verifying certain sum and product formulas as well as specific inequalities. They can also be used for particular combinatorial and geometrical problems.

1.1 Sets

In this section, we introduce the basic notion of a set and discuss operations on sets. A **set** is a fundamental notion of mathematics, and so there is no definition of a set by other basic mathematical notions for a simplification. A set may be considered as a collection of **distinct objects** which are called the **elements** of the set. For each object, it can be uniquely decided whether it is an element of the set or not. We write:

$$\begin{array}{ll} a \in A: & a \text{ is an element of the set } A; \\ b \notin A: & b \text{ is not an element of the set } A. \end{array}$$

A set can be given either by **enumeration** or by **description**. In the first case, we give explicitly the elements of the set, e.g.

$$A = \{1, 3, 5, 7, 9, 11, 13, 15\},$$

i.e., A is the set that contains the eight elements $1, 3, 5, 7, 9, 11, 13, 15$. Alternatively, we can describe a set in the form

$$B = \{b \mid b \text{ has property P }\}.$$

The set B is the set of all elements b which have the property P. In this way, we can describe the above set A e.g. as follows:

$$A = \{a \mid a \text{ is an odd integer and } 1 \le a \le 15\}.$$

For special subsets of real numbers, one often uses also the **interval notation**. In particular, we have

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

where $\mathbb{R}$ is the set of all real numbers. The interval $[a,b]$ is called a **closed** interval whereas (a,b) is called an **open** interval. Accordingly, the intervals $(a,b]$ and $[a,b)$ are called **half-open** (left-open and right-open, respectively) intervals. The set $\mathbb{R}$ of real numbers can also be described by the interval $(-\infty,\infty)$. Moreover, we use the following abbreviations:

$$\mathbb{R}_{\ge a} = [a,\infty), \quad \mathbb{R}_{>a} = (a,\infty), \quad \mathbb{R}_{\le a} = (-\infty,a] \quad \text{and} \quad \mathbb{R}_{<a} = (-\infty,a)\,.$$

Thus, $\mathbb{R}_{\ge 0}$ denotes the set of all non-negative real numbers.

Next, we introduce operations on sets: the union, the intersection and the difference of two sets.

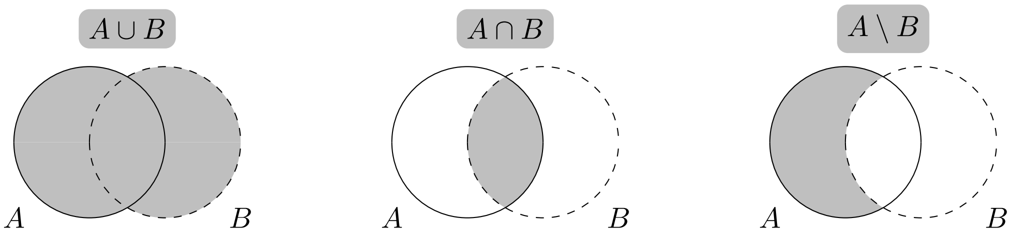


Figure 1.1: Union, intersection and difference of two sets

Union of two sets:
The set of all elements which belong either only to a set A or only to a set B or to both sets A and B is called the **union** of the two sets A and B (in symbols $A \cup B$, read: A union B):

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\}.$$

The set $A \cup B$ contains all elements that belong **at least** to one of the sets A and B.

Intersection of two sets and disjoint sets:
The set of all elements belonging to both sets A and B is called the **intersection** of the two sets A and B (in symbols $A \cap B$, read: A intersection B):

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}.$$

Two sets A and B are called **disjoint**, if $A \cap B = \emptyset$.

Thus, the sets A and B are disjoint if they have no common elements.

Difference of two sets:
The set of all elements belonging to a set A but not to a set B is called the **difference set** of A and B (in symbols $A \setminus B$, read: A minus B):

$$A \setminus B = \{x \mid x \in A \text{ and } x \notin B\}.$$

The union, intersection and difference of two sets A and B are illustrated in Fig. 1.1.

Next, we summarize some basic rules for working with sets.

Rules for sets:
Let A, B, C be arbitrary sets. Then:

1. $A \cap B = B \cap A, \qquad A \cup B = B \cup A$
(**commutative laws** of intersection and union);
2. $(A \cap B) \cap C = A \cap (B \cap C), \qquad (A \cup B) \cup C = A \cup (B \cup C)$
(**associative laws** of intersection and union);
3. $(A \cap B) \cup C = (A \cup C) \cap (B \cup C), \qquad (A \cup B) \cap C = (A \cap C) \cup (B \cap C)$
(**distributive laws** of intersection and union).

Note that the difference of two sets is not a commutative operation, i.e., in general we have

$$A \setminus B \neq B \setminus A.$$

Example 1.1 *Let* $A = \{2, 3, 5, 8, 11, 14\}$ *and* $B = \{1, 5, 11, 12, 14, 16\}$. *Then*

$$A \cup B = \{1, 2, 3, 5, 8, 11, 12, 14, 16\} \quad \textit{and} \quad A \cap B = \{5, 11, 14\}.$$

Thus, the union of the sets A and B contains nine elements and the intersection contains three elements. Moreover, we get

$$A \setminus B = \{2, 3, 8\} \quad \textit{and} \quad B \setminus A = \{1, 12, 16\}.$$

Notice that the elements of a set can be given in **arbitrary order**.

Example 1.2 *Let E be the set containing the cities Berlin, London, Magdeburg, Madrid, Moscow, Paris, Rom and Stockholm and A be the set containing the cities Chicago, Montreal, New York, San Francisco, Toronto and Vancouver. Thus, the set A contains six North American cities and E contains eight European cities. Thus, there is no city contained in both sets and therefore, we have $A \cap B = \emptyset$. Moreover,*

$$A \cup E = \{\textit{Chicago, Montreal, New York, San Francisco, Toronto, Vancouver, Berlin, London, Magdeburg, Madrid, Moscow, Paris, Rom, Stockholm}\}$$

Moreover, we have $A \setminus E = A$ and $E \setminus A = E$ in this example.

Example 1.3 *Let*

$$A = \{a \mid 1 \leq a \leq 100,\ a \textit{ is integer and divisable by } 3\}$$

and

$$B = \{b \mid 1 \leq b \leq 100,\ b \textit{ is integer and divisable by } 5\},$$

i.e., the set A contains 33 integers and the set B contains 20 integers. Obviously, we can rewrite the sets A and B as follows:

$$A = \{3, 6, 9, 12, \ldots, 90, 93, 96, 99\} \quad \textit{and} \quad B = \{5, 10, 15, 20, \ldots, 85, 90, 95, 100\}.$$

Then we obtain

$$\begin{aligned} A \cup B &= \{a \mid 1 \leq a \leq 100,\ a \textit{ is integer and divisible by } 3 \textit{ or } 5\} \\ &= \{3, 5, 6, 9, 10, 12, 15, \ldots, 95, 96, 99, 100\} \end{aligned}$$

and

$$\begin{aligned} A \cap B &= \{a \mid 1 \leq a \leq 100,\ a \textit{ is integer and divisible by } 3 \textit{ and } 5\} \\ &= \{15, 30, 45, 60, 75, 90\}. \end{aligned}$$

This means that $A \cap B$ contains all integers between 1 and 100 which are divisible by $3 \cdot 5 = 15$. Moreover,

$$\begin{aligned} A \setminus B &= \{a \mid 1 \leq a \leq 100,\ a \textit{ is integer and divisible by } 3 \textit{ but not by } 5\} \\ &= \{3, 6, 9, 12, 18, 21, \ldots, 87, 93, 96, 99\} \end{aligned}$$

and

$$\begin{aligned} B \setminus A &= \{a \mid 1 \leq a \leq 100,\ a \textit{ is integer and divisible by } 5 \textit{ but not by } 3\} \\ &= \{5, 10, 20, 25, 35, 40, \ldots, 70, 80, 85, 95, 100\}. \end{aligned}$$

To find $A \setminus B$, we exclude from A the numbers 15, 30, 45, 60, 75 and 90. To determine $B \setminus A$, we exclude from B the same numbers.

Thus, the union $A \cup B$ contains 47 integers, the intersection $A \cap B$ contains 6 integers and the difference sets $A \setminus B$ and $B \setminus A$ contain 27 and 14 integers, respectively.

Example 1.4 *Consider the following intervals*

$$I_1 = [2, 6), \quad I_2 = [4, 8] \quad \textit{and} \quad I_3 = (2, 9].$$

Then we obtain

$$\begin{array}{lll} I_1 \cup I_2 = [2, 8], & I_1 \cup I_3 = [2, 9], & I_2 \cup I_3 = I_3, \\ I_1 \cap I_2 = [4, 6), & I_1 \cap I_3 = (2, 6), & I_2 \cap I_3 = I_2, \\ I_1 \setminus I_2 = [2, 4), & I_1 \setminus I_3 = \{2\}, & I_2 \setminus I_3 = \emptyset, \\ I_2 \setminus I_1 = [6, 8], & I_3 \setminus I_1 = [6, 9], & I_3 \setminus I_2 = (2, 4) \cup (8, 9]. \end{array}$$

Example 1.5 *In this example, we illustrate the distributive law for set operations. Let*

$$A = \{1, 2, 4, 7, 8, 10\}, \quad B = \{2, 3, 4, 9\} \quad \textit{and} \quad C = \{3, 8\}.$$

One the one side, we obtain

$$(A \cap B) \cup C = \{2, 4\} \cup \{3, 8\} = \{2, 3, 4, 8\}.$$

On the other side, we get

$$(A \cup C) \cap (B \cup C) = \{1, 2, 3, 4, 7, 8, 10\} \cap \{2, 3, 4, 8, 9\} = \{2, 3, 4, 8\}.$$

This illustrates the first distributive law given above. Moreover,

$$(A \cup B) \cap C = \{1, 2, 3, 4, 7, 8, 9, 10\} \cap \{3, 8\} = \{3, 8\}$$

and

$$(A \cap C) \cup (B \cap C) = \{8\} \cup \{3\} = \{3, 8\},$$

which illustrates the validity of the second distributive law given above.

1.2 Sum and Product Notation

For writing a sum or product in short form, one often uses the Greek letters $\sum$ and $\prod$, respectively. In particular, we have

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \ldots + a_n \qquad \text{and} \qquad \prod_{i=1}^{n} a_i = a_1 \cdot a_2 \cdot \ldots \cdot a_n \,.$$

The letter i below from the sum (product) symbol is denoted as **summation (multiplication) index**. This index can be arbitrarily chosen. The number under the product symbol is denoted as **lower limit** of summation and multiplication, respectively, and the number above the symbol is denoted as the **upper limit**. So, the summation above is made from $i = 1$ to $i = n$. Note that a sum and product, respectively, is defined to be equal to zero if the lower limit of summation (multiplication) is greater than the upper limit of summation (multiplication).

Example 1.6 *We illustrate the use of sums and products.*

a) We have

$$\sum_{i=1}^{4} i^2 = 1^2 + 2^2 + 3^2 + 4^2 = 1 + 4 + 9 + 16 = 30$$

and

$$1^3 + 2^3 + 3^3 + \ldots + 50^3 = \sum_{i=1}^{50} i^3 = \sum_{k=1}^{50} k^3 \,.$$

b) $\displaystyle\sum_{k=-2}^{2} \frac{1}{2k+3} = \frac{1}{-1} + \frac{1}{1} + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} = \frac{35}{105} + \frac{21}{105} + \frac{15}{105} = \frac{71}{105}$ *;*

c) $\displaystyle\sum_{j=1}^{4} (3j-1)^{j-1} = 2^0 + 5^1 + 8^2 + 11^3 = 1 + 5 + 64 + 1331 = 1401$ *;*

d) $\displaystyle\prod_{n=0}^{3} (a_n + b_n) = (a_0 + b_0) \cdot (a_1 + b_1) \cdot (a_2 + b_2) \cdot (a_3 + b_3)$ *;*

e) $\displaystyle\prod_{l=-1}^{3} (2l-1) = (-3) \cdot (-1) \cdot 1 \cdot 3 \cdot 5 = 45$ *;*

f) $\displaystyle\prod_{k=1}^{5} k^{k-3} = 1^{-2} \cdot 2^{-1} \cdot 3^0 \cdot 4^1 \cdot 5^2 = 1 \cdot \frac{1}{2} \cdot 1 \cdot 4 \cdot 25 = 50 \,.$

We summarize the following rules when using sum signs:

Working with sums:

$$\begin{aligned}
\sum_{i=1}^{n}(a_i \pm b_i) &= \sum_{i=1}^{n} a_i \pm \sum_{i=1}^{n} b_i \;; \\
\sum_{i=1}^{n} c\,a_i &= c\sum_{i=1}^{n} a_i \;; \\
\sum_{i=1}^{n} c &= n\,c\;; \\
\sum_{i=1}^{m} a_i + \sum_{m+1}^{n} a_i &= \sum_{i=1}^{n} a_i \;.
\end{aligned}$$

Similarly, we get the following rules when using product signs:

Working with products:

$$\begin{aligned}
\prod_{i=1}^{n}(a_i\, b_i) &= \prod_{i=1}^{n} a_i \cdot \prod_{i=1}^{n} b_i \;; \\
\prod_{i=1}^{n} c\,a_i &= c^n \prod_{i=1}^{n} a_i \;; \\
\prod_{i=1}^{n} c &= c^n\;; \\
\prod_{i=1}^{m} a_i \cdot \prod_{m+1}^{n} a_i &= \prod_{i=1}^{n} a_i \;.
\end{aligned}$$

Example 1.7 *We rewrite*

$$S = \sum_{k=1}^{20}(k^3 + 3k - 2)$$

in separate terms and obtain

$$S = \sum_{k=1}^{20} k^3 + 3\sum_{k=1}^{20} k - 2 \cdot 20.$$

For a sum or a product, there exist several representations. So, one can transform the summation (multiplication) index. For instance, we have

$$a_5 + a_6 + a_7 + a_8 = \sum_{i=5}^{8} a_i = \sum_{k=1}^{4} a_{k+4} = \sum_{n=0}^{3} a_{n+5}$$

or as another example,

$$3 + 5 + 7 + 9 + 11 = \sum_{k=1}^{5}(2k+1) = \sum_{n=0}^{4}(2n+3)\;.$$

Observe again that the summation index can be chosen arbitrarily. In general, when shifting the summation or multiplication index, we have the following rules for all integers j and n:

$$\sum_{i=1}^{n} a_i = \sum_{k=-j+1}^{n-j} a_{k+j} ;$$
$$\prod_{i=1}^{n} a_i = \prod_{k=-j+1}^{n-j} a_{k+j} .$$

The above transformation of the summation and multiplication indices corresponds to a shifting according to $i = k + j$.

Example 1.8 *We compute*

$$S = \sum_{k=1}^{8} k^3 - \sum_{i=4}^{8} (i-2)^3.$$

This yields

$$\begin{aligned} S &= \sum_{k=1}^{8} k^3 - \sum_{j=2}^{6} j^3 \\ &= (1^3 + 2^3 + \ldots + 8^3) - (2^3 + 3^3 + \ldots + 6^3) \\ &= 1^3 + 7^3 + 8^3 = 1 + 343 + 512 = 856. \end{aligned}$$

It is worth noting that for $n > 1$, we have in general

$$\sum_{i=1}^{n} a_i\, b_i \neq \sum_{i=1}^{n} a_i \cdot \sum_{i=1}^{n} b_i. \tag{1.1}$$

Let e.g. $n = 3$. Then (1.1) can be written in detail as follows:

$$a_1 b_1 + a_2 b_2 + a_3 b_3 \neq (a_1 + a_2 + a_3) \cdot (b_1 + b_2 + b_3).$$

Moreover, we have in general

$$\prod_{i=1}^{n} (a_i + b_i) \neq \prod_{i=1}^{n} a_i + \prod_{i=1}^{n} b_i. \tag{1.2}$$

To illustrate (1.1) and (1.2), let again $n = 3$ and $a_1 = 2, a_2 = 5, a_3 = 3$ as well as $b_1 = 1, b_2 = 2, b_3 = 7$. Then we obtain

$$\sum_{i=1}^{3} a_i\, b_i \neq 2 + 10 + 21 = 33 \neq \sum_{i=1}^{3} a_i \cdot \sum_{i=1}^{3} b_i = (2+5+3) \cdot (1+2+7) = 10 \cdot 10 = 100$$

and

$$\prod_{i=1}^{n} (a_i + b_i) = (2+1) \cdot (5+2) \cdot (3+7) = 210 \neq \prod_{i=1}^{n} a_i + \prod_{i=1}^{n} b_i = (2 \cdot 5 \cdot 3) + (1 \cdot 2 \cdot 7) = 30 + 14 = 44.$$

Often one uses also **double sums** of the form

$$\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij},$$

where the inner summation is made over the second index j and the outer summation is made over the first index i. Thus, we have

$$\begin{aligned}\sum_{i=1}^{3}\sum_{j=1}^{4} a_{ij} \quad = \quad & a_{11} + a_{12} + a_{13} + a_{14} \\ & + a_{21} + a_{22} + a_{23} + a_{24} \\ & + a_{31} + a_{32} + a_{33} + a_{34}\,.\end{aligned}$$

Note that we can interchange the order of the summations and have the following:

$$\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} = \sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij}.$$

1.3 Proof by Induction

In this section, we review **proofs by induction** which can be used e.g. to prove specific sum or product formulas or also inequalities. This type of proof can e.g. be used when proving a particular property for all natural numbers $n \geq n_0$, i.e., for infinitely many numbers n. This proof consists of two steps.

In the **initial step** (also denoted as **base step**), we prove that this property holds for the first natural number $n = n_0$ considered. Often, one has $n_0 = 1$. Then, in the **inductive step**, we prove that, if the property holds for a particular value $n = k$ (this is the **induction assumption**), then it also holds for the succeeding value $n = k + 1$ (this is the **induction hypothesis**). This has to be proven only once. Now one can argue as follows. Since the property holds for the initial number n_0, it also holds by the inductive step for the next number $n_0 + 1$. Applying now again the inductive step, we can conclude that, since the property holds for $n_0 + 1$, it also holds for the next number $n_0 + 2$, and so on.

Example 1.9 *We prove that the equality*

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

holds for all natural numbers. In the initial step, for $n = 1$, we have

$$\sum_{i=1}^{1} i = 1 = \frac{1 \cdot 2}{2} = 1,$$

i.e., the above formula is correct. Using now the induction assumption that the formula is correct for some number $n = k$:

$$\sum_{i=1}^{k} i = \frac{k(k+1)}{2}, \tag{1.3}$$

we prove in the inductive step that the formula is also correct for the succeeding number $n = k+1$, i.e., we replace in the given formula k at every occurrence by $k + 1$ which means that we have to show that the equality

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

holds. Using equality (1.3), we obtain

$$\begin{aligned}
\sum_{i=1}^{k} i + (k+1) &= \frac{k(k+1)}{2} + k + 1 \\
\sum_{i=1}^{k+1} i &= \frac{k(k+1) + 2(k+1)}{2} \\
\sum_{i=1}^{k+1} i &= \frac{(k+1)(k+2)}{2}.
\end{aligned}$$

Thus, the given formula is correct for all natural numbers n.

Example 1.10 *We prove by induction that the equality*

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

holds for all natural numbers n. We begin with the initial step. For $n = 1$, we have

$$\sum_{i=1}^{1} q^{i-1} = q^{1-1} = q^0 = 1 = \frac{1-q^1}{1-q} = 1,$$

i.e., the above formula is correct. Thus, we can use the induction assumption that the above formula is correct for some $n = k$. Hence, we can use

$$\sum_{i=1}^{k} q^{i-1} = \frac{1-q^k}{1-q}. \tag{1.4}$$

In the inductive step, we have to show that this formula is also correct for the next natural number $n = k+1$, i.e., we prove that

$$\sum_{i=1}^{k+1} q^{i-1} = \frac{1-q^{k+1}}{1-q}.$$

Using equality (1.4), we obtain

$$\begin{aligned} \sum_{i=1}^{k} q^{i-1} + q^{(k+1)-1} &= \frac{1-q^k}{1-q} + q^{(k+1)-1} \\ \sum_{i=1}^{k+1} q^{i-1} &= \frac{1-q^k+q^k(1-q)}{1-q} \\ \sum_{i=1}^{k+1} q^{i-1} &= \frac{1-q^k+q^k-q^{k+1}}{1-q} \\ \sum_{i=1}^{k+1} q^{i-1} &= \frac{1-q^{k+1}}{1-q}. \end{aligned}$$

Therefore, the given formula is correct for all natural numbers n.

Example 1.11 *We prove by induction that the equality*

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6} \tag{1.5}$$

holds for all natural numbers. In the initial step, we obtain for $n = 1$

$$\sum_{i=1}^{1} i^2 = 1^2 = 1 = \frac{1 \cdot 2 \cdot 3}{6} = 1.$$

Thus, for the inductive step, we can assume that the above formula is correct for a specific natural number $n = k$:

$$\sum_{i=1}^{k} i^2 = \frac{k(k+1)(2k+1)}{6}, \tag{1.6}$$

and we have to show that it is then also correct for the succeeding natural number $k+1$, *i.e., we have to show:*

$$\sum_{i=1}^{k+1} i^2 = \frac{(k+1)[(k+1)+1][2(k+1)+1]}{6} = \frac{(k+1)(k+2)(2k+3)}{6}.$$

In the above equality, we replaced k *in formula (1.6) at every place it occurs by* $k+1$ *and simplified the resulting term. In the inductive step, using the correctness of equality (1.6), we obtain*

$$\begin{aligned}\sum_{i=1}^{k} i^2 + (k+1)^2 &= \frac{k(k+1)(2k+1)}{6} + (k+1)^2\\ \sum_{i=1}^{k+1} i^2 &= \frac{k(k+1)(2k+1) + 6(k+1)^2}{6}\\ \sum_{i=1}^{k+1} i^2 &= \frac{(k+1)\,[k(2k+1) + 6(k+1)]}{6}\\ \sum_{i=1}^{k+1} i^2 &= \frac{(k+1)\,(2k^2 + k + 6k + 6)}{6}\,.\end{aligned}$$

Now we rewrite the expression $k+6k$ *as* $4k+3k$ *so that we are able to factor out* $k+2$ *in two steps (see what we want to prove above). This yields:*

$$\begin{aligned}\sum_{i=1}^{k+1} i^2 &= \frac{(k+1)(2k^2 + 4k + 3k + 6)}{6}\\ \sum_{i=1}^{k+1} i^2 &= \frac{(k+1)[(k+2)2k + (k+2)3]}{6}\\ \sum_{i=1}^{k+1} i^2 &= \frac{(k+1)(k+2)(2k+3)}{6}\,.\end{aligned}$$

Thus, the given formula is correct for all natural numbers n.

Example 1.12 *We prove by induction that the inequality*

$$3^{n-1} \geq n$$

holds for all natural numbers n. *In the initial step, we get for* $n = 1$

$$3^{1-1} = 1 \geq 1.$$

In the inductive step, we can assume that

$$3^{k-1} \geq k \tag{1.7}$$

is correct for some natural number $n = k$, *and we have to show that it is also correct for* $n = k+1$*:*

$$3^k \geq k+1\,. \tag{1.8}$$

From inequality (1.7), it follows that

$$3^{k-1} \cdot 3 \geq k \cdot 3$$

and, due to $3k \geq k+1$, *we get*

$$3^k \geq 3k \geq k+1,$$

i.e., we obtain inequality (1.8).

EXERCISES

1.1 Let

$$A = \{x \mid x \text{ is a prime number }\} \quad \text{and} \quad B = \{x \mid x \text{ is integer and } 10 \leq x \leq 30\}.$$

Determine $A \cap B$; $A \setminus B$ and $B \setminus A$.

1.2 Let

$$I_1 = [-3, 4], \quad I_2 = (-1, 4) \quad \text{and} \quad I_3 = (3, 5).$$

Determine all possible unions, intersections and differences of any two of the above sets.

1.3 Let

$$X = \{a, b, d, e, f\}, \quad Y = \{b, e, g, h\} \quad \text{and} \quad Z = \{a, h, i\}$$

Determine $X \cup Y$; $X \cap Y$; $X \setminus Y$; $(X \cup Y) \cap Z$; $(Y \setminus Z) \cup X$.

1.4 Let X and Y be arbitrary sets. Determine

(a) $X \cup (X \cap Y)$; (b) $Y \cap (X \cup Y)$; (c) $(X \setminus Y) \cup (Y \setminus X) \cup (X \cap Y)$.

1.5 Let $a_i = 2i + 1$ and $b_i = (i+1)^2$ for $i = 1, 2, \ldots, 10$. Calculate

(a) $\displaystyle\sum_{i=1}^{10} a_i$; *(b)* $\displaystyle\sum_{i=3}^{8} (a_i + b_{i-2})$; *(c)* $\displaystyle\prod_{i=1}^{6} (a_i + b_{i+3})$.

1.6 Prove by induction that

$$\sum_{i=1}^{n} (2i - 1) = n^2$$

holds for all natural numbers n.

1.7 Prove by induction that

$$\sum_{i=0}^{n} (m + i) = \frac{(n+1)(2m+n)}{2}$$

holds for all integers $n \geq 0$, where m is an arbitrary fixed integer.

1.8 Prove by induction that the equality

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

holds for all natural numbers n.

1.9 Prove by induction that the inequality

$$nx + 1 \leq (x+1)^n$$

holds for all natural numbers n, where x is a positive number.

Chapter 2

Real Numbers and Arithmetic Operations

This chapter reviews the major arithmetic operations for working with real numbers. In particular, we give an overview about

- number systems,
- rules for working with absolute values and fractions and
- rules for working with powers, roots and logarithms.

2.1 Real Numbers

In this section, we start with the set $\mathbb{N}$ of **natural numbers**, i.e., $\mathbb{N} = \{1, 2, \ldots\}$. This set can be illustrated on a straight line known as the number line (see Fig. 2.1). If we union this set with number 0, we get the set

$$\mathbb{N}_0 = \mathbb{N} \cup \{0\} = \{0, 1, 2, 3, \ldots\}.$$

We can perform the operations of addition and multiplication within the set $\mathbb{N}$ of natural numbers, i.e., for $a, b \in \mathbb{N}$, we get that the number $a + b$ belongs to set $\mathbb{N}$ and the number $a \cdot b$ belongs to set $\mathbb{N}$. In other words, the sum and product of any two natural numbers is again a natural number. However, the difference and the quotient of two natural numbers are not necessarily a natural number.

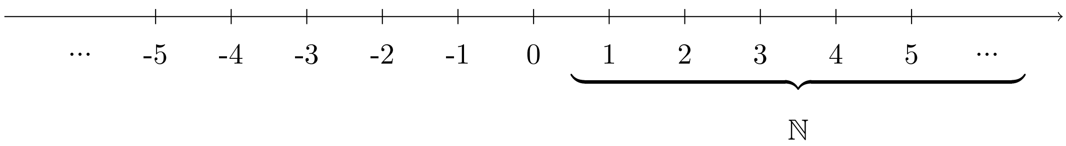


Figure 2.1: Integers on the number line

The first extension which we perform is to union the set of negative integers $\{-1, -2, -3, \ldots\}$ with the set $\mathbb{N}_0$ which yields the set $\mathbb{Z}$ of **integers** (see Fig. 2.1), i.e.,

$$\mathbb{Z} = \mathbb{N}_0 \cup \{-1, -2, -3, \ldots\}.$$

This allows us now to perform the three operations of addition, subtraction and multiplication within the set of integers, i.e., for $a, b \in \mathbb{Z}$, we get that the number $a + b$ belongs to set $\mathbb{Z}$, the number $a - b$ belongs to set $\mathbb{Z}$ and the number $a \cdot b$ belongs to set $\mathbb{Z}$.

To be able to perform the division of two integers, we introduce the set of all **fractions** p/q with $p \in \mathbb{Z}, q \in \mathbb{N}$. The union of the integers and the fractions is denoted as the set $\mathbb{Q}$ of **rational numbers**, i.e., we have

$$\mathbb{Q} = \mathbb{Z} \cup \left\{ \frac{p}{q} \mid p \in \mathbb{Z},\ q \in \mathbb{N} \right\}.$$

Now all the four elementary operations of addition, subtraction, multiplication and division (except by zero) can be performed within the set $\mathbb{Q}$ of rational numbers.

Consider next the equation $x^2 = 2$. This equation cannot be solved within the set of rational numbers, i.e., there exists no rational number p/q such that $(p/q)^2 = 2$. This leads to the extension of the set $\mathbb{Q}$ of rational numbers by the **irrational numbers**. These are numbers which cannot be written as the quotient of two integers. There are infinitely many irrational numbers, e.g.

$$\sqrt{2} \approx 1.41421, \quad \sqrt{3} \approx 1.73205, \quad e \approx 2.71828 \quad \text{and} \quad \pi \approx 3.14159.$$

Irrational numbers are characterized by decimal expansions that never end and by the fact that their digits have no repeating pattern (i.e., any irrational number cannot be presented as a periodic decimal number).

The union of the set $\mathbb{Q}$ of rational numbers and the set of all irrational numbers is denoted as the set $\mathbb{R}$ of **real numbers**. We have the following property: There is a one-to-one correspondence between real numbers and points on the number line, i.e., any real number corresponds to a point on the number line and vice versa. Within the set of real numbers, we can perform the operations of additions, subtraction, multiplication, division (except by zero), and we can also compute logarithms of positive real numbers and roots of non-negative real numbers. The stepwise extension of the set $\mathbb{N}$ of natural numbers to the set $\mathbb{R}$ of real numbers is illustrated in Fig. 2.2.

2.2 Basic Arithmetic Rules and the Absolute Value

In the following, we often deal with terms. A mathematical **term** is composed of letters, numbers and operational signs such as $+, -, \cdot, :$ or $\sqrt{\ }$. Mathematical terms are e.g.

$$ax + by - c, \qquad 17, \qquad (x+y)^3, \qquad 2x^3 - 4x + 7 \qquad \text{or} \qquad \sqrt{2x+5}.$$

We start this section with summarizing some well-known properties of real numbers. These properties are needed e.g. to transform mathematical terms or to solve equations.

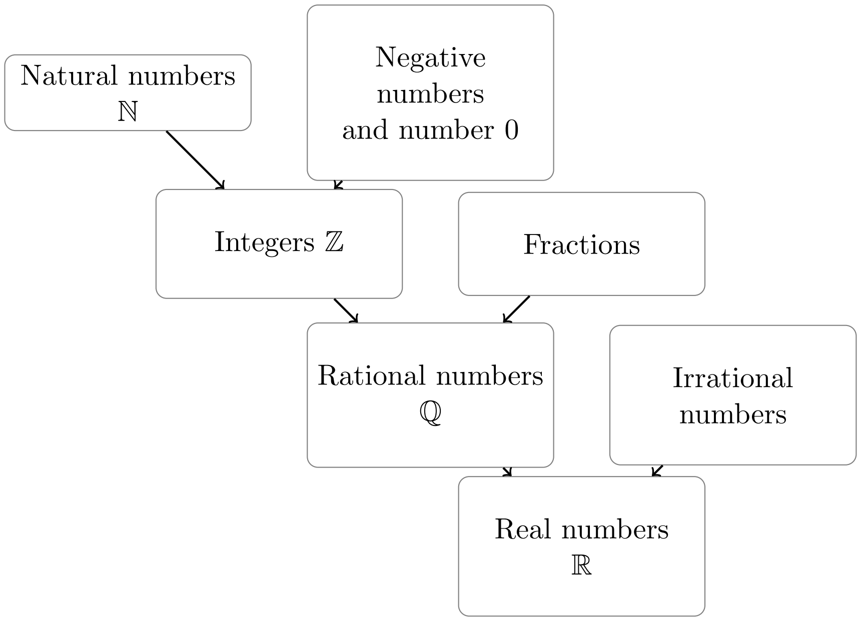


Figure 2.2: Number systems

Properties of real numbers with respect to addition ($a, b, c \in \mathbb{R}$):

1. $a + b = b + a$ (**commutative law of addition**);
2. there exists a number $0 \in \mathbb{R}$ such that for all a
   $a + 0 = 0 + a = a$;
3. for all a, b, there exists a number $x \in \mathbb{R}$ with
   $a + x = x + a = b$;
4. $a + (b + c) = (a + b) + c$ (**associative law of addition**).

Properties of real numbers with respect to multiplication ($a, b, c \in \mathbb{R}$):

1. $a \cdot b = b \cdot a$ (**commutative law of multiplication**);
2. there exists a number $1 \in \mathbb{R}$ such that for all a
   $a \cdot 1 = 1 \cdot a = a$;
3. for all a, b with $a \neq 0$, there exists a real number $x \in \mathbb{R}$ such that
   $a \cdot x = x \cdot a = b$;
4. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ (**associative law of multiplication**).

The number 0 is the neutral element with respect to addition (i.e., for any $a \in \mathbb{R}$ we have $a + 0 = a$) and number 1 is the neutral element with respect to multiplication (i.e., for any

$a \in \mathbb{R}$ we have $a = 1 \cdot a$).

Moreover, we have the following rules (which can e.g. be used to transform a given sum or difference into a product):

Distributive law and generalizations:

1. $(a+b) \cdot c = a \cdot c + b \cdot c$ (**distributive law**);
2. $a \cdot (b+c+d) = ab + ac + ad$;
3. $(a+b) \cdot (c+d) = ac + ad + bc + bd$;
4. $(a+b) \cdot (c-d) = ac - ad + bc - bd$.

We illustrate the process of **factorizing** a term by the following examples.

Example 2.1 *We get the following factor representations:*

a) $3a^4 + 9a^3 = 3a^3 \cdot (a+3)$*;*

b) $(a+2b) \cdot (s+3t) - (2s-t) \cdot (a+2b) = (a+2b) \cdot [(s+3t) - (2s-t)]$
$= (a+2b) \cdot (4t - s)$*;*

c) $4as - 10at + 12bs - 30bt = 4s(a+3b) - 10t(a+3b) = (a+3b) \cdot (4s - 10t)$
$= 2 \cdot (a+3b) \cdot (2s - 5t)$.
First, we have combined the first and third terms as well as the second and fourth terms. Then we have factored out the term $a + 3b$*.*

Moreover, for all real numbers a and b, we get the following **binomial formulas**:

Binomial formulas:

1. $(a+b)^2 = a^2 + 2ab + b^2$;
2. $(a-b)^2 = a^2 - 2ab + b^2$;
3. $(a+b) \cdot (a-b) = a^2 - b^2$.

Example 2.2 *Using the above binomial formulas, we get the following equalities:*

a) $(3x+4y)^2 = (3x)^2 + 2 \cdot 3x \cdot 4y + (4y)^2 = 9x^2 + 24xy + 16y^2$.

b) $(2x - ay)^2 = (2x)^2 - 2 \cdot 2x \cdot ay + (ay)^2 = 4x^2 - 4axy + a^2y^2$*;*

c) We wish to represent the difference

$$25s^2 - 3t^2$$

in product form. Using the third binomial formula with $a = 5s$ *and* $b = \sqrt{3}\,t$*, we get*

$$25s^2 - 3t^2 = (5s + \sqrt{3}\,t) \cdot (5s - \sqrt{3}\,t).$$

Next, we introduce the notion of the absolute value of a number, which can be used to express the distance of numbers or more general terms.

Absolute value:
Let $a \in \mathbb{R}$. Then

$$|a| = \begin{cases} a & \text{for } a \geq 0 \\ -a & \text{for } a < 0 \end{cases}$$

is called the **absolute value** of a.

From the above definition, it follows that the absolute value $|a|$ is always a non-negative number (note that, if $a < 0$, then $-a > 0$). The absolute value of a real number represents the distance of this number from point zero on the number line. For instance, we get $|3| = 3$, $|-5| = 5$ and $|0| = 0$. In the following, we review some properties of absolute values.

Properties of absolute values:
Let $a, b \in \mathbb{R}$ and $c \in \mathbb{R}_{\geq 0}$. Then:

1. $|-a| = |a|\,;$
2. $|a| \leq c \iff -c \leq a \leq c\,;$
3. $|a| > c \iff (a < -c) \text{ or } (a > c)\,;$
4. $\Big|\, |a| - |b| \,\Big| \leq |a+b| \leq |a| + |b|\,;$
5. $|a \cdot b| = |a| \cdot |b|\,;$
6. $\left|\dfrac{a}{b}\right| = \dfrac{|a|}{|b|}\,.$

According to the definition of the absolute value of a real number, we can give the following generalization.

Let $x, a \in \mathbb{R}$. Then

$$|x - a| = \begin{cases} x - a & \text{for } x \geq a \\ -(x-a) = a - x & \text{for } x < a \end{cases}$$

Thus, the non-negative value $|x - a|$ gives the **distance** of the number x from the number a. For instance, if $a = -2$ and $x = 6$, we get

$$|x - a| = |6 - (-2)| = 8,$$

i.e., the number 6 has a distance of eight units from the number -2. Moreover, if $a = -2$ and $x = -4$, we get

$$|x - a| = |-4 - (-2)| = |-4 + 2| = |-2| = 2,$$

i.e., the number -4 has a distance of two units from the number -2.

Example 2.3 *We simplify the term*

$$T = 6x + 2 \cdot |2a - 3x|$$

by 'removing' the absolute values and obtain

$$T = \begin{cases} 6x + 2 \cdot |2a - 3x| = 4a & \text{for } 2a - 3x \geq 0 \\ 6x - 2 \cdot (2a - 3x) = -4a + 12x & \text{for } 2a - 3x < 0 \end{cases} .$$

2.3 Calculations with Fractions

We recall that the term

$$\frac{a}{b}$$

is denoted as a **fraction** (note that in the text, we write a fraction also as a/b). Here $a \in \mathbb{Z}$ is known as the **numerator** and $b \in \mathbb{N}$ as the **denominator**. Thus, a fraction is only defined for $b \neq 0$, but we can also write a negative integer in the denominator. So we have

$$-\frac{3}{5} = \frac{-3}{5} = \frac{3}{-5} .$$

If $a < b$, then it is a **proper** fraction. If $a \geq b$, it is an **improper** fraction. Obviously, we have for $c \neq 0$ (and $b \neq 0$) the following equality:

$$\frac{a}{b} = \frac{a \cdot c}{b \cdot c}.$$

If we transform the left representation into the right one, we **expand** the fraction. If we transform the right representation into the left one, we **reduce** the fraction. Next, we review the rules for fractions.

Rules for working with fractions $(a, c \in \mathbb{Z},\ b, d \in \mathbb{N}\})$**:**

1. $\dfrac{a}{b} \pm \dfrac{c}{b} = \dfrac{a \pm c}{b}$;
2. $\dfrac{a}{b} \pm \dfrac{c}{d} = \dfrac{a\,d \pm b\,c}{b\,d}$;
3. $\dfrac{a}{b} \cdot \dfrac{c}{d} = \dfrac{a\,c}{b\,d}$;
4. $\dfrac{a}{b} : \dfrac{c}{d} = \dfrac{a\,d}{b\,c} \qquad (c \neq 0)$.

The term

$$\frac{\frac{a}{b}}{\frac{c}{d}}$$

is known as a **double fraction**. According to the above rules, we obtain for a double fraction

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

We illustrate the above rules by some examples.

Example 2.4 *We determine the following difference D of the given two fractions:*

$$D = \frac{3u-5v}{u+2v} - \frac{2u-v}{u+2v}.$$

Both denominators are equal and we obtain

$$D = \frac{(3u-5v)-(2u-v)}{u+2v} = \frac{u-4v}{u+2v}.$$

Adding or subtracting two fractions requires that they have the same denominator (see rule (1) above). If this is not the case, one has to determine a common denominator for both fractions. A simple way is to take as the denominator the product of both denominators (see rule (2) above). However, if several fractions have to be added or subtracted, the new dominator can be rather large. Another way is to take the **least common multiple** of the fractions in order to apply an analogue rule. We illustrate this by the following examples.

Example 2.5 *We determine the number*

$$N = \frac{8}{15} - \frac{5}{16} + \frac{1}{36}.$$

For all fractions, we determine the product representation by prime factors and the corresponding expansion factors (by which the corresponding numerator has to be multiplied so that the denominator is equal to the last common multiple). This yields:

prime factor representation	*expansion factors*
$15 = 3 \cdot 5$	$2^4 \cdot 3 = 48$
$16 = 2^4$	$3^2 \cdot 5 = 45$
$36 = 2^2 \cdot 3^2$	$2^2 \cdot 5 = 20$

The least common multiple (i.e., the common denominator for all fractions) is obtained as the product of the prime factors with the largest exponents occurring in the above prime factor representation. Thus, the least common multiple is

$$2^4 \cdot 3^2 \cdot 5 = 720,$$

from which the expansion factors for each of the three fractions given in the last column are obtained. Hence, we get

$$N = \frac{8 \cdot 48}{720} - \frac{5 \cdot 45}{720} + \frac{10 \cdot 20}{720} = \frac{384 - 225 + 200}{720} = \frac{359}{720}.$$

Example 2.6 *We want to compute the fraction*

$$F = \frac{1}{2ab} + \frac{a-2}{a^2+2ab} - \frac{b-1}{ab+2b^2},$$

where $a, b \neq 0$ and $a \neq -2b$. For all fractions, we again determine product representations and the corresponding expansion factors. This yields:

prime factor representation	*expansion factors*
$2ab$	$a+2b$
$a^2+2ab=a(a+2b)$	$2b$
$ab+2b^2=b(a+2b)$	$2a$

This yields the least common multiple

$$2\cdot a\cdot b\cdot(a+2b)$$

from which the above expansion factors result (see the right column). Hence, we obtain

$$\begin{aligned} F &= \frac{1\cdot(a+2b)}{2ab(a+2b)}+\frac{(a-2)\cdot 2b}{2ab(a+2b)}-\frac{(b-1)\cdot 2a}{2ab(a+2b)} \\ &= \frac{(a+2b)+(2ab-4b)-(2ab-2a)}{2ab(a+2b)} \\ &= \frac{3a-2b}{2ab(a+2b)} \end{aligned}$$

We note that the **greatest common divisor** can be found by means of these so-called prime factor representations. Another possibility is to use the **Euclidean algorithm**.

Example 2.7 *We compute*

$$DF=\frac{\frac{2}{3}+\frac{1}{6}}{\frac{4}{3}-\frac{1}{4}}$$

and transform the above term first into one double fraction. This yields

$$DF=\frac{\frac{4}{6}+\frac{1}{6}}{\frac{16}{12}-\frac{3}{12}}=\frac{\frac{5}{6}}{\frac{13}{12}}\,.$$

Thus, we obtain

$$DF=\frac{5}{6}\cdot\frac{12}{13}=\frac{10}{13}\,.$$

Example 2.8 *We want to simplify the double fraction*

$$DF=\frac{\frac{2}{x}-\frac{3}{y}}{\frac{1}{x}+\frac{2}{y}}$$

assuming that $x\neq 0, y\neq 0$ and $y\neq -2x$. This term can be transformed as follows:

$$DF=\frac{\frac{2y-3x}{xy}}{\frac{y+2x}{xy}}\,.$$

This yields

$$DF = \frac{2y-3x}{xy} \cdot \frac{xy}{y+2x} = \frac{2y-3x}{y+2x}\,.$$

Thus, if a double fraction has the same denominators, we get the fraction including both numerators of the double fraction.

Example 2.9 *We simplify the term*

$$DF = \frac{\dfrac{x+1}{x-1} - 1}{\dfrac{x+1}{x-1} + 1},$$

where $x \notin \{0,1\}$. *This yields*

$$DF = \frac{\dfrac{x+1-(x-1)}{x-1}}{\dfrac{x+1+(x-1)}{x-1}} = \frac{2}{2x} = \frac{1}{x}\,.$$

2.4 Calculations with Powers and Roots

First, we review some important rules for working with powers and roots. These rules are important e.g. for transforming terms or solving equations.

Rules for working powers $(a, b, m, n \in \mathbb{R})$**:**

1. $a^m \cdot a^n = a^{m+n}\,;$
2. $a^n \cdot b^n = (a\,b)^n\,;$
3. $\dfrac{a^m}{a^n} = a^{m-n} \qquad (a \neq 0);$
4. $\dfrac{a^n}{b^n} = \left(\dfrac{a}{b}\right)^n \qquad (b \neq 0)\,;$
5. $a^{-n} = \dfrac{1}{a^n} \qquad (a \neq 0)\,;$
6. $(a^n)^m = a^{n \cdot m} = (a^m)^n\,.$

Rules for working with roots $(a, b \in \mathbb{R}_{\geq 0}, m \in \mathbb{Z}, n \in \mathbb{N})$**:**

1. $\sqrt[n]{a \cdot b} = \sqrt[n]{a} \cdot \sqrt[n]{b}\,;$
2. $\sqrt[n]{\dfrac{a}{b}} = \dfrac{\sqrt[n]{a}}{\sqrt[n]{b}} \qquad (b \neq 0)\,;$
3. $\sqrt[n]{a^m} = a^{m/n} \qquad (a \neq 0 \text{ or } m/n > 0)\,.$

Using formula (3) for roots with $m = 1$, we obtain

$$\sqrt[n]{a} = a^{1/n} ,$$

i.e., a root can be written as a power with a rational exponent. Moreover, we can derive some further rules for working with roots by using the corresponding rules for powers applied to rational exponents:

1. $\sqrt[m]{a} \cdot \sqrt[n]{a} = a^{1/m} \cdot a^{1/n} = a^{1/m+1/n} = a^{(n+m)/mn} = \sqrt[mn]{a^{n+m}}$;
2. $\sqrt[m]{\sqrt[n]{a}} = \sqrt[m]{a^{1/n}} = (a^{1/n})^{1/m} = a^{1/nm} = \sqrt[nm]{a}$;
3. $\sqrt[n]{a^m} = (a^m)^{1/n} = a^{m/n} = (a^{1/n})^m = (\sqrt[n]{a})^m$.

Example 2.10 *First, we illustrate the use of powers and roots by the following computations.*

a) $2^4 \cdot 2^{-5} = 2^{4-5} = 2^{-1} = \frac{1}{2}$;

b) $4^3 \cdot 6^3 = (4 \cdot 6)^3 = 24^3 = 13,824$;

c) $(a^2x^3)^4 = (a^2)^4 \cdot (x^3)^4 = a^8 \cdot x^{12}$;

d) $\frac{x^{n+2} \cdot y^{3n-1}}{z^{2n+1}} \cdot \frac{x^{n-4} \cdot z^{2n-3}}{y^{2n+1}} = x^{(n+2)+(n-4)} \cdot y^{(3n-1)-(2n+1)} \cdot z^{(2n-3)-(2n+1)}$
$= x^{2n-2} \cdot y^{n-2} \cdot z^{-4} = \frac{x^{2n-2} \cdot y^{n-2}}{z^4}$;

e) $a^{-2/3} = \frac{1}{a^{2/3}} = \frac{1}{\sqrt[3]{a^2}}$;

f) $\sqrt{50} \cdot \sqrt{2} = \sqrt{50 \cdot 2} = \sqrt{100} = 10$;

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

h) $\sqrt{x^2} = |x|$
(notice that both x and $-x$ satisfy the above equality, but $|x| \geq 0$ by definition);

i) $(2\sqrt{3} + 3\sqrt{2}) \cdot (5\sqrt{2} - 4\sqrt{3}) = 10\sqrt{2}\sqrt{3} - 8(\sqrt{3})^2 + 15(\sqrt{2})^2 - 12\sqrt{2}\sqrt{3}$
$= 15 \cdot 2 - 8 \cdot 3 - 2\sqrt{2}\sqrt{3} = 6 - 2\sqrt{6}$;

j) $(\sqrt{14} + \sqrt{8})^2 = (\sqrt{14})^2 + (\sqrt{8})^2 + 2\sqrt{14} \cdot \sqrt{8} = 14 + 8 + 2\sqrt{14 \cdot 8} = 22 + 2\sqrt{112}$
$= 22 + 2\sqrt{16 \cdot 7} = 22 + 2 \cdot 4 \cdot \sqrt{7} = 22 + 8\sqrt{7}$;

k) $\sqrt[3]{4 + \sqrt{2}} \cdot \sqrt[3]{4 - \sqrt{2}} = \sqrt[3]{(4 + \sqrt{2}) \cdot (4 - \sqrt{2})} = \sqrt[3]{4^2 - (\sqrt{2})^2} = \sqrt[3]{16 - 2} = \sqrt[3]{14}$.

l) *We wish to write $a^{b+3} + a^{b-2} + a^b$ as a product. We can factor out the power with the lowest exponent. This gives $a^{b+3} + a^{b-2} + a^b = a^{b-2} \cdot (a^5 + 1 + a^2)$.*

Example 2.11 *We simplify the term*

$$T = \frac{(6a^2x)^3}{(5ax^2y)^4} \cdot \frac{(12ax^2y)^5}{(4a^3x)^2} \,.$$

Transforming the numbers occurring above into a product representation of prime numbers and combining both fractions into one fraction, we get

$$T = \frac{(2 \cdot 3 \cdot a^2 \cdot x)^3 \cdot (2^2 \cdot 3 \cdot a \cdot x^2 \cdot y)^5}{(5 \cdot a \cdot x^2 \cdot y)^4 \cdot (2^2 \cdot a^3 \cdot x)^2} \,.$$

Using now power rules and ordering the resulting terms according to common bases, we obtain

$$T = \frac{\left[2^3 \cdot 3^3 \cdot (a^2)^3 \cdot x^3\right] \cdot \left[(2^2)^5 \cdot 3^5 \cdot a^5 \cdot (x^2)^5 \cdot y^5\right]}{\left[(5^4 \cdot a^4 \cdot (x^2)^4 \cdot y^4\right] \cdot \left[(2^2)^2 \cdot (a^3)^2 \cdot x^2\right]}$$

Summarizing further all powers with the same base, we get

$$\begin{aligned} T &= \frac{2^{3+2\cdot 5} \cdot 3^{3+5} \cdot a^{2\cdot 3+5} \cdot x^{3+2\cdot 5} \cdot y^5}{2^{2\cdot 2} \cdot 5^4 \cdot a^{4+3\cdot 2} \cdot x^{2\cdot 4+2} \cdot y^4} \\ &= \frac{2^{13} \cdot 3^8 \cdot a^{11} \cdot x^{13} \cdot y^5}{2^4 \cdot 5^4 \cdot a^{10} \cdot x^{10} \cdot y^4} \\ &= \frac{2^{13-4} \cdot 3^8 \cdot a^{11-10} \cdot x^{13-10} \cdot y}{5^4} \\ &= \frac{2^9 \cdot 3^8}{5^4} \cdot a \cdot x^3 \cdot y \,. \end{aligned}$$

Example 2.12 *We simplify the following term:*

$$T = \left(\frac{a^2 - b^2}{x^2 - y^2}\right)^n \cdot \left(\frac{x - y}{a + b}\right)^n \,.$$

Applying power rules and the third binomial formula, we get

$$\begin{aligned} T &= \left(\frac{(a^2 - b^2) \cdot (x - y)}{(x^2 - y^2) \cdot (a + b)}\right)^n \\ &= \left(\frac{(a - b) \cdot (a + b) \cdot (x - y)}{(x - y) \cdot (x + y) \cdot (a + b)}\right)^n \\ &= \left(\frac{a - b}{x + y}\right)^n \,. \end{aligned}$$

Example 2.13 *We simplify the term*

$$T = \sqrt{\sqrt[5]{(a^2 - 2ab + b^2)}} \,.$$

Using the second binomial formula and the rules for working with roots, we obtain

$$T = \sqrt[5]{\sqrt{(a - b)^2}} = \sqrt[5]{|a - b|} \,.$$

If in a fraction the denominator includes roots, one usually transforms the corresponding terms in such a way that the denominator becomes rational. For a fraction N/D with

$$\frac{N}{D} = \frac{N}{\sqrt{a} \pm \sqrt{b}},$$

this can be done as follows:

$$\frac{N}{D} = \frac{N \cdot (\sqrt{a} \mp \sqrt{b})}{\sqrt{a^2} - \sqrt{b^2}} = \frac{N \cdot (\sqrt{a} \mp \sqrt{b})}{a - b}\ .$$

Example 2.14 *Let us consider the following examples for rationalizing the denominator.*

a) $\dfrac{a}{\sqrt{b}} = \dfrac{a\sqrt{b}}{(\sqrt{b})^2} = \dfrac{a}{b} \cdot \sqrt{b};$

b) $\dfrac{3}{5+\sqrt{2}} = \dfrac{3 \cdot (5-\sqrt{2})}{(5+\sqrt{2}) \cdot (5-\sqrt{2})} = \dfrac{15 - 3\sqrt{2}}{5^2 - (\sqrt{2})^2} = \dfrac{15 - 3\sqrt{2}}{23}\ ;$

c)
$$\begin{aligned}
\frac{10}{\sqrt{2}+\sqrt{3}-\sqrt{7}} &= \frac{10}{(\sqrt{2}+\sqrt{3})-\sqrt{7}} \\
&= \frac{10 \cdot (\sqrt{2}+\sqrt{3}+\sqrt{7})}{\left[(\sqrt{2}+\sqrt{3})-\sqrt{7}\right] \cdot \left[(\sqrt{2}+\sqrt{3})+\sqrt{7}\right]} \\
&= \frac{10 \cdot (\sqrt{2}+\sqrt{3}+\sqrt{7})}{\left(\sqrt{2}+\sqrt{3}\right)^2 - 7} = \frac{10 \cdot (\sqrt{2}+\sqrt{3}+\sqrt{7})}{(2+2\sqrt{6}+3)-7} = \frac{10 \cdot (\sqrt{2}+\sqrt{3}+\sqrt{7})}{2\sqrt{6}-2}; \\
&= \frac{10 \cdot (\sqrt{2}+\sqrt{3}+\sqrt{7}) \cdot (\sqrt{6}+1)}{2 \cdot (\sqrt{6}-1) \cdot (\sqrt{6}+1)} = \frac{10 \cdot (\sqrt{2}+\sqrt{3}+\sqrt{7}) \cdot (\sqrt{6}+1)}{2 \cdot (6-1)} \\
&= \frac{10}{2 \cdot 5} \cdot (\sqrt{2}+\sqrt{3}+\sqrt{7}+\sqrt{12}+\sqrt{18}+\sqrt{42}) \\
&= \sqrt{2}+\sqrt{3}+\sqrt{7}+3\sqrt{2}+2\sqrt{3}+\sqrt{42} = 4\sqrt{2}+3\sqrt{3}+\sqrt{7}+\sqrt{42}\ .
\end{aligned}$$

2.5 Calculations with Logarithms

Logarithm:
Let $a^x = b$ with $a, b > 0$ and $a \neq 1$. Then

$$x = \log_a b$$

is defined as the **logarithm** of (number) b to the base a.

Thus, the logarithm of b to the base a is the power to which one must raise a to yield b. As a consequence from the above definition of the logarithm, we have

$$a^{\log_a b} = b.$$

> **Rules for working with logarithms** $(a > 0,\ a \neq 1,\ x > 0,\ y > 0, n \in \mathbb{R})$:
>
> 1. $\log_a 1 = 0; \qquad \log_a a = 1;$
> 2. $\log_a(x \cdot y) = \log_a x + \log_a y;$
> 3. $\log_a \left(\frac{x}{y}\right) = \log_a x - \log_a y;$
> 4. $\log_a (x^n) = n \cdot \log_a x;$

When using a pocket calculator, often only logarithms to base e and 10 can be computed. If a logarithm to another base should be computed, one can use the change-of-base formula for logarithms:

> **Change-of-base formula for logarithms:**
>
> $$\log_b c = \frac{\log_a c}{\log_a b}.$$

As an example, if we wish to determine $\log_4 12$, we get

$$\log_4 12 = \frac{\lg 12}{\lg 4} \approx \frac{1.079181246}{0.602059991} \approx 1.792481251\ .$$

When using base e instead of base 10, we get

$$\log_4 12 = \frac{\ln 12}{\ln 4} \approx \frac{2.48490665}{1.386294361} \approx 1.792481251\ .$$

In the above computations, we have used the 9-digit-numbers obtained by a pocket calculator. These numbers have already been rounded because they are all irrational numbers. We illustrate the use of logarithms by some small examples.

Example 2.15 *The following logarithms can be immediately found by applying the definition of the logarithm.*

a) We get

$$\log_4 64 = 3$$

since

$$4^3 = 64\ .$$

b) We get

$$\log_{16} 4 = \frac{1}{2}$$

since

$$16^{1/2} = \sqrt{16} = 4\ .$$

c) We get

$$\log_{1/4}\ 4^3 = -3$$

since

$$\left(\frac{1}{4}\right)^{-3} = \frac{1}{4^{-3}} = 4^3\ .$$

Example 2.16 *We apply the rules for logarithms to write the following terms in an equivalent form:*

a) $\log_2(ab)^4 = 4 \cdot \log_2(ab) = 4 \cdot (\log_2 a + \log_2 b)$;

b) $\log_2(2a\sqrt{b+c}) = \log_2 2 + \log_2 a + \log_2 \sqrt{b+c} = 1 + \log_2 a + \log_2\ (b+c)^{1/2}$
$= 1 + \log_2 a + \frac{1}{2}\log_2(b+c)$;

c) $\ln \frac{1}{\sqrt[3]{ab}} = \ln 1 - \ln(ab)^{1/3} = 0 - \frac{1}{3}\ \ln(ab) = -\frac{1}{3}(\ln a + \ln b)$;

d) $\log_{10} \sqrt[4]{\frac{a^2}{b^3c}} = \log_{10}\left(\frac{a^2}{b^3c}\right)^{1/4} = \frac{1}{4} \cdot \left[\log_{10}\ a^2 - (\log_{10}\ b^3 + \log_{10}\ c)\right]$
$= \frac{1}{4} \cdot (2\log_{10} a - 3\log_{10} b - \log_{10} c)$;

e) The term

$$\log_a s + \log_b t$$

cannot be transformed since both logarithms have different bases. The term

$$\log_a(s + t^2)$$

cannot be transformed since there is no rule for the logarithm of a sum.

Example 2.17 *We transform the term*

$$L = \ln(a+b) + \ln(a-b) - \frac{1}{2} \cdot \ln(a^2 - b^2)$$

into one logarithmic term. Using the rules for logarithms, we get

$$\begin{aligned} L &= \ln \frac{(a+b)\cdot(a-b)}{(a^2-b^2)^{1/2}} = \ln \frac{a^2-b^2}{\sqrt{a^2-b^2}} = \ln \frac{\sqrt{a^2-b^2}\cdot\sqrt{a^2-b^2}}{\sqrt{a^2-b^2}} = \\ &= \ln \sqrt{a^2-b^2} = \ln(a^2-b^2)^{1/2} = \frac{1}{2}\cdot \ln(a^2-b^2)\,. \end{aligned}$$

EXERCISES

2.1 Calculate:
(a) $\left(2a - 5b^2\right)^2$; (b) $(x^2 - 3y)(x^2 + 3y)$.

2.2 Determine product representations for the following terms:
(a) $12a^3 + 4a^2$; (b) $3ax - 6ay - 15bx + 30by$; (c) $x^a - x^{a+3} - x^{a-1}$.

2.3 Determine all real numbers x with
(a) $|3x - 8| = 5$; (b) $|2x + 7| = 10$.

2.4 Calculate the fraction

$$\frac{97}{42} - \frac{5}{8} - \frac{43}{36}\,.$$

2.5 Combine into one fraction and simplify:

$$DF = \frac{1}{\frac{1}{a} - \frac{1}{a-1} - \frac{1}{a-2}}\,.$$

2.6 Combine into one fraction and simplify:

$$\frac{x}{x+y} + \frac{y}{x-y} + \frac{2xy}{x^2-y^2}\,.$$

2.7 Combine into one fraction and simplify:

$$F = \frac{u+6v}{u^2-3uv} - \frac{1}{2v} - \frac{9y-x}{2xy-6y^2}\,.$$

2.8 Combine into one fraction and simplify:

$$F = \frac{5(12-3x)}{9x^2-15xy} + \frac{5y-12}{3xy-5y^2}\,.$$

2.9 Calculate and simplify as much as possible:

$$\left(\frac{3}{x} - \frac{5}{y}\right) : \left(\frac{5}{x} - \frac{3}{y}\right)\,.$$

2.10 Calculate the number (without the use of a calculator)

$$\frac{(6^3)^3 \cdot (8^4)^2}{(12)^{12}}\,.$$

2.11 Calculate:

(a) $(x^2y^3z)^5$;　　(b) $\dfrac{(3ax^2)^2}{(4a^2x)^3} \cdot \dfrac{(ax)^3}{(5a^3x)^2}\,.$

2.12 Simplify

$$\frac{x^{n+1}y^{2n-1}}{z^{n-1}} : \frac{x^{n-2}y^{n+1}}{z^{n+3}}\,.$$

2.13 Simplify the following terms (without use of a calculator):

(a) $\dfrac{\sqrt{27^3}}{\sqrt[6]{3^5}}$;　　(b) $\sqrt{35}+\sqrt{21})^2$;　　(c) $\sqrt[3]{2+\sqrt{2}} \cdot \sqrt[3]{2-\sqrt{2}}\,.$

2.14 Rationalize the following fractions:

(a) $\dfrac{a}{\sqrt[5]{a}}$;　　(b) $\dfrac{2}{\sqrt{50}}$;　　(c) $\dfrac{1+x}{\sqrt{1-x}}$;　　(d) $\dfrac{a}{1-\sqrt{ax}}\,.$

2.15 Calculate the following logarithms:
(a) $\log_{1.03}\ 12$;　　(b) $\log_{1.1}\ 10$;　　(c) $\log_{1/2}\ 16\,.$

2.16 Calculate the following logarithms (without the use of a calculator):
(a) $\lg\ 20 + \lg\ 50$;　　(b) $\ln\ e^4 - \ln\ \sqrt{e}\,.$

Chapter 3

Equations

Equation:
Let T_1 and T_2 be two mathematical terms. Then

$$T_1 = T_2$$

is called an **equation**.

In general, finding all solutions of a given equation is a difficult but important problem in mathematics. For instance, often the zero of a function has to be determined. Here we consider several special classes of such equations and discuss some of their properties as well as how they can be solved. In particular, we review

- linear equations and systems of two linear equations with two variables;
- quadratic, root, logarithmic, exponential equations and proportions and
- the approximate solution of equations (without use of differentiation).

First, we introduce the notion of an equivalent transformation.

Equivalent transformations:
Equivalent transformations of mathematical terms are characterized by

- combining several terms on one side,
- adding or subtracting the same number or term on both sides or
- multiplying or dividing both sides by the same non-zero number or term.

3.1 Linear Equations

Linear equation:
An equation is called **linear** if it can be transformed by equivalent transformations into the form $ax + b = 0$ with $a \neq 0$.

For instance

$$3x + 7 = 0, \qquad \frac{1}{3}x = \frac{7}{11} \qquad \text{or} \qquad \frac{2}{x-1} = 5$$

are linear equations. In the latter case, it is assumed that $x - 1 \neq 0$. To transform this equation into the form $ax + b = 0$, we have to multiply both sides of the equation by the term $x - 1$. In the following, we often write the corresponding transformation after a vertical line on the right-hand side (i.e., $|\cdot(x-1)$ in the above case).

A linear equation is characterized by the occurrence of the first power of x, which is denoted as the **variable** or **unknown**. a and b are called the **parameters** or **constants**, and they are given fixed values. Since $a \neq 0$, we can solve the given linear equation for x and obtain the unique solution

$$x = -\frac{b}{a}.$$

Example 3.1 *Consider the equation*

$$27x - 10 = 15x + 6.$$

Solving for x, we obtain

$$\begin{aligned} 27x - 10 &= 15x + 6 && \mid -15x + 10 \\ 12x &= 16 && \mid : 12 \\ x &= \frac{16}{12} = \frac{4}{3}, \end{aligned}$$

i.e., $x = 4/3$ is the only solution of the given equation.

Example 3.2 *Let the equation*

$$4ax - 2ab^2 = 3ax - 3dx + 6b^2d$$

be given, where a, b, d are real parameters and it is assumed that $a \neq -3d$. This is a linear equation in the variable x because x occurs only in the first power (while parameter b occurs in the second power). We write all terms including the variable x on the left-hand side and all other terms on the right-hand side. This gives

$$\begin{aligned} 4ax - 3ax + 3dx &= 2ab^2 + 6b^2d \\ ax + 3dx &= 2ab^2 + 6b^2d \\ x(a + 3d) &= 2b^2(a + 3d) && \mid : (a + 3d) \neq 0 \\ x &= 2b^2. \end{aligned}$$

According to the given assumption $a \neq -3d$, the term $a+3d$ is different from zero and therefore, we can divide both sides in the second-to-last row by this non-zero number. We can easily test the correctness of our computations by substituting the solution obtained into the given equation. This yields

$$\begin{aligned} 4a \cdot 2b^2 - 2ab^2 &= 3a \cdot 2b^2 - 3d \cdot 2b^2 + 6b^2 d \\ 6ab^2 &= 6ab^2 - 6b^2 d + 6b^2 d. \end{aligned}$$

The above identity $6ab^2 = 6ab^2$ confirms the result obtained.

We continue with a few linear equations including fractions.

Example 3.3 *Consider the equation*

$$\frac{x+3}{8} - 1 = \frac{x+1}{3} + \frac{x-2}{4}.$$

First, we rewrite the above equation using the least common multiple of the numbers 3, 4 and 8, which is 24. This yields

$$\frac{3x+9}{24} - \frac{24}{24} = \frac{8x+8}{24} + \frac{6x-12}{24}$$

(see also Section 2.3). Combining the fractions on both sides and solving for x, we obtain:

$$\begin{aligned} \frac{3x-15}{24} &= \frac{14x-4}{24} && | \cdot 24 \\ 3x-15 &= 14x-4 && | -14x+15 \\ -11x &= 11 && |: (-1) \\ x &= -1. \end{aligned}$$

In the first row above, we can also use the property that two fractions with the same denominator are equal if the numerators are equal (this gives the second row). We can easily test that our computations are correct by substituting the result into the given equation:

$$\begin{aligned} \frac{-1+3}{8} - 1 &= \frac{-1+1}{3} + \frac{-1-2}{4} \\ \frac{2}{8} - \frac{8}{8} &= 0 - \frac{3}{4} \\ -\frac{3}{4} &= -\frac{3}{4}. \end{aligned}$$

Example 3.4 *Consider the linear equation*

$$\frac{x}{a} = \frac{b}{c}x + \frac{d}{c}.$$

with $a \neq 0$ and $c \neq 0$. We can combine the right-hand side into one fraction, which gives

$$\frac{x}{a} = \frac{bx+d}{c}.$$

After cross multiply, we obtain

$$cx = a \cdot (bx+d) = abx + ad.$$

Then, we get

$$\begin{aligned} cx - abx &= ad \\ (c - ab)x &= ad. \end{aligned}$$

When solving now for x, we have to distinguish three cases.

Case 1: $c - ab = 0$ *and* $ad \neq 0$.

In this case, there does not exist a solution because $0 \cdot x$ cannot be different from zero.

Case 2: $c - ab = 0$ *and* $ad = 0$.

In this case, any real number x satisfies the given equation, i.e., there exist infinitely many solutions.

Case 3: $c - ab \neq 0$.

In this case, there exists a unique solution, and we obtain

$$x = \frac{ad}{c - ab}\,.$$

Example 3.5 *If one adds a certain number to the numerator of the fraction $11/14$ and subtracts the half of the same number from the denominator, we obtain the fraction $17/11$. Which is the number we are looking for? Let us denote this number by x. We can setup an equation*

$$\frac{11 + x}{14 - \frac{x}{2}} = \frac{17}{11}.$$

Using cross multiply, this yields

$$\begin{aligned} 11 \cdot (11 + x) &= 17 \cdot \left(14 - \frac{x}{2}\right) \\ 121 + 11x &= 238 - 8.5x \\ 19.5x &= 117 \\ x &= 6\,, \end{aligned}$$

i.e., the number $x = 6$ satisfies the above property.

The next three examples consider some applications from physics, where linear equations occur.

Example 3.6 *A swimming pool has two inflows and one outflow. It is known that, if only the first inflow operates, the pool would be filled after 2 h. Moreover, it would also be filled after 3 h, if only the second inflow operates. On the other side, by means of the outflow the pool would be empty after 1.5 h when there would be no inflow.*

If both inflows and the outflow operate, what time is necessary to fill the pool completely? Let F denote the inflow per hour, V be the volume of the pool, and t the required time to fill the pool (in hours). The amount F_1 of the first inflow per hour is $F_1 = V/2$, the amount F_2 of the second inflow per hour is $F_2 = V/3$, and the amount F_3 of the outflow per hour is $F_3 = V/1.5$. Thus, the total inflow per hour into the pool is given by

$$F = F_1 + F_2 - F_3 = \frac{V}{2} + \frac{V}{3} - \frac{V}{1.5} = \frac{V}{2} + \frac{V}{3} - \frac{2V}{3} = \frac{3V + 2V - 4V}{6} = \frac{V}{6}.$$

Moreover, we have $F \cdot t = V$ *which gives*

$$\frac{V}{6} \cdot t = V$$

from which we obtain the solution $t = 6\ h$*, i.e., if both inflows and the outflow of the pool operate, it is completely filled after 6 h.*

Example 3.7 *A bus starts from Magdeburg to Munich with a constant speed of 95 km/h. At the same time, a truck starts from Munich to Magdeburg with a constant speed of 80 km/h. When and where will they meet?*

It is known from physics that the velocity v is defined as the quotient of distance s and time t:

$$v = \frac{s}{t}.$$

The distance between Magdeburg and Munich is equal to $s = 525\ km$*. Let* t *denote the time from the start of the bus and the truck until they meet,* s_1 *the driven distance of the bus starting from Magdeburg and* s_2 *be the distance of the truck starting from Munich till the meeting point. Then we have*

$$s_1 = v_1 \cdot t \qquad \text{and} \qquad s_2 = v_2 \cdot t\,.$$

Since $s = s_1 + s_2$*, we get*

$$s = v_1 \cdot t + v_2 \cdot t = (v_1 + v_2)\, t.$$

Solving for t*, we get*

$$t = \frac{s}{v_1 + v_2} = \frac{525\ km}{(95 + 80)\ km/h} = 3\ h.$$

Thus, the distances driven by the bus and the truck are

$$s_1 = 95\,\frac{km}{h} \cdot 3\ h\ = 285\ km \qquad \text{and} \qquad s_2 = 80\,\frac{km}{h} \cdot 3\ h\ = 240\ km\ .$$

This means that they meet on the highway near the German town Bayreuth which is approximately 230 km far from Munich.

Example 3.8 *We determine the time necessary for a car of 4.5 m length to drive past a truck of 15.5 m length when the car has a constant speed of 100 km/h and the truck has a constant speed of 80 km/h. Let* $v_C = 100\ km/h$ *be the speed of the car and* $l_C = 4.5\ m$ *be the length of the car,* $v_T = 80\ km/h$ *be the speed of the truck and* $l_T = 15.5\ m$ *be the length of the truck. During driving past, the distance* s_C *of the car is given by*

$$s_C = s_T + l_C + l_T = s_T + 4.5\ m + 15.5\ m = s_T + 20\ m,$$

where s_T *is the distance driven by the truck in time* t *required for driving past the truck. Then we have*

$$\frac{s_C}{v_C}\ = t\ =\ \frac{s_T}{v_T},$$

which corresponds to

$$\frac{s_T + 20\ m}{v_C} = \frac{s_T}{v_T}.$$

We determine s_T and obtain

$$\begin{aligned} \frac{s_T}{v_T} - \frac{s_T}{v_C} &= \frac{20m}{v_C} \\ s_T \cdot \left(\frac{1}{v_T} - \frac{1}{v_C} \right) &= \frac{20m}{v_C} \\ s_T \cdot \frac{v_C - v_T}{v_T \cdot v_C} &= \frac{20m}{v_C}, \end{aligned}$$

which yields

$$s_T = \frac{20\ m \cdot v_T}{v_C - v_T} = \frac{20\ m \cdot 80\ km/h}{20\ km/h} = 80\ m\ .$$

Thus, the distance driven by the car in time t is

$$s_T = 80\ m\ + 20\ m = 100\ m\ .$$

Consequently, the required time for diving past is

$$t = \frac{s_T}{v_T} = \frac{80\ m}{80\ km/h} = \frac{0.08\ km}{80\ km/h} = 0.001\ h\ .$$

Using that 1 h corresponds to 3,600 s, we get the time

$$t = 0.001 \cdot 3,600\ s = 3.6\ s,$$

i.e., 3.6 s are required for driving past the truck.

Finally, we generalize our considerations to the case of two linear equations with two variables x and y.

System of two linear equations with two variables:
Let the equations

$$\begin{aligned} a_1x + b_1y &= c_1 \\ a_2x + b_2y &= c_2 \end{aligned}$$

be given. This is called a **system of two linear equations with two variables**.

Such a system can be solved by eliminating one of the variables. If one wants to eliminate variable x, we can multiply the first equation by $-a_2/a_1$ and add it to the second equation. We get

$$\left(-b_1 \cdot \frac{a_2}{a_1} + b_2 \right) y = -c_1 \cdot \frac{a_2}{a_1} + c_2$$

from which we can determine y. Then we can **substitute** this result into one of the original equations, which gives a linear equation with the only variable x. For illustration, we consider the following three examples:

Example 3.9 *Let the system*

$$\begin{aligned} -x+3y &= 4 \\ 3x-y &= 12 \end{aligned}$$

be given. In order to eliminate x, we multiply the first equation by $-a_2/a_1 = -3/(-1) = 3$ and add it to the second equation, which yields

$$8y = 24$$

and therefore,

$$y = 3.$$

Substituting this result e.g. into the first equation, we get

$$\begin{aligned} -x+3\cdot 3 &= 4 \\ x &= 5. \end{aligned}$$

Example 3.10 *Consider the system*

$$\begin{aligned} 4x+3y &= -2 \\ 7x-5y &= 17. \end{aligned}$$

Instead of multiplying the first equation by $-a_2/a_1 = -7/4$ and adding it to the second equation, we can multiply the first equation by -7 and the second one by 4 (in this case we avoid fractions in the resulting equation) and after adding the resulting equations, we get

$$(-21-20)y = 14+68\,,$$

i.e.,

$$-41y = 82$$

and thus

$$y = -2.$$

Substituting this result for instance into the first equation, we get

$$4x = -2 - 3y = -2 - 3 \cdot (-2) = 4$$

and thus

$$x = 1.$$

A test confirms the correctness of our computations:

$$\begin{aligned} 4 \cdot 1 + 3 \cdot (-2) &= -2 \\ 7 \cdot 1 - 5 \cdot (-2) &= 17. \end{aligned}$$

Example 3.11 *Consider the system*

$$\begin{aligned} 2x - 3y &= 4 \\ -6x + 9y &= 6. \end{aligned}$$

Multiplying the first equation by $-a_2/a_1 = -(-6)/2 = 3$ *and adding it to the second one, we get*

$$0x + 0y = 18$$

This is a contradiction because the left-hand side is equal to zero while the right-hand side is different from zero. Therefore, the given system of linear equations has no solution.

In a similar solution approach, one can transform the equations so that both coincide in one side and then uses that the other side of both equations must also coincide. We only mention that a general procedure for solving systems of m linear equations with n variables is **Gaussian elimination**.

3.2 Quadratic Equations

Quadratic equation and normal form:
An equation of the form

$$ax^2 + bx + c = 0 \tag{3.1}$$

with $a \neq 0$ is called a **quadratic equation**. Dividing both sides by a gives the **normal form**

$$\mathbf{1}x^2 + px + q = 0$$

with $p = b/a$ and $q = c/a$.

For presenting all solutions, we have to distinguish three cases.

Case 1: $p^2/4 - q > 0$.

In this case, there exist two distinct real solutions

$$x_1 = -\frac{p}{2} + \sqrt{\frac{p^2}{4} - q} \qquad \text{and} \qquad x_2 = -\frac{p}{2} - \sqrt{\frac{p^2}{4} - q}.$$

Case 2: $p^2/4 - q = 0$.

In this case, there exists a real double solution

$$x_1 = x_2 = -\frac{p}{2}.$$

Case 3: $p^2/4 - q < 0$.

In this case, there does not exist a real solution of the given quadratic equation.

Example 3.12 *Consider the quadratic equation*

$$3x^2 + 6x - 9 = 0.$$

After dividing both sides by 3*, we get a quadratic equation in normal form with* $p = 2$ *and* $q = -3$*:*

$$x^2 + 2x - 3 = 0.$$

Since $p^2/4 - q = 1 - (-3) = 4 > 0$*, we get two distinct solutions*

$$x_1 = -1 + \sqrt{1+3} = -1 + \sqrt{4} = 1 \qquad \textit{and} \qquad x_2 = -1 - \sqrt{1+3} = -1 - \sqrt{4} = -3.$$

Example 3.13 *Let the quadratic equation*

$$x^2 + 10x + 25 = 0$$

be given. Since $p^2/4 - q = 100/4 - 25 = 0$*, there exists a double solution*

$$x_1 = x_2 = -5.$$

Example 3.14 *We solve the quadratic equation*

$$x^2 + 2x + 5 = 0.$$

Since $p = 2$ *and* $q = 5$, *we get* $p^2/4 - q = 1 - 5 = -4 < 0$ *and therefore, the above equation has no real solution.*

We only mention that in several books the solution formula for a quadratic equation refers to equation (3.1). In this case, the sign of the term $D = b^2 - 4ac$ decides about the number of real solutions. For instance, if $D > 0$, we get the two distinct real solutions

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

We finish with some quadratic equations, where the application of the above solution formula is not necessary.

Example 3.15 *Let*

$$4x^2 - 49 = 0$$

We can use the following transformations:

$$\begin{aligned} x^2 &= \frac{49}{4} \qquad\qquad \mid \sqrt{\ } \\ \sqrt{x^2} &= \sqrt{\frac{49}{4}} \\ |x| &= \frac{5}{2}\,. \end{aligned}$$

Note that we have to use the absolute value on the left-hand side since otherwise one solution gets lost. Therefore, we get the two solutions

$$x_1 = \frac{5}{2} \qquad \textit{and} \qquad x_2 = -\frac{5}{2}.$$

Example 3.16 *We consider the equation*

$$3x^2 + 12x = 0.$$

Factoring out $3x$, *we get*

$$3x(x + 4) = 0$$

which has the two solutions

$$x_1 = 0 \qquad \textit{and} \qquad x_2 = -4.$$

Example 3.17 *Consider the quadratic equation*

$$(3x + 7)^2 = 25$$

Taking the square roots on both sides we get

$$|3x + 7| = 5.$$

According to the definition of the absolute value, we have to consider two cases.

Case 1: $3x + 7 \geq 0$.

Then we have

$$|3x + 7| = 3x + 7 = 5$$

which gives the first solution

$$x_1 = -\frac{2}{3}.$$

Case 2: $3x + 7 < 0$.

Then we have

$$|3x + 7| = -(3x + 7) = -3x - 7 = 5$$

which gives the second solution

$$x_2 = -4.$$

A test confirms our computations. For $x_1 = -2/3$, *we obtain*

$$\left[3 \cdot \left(-\frac{2}{3}\right) + 7\right]^2 = 5^2 = 25.$$

For $x_2 = -4$, *we obtain*

$$[3 \cdot (-4) + 7]^2 = (-5)^2 = 25.$$

We continue with a generalization of quadratic equations, called a **biquadratic equation**. Such an equation is characterized by the occurrence of only the fourth and second powers of the variable x, and it has the form:

$$ax^4 + bx^2 + c = 0 \qquad (a \neq 0)\,.$$

In this case, we introduce a new variable z by means of the substitution $z = x^2$ and solve the resulting quadratic equation for the variable z. Then we substitute back and determine all solutions for the variable x. We consider the following example.

Example 3.18 *Consider the biquadratic equation*

$$x^4 - 2x^2 - 3 = 0.$$

After substituting $z = x^2$, *we obtain the quadratic equation*

$$z^2 - 2z - 3 = 0.$$

Applying the solution formula for a quadratic equation, we obtain

$$z_1 = 1 + \sqrt{1+3} = 3 \qquad \text{and} \qquad z_2 = 1 - \sqrt{1+3} = -1.$$

After substituting back, we obtain from $x^2 = z_1 = 3$ *the two solutions*

$$x_1 = \sqrt{3} \qquad \text{and} \qquad x_2 = -\sqrt{3}.$$

The equation $x^2 = z_2 = -1$ *has no real solution and therefore, there exist only the real solutions* x_1 *and* x_2.

Next, we consider an example which can be also reduced to a quadratic equation.

Example 3.19 *Given is the equation*

$$x + 4\sqrt{x} - 12 = 0\,.$$

Using the substitution $z = \sqrt{x}$*, we get the quadratic equation*

$$z^2 + 4z - 12 = 0$$

which has the two solutions

$$z_1 = -2 + \sqrt{4+12} = 2 \qquad \textit{and} \qquad z_2 = -2 - \sqrt{4+12} = -6\,.$$

Substituting back, for z_1 *we get* $\sqrt{x} = 2$ *which yields the solution* $x_1 = 4$*. For* z_2*, there does not exist a solution* x_2 *since we always have* $\sqrt{x} \neq -6$*. A test confirms that*

$$4 + 4 \cdot \sqrt{4} - 12 = 4 + 8 - 12 = 0.$$

Finally, we consider two systems of equations with two variables, where in one equation one variable can be eliminated and plugged in the other equation (so that a quadratic equation results).

Example 3.20 *Let the system of equations*

$$\begin{aligned} 2x + 5y &= -12 \\ x\,y &= 2 \end{aligned}$$

be given. From the second equation, we get for $x \neq 0$

$$y = \frac{2}{x}. \tag{3.2}$$

Substituting this into the first equation we get

$$2x + 5 \cdot \frac{2}{x} = -12.$$

Multiplying both sides by x *and dividing the resulting equation by 2, we get*

$$x^2 + 6x + 5 = 0.$$

This quadratic equation has the two solutions

$$x_1 = -3 + \sqrt{9-5} = -1 \qquad \textit{and} \qquad x_2 = -3 - \sqrt{9-5} = -5.$$

Using equation (3.2), we get the corresponding solutions

$$y_1 = \frac{2}{x_1} = -2 \qquad \textit{and} \qquad y_2 = \frac{2}{x_2} = -\frac{2}{5}.$$

Thus, the given system has the two solutions $(x_1, y_1) = (-1, -2)$ *and* $(x_2, y_2) = (-5, -2/5)$.

Example 3.21 *Let us consider the system*

$$\begin{aligned} x - 2y &= 0 \\ x^2 - xy + y^2 &= 48. \end{aligned}$$

From the first equation, we immediately get

$$x = 2y. \tag{3.3}$$

Substituting this into the second equation, we obtain

$$\begin{aligned}(2y)^2 - 2y \cdot y + y^2 &= 48\\ 4y^2 - 2y^2 + y^2 &= 48\\ y^2 &= 16\end{aligned}$$

and thus the solutions

$$y_1 = 4 \qquad \text{and} \qquad y_2 = -4.$$

Using equation (3.3), we get

$$x_1 = 2y_1 = 8 \qquad \text{and} \qquad x_2 = 2y_2 = -8.$$

Thus, the given system has the two solutions $(x_1, y_1) = (8, 4)$ *and* $(x_2, y_2) = (-8, -4)$.

3.3 Root Equations

Root equation:
An equation in which the variable x occurs under the root sign is called a **root equation**.

For instance,

$$\sqrt{2x+4}=10, \qquad \sqrt{3x}+\sqrt{6x-4}-\sqrt{x+2}=20 \qquad \text{or} \qquad \sqrt[3]{4x+3}-27=0$$

are root equations. One can solve such an equation by taking an appropriate power such that the corresponding roots are removed. For a root equation, we always have to test whether the solution(s) obtained indeed satisfy the given equation.

Example 3.22 *Given is the root equation*

$$\sqrt{3x+4}-7=0.$$

We can transform this equation as follows:

$$\begin{aligned} \sqrt{3x+4} &= 7 & &|\ (\ )^2 \\ 3x+4 &= 49 & &|\ -4 \\ 3x &= 45 & &|: 3 \\ x &= 15. \end{aligned}$$

Substituting the result into the given root equation, we get

$$\sqrt{3\cdot 15+4}-7=\sqrt{49}-7=0,$$

i.e., $x=15$ *is indeed a solution of the given root equation.*

Let us now consider the equation

$$\sqrt{3x+4}+7=0. \tag{3.4}$$

We get

$$\sqrt{3x+4}=-7.$$

After taking the square on both sides, we get the same result as in the calculations before. However, $x=15$ *is not a solution of the given equation (3.4) since*

$$\sqrt{3\cdot 15+4}+7=\sqrt{49}+7=14\neq 0.$$

It is immediately clear that equation (3.4) cannot have a solution since $\sqrt{3x+4}\geq 0$ *for all real numbers* x*.*

Example 3.23 *Next, we consider the root equation*

$$\sqrt{x-1}-\sqrt{2x-3}=\sqrt{3x-4}.$$

Taking the square on both sides, we get by applying the second binomial formula

$$(x-1)+(2x-3)-2\cdot\sqrt{x-1}\cdot\sqrt{2x-3}=3x-4,$$

which corresponds to

$$3x-4-2\cdot\sqrt{x-1}\cdot\sqrt{2x-3}=3x-4.$$

Subtracting $3x-4$ *on both sides, dividing the resulting equation by* -2 *and applying the rules for roots, we get*

$$\sqrt{(x-1)\cdot(2x-3)}=0$$

which has the two solutions

$$x_1 = 1 \qquad \textit{and} \qquad x_2 = \frac{3}{2}\,.$$

For $x_1 = 1$*, the term* $\sqrt{2x-3}$ *is not defined since* $2 \cdot 1 - 3 = -1 < 0$*. For* $x_2 = 3/2$*, we get the correct identity*

$$\sqrt{\frac{3}{2} - 1} - \sqrt{2 \cdot \frac{3}{2} - 3} = \sqrt{3 \cdot \frac{3}{2} - 4},$$

i.e.,

$$\sqrt{\frac{1}{2}} = \sqrt{\frac{1}{2}}.$$

Thus, $x_2 = 3/2$ *is the only solution of the given root equation.*

Example 3.24 *We determine all solutions of the root equation*

$$\sqrt{x + \sqrt{2x+11} + 2} = 2.$$

Taking the square on both sides, we get

$$x + \sqrt{2x+11} + 2 = 2^2 = 4,$$

which can be written as

$$\sqrt{2x+11} = 2 - x.$$

Taking now again the square on both sides, we obtain

$$2x + 11 = (2-x)^2 = 4 - 4x + x^2,$$

which corresponds to

$$x^2 - 6x - 7 = 0.$$

This quadratic equation has the two solutions

$$x_1 = 3 + \sqrt{9+7} = 7 \qquad \textit{and} \qquad x_2 = 3 - \sqrt{9+7} = -1.$$

We test whether both solutions indeed satisfy the given root equation. For $x_1 = 7$*, we obtain*

$$\sqrt{7 + \sqrt{2 \cdot 7 + 11} + 2} = \sqrt{7 + 5 + 2} = \sqrt{14} \neq 2.$$

For $x_2 = -1$*, we obtain*

$$\sqrt{-1 + \sqrt{2 \cdot (-1) + 11} + 2} = \sqrt{-1 + \sqrt{9} + 2} = \sqrt{4} = 2.$$

This, only $x_2 = -1$ *satisfies the given equation.*

Example 3.25 *Let us consider consider the root equation*

$$\sqrt[3]{17 - 3\sqrt{5x-11}} = 2.$$

Transforming this equation, we get

$$\begin{aligned} 17 - 3 \cdot \sqrt{5x - 11} &= 2^3 = 8 \\ -3 \cdot \sqrt{5x - 11} &= -9 \qquad | \, (\,)^2 \\ 9 \cdot (5x - 11) &= 81 \\ 5x - 11 &= 9 \\ x &= 4 \end{aligned}$$

A test confirms the correctness of our computations:

$$\sqrt[3]{17 - 3 \cdot \sqrt{5 \cdot 4 - 11}} = \sqrt[3]{17 - 3 \cdot \sqrt{9}} = \sqrt[3]{8} = 2.$$

Example 3.26 *Finally, we determine all solutions of the root equation*

$$\sqrt[4]{18 - 2 \cdot \sqrt[3]{3x - 2}} = 2.$$

We consecutively apply the rules for working with powers and obtain

$$\begin{aligned} 18 - 2 \cdot \sqrt[3]{3x - 2} &= 2^4 \\ 2 \cdot \sqrt[3]{3x - 2} &= 2 \qquad |: 2 \\ \sqrt[3]{3x - 2} &= 1 \qquad | \, (\,)^3 \\ 3x - 2 &= 1^3 = 1 \\ x &= 1. \end{aligned}$$

A test gives

$$\sqrt[4]{18 - 2 \cdot \sqrt[3]{3 \cdot 1 - 2}} = \sqrt[4]{18 - 2 \cdot \sqrt[3]{1}} = \sqrt[4]{16} = 2\,.$$

3.4 Logarithmic and Exponential Equations

Logarithmic equation:
An equation in which the variable x occurs either in the base or the numerus of a logarithm is called a **logarithmic equation**.

For instance,

$$\ln(2x - 7) = 0.5, \qquad \log_{2x} = 32 \qquad \text{or} \qquad \lg 3^{2x} + \lg 5^{x-1} - 3 = 0$$

are logarithmic equations. Simple logarithmic equations can be solved by writing the left-hand and right-hand sides as exponent with an appropriately chosen base.

Example 3.27 *Consider the logarithmic equation*

$$\log_2 \, 4x = 5$$

Due to $a^{\log_a b} = b$ *and the property that for an exponential term the equality* $a^{T_1} = a^{T_2}$ *is equivalent to the equality* $T_1 = T_2$ *(in the subsequent examples, we use such a property also for* *other terms such as logarithmic terms), we transform the above equation into an exponential term with base 2 and obtain:*

$$\begin{aligned} \mathbf{2}^{log_2\ 4x} &= \mathbf{2}^5 \\ 4x &= 32 \\ x &= 8. \end{aligned}$$

We can test the correctness of the result by substituting into the original equation and obtain

$$\log_2\ (4 \cdot 8) = \log_2\ 32 = \log_2\ 2^5 = 5.$$

Example 3.28 *We consider the logarithmic equation*

$$\lg 20^x + \lg 5^x = 4.$$

We apply the rules for working with logarithms and obtain

$$\begin{aligned} x \cdot \lg 20 + x \cdot \lg 5 &= 4 \\ x \cdot (\lg 20 + \lg 5) &= 4 \\ x \cdot \lg(20 \cdot 5) &= 4 \\ x \cdot \lg(10^2) &= 4 \\ x \cdot 2 \cdot \lg 10 &= 4 \\ x \cdot 2 &= 4 \\ x &= 2. \end{aligned}$$

A test confirms the correctness of our result:

$$\lg 20^2 + \lg 5^2 = \lg 400 + \lg 25 = \lg 10,000 = \lg 10^4 = 4.$$

Example 3.29 *Let us consider the logarithmic equation*

$$\ln x^2 = \ln(a^3 - b^3) - \ln(a - b) - \ln(a^2 + ab + b^2).$$

We assume that the positive parameters a *and* b *satisfy the inequalities* $a < b$ *and* $a^3 > b^3$*. Under these assumptions, all logarithms above are defined. We simplify the right-hand side and obtain*

$$\begin{aligned} \ln x^2 &= \ln \frac{a^3 - b^3}{(a-b) \cdot (a^2 + ab + b^2)} \\ x^2 &= \frac{a^3 - b^3}{(a-b) \cdot (a^2 + ab + b^2)} \\ x^2 &= \frac{a^3 - b^3}{a^3 - b^3} \\ x^2 &= 1. \end{aligned}$$

From the last equation we obtain the two solutions $x_1 = 1$ *and* $x_2 = -1$*. Both solutions satisfy the given equation: In both cases the left-hand side has the value* $\ln 1 = 0$ *which is equal to the value of the right-hand side.*

Example 3.30 *Consider the logarithmic equation*

$$2\ln(x-1) = \ln(x+11).$$

We can use the rules for logarithms and rewrite the equation as an exponential term with base e. This yields:

$$\begin{aligned} \ln(x-1)^2 &= \ln(x+11) \\ (x-1)^2 &= x+11 \\ x^2-2x+1 &= x+11 \\ x^2-3x-10 &= 0\,. \end{aligned}$$

We note that the second equation above can also be skipped since from $\ln y = \ln z$*, it immediately follows that* $y = z$*. The quadratic equation above has the two solutions*

$$x_1 = \frac{3}{2} + \sqrt{\frac{9}{4}+10} = \frac{3}{2} + \sqrt{\frac{49}{4}} = \frac{3}{2} + \frac{7}{2} = 5$$

and

$$x_2 = \frac{3}{2} - \sqrt{\frac{9}{4}+10} = \frac{3}{2} - \sqrt{\frac{49}{4}} = \frac{3}{2} - \frac{7}{2} = -2.$$

We have to test whether x_1 *and* x_2 *are indeed solutions of the given equation. For* $x_1 = 5$*, we obtain*

$$2\ln(5-1) = \ln(5-1)^2 = \ln 16 = \ln(5+11).$$

However, for $x_2 = -2$ *the term* $\ln(x-1) = \ln(-3)$ *is not defined. Therefore,* $x_1 = 5$ *is the only solution of the given logarithmic equation.*

Exponential equation:
An equation in which the variable x occurs in the exponent is called an **exponential equation**.

For instance,

$$3^{x+4} = 20, \qquad a^x = b^{3x+1} \qquad \text{or} \qquad \sqrt[x]{2.5} = 12$$

are exponential equations. Only simple exponential equations can be solved, often one has to look for approximative solutions. Simple exponential equations can be solved by taking on both sides of the equation the logarithm to an appropriate base.

Example 3.31 *Consider the exponential equation*

$$-3 \cdot 2^x + 4^{x/2+1} = 2 \cdot 3^{x+1} - 3^x.$$

We rewrite the above equation in such a form that on the left-hand side only exponential terms with power 2 *and on the right-hand side only exponential terms with power* 3 *appear. Using* $2^2 = 4$, *we get*

$$-3 \cdot 2^x + (2^2)^{x/2+1} = 2 \cdot 3 \cdot 3^x - 3^x,$$

which gives

$$-3 \cdot 2^x + 2^{2 \cdot (x/2+1)} = 5 \cdot 3^x.$$

Moreover, using

$$2^{2(x/2+1)} = 2^{x+2} = 2^x \cdot 2^2 = 4 \cdot 2^x,$$

we get

$$\begin{aligned}
-3\cdot 2^x + 4\cdot 2^x &= 5\cdot 3^x \\
1\cdot 2^x &= 5\cdot 3^x \\
\frac{2^x}{3^x} &= 5 \\
\left(\frac{2}{3}\right)^x &= 5 \qquad \mid \ln(\,) \\
\ln\left(\frac{2}{3}\right)^x = x\ln\frac{2}{3} &= \ln 5 \\
x &= \frac{\ln 5}{\ln 2 - \ln 3} \\
x &= -3.96936,
\end{aligned}$$

which indeed satisfies (approximately) the original equation (note that rounding differences may occur):

$$\begin{aligned}
-3\cdot 2^{-3.96936} + 4^{-3.96936/2+1} &\approx 2\cdot 3^{-3.96936+1} - 3^{-3.96936} \\
-0.191525 + 0.255366 &\approx 0.0776610 - 0.012768 \\
0.063841 &\approx 0.063842.
\end{aligned}$$

Example 3.32 *We determine all solutions of the exponential equation*

$$\sqrt[x+1]{a^{x-2}} = \sqrt[x+5]{a^{x-1}}.$$

Transforming this equation, we get

$$\begin{aligned} a^{(x-2)/(x+1)} &= a^{(x-1)/(x+5)} \\ \frac{x-2}{x+1} &= \frac{x-1}{x+5} \\ (x-2)\cdot(x+5) &= (x-1)\cdot(x+1) \\ x^2+3x-10 &= x^2-1 \\ 3x &= 9 \\ x &= 3. \end{aligned}$$

A test confirms

$$\sqrt[3+1]{a^{3-2}} = \sqrt[4]{a} = a^{1/4}$$

and

$$\sqrt[3+5]{a^{3-1}} = \sqrt[8]{a^2} = a^{2/8} = a^{1/4},$$

i.e., both sides of the given equation have the same value $a^{1/4}$.

Example 3.33 *Let the exponential equation*

$$12 - 3e^{2x} - 9e^x = 0$$

be given. We can transform this equation into a quadratic equation by means of the substitution $z = e^x$ *by using that* $e^{2x} = (e^x)^2$. *Dividing the resulting equation by* (-3), *we get*

$$z^2 + 3z - 4 = 0.$$

This quadratic equation has the two solutions

$$z_1 = -\frac{3}{2} + \sqrt{\frac{9}{4} + 4} = -\frac{3}{2} + \frac{5}{2} = 1 \qquad \text{and} \qquad z_2 = -\frac{3}{2} - \sqrt{\frac{9}{4} + 4} = -\frac{3}{2} - \frac{5}{2} = -4.$$

Substituting back, we get for $z_1 = 1$ *the equation* $e^x = 1$ *which gives the solution*

$$x_1 = \ln 1 = 0.$$

For $z_2 = -4$*, the resulting equation* $e^x = -4$ *has no solution since* e^x *is positive for all real numbers* x*. Therefore,* $x_1 = 0$ *is the only solution of the given exponential equation.*

3.5 Proportions

Proportion:
If the two ratios $a : b$ and $c : d$ have the same value, then

$$a : b = c : d \tag{3.5}$$

with $b \neq 0$ and $d \neq 0$ is called a **proportion**.

If proportion (3.5) holds, then also the proportions

$$d : c = b : a, \qquad a : c = b : d \qquad \text{and} \qquad c : a = d : b$$

hold. Moreover, proportion (3.5) is equivalent to the equalities

$$a = c \cdot f \qquad \text{and} \qquad b = d \cdot f,$$

where f is denoted as the **factor of proportionality**. We illustrate the use of proportions by the following examples.

Example 3.34 *A car consumes 6.25 litres of fuel per 100 km. How many kilometres can the car go if the tank of this car is filled up with 48 litres. Thus, the ratios that the car consumes 6.25 liters per 100 km and 49 litres per x km are equal. We can establish the following proportion:*

$$6.25 : 100 = 48 : x.$$

This gives

$$\begin{aligned} 6.25 \cdot x &= 48 \cdot 100 \\ x &= \frac{4,810}{6.25} = 768. \end{aligned}$$

This means that the car can go 768 kilometres with a tank filled up with 48 litres.

Example 3.35 *In an examination 6 students of a group fail. This corresponds to 18.75 % of the participants in the examination. How many students took part in the examination? This means that the ratio* $6 : x$ *is the same as* $18.75 : 100$*. We obtain the following proportion:*

$$6 : x = 18.75 : 100$$

Solving for x, we get

$$\begin{aligned} 18.75 \cdot x &= 6 \cdot 100 \\ x &= \frac{600}{18.75} = 32, \end{aligned}$$

i.e., 32 students took part in the examination.

Example 3.36 *A way of a length of 6.56 m should be divided according to the ratio 5 : 3. Let x and y denote the lengths of the two parts, i.e., x and y have to satisfy the proportion*

$$x : y = 5 : 3 \ .$$

Using $x + y = 6.56\ m$*, we get*

$$\frac{x}{y} = \frac{6.56\ m\ - y}{y} = \frac{5}{3}\ .$$

Now, by cross multiply, we get

$$5y = 3 \cdot (6.56\ m\ - y)$$

from which we obtain

$$y = \frac{3 \cdot 6.56\ m}{8} = 2.46\ m$$

and

$$x = 6.56\ m\ - 2.46\ m\ = 4.10\ m\ .$$

Example 3.37 *A wire with a length $l_1 = 100\ m$ and a diameter $d_1 = 3\ mm$ has the weight $w_1 = 8\ kg$. A second wire made from the same material has the diameter $d_2 = 5\ mm$ and the weight $w_2 = 20\ kg$. How long is the second wire?*

From physics it is known that the ratio of the weights of the wires is identical to the ratio of the volumes $V_1 : V_2$, i.e., we obtain

$$\begin{aligned} w_1 : w_2 &= V_1 : V_2 \\ w_1 : w_2 &= \frac{d_1^2}{4}\,\pi\, l_1 \ : \ \frac{d_2^2}{4}\,\pi\, l_2 \\ \frac{w_1}{w_2} &= \frac{d_1^2 \cdot l_1}{d_2^2 \cdot l_2} \end{aligned}$$

Solving for l_2, we get

$$l_2 = \frac{w_2\, d_1^2\, l_1}{w_1\, d_2^2} = \frac{20\ kg\ \cdot 3^2\ mm^2 \cdot 100\ m}{25\ mm^2 \cdot 8\ kg} = 90\ m\ ,$$

i.e., the second wire has a length of $90\ m$.

3.6 Approximate Solution of Equations

Finding the zeroes of a function (see also Chapter 8) is often a hard problem in mathematics which leads to the solution of an equation. If one is unable to find exactly the zeroes, one can apply numerical procedures for finding a zero approximately. There are several procedures for finding zeroes approximately without use of differential calculus, e.g. **interval nesting** and the use of the **fixed-point theorem**. Here we describe a procedure known as **Regula falsi**, which also does not use derivatives of a function. In this case, we approximate the continuous function f between $x_0 = a$ and $x_1 = b$ with $f(a) \cdot f(b) < 0$ by a straight line through the points $(a, f(a))$ and $(b, f(b))$. This yields:

$$y - f(a) = \frac{f(b) - f(a)}{b - a} \cdot (x - a).$$

Since we look for a zero $\overline{x}$ of function f, we set $y = 0$, solve the equation for x and use the result as an approximate value $(x_2 = x)$

$$x_2 = a - f(a) \cdot \frac{b - a}{f(b) - f(a)}$$

for the zero. In general, we have $f(x_2) \neq 0$ (in the other case, we have found the exact value of the zero). Now we check which of both closed intervals $[a, x_2]$ and $[x_2, b]$ contains the zero. If $f(a) \cdot f(x_2) < 0$, then there exists a zero in the interval $[a, x_2]$. We replace b by x_2, and determine a new approximate value x_3. Otherwise, i.e., if $f(x_2) \cdot f(b) < 0$, then interval $[x_2, b]$ contains a zero. In this case, we replace a by x_2 and determine then an approximate value x_3, too. Continuing in this way, we get a sequence $\{x_n\}$ converging to a zero $\overline{x} \in [a, b]$. This procedure is illustrated in Fig. 3.1.

Example 3.38 *Let a function f be given with*

$$f(x) = \lg x - x + 3\,.$$

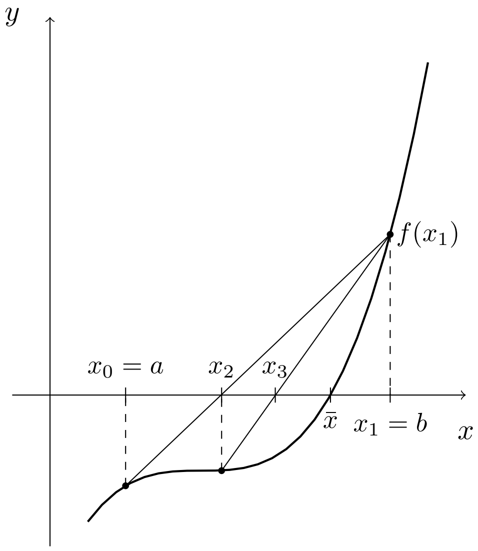


Figure 3.1: Illustration of Regula falsi

We are looking for one real zero (i.e., we wish to solve the equation $\lg x - x + 3 = 0$*). Since* $f(3) = 0.47712 > 0$ *and* $f(4) = -0.39794$ *(and function* f *is continuous), there must be a zero in the interval* $(3, 4)$*. Letting* $x_0 = a = 3$ *and* $x_1 = b = 4$*, we get the results presented in Table 3.1 with a precision of three decimal places.*

Table 3.1: Application of Regula falsi

n	x_n	$f(x_n)$
0	3	0.47712
1	4	−0.39794
2	3.5452	0.0044
3	3.5502	0.00005
4	**3.550**	

EXERCISES

3.1 Determine the solutions of the following linear equations:

(a) $2x + \frac{5}{4} = \frac{3x+5}{2}\,;$ (b) $\frac{3x+1}{4} - 2 = 3 - \frac{x+2}{3} + \frac{x+4}{5}\,;$

(c) $\frac{3x+4}{5} + \frac{x+6}{4} - 4 = 0\,;$ (d) $a - 1 + \frac{1}{x} = \frac{a}{x}\,;$

(e) $\frac{2}{3} + \frac{1}{x} = \frac{1}{\frac{3}{2} - \frac{1}{6}}\,;$ (f) $\frac{a^2x+b}{a} - \frac{a(1-x)}{b} = 2 \quad (ab \neq 0)\,.$

3.2 Solve the following equation for g:

$$s = v_0\,t - \frac{1}{2}\,g\,t^2\,.$$

3.3 Determine all solutions of the following systems of linear equations:

(a) $\begin{array}{rcrcr} x & + & 2y & = & -5 \\ 4x & - & y & = & 7 \end{array}$;

(b) $\begin{array}{rcrcr} x & - & 3y & = & 4 \\ -2x & + & 6y & = & -2 \end{array}$.

3.4 Determine all real solutions of the following equations:

(a) $x^2 + x - 12 = 0$;

(b) $(ax + 2b)(a^2 - bx) = 0$;

(c) $x^2 + 2x + 17 = 0$;

(d) $5x^2 - 10x = 40$;

(e) $ax^2 - 1 = b + x^2$;

(f) $a(b - x) + x^2 - bx = (a - x) \cdot x$;

(g) $x^4 + 3x^2 - 10 = 0$;

(h) $x - \frac{21}{2}\sqrt{x} + 5 = 0$.

3.5 Find the quadratic equation having the following roots:

(a) $x_1 = 5$ and $x_2 = -6$;

(b) $x_1 = 2 + \sqrt{5}$ and $x_2 = 2 - \sqrt{5}$.

3.6 Determine all solutions of the equation

$$x^2 + \frac{a}{x} + a = 0 \qquad (a \neq 0)\,.$$

For which values of a does the equation have real solutions?

3.7 Determine all solutions of the following systems of equations:

(a) $\begin{array}{rcrcr} x^2 & + & y^2 & = & 13 \\ & & \dfrac{x}{y} & = & \dfrac{2}{3} \end{array}$;

(b) $\begin{array}{rcr} (x-2)\cdot(y-2) & = & 3 \\ x + y & = & 8 \end{array}$.

3.8 Determine all solutions of the following root equations:

(a) $\sqrt{x-2} + \sqrt{x+14} = 8$;

(b) $\frac{1 - \sqrt{x}}{1 + \sqrt{x}} = \frac{1}{2}$;

(c) $(13 - \sqrt{x}) \cdot (6 - \sqrt{x}) = x + 2$;

(d) $\sqrt{x} + 1 = \frac{a}{b} \cdot (\sqrt{x} - 1) \quad (b \neq 0)$;

(e) $\sqrt{a - x} + \frac{b - a}{\sqrt{b - x}} = \sqrt{b - x} \qquad (x \leq a,\ x < b)$.

3.9 Determine all solutions of the following logarithmic equations:

(a) $\log_4(3x+1) = 2\,;$

(b) $2\ln x = 4\ln 3\,;$

(c) $\lg x^7 - 12 = \lg x^3\,;$

(d) $\lg 25^x + \lg 4^x = 6\,;$

(e) $\log_2(x+4) - \log_2(2x-1) = 1\,;$

(f) $\ln\left(\frac{8}{x+4} - 1\right) + \ln(x+4) = \ln\left(\frac{16}{4-x}\right).$

3.10 Determine all solutions of the following exponential equations:

(a) $2^{4x-5} - 128 = 0\,;$

(b) $4^x - \frac{1}{256} = 0\,;$

(c) $2^{3(x-2)} = 8^{1-2x}\,;$

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

(e) $a^4 \cdot a^{2x-1} = \frac{a \cdot a^{x+2}}{a^{2x}}\,;$

(f) $\sqrt{a^{3-x}} = a^{3x+4}\,;$

(g) $\sqrt[x+1]{a^{3x-1}} = \sqrt[5-x]{a^{x+3}} \quad (a > 0)\,;$

(h) $\sqrt[x]{a} = 10b \quad (a > 0)\,.$

3.11 What is the interest payment if an amount of 2,500 EUR is given to a bank for a period of 110 days when the interest rate per year, i.e., for 365 days, is equal to 5 %?

3.12 A bus with a constant speed required 75 minutes to drive a distance of 120 km. If the speed does not change, what time is needed for a distance of 280 km?

3.13 A 3-day trip of a length of 456 km should be divided into three daily sub-trips according to the ratio 5 : 4 : 3. What are the corresponding distances for the particular days?

3.14 Determine a solution of the equation

$$x^4 + 2x^3 - 4x^2 - 3 = 0$$

between $a = 1$ and $b = 2$ approximately by applying Regula falsi.

3.15 Determine a solution of the equation

$$3x - \ln x - 14 = 0$$

between $a = 4$ and $b = 6$ approximately by applying Regula falsi.

Chapter 4

Inequalities

Inequalities are often obtained when comparing the values of different mathematical terms. Often practical problems do not lead to equations but to inequalities. For instance, the amount of an available resource or a financial budget may not be exceeded by the chosen production program but it is not required to fully use the resource or the budget. Inequalities play also often a role in estimations e.g. of the maximal possible error, or they may occur when determining the domain of a function. The learning objectives of this chapter are

- to review the rules for working with inequalities and
- to discuss the most important types of inequalities.

Similar to equations, the main focus is on the determination of all feasible solutions of linear and quadratic inequalities as well as inequalities with absolute values. We will also discuss the solution of some more general inequalities.

4.1 Basic Rules

Inequality:
An **inequality** compares two mathematical terms by one of the inequality signs
$<$ (smaller than), $>$ (greater than), $\leq$ (less than or equal to) and $\geq$ (greater than or equal to), respectively.

For instance, the mathematical terms T_1 and T_2 must satisfy the inequality

$$T_1 \leq T_2.$$

For working with such inequalities, we first review the most important rules. We assume that a, b, c and d are real numbers.

Rules for inequalities:

1. If $a < b$, then
$$a + c < b + c, \qquad a - c < b - c;$$
2. if $a < b$ and $b < c$, then $a < c$;
3. if $a < b$ and $c > 0$, then
$$ac < bc, \qquad \frac{a}{c} < \frac{b}{c};$$
4. if $a < b$ and $c < 0$, then
$$ac > bc, \qquad \frac{a}{c} > \frac{b}{c};$$
5. if $a < b$ and $c < d$, then $a + c < b + d$;
6. if $0 < a < b$, then
$$\frac{1}{a} > \frac{1}{b};$$
7. if $a^2 \leq b$ and $b > 0$, then
$$a \geq -\sqrt{b} \quad \text{and} \quad a \leq \sqrt{b} \qquad \text{(or correspondingly, } -\sqrt{b} \leq a \leq \sqrt{b}\text{)}.$$

As a special case of rule (4) above, we have for $c = -1$:

$$\text{If } a < b, \text{ then } -a > -b.$$

Rule (4) says that by multiplying or dividing both sides of an inequality by a negative number c, the inequality sign between the terms changes. For a positive number c, this is not the case (see rule (3)). Rule (5) cannot be extended to the subtraction, i.e., if $a < b$ and $c < d$, then the inequality sign between the terms $a - c$ and $b - d$ may be $<, >, \leq$ or $\geq$.

4.2 Linear Inequalities

Linear inequality:
An inequality of the form

$$ax + b > 0 \qquad \text{and} \qquad ax + b \geq 0,$$

respectively, is called a **linear inequality**.

Notice that this includes the cases of the inequality signs $<$ or $\leq$ since by multiplying the inequality by -1, we get an inequality of the above type. We consider a few examples for solving linear inequalities by applying the above rules.

Example 4.1 *Let us consider the linear inequality*

$$4(x + 2) < 6(x - 5) + 16$$

Applying the above rules for transforming inequalities, we can eliminate x and obtain

$$\begin{aligned} 4x + 8 &< 6x - 30 + 16 && \mid -6x - 8 \\ -2x &< -22 && \mid : (-2) \\ x &> 11, \end{aligned}$$

i.e., every real number x from the interval $(11, \infty)$ satisfies the given inequality. Alternatively, we can describe the set S of solutions in interval notation as

$$S = (11, \infty).$$

Example 4.2 *We wish to determine all integers satisfying the inequality*

$$x + 4 < 2x + \frac{11}{2}.$$

We disregard the integer requirement and solve the above inequality for x. This gives

$$\begin{aligned} x + 4 &< 2x + \frac{11}{2} && \mid -2x - 4 \\ -x &< \frac{3}{2} && \mid \cdot(-1) \\ x &> -\frac{3}{2}. \end{aligned}$$

We emphasize that in the last row, the inequality sign changes because we have multiplied both sides by the negative number -1. *Hence, the set* S *of solutions is the set of all integers greater than* $-3/2$, *i.e.,*

$$S = \{-1, 0, 1, 2, \ldots\}.$$

Example 4.3 *We determine all solutions satisfying the inequality*

$$\frac{2-x}{x+3} < 4$$

To solve for x, *we want to multiply both sides of the inequality by* $x+3$. *Since this term can be positive or negative, we have to consider the two cases:* $x < -3$ *and* $x > -3$ *(note that* $x = -3$ *has to be excluded).*

Case 1: $x < -3$: *In this case, the term* $x+3$ *is negative, and we obtain*

$$2 - x > 4\,(x+3)$$

from which we get

$$\begin{aligned} -10 &> 5x \\ x &< -2 \end{aligned}$$

This, in this case a solution must satisfy $x < -3$ *and* $x < -2$ *which gives the set of solutions*

$$S_1 = (-\infty, -3)\,.$$

Case 2: $x > -3$: *In this case, the term* $x+3$ *is positive, and we obtain*

$$2 - x < 4\,(x+3)$$

from which we get

$$\begin{aligned} -10 &< 5x \\ x &> -2 \end{aligned}$$

Thus, in this case a solution must satisfy $x > -3$ *and* $x > -2$ *which gives the set of solutions*

$$S_2 = (-2, \infty)\,.$$

The set S *of solutions is the union of the sets obtained in the particular cases:*

$$S = S_1 \cup S_2 = (-\infty, -3) \cup (-2, \infty)\,.$$

Example 4.4 *We determine all real numbers satisfying the inequality*

$$\frac{1}{x+4} \le \frac{1}{2x-2}\,. \tag{4.1}$$

In order to transform this inequality, we need to distinguish the cases, where the denominators are positive and negative, respectively. We determine the real numbers for which the denominators are equal to zero and obtain:
From the equation $x + 4 = 0$, *we get* $x = -4$.
From the equation $2x - 2 = 0$, *we get* $x = 1$.
Therefore, we have to distinguish the following three cases:

- *Case 1:* $x > 1$.

  In this case, both denominators in the given inequality are positive. After multiplying both sides of inequality (4.1) by $(x+4)\cdot(2x-2)$*, we obtain the inequality*

$$2x - 2 \leq x + 4,$$

  which can be rewritten after subtracting x *on both sides and then adding 2 on both sides as*

$$x \leq 6.$$

  Since both inequalities $x > 1$ *and* $x \leq 6$ *must be satisfied, we get the following set* S_1 *of solutions in case 1:*

$$S_1 = \{x \in \mathbb{R} \mid 1 < x \leq 6\} = (1, 6].$$

- *Case 2:* $-4 < x < 1$.

  In this case, the denominator $x + 4$ *is positive, but the denominator* $2x - 2$ *is negative. Therefore, when multiplying both sides of the given inequality by* $(x+4)\cdot(2x-2)$*, we multiply by a negative number and thus, the inequality sign changes. Hence, we get*

$$2x - 2 \geq x + 4$$

  which can be rewritten as

$$x \geq 6$$

  Since we must have $-4 < x < 1$ *and* $x \geq 6$*, no real number is a solution of the given inequality in case 2, i.e., the set* S_2 *of the solutions is the empty set:* $S_2 = \emptyset$.

- *Case 3:* $x < -4$.

  In this case, both denominators of the given inequality are negative. Thus, the product $(x+4)\cdot(2x-2)$ *is positive and therefore, we can transform inequality (4.1) as in the first case, i.e., we obtain*

$$2x - 2 \leq x + 2$$

  which again gives

$$x \leq 6.$$

  Since both inequalities $x < -4$ *and* $x \leq 6$ *must be satisfied, we get the set of solutions*

$$S_3 = \{x \in \mathbb{R} \mid x < -4\}.$$

In order to give the complete set of solutions, we have to take into account that in each of the above cases we get a solution. Therefore, the set S *of all solutions is the union of the sets of solutions for the particular cases, i.e., we obtain*

$$S = S_1 \cup S_2 \cup S_3 = \{x \in \mathbb{R} \mid x \leq -4 \ \ or \ \ 1 < x \leq 6\} = (-\infty, -4) \cup (1, 6].$$

4.3 Inequalities with Absolute Values

Next, we consider inequalities including absolute values. Based on the definition of the absolute value of a real number, we have the following properties:

Properties of absolute values:

1. $-|a| \leq a \leq |a|$;
2. Inequality $|x| < a$ is equivalent to the inclusion $x \in (-a, a)$, $\quad (a \geq 0)$;
3. Inequality $|x| \geq a$ is equivalent to the inclusions $x \in (-\infty, -a)$ or $x \in [a, \infty)$, $\quad (a \geq 0)$;
4. $|a + b| \leq |a| + |b|$;
5. $|a| - |b| \leq |a - b| \leq |a| + |b|$.

Inequality (4) is also known as the **triangle inequality**. It says that the length of the third side of a triangle (i.e., $|a + b|$) is never greater than the sum of the lengths of the other sides (i.e., $|a| + |b|$).

Consider the inequality

$$|x - a| \leq b. \tag{4.2}$$

Here the set of solutions is the set of all real numbers x with a distance from the number a not greater than b. This means that the set S of solutions contains the set of all real numbers x not smaller than $a - b$ but not greater than $a + b$, i.e.,

$$S = [a - b, a + b]. \tag{4.3}$$

Formally, we can confirm this by using the definition of the absolute value of a number x. We distinguish the following two cases.

Case 1: $x \geq a$.

Case 2: $x < a$.

Case 1: If $x \geq a$, then $|x-a| = x-a$. Thus, inequality (4.2) turns into

$$x - a \leq b,$$

which gives

$$x \leq a + b.$$

Thus, in this case, we get the set S_1 of solutions with

$$S_1 = [a, a+b].$$

Case 2: If $x < a$, then $|x-a| = -(x-a) = a-x$. Thus, inequality (4.2) turns into

$$a - x \leq b,$$

which gives

$$x \geq a - b.$$

Thus, in this case, we get the set S_2 of solutions with

$$S_2 = [a-b, a].$$

To get the complete set S of solutions, we have to consider the union of both sets, which gives the set $S = S_1 \cup S_2 = [a-b, a+b]$ (see (4.3)).

Example 4.5 *Let the inequality*

$$|x-3| \leq 5$$

be given. The set of solutions contains the set of all real numbers having from number -3 *a distance not greater than 5, i.e., we immediately get*

$$S = [-3-5, -3+5] = [-8, 2].$$

Example 4.6 *Consider the inequality*

$$|x+1| > 2.$$

The set of solutions is obtained as the set of all real numbers x *having from the number* -1 *(notice that* $x+1 = x-(-1)$*) a distance greater than 2. Thus, all real numbers* x *smaller than* $-1-2=-3$ *or greater than* $-1+2=1$ *satisfy the given inequality:*

$$S = (-\infty, -3) \cup (1, \infty).$$

Example 4.7 *Let us consider the inequality*

$$\frac{|3x-6|}{x+1} < 2.$$

Here we also have to consider several cases. In order to transform the given inequality, we have to distinguish the cases when the term within the absolute-value-signs is non-negative and positive and when the denominator is positive and negative, respectively.

From $3x-6=0$*, we get* $x=2$ *and from* $x+1=0$*, we get* $x=-1$*.*

Therefore, we have to consider the following three cases.

Case 1: $x \geq 2$.

Case 2: $-1 < x < 2$.

Case 3: $x < -1$.

Note that $x=-1$ *must be excluded because the denominator would then be equal to zero. We determine the set of all solutions for the individual cases.*

Case 1: If $x \geq 2$*, then* $3x-6$ *is positive for each* x *under consideration, which yields*

$$|3x-6| = 3x-6.$$

Moreover, the denominator is positive. Thus, we can transform the given inequality as follows:

$$\begin{array}{rcll} \dfrac{3x-6}{x+1} & < & 2 & \quad | \cdot (x+1) > 0 \\ 3x-6 & < & 2\cdot(x+1) & \quad | -2x+6 \\ x & < & 8. & \end{array}$$

Since we must have $x \geq 2$ *and* $x < 8$*, we get as the set* S_1 *of all solutions in this case*

$$S_1 = [2, 8).$$

Case 2: For all numbers x considered in case 2, the term $3x - 6$ is negative and therefore,

$$|3x - 6| = -(3x - 6) = 6 - 3x.$$

Consequently, since $x + 1$ is positive, we obtain

$$\begin{aligned} \frac{6-3x}{x+1} &< 2 && \mid \cdot(x+1) > 0 \\ 6 - 3x &< 2 \cdot (x+1) && \mid +3x - 2 \\ 4 &< 5x \\ x &> \frac{4}{5}. \end{aligned}$$

From $-1 < x < 2$ and $x > 4/5$, we obtain the set of solutions

$$S_2 = \left(\frac{4}{5}, 2\right).$$

Case 3: In this case, we always have

$$|3x - 6| = -(3x - 6) = 6 - 3x,$$

and we can transform the given inequality as follows:

$$\begin{aligned} \frac{6-3x}{x+1} &< 2 && \mid \cdot(x+1) < 0 \\ 6 - 3x &> 2 \cdot (x+1) \\ 4 &> 5x \\ x &< \frac{4}{5}. \end{aligned}$$

From $x < -1$ and $x < 4/5$, we get the set S_3 of solutions in case 3 as follows:

$$S_3 = (-\infty, -1).$$

In order to find the complete set of solutions of the given inequality, we have to combine the individual sets of solutions since any of the three cases is possible, and so we obtain

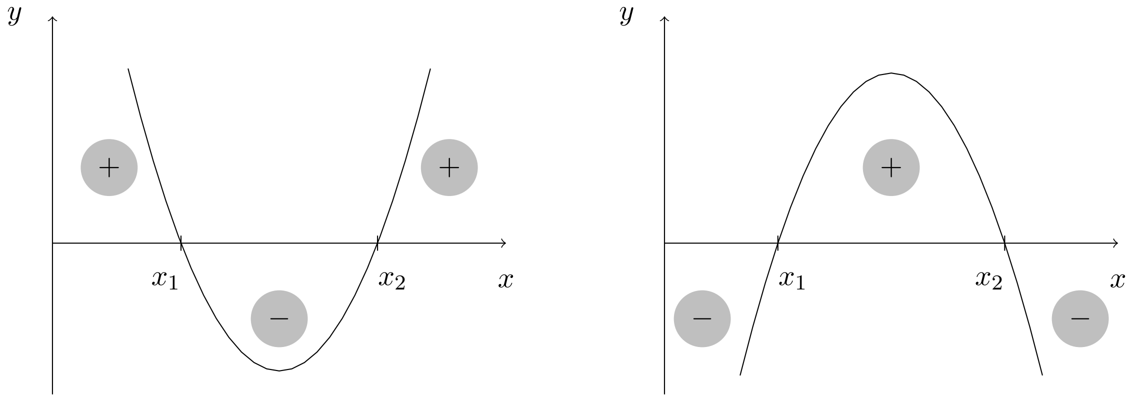
$$S = (-\infty, -1) \cup \left(\frac{4}{5}, 8\right).$$

4.4 Quadratic Inequalities

Quadratic inequality:
An inequality which can be transformed into one of the forms

$$ax^2 + bx + c > 0 \qquad \text{or} \qquad ax^2 + bx + c \geq 0$$

weith $a \neq 0$ is called **quadratic**.



(a) $a > 0$ (b) $a < 0$

Figure 4.1: Solving a quadratic inequality

Again, the case of opposite inequality signs can be reduced to the latter type of inequality by multiplying both sides by the number -1, which changes the inequality sign.

We discuss two variants of solving such quadratic inequalities. The first way is to consider the corresponding quadratic equation with the inequality sign replaced by the equality sign. We determine the roots of this equation and check which of the resulting interval(s) satisfies(y) the corresponding inequality (see Fig. 4.1).

The second way of solving a quadratic inequality is to transform the quadratic and linear terms $x^2 + ax$ into a **complete square** by using the binomial formula and taking then the square root on both sides. For illustration, we first consider the following example, where we apply both solution variants.

Example 4.8 *We consider the inequality*

$$2x^2 - 6x - 20 \leq 0$$

and apply both solution strategies. Considering the corresponding quadratic equation and dividing both sides by 2, we obtain

$$x^2 - 3x - 10 = 0$$

which has the two solutions

$$x_1 = \frac{3}{2} + \sqrt{\frac{9}{4} + 10} = \frac{3}{2} + \sqrt{\frac{49}{4}} = 5$$

and

$$x_2 = \frac{3}{2} - \sqrt{\frac{9}{4} + 10} = \frac{3}{2} - \sqrt{\frac{49}{4}} = -2$$

Now we take some arbitrary value, e.g. $x = 0$ *and check whether it satisfies the given inequality. Since* $2 \cdot 0^2 - 6 \cdot 0 - 10 \leq 0$, *all real numbers* x *between* -2 *and* 5 *satisfy the given inequality.*

Applying the second approach, we divide the given inequality by 2 *and rewrite it as follows:*

$$x^2 - 3x \ \leq \ 10.$$

Transforming the left-hand side into a complete square, we use the binomial formula

$$(x-a)^2 = x^2 - 2ax + a^2 \ .$$

From $-2ax = -3x$*, we get* $a = 3/2$ *and then*

$$x - 3x + \left(\mathbf{\frac{3}{2}}\right)^{\mathbf{2}} \ \le \ 10 + \left(\mathbf{\frac{3}{2}}\right)^{\mathbf{2}},$$

which can be rewritten as

$$\left(x - \frac{3}{2}\right)^2 \ \le \ \frac{49}{4}.$$

Taking on both sides the square root, we get

$$\left| x - \frac{3}{2} \right| \ \le \ \frac{7}{2}.$$

Thus, the set of solutions is given by the set of all real numbers x *having from* $3/2$ *a distance not greater than* $7/2$*:*

$$S = \{x \in \mathbb{R} \mid -2 \le x \le 5\} = [-2, 5].$$

Next, we present one other example where we apply the second solution variant while for all other examples, the first solution variant will be used.

Example 4.9 *We determine all real numbers* x *satisfying the inequality*

$$x^2 + 4x \ \le \ 5.$$

We write the left-hand side as a complete square $(x+a)^2$ *by using the binomial formula*

$$(x+a)^2 = x^2 + 2ax + a^2.$$

From $4x = 2ax$*, we obtain* $a = 2$*. Therefore, we add* $a^2 = 2^2 = 4$ *on both sides:*

$$x^2 + 4x + \mathbf{2^2} \ \le \ 5 + \mathbf{2^2},$$

which yields

$$(x+2)^2 \ \le \ 9.$$

Now we take the square root on both sides and obtain

$$|x+2| \ \le \ 3.$$

Note that, as mentioned before, we have to use the absolute value of $x+2$ *on the left-hand side because there are two solutions when taking the square root of some number* z^2*: namely* $+z$ *and* $-z$*, and these cases can be summarized to* $|z|$*. The set of solutions* S *of the inequality is given by the set of all real numbers* x *having from the number* -2 *a distance not greater than 3, i.e.,*

$$S = \{x \in \mathbb{R} \mid -5 \le x \le 1\},$$

or in interval notation

$$S = [-5, 1].$$

Example 4.10 *Consider the quadratic inequality*

$$-2x^2 - 4x + 16 < 0.$$

Taking the resulting quadratic equation and dividing both sides by -2, *we obtain*

$$x^2 + 2x - 8 = 0,$$

which gives the solutions

$$x_1 = -1 + \sqrt{1+8} = 2 \qquad \textit{and} \qquad x_2 = -1 - \sqrt{1+8} = -4.$$

Now we test whether the given inequality is satisfied for an arbitrary trial value. For instance, for $x = 0$ *we obtain that* $-2 \cdot 0^2 - 4 \cdot 0 + 16$ *is equal to* 16 *and therefore, the trial value* $x = 0$ *does not satisfy the given inequality. Therefore, all* x *from the interval* $[x_2, x_1] = [-4, 2]$ *do not satisfy the given inequality. Consequently, the set* S *of solutions satisfying the given inequality is given by the set of all real numbers that are smaller than* -4 *or greater than* 2:

$$S = (-\infty, -4) \cup (2, \infty).$$

Example 4.11 *Let us find all real numbers satisfying the quadratic inequality*

$$x^2 + 6x + 9 > 0.$$

Considering the corresponding quadratic equation, we have

$$x^2 + 6x + 9 = 0,$$

which has the solutions

$$x_1 = -3 + \sqrt{9-9} = -3 \qquad \textit{and} \qquad x_2 = -3 - \sqrt{9-9} = -3,$$

i.e., there is a real double solution $x_1 = x_2$*. As a consequence, the sign of the quadratic term* $x^2 + 6x + 9$ *does not change at the root* -3*. We test the sign of the quadratic term for an arbitrary value of* x*. Using for instance* $x = 0$*, we get* $0^2 + 6 \cdot 0 + 9 > 0$*, i.e., the given inequality is satisfied for* $x = 0$ *and therefore also for all real numbers except* -3*, i.e., we get*

$$S = (-\infty, -3) \cup (-3, \infty).$$

Example 4.12 *We determine all real numbers satisfying the inequality*

$$|x^2 + 3x - 4| \leq x + 4.$$

First, we find the zeroes of the term within the absolute-value signs: From $x^2 + 3x - 4 = 0$*, we get the solutions*

$$x_1 = -\frac{3}{2} + \sqrt{\frac{9}{4} + 4} = -\frac{3}{2} + \frac{5}{2} = 1$$

and

$$x_2 = -\frac{3}{2} - \sqrt{\frac{9}{4} + 4} = -\frac{3}{2} - \frac{5}{2} = -4.$$

For $x = 0$*, we obtain* $0^2 + 3 \cdot 0 - 4 < 0$*. Since* $x = 0$ *is contained in the interval* $(-4, 1)$*, the inequality* $x^2 + 3x - 4 < 0$ *holds if and only if* $x \in (-4, 1)$*, and* $x^2 + 3x - 4 \geq 0$ *holds if and only if* $x \leq -4$ *or* $x \geq 1$*. Therefore, we consider the following two cases:*

Case 1: $x \le -4$ *or* $x \ge 1$;

Case 2: $-4 < x < 1$.

Case 1: For $x \le -4$ *or* $x \ge 1$, *we have*

$$|x^2 + 3x - 4| = x^2 + 3x - 4$$

and the given inequality turns into

$$x^2 + 3x - 4 \le x + 4$$

which can be rewritten as

$$x^2 + 2x - 8 \le 0.$$

Replacing the inequality sign the the equality sign and solving the resulting quadratic equation, we get

$$x_3 = -1 + \sqrt{1+8} = -1 + 3 = 2$$

and

$$x_4 = -1 - \sqrt{1+8} = -1 - 3 = -4.$$

Since the inequality $x^2 + 3x - 4 \le 0$ *is satisfied for instance for* $x = 0 \in [-4, 2]$, *the solutions in case 1 must satisfy* $x \le -4$ *or* $x \ge 1$ *(according to the assumption) and in addition,* $-4 \le x \le 2$. *Therefore, the set* S_1 *of solutions for case 1 is given by*

$$S_1 = \{-4\} \cup [1, 2]\ .$$

Case 2: For $-4 < x < 1$*, we have*

$$|x^2 + 3x - 4| = -(x^2 + 3x - 4) = -x^2 - 3x + 4.$$

Thus, the given inequality turns into

$$-x^2 - 3x + 4 \leq x + 4$$

which can be rewritten as

$$x^2 + 4x = x\,(x + 4) \geq 0.$$

This inequality is satisfied for $x \leq -4$ *(both factors are non-positive) and* $x \geq 0$ *(both factors are non-negative). Therefore, a solution in case 2 must satisfy* $x \leq -4$ *or* $x \geq 0$*, but according to the assumption, additionally* $-4 < x < 1$ *must hold. Thus, the set* S_2 *of solutions for case 2 is given by*

$$S_1 = [0, 1).$$

Thus, the complete set S *of real numbers satisfying the given inequality is given by*

$$S = S_1 \cup S_2 = \{-4\} \cup [0, 2]\,.$$

Example 4.13 *We determine all real numbers satisfying the inequality*

$$(1 - x) \cdot (x^2 - x - 6) \leq 0.$$

Although this is not a quadratic inequality (the term with the highest occurring power is $-x^3$ *if we multiply out), we can nevertheless find all solutions using our knowledge about quadratic inequalities. A solution* x *has to satisfy one of the following cases:*

Case 1: $1 - x \geq 0$ *and* $x^2 - x - 6 \leq 0$;

Case 2: $1 - x \leq 0$ *and* $x^2 - x - 6 \geq 0$.

Case 1: From $1 - x \geq 0$, *we get* $x \leq 1$. *Setting now* $x^2 - x - 6 = 0$, *we obtain the two solutions*

$$x_1 = \frac{1}{2} + \sqrt{\frac{1}{4} + 6} = \frac{1}{2} + \frac{5}{2} = 3$$

and

$$x_2 = \frac{1}{2} - \sqrt{\frac{1}{4} + 6} = \frac{1}{2} - \frac{5}{2} = -2\ .$$

We have $x^2 - x - 6 \leq 0$ *for* $x \in [-2, 3]$. *Since* $x \leq 1$ *by assumption, we get the set of solutions*

$$S_1 = [-2, 1]$$

for case 1.

Case 2: In this case, we have $x^2 - x - 6 \geq 0$ *for all* $x \in (-\infty, -2] \cup [3, \infty)$. *Since now* $x \geq 1$ *by assumption, we get the set of solutions*

$$S_2 = [3, \infty)$$

for case 2.

Combining both cases, we get the complete set S *of real numbers satisfying the given inequality:*

$$S = S_1 \cup S_2 = [-2, 1] \cup [3, \infty)\ .$$

Example 4.14 *We determine all real numbers satisfying the inequality*

$$|x^2 - 9| \leq 2\,.$$

We have to distinguish the cases $x^2 - 9 \geq 0$ *and* $x^2 - 9 < 0$. *This gives the following two cases:*

Case 1: $x \leq -3$ *or* $x \geq 3$;

Case 2: $-3 < x < 3$.

Case 1: In this case, we have $|x^2 - 9| = x^2 - 9$ *and thus, the given inequality turns into*

$$x^2 - 9 \leq 2$$

which can be written as

$$x^2 \leq 11$$

This inequality is satisfied for all $x \in [-\sqrt{11}, \sqrt{11}]$. *Thus, in case 1, we get the set of solutions*

$$S_1 = [-\sqrt{11}, -3] \cup [3, \sqrt{11}]\,.$$

Case 2: In this case, we have $|x^2 - 9| = -(x^2 - 9) = 9 - x^2$ *and thus, the given inequality turns into*

$$9 - x^2 \leq 2$$

which can be written as

$$x^2 \geq 7\,.$$

This inequality is satisfied for all $x \in (-\infty, -\sqrt{7}] \cup [\sqrt{7}, \infty)$. *Taking the assumption of case 2 into account, we get the set of solutions*

$$S_2 = (-3, -\sqrt{7}] \cup [\sqrt{7}, \sqrt{3})\,.$$

For the complete set of all real numbers satisfying the given inequality, we get

$$S = S_1 \cup S_2 = [\sqrt{11}, -\sqrt{7}] \cup [\sqrt{7}, \sqrt{11}]\,.$$

4.5 Further Inequalities

In this section, we consider some inequalities involving logarithmic, exponential or root terms. For solving them, elementary knowledge about the monotonicity of the corresponding functions is required (see also Chapter 8).

Example 4.15 *We determine all real numbers satisfying the inequality*

$$\lg(x - 1) < 2.$$

Since the logarithm is defined only for positive numbers, we first must require $x > 1$. *We can rewrite this inequality as*

$$x - 1 < 10^2 = 100.$$

Formally we have raised both sides of the given inequality to the exponent of an exponential term with base 10. This can be done due to the monotonicity of the corresponding power function: From $x_1 < x_2$, *we obtain* $10^{x_1} < 10^{x_2}$. *Now we get* $x < 101$ *and thus, the set of all real numbers satisfying the given inequality is given by*

$$S = (1, 101)\,.$$

Example 4.16 *We determine all real numbers satisfying the inequality*

$$\ln(x^2 - x - 5) \leq 0\,.$$

First, we must have $x^2 - x - 5 > 0$*. Solving the corresponding equation* $x^2 - x - 5 = 0$*, we get*

$$x_1 = \frac{1}{2} + \sqrt{\frac{1}{4} + 5} = \frac{1}{2} + \frac{1}{2} \cdot \sqrt{21}$$

and

$$x_2 = \frac{1}{2} - \sqrt{\frac{1}{4} + 5} = \frac{1}{2} - \frac{1}{2} \cdot \sqrt{21}\,.$$

Note that $x_1 > 3$ *while* $x_2 < -2$*. Next, similar to the previous example, we can rewrite the given inequality as*

$$x^2 - x - 5 \leq e^0 = 1\,.$$

This is equivalent to

$$x^2 - x - 6 \leq 0\,.$$

Considering the corresponding equation and finding the zeroes, we obtain

$$x_1 = \frac{1}{2} + \sqrt{\frac{1}{4} + 6} = \frac{1}{2} + \frac{5}{2} = 3$$

and

$$x_2 = \frac{1}{2} - \sqrt{\frac{1}{4} + 6} = \frac{1}{2} - \frac{5}{2} = -2\,.$$

Since $0^2 - 0 - 5 \leq 0$*, the latter inequality is satisfied for* $x \in [-2, 3]$*. At the same time, the definition of the logarithm requires that*

$$x < \frac{1}{2} - \sqrt{\frac{1}{4} + 5} = \frac{1}{2} - \frac{1}{2} \cdot \sqrt{21} < -2$$

or

$$x > \frac{1}{2} + \sqrt{\frac{1}{4} + 5} = \frac{1}{2} + \frac{1}{2} \cdot \sqrt{21} > 3\,.$$

Thus, the set of all real numbers satisfying the given inequality is empty: $S = \emptyset$*.*

Example 4.17 *We determine all real numbers satisfying the inequality*

$$e^{4-2x} - 1 > 0\,.$$

Adding one to both sides and taking the natural logarithm on both sides, we obtain

$$4 - 2x > \ln 1 = 0$$

which gives $x < 2$*, i.e., the set of all real numbers satisfying the given inequality is given by*

$$S = (-\infty, 2)\,.$$

Example 4.18 *We determine all real numbers satisfying the inequality*

$$\sqrt[4]{x+1} < 2\,.$$

First, we must require $x+1 \geq 0$, *i.e.,* $x \geq -1$. *By taking both sides to the fourth power (again it does not change the inequality sign due to the monotonicity of this function for non-negative values), we obtain*

$$x+1 < 2^4 = 16$$

which gives $x < 15$. *Thus, the set of all real numbers satisfying the given inequality is obtained as*

$$S = [-1, 15)\,.$$

EXERCISES

4.1 Determine all real numbers satisfying the following linear inequalities:

(a) $3(x+4) \leq 2(x-2) - 1\,;$ (b) $\frac{1}{3}x + 1 \leq 4 - \frac{3}{2}x$; (c) $\frac{4}{1+x} \geq 3$;

(d) $\frac{x+3}{2x+2} \leq 1\,;$ (e) $\frac{3x+2}{3-2x} < 2$.

4.2 Determine all real numbers satisfying the following inequalities:

(a) $(x-3)(x+4) \geq 0\,;$ (b) $4x^2 + x \geq \frac{15}{2}$; (c) $\frac{3x+10}{x-2} \leq 2x+3\,;$

(d) $\frac{2-x}{x+3} > -2x\,;$ (e) $\frac{9}{x-7} > x+3$.

4.3 Determine all real numbers satisfying the following inequalities:

(a) $(x+3)(x^2+2x-8) > 0$; (b) $x^3 - 5x^2 + 6x \geq 0\,.$

4.4 Determine all numbers satisfying the following inequalities:

(a) $|x^2 - 25| \geq 11\,;$ (b) $|x^2 - 7| \leq 9\,;$ (c) $|2x+1| - 2 > x\,;$

(d) $|2x-2| \leq |x|$; (e) $|x+3| - |2x-5| \geq 2x\,;$ (f) $|x-2| + |x+3| \leq 5\,.$

4.5 Determine all real numbers satisfying the following inequalities:

(a) $|x^2 - x - 6| \leq 6\,;$ (b) $|3-x| > |x^2 - 2x - 3|$.

4.6 Determine all real numbers satisfying the following inequalities:

(a) $\sqrt{x-2} \leq \frac{1}{4}\,;$ (b) $\lg(3+x) < 2$; (c) $e^{3x-2} \leq 1\,;$

(d) $\ln(x^2 + 4x + 3) \geq 0\,.$

Chapter 5

Trigonometry and Goniometric Equations

Trigonometric terms and goniometric equations often play a role when analyzing geometric problems. In trigonometry, triangles and other plane figures (which can be partitioned into triangles) are investigated. In goniometry, one works with terms depending on angles. The learning objectives of this chapter are to review

- some basic facts about trigonometric relationships and
- the solution of simple goniometric equations including trigonometric terms.

5.1 Trigonometry

Consider the right-angled triangle given in Fig. 5.1. Then we define the following trigonometric terms (note that the angle α is less than 90^o):

$$\sin\alpha = \frac{a}{c}\,, \qquad \cos\alpha = \frac{b}{c}\,, \qquad \tan\alpha = \frac{a}{b} \qquad \text{and} \qquad \cot\alpha = \frac{b}{a}\,.$$

The side c is called the **hypotenuse** and the sides a and b are the **legs** of the right-angled triangle. One can extend the definition of trigonometric terms for angles $\alpha \geq 90^0$ on a unit circle, i.e., we have $r = 1$. In Fig. 5.2, we illustrate this for the case $90^o \leq \alpha \leq 180^o$.

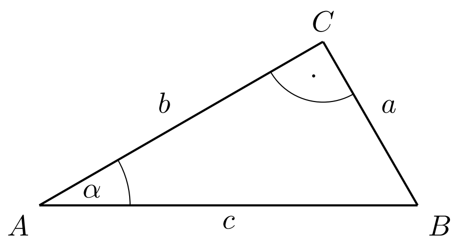


Figure 5.1: Definition of trigonometric terms in a right-angled triangle

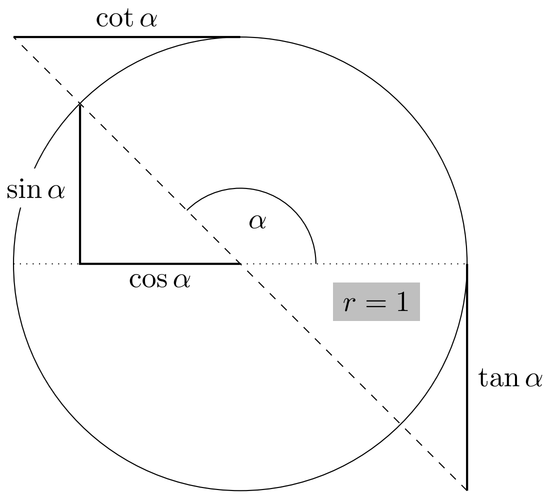


Figure 5.2: Trigonometric terms in a unit circle for $90^o \leq \alpha \leq 180^o$

In the following, we review some trigonometric formulas when the sum or difference of two angles is considered. These formulas are often useful when simplifying terms in problems from differential calculus or integration.

Addition theorems for trigonometric terms:

1. $\sin(\alpha \pm \beta) = \sin\alpha\cos\beta \pm \cos\alpha\sin\beta\,;$
2. $\cos(\alpha \pm \beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta\,;$
3. $\tan(\alpha \pm \beta) = \dfrac{\tan\alpha \pm \tan\beta}{1 \mp \tan\alpha\tan\beta}\,;$
4. $\cot(\alpha \pm \beta) = \dfrac{\cot\alpha\cot\beta \mp 1}{\cot\beta \pm \cot\alpha}\,.$

For the special case of $\alpha = \beta$, properties (1) and (2) turn into

$$\sin 2\alpha = 2\sin\alpha\cos\alpha \qquad \text{and} \qquad \cos 2\alpha = \cos^2\alpha - \sin^2\alpha\,. \tag{5.1}$$

Accordingly, properties (3) and (4) turn into

$$\tan 2\alpha = \frac{2\tan\alpha}{1-\tan^2\alpha} \qquad \text{and} \qquad \cot 2\alpha = \frac{\cot^2\alpha - 1}{2\cot\alpha}\,. \tag{5.2}$$

Moreover, using $\beta = \pi/2$ and $\beta = -\pi/2$, respectively, we get

$$\sin\left(\alpha \pm \frac{\pi}{2}\right) = \pm\cos\alpha, \qquad \cos\left(\alpha \pm \frac{\pi}{2}\right) = \mp\sin\alpha\,.$$

Further trigonometric formulas:

1. $\sin\alpha + \sin\beta = 2 \cdot \sin\dfrac{\alpha+\beta}{2} \cdot \cos\dfrac{\alpha-\beta}{2}\,;$

   $\sin\alpha - \sin\beta = 2 \cdot \cos\dfrac{\alpha+\beta}{2} \cdot \sin\dfrac{\alpha-\beta}{2}\,;$

2. $\cos\alpha + \cos\beta = 2 \cdot \cos\dfrac{\alpha+\beta}{2} \cdot \cos\dfrac{\alpha-\beta}{2}\,;$

   $\cos\alpha - \cos\beta = -2 \cdot \sin\dfrac{\alpha+\beta}{2} \cdot \sin\dfrac{\alpha-\beta}{2}\,;$

3. $\sin^2\alpha + \cos^2\alpha = 1\,;$
4. $1 + \tan^2\alpha = \dfrac{1}{\cos^2\alpha}\,;$
5. $\tan x \cdot \cot x = 1\,.$

Property (3) is also denoted as **Pythagorean theorem** for trigonometric terms. We can easily convince that the identity given in property (4) above is correct and obtain

$$1 + \tan^2\alpha = 1 + \frac{\sin^2\alpha}{\cos^2\alpha} = \frac{\cos^2\alpha + \sin^2\alpha}{\cos^2\alpha} = \frac{1}{\cos^2\alpha}.$$

The last equality follows from property (3) above.

Moreover, we can present formulas for the product of two sine and cosine terms:

Product formulas:

1. $\sin\alpha \cdot \sin\beta = \dfrac{1}{2} \cdot [\cos(\alpha-\beta) - \cos(\alpha+\beta)]\,;$
2. $\cos\alpha \cdot \cos\beta = \dfrac{1}{2} \cdot [\cos(\alpha-\beta) + \cos(\alpha+\beta)]\,;$
3. $\sin\alpha \cdot \cos\beta = \dfrac{1}{2} \cdot [\sin(\alpha-\beta) + \sin(\alpha+\beta)]\,.$

Example 5.1 *Consider the trigonometric term*

$$T = \frac{\sin\alpha + \sin\beta}{\cos\alpha + \cos\beta}.$$

We want to simplify term T by using only one trigonometric expression. We obtain

$$T = \frac{2\sin\frac{\alpha+\beta}{2} \cdot \cos\frac{\alpha-\beta}{2}}{2 \cdot \cos\frac{\alpha+\beta}{2} \cdot \cos\frac{\alpha-\beta}{2}} = \tan\frac{\alpha+\beta}{2}\ .$$

Consider now the oblique triangle given in Fig. 5.3. Then we can summarize the following properties:

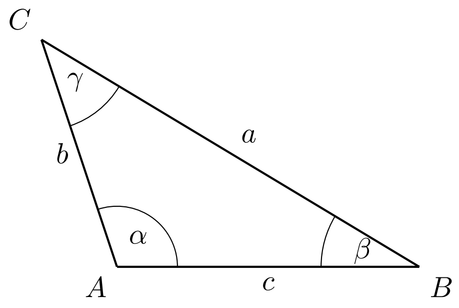


Figure 5.3: Oblique triangle

Relationships in a triangle:

1. $\alpha + \beta + \gamma = 180^o$;
2. $a : b : c = \sin\alpha : \sin\beta : \sin\gamma$;
3. $a^2 = b^2 + c^2 - 2bc\cos\alpha$.

Equality (1) expresses that in any triangle, the sum of the three angles is 180^o. Equality (2) is also known as the **sine theorem**. Equality (3) above is denoted as the **cosine theorem**.

Example 5.2 *In an oblique triangle, the side $b = 12$ cm and the angles $\beta = 32^o$ and $\gamma = 85^o$ are given. We determine the angle α and the sides a and c. First, we obtain from relationship (1) above*

$$\alpha = 180^o - (\beta + \gamma) = 180^o - (32^0 - 85^o) = 63^o .$$

From the sine theorem, we obtain

$$a = \frac{b \cdot \sin\alpha}{\sin\beta} = \frac{12\ cm\ \cdot \sin 85^o}{\sin 32^o} = \frac{12 \cdot 0.9962}{0.5299}\ cm\ = 22.56\ cm .$$

Example 5.3 *In an oblique triangle, the two side lengths $b = 14$ cm and $c = 17$ cm are given. The angle α between the two sides is given as $\alpha = 80^o$. We determine the length of the remaining side a. We use the cosine theorem and obtain:*

$$a = \sqrt{b^2 + c^2 - 2bc\cos\alpha} = \sqrt{14^2 + 17^2 - 2 \cdot 14 \cdot 17 \cdot \cos 32^o}\ cm\ = \sqrt{232.7676}\ cm\ = 15.24\ cm\ .$$

Finally, we present a possibility to compute the area A of a triangle by means of the lengths of two sides and the angle formed by the two sides. We have the following formulas for determining the **area** A (see Fig. 5.3):

$$A = \frac{1}{2}ab\sin\gamma = \frac{1}{2}bc\sin\alpha = \frac{1}{2}ac\sin\beta .$$

Example 5.4 *In a triangle, the lengths of the sides $a = 16$ cm and $b = 22$ cm are known together with the angle $\gamma = 98^o$ of the angle formed by the two sides. We obtain the area A as follows:*

$$A = \frac{1}{2}ab\sin\gamma = \frac{1}{2} \cdot 16 \cdot 22 \cdot \sin 98^o\ cm^2 = 174.29\ cm^2 .$$

5.2 Goniometric Equations

An equation that includes trigonometric terms depending on an angle or a multiple of this angle is denoted as a **goniometric equation**. Examples of such equations are e.g.

$$\cos\alpha + 2\cos\ 2\alpha = \cos\ 3\alpha \qquad \text{or} \qquad \sin^2\beta + 2\tan^2\beta = 3.$$

In order to find all solutions of a goniometric equation, one tries to express all trigonometric terms by means of one or several of the above formulas by **one** trigonometric term. Without loss of generality, assume that it is some sine term of the form $\sin\alpha x$.

Now, we substitute $z = \sin\alpha x$ and determine all solutions of the equation $g(z) = g(\sin\alpha x) = 0$. Let $z_1, z_2, \ldots, z_n$ be the solutions. By substituting back, we get

$$z_1 = \sin\alpha x, \qquad z_2 = \sin\alpha x, \ldots, \qquad z_n = \sin\alpha x\ .$$

Finally, one has to check whether all solutions x obtained satisfy indeed the given equation.

We illustrate the solution of goniometric equations by the following examples, where we only look for solutions in the interval $[0, 2\pi]$ in all subsequent problems.

Example 5.5 *Let the goniometric equation*

$$\sin x + \cos x = 1$$

be given. First, we substitute the cosine term by the sine term using $sin^2 x + \cos^2 x = 1$*. This gives*

$$\cos x = \pm\sqrt{1 - \sin^2 x}.$$

Substituting now $z = \sin x$*, we get*

$$z \pm \sqrt{1 - z^2} = 1.$$

After eliminating the square root on one side and taking the square on both sides, we get a quadratic equation in the new variable z:

$$\begin{aligned} \pm\sqrt{1 - z^2} &= 1 - z \\ 1 - z^2 &= 1 - 2z + z^2 \\ 2z^2 - 2z &= 0. \end{aligned}$$

Factoring out, we obtain

$$2z \cdot (z - 1) = 0$$

which has the two solutions

$$z_1 = 0 \qquad \textit{and} \qquad z_2 = 1\,.$$

From $z_1 = \sin x = 0$*, we get the solutions*

$$x_1 = 0, \quad x_2 = \pi, \quad x_3 = 2\pi\,.$$

From $z_2 = \sin x = 1$, *we get the solutions*

$$x_4 = \frac{\pi}{2}, \quad x_5 = \frac{3}{2}\pi .$$

A test confirms that x_1, x_3 *and* x_4 *are indeed a solution of the given equation while for* x_2 *and* x_5, *we get* $\sin x + \cos x = -1$ *and thus,* x_2 *and* x_5 *are not a solution. Note that, if we would be looking for all solutions, we would get*

$$x_{1,k} = 2k\pi, \qquad x_{2,k} = \frac{\pi}{2} + 2k\pi, \qquad k \in \mathbb{Z} .$$

Example 5.6 *We consider the goniometric equation*

$$4\sin^2 x + 4\cos x = 1 .$$

Using $\sin^2 x + \cos^2 x = 1$, *we obtain*

$$\begin{aligned} 4(1-\cos^2 x) + 4\cos x &= 1 \\ 4\cos^2 x - 4\cos x - 3 &= 0 \end{aligned}$$

Substituting now $z = \cos x$ *and dividing the equation by 4, we get*

$$z^2 - z - \frac{3}{4} = 0$$

which has the two solutions

$$z_1 = \frac{1}{2} + 1 = \frac{3}{2} \qquad \textit{and} \qquad z_2 = \frac{1}{2} - 1 = -\frac{1}{2}\,.$$

For the first solution z_1*, we do not obtain a solution since the cosine values are in the interval* $[-1,1]$*. After substituting back, we get from the second solution* z_2

$$\cos x = -\frac{1}{2}$$

which has the two solutions

$$x_1 = 120^o \qquad \textit{and} \qquad x_2 = 240^o\,.$$

Inserting x_1 *and* x_2 *into the given goniometric equation, we confirm that both solutions satisfy the original equation:*

$$4 \cdot \left(\pm\frac{1}{2}\sqrt{3}\right)^2 + 4 \cdot \left(-\frac{1}{2}\right) = 4 \cdot \frac{3}{4} - 2 = 1\,.$$

Example 5.7 *We determine all solutions of the equation*

$$\tan^2 x + \frac{5}{\tan x} = 6\,.$$

Multiplying both sides by $\tan x$ *and putting all terms on the left-hand side, we obtain*

$$\tan^2 x - 6\tan x + 5 = 0\,.$$

Substituting now $z = \tan x$*, we get the quadratic equation*

$$z^2 - 6z + 5 = 0$$

which has the two solutions

$$z_1 = 3 + \sqrt{9-5} = 5 \qquad \textit{and} \qquad z_2 = 3 - \sqrt{9-5} = 1\,.$$

Substituting back, we get for z_1 *the equation*

$$\tan x = 5$$

which gives the solutions

$$x_1 = 78.69^o \qquad \textit{and} \qquad x_2 = 258.69^o\,.$$

They do not satisfy the given equation:

$$5^2 + \frac{5}{5} \neq 6\,.$$

Moreover, for z_2 *we get*

$$\tan x = 1$$

which gives the solutions

$$x_3 = 45^o \qquad \textit{and} \qquad x_4 = 225^o\,.$$

Inserting these solutions into the original equation, we can confirm that they satisfy the given equation:

$$1^2 + \frac{5}{1} = 6\,.$$

Example 5.8 *Consider the equation*

$$\cos x - \sin 2x = 0\ .$$

Using the double angle formula for the sine term, we get

$$\cos x - 2\sin x \cos x = 0$$

which gives

$$\cos x \cdot (1 - 2\sin x) = 0\,.$$

The product is equal to 0 if at least one of the factors is equal to zero. From $\cos x = 0$*, we obtain the solutions*

$$x_1 = 90^o \qquad \textit{and} \qquad x_2 = 270^o\,.$$

From $1 - 2\sin 2x = 0$*, we get*

$$\sin 2x = \frac{1}{2}$$

and therefore,

$$2x \in \{30^o, 150^o, 390^o, 510^o, 750^o, 870^o, \ldots\}$$

Therefore, in $[0, 2\pi]$*, we get the solutions*

$$x_3 = 15^o, \quad x_4 = 75^o \quad x_5 = 195^o, \quad \textit{and} \quad x_6 = 255^o\,.$$

A test confirms that all angles satisfy the given equation.

Example 5.9 *We find all solutions of the equation*

$$3\sin x - \sqrt{3}\cos x = 0\,.$$

One way to transform the above equation such that only one trigonometric term occurs is to divide both sides by $\cos x$*. In this case, we have to discuss* $\cos x = 0$ *separately. This equation is satisfied for*

$$x_1 = 90^o \qquad \textit{and} \qquad x_2 = 270^0\,.$$

These angles do not satisfy the given equation. For $\cos x \neq 0$*, we obtain*

$$\begin{aligned} \sqrt{3} \cdot \frac{\sin x}{\cos x} - 1 &= 0 \\ \tan x &= \frac{1}{3}\sqrt{3}\,. \end{aligned}$$

This gives the solution

$$x_3 = 30^o \qquad \textit{and} \qquad x_4 = 210^o\,.$$

A test confirms that these angles satisfy the original equation.

EXERCISES

5.1 Determine all angles α between 0^0 and 360^o for which the following equalities hold:

(a) $\sin\ \alpha = 0.25\,;$ (b) $\cos\ \alpha = -0.55\,;$

(c) $|\tan\alpha| = 2\,;$ (d) $|\cot\alpha| = \sqrt{3}\,.$

5.2 In a right-angled triangle, the length of the hypotenuse $c = 20$ cm and the angle $\alpha = 18^o$ are given. Determine the lengths of the remaining sides and angles.

5.3 In a right-angled triangle, the lengths of the legs $a = 2.5$ m and $b = 2.1$ m are given. Determine the angles α and β and the length of the hypotenuse.

5.4 In an oblique triangle, the lengths of the sides are $a = 9$ cm, $b = 12$ cm and $c = 17$ cm are known. Determine the angles of the triangle.

5.5 Determine the area of the triangle with the lengths of the sides $a = 12$ cm and $b = 14$ cm and the angle between the two sides $\gamma = 111^o$.

5.6 Find the solutions of the following goniometric equations:

(a) $2\cos^2 x = \frac{1}{2} + 2\sin x\,;$ (b) $8\cos 2x = \sin x\,;$

(c) $4\sin 2x + \cos x = 0\,;$ (d) $\dfrac{3}{\tan x} = 4 - \tan x\,.$

Chapter 6

Analytic Geometry in the Plane

In analytic geometry, geometric investigations are done by means of analytic methods. It has applications e.g. when working with vectors. Knowledge in geometry is also required for several applications of differential calculus and integration when considering functions of one (or several) variable(s).

The learning objectives of this chapter are to review

- properties of lines and their representations and
- the different forms of a curve of second order.

6.1 Lines

A line is determined by two points $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ (see Fig. 6.1). For the angle α, we have the equality

$$\tan\alpha = \frac{y_2 - y_1}{x_2 - x_1} .$$

The term $\tan\alpha$ is the **slope** of the line. By means of the two points P and Q, we get the **two-point equation** of a line:

$$y - y_1 = \frac{y_2 - y_1}{x_2 - x_1} \cdot (x - x_1) .$$

Using the points $P = (0, b)$ and $Q = (x, y)$ and setting $a = \tan\alpha$, we obtain the **normal form** (or **point-slope form**) of a line:

$$y = ax + b .$$

Thus, the value b gives the y-coordinate, where the line intersects the y-axis. If $b = 0$, the line $y = ax$ goes through the origin of the coordinate system. If $a = 0$, the line $y = b$ is parallel to the x-axis of the coordinate system. All equations of a line can be put into the form

$$Ax + By + C = 0,$$

where parameters A and B are not both equal to zero. This form is also known as the **general equation** of a line.

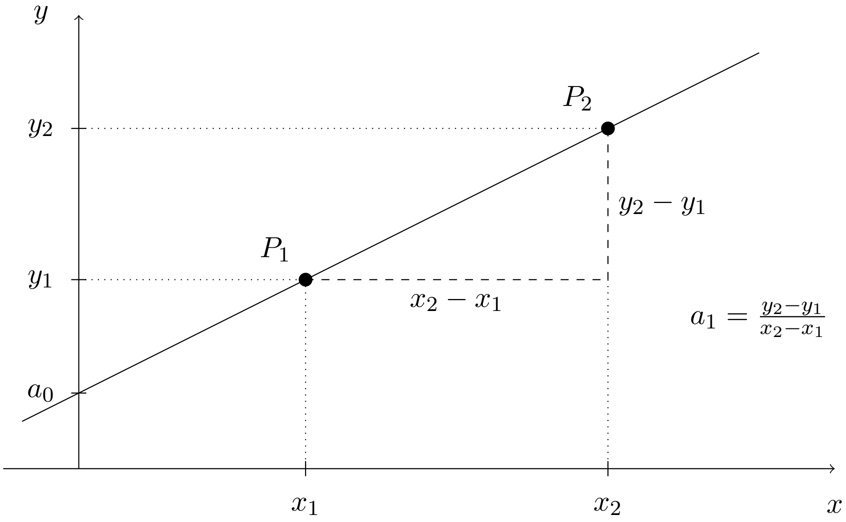


Figure 6.1: Definition of a line

Next, we introduce the **point-direction equation** of a line in the plane. It uses one point $P = (x_1, x_1)$ and a directional vector $\mathbf{a}$ (see also Chapter 11) which can be obtained by means of a second point $Q = (x_2, y_2)$:

$$\mathbf{a} = \begin{pmatrix} a_x \\ a_y \end{pmatrix} = \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix} .$$

Then the equation of the line is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + t \cdot \begin{pmatrix} a_x \\ a_y \end{pmatrix} , \qquad t \in \mathbb{R} .$$

If a line does not pass through the origin, its equation can be written in the **intercept form**:

$$\frac{x}{a} + \frac{y}{b} = 1 .$$

Here the parameters a and b give the intercepts with the x- and y-axis, respectively (see Fig. 6.2).

Example 6.1 *We put a line through the point $P = (3, 4)$ which has an angle of 45^o with the x-axis. Thus, we have*

$$a = \tan \alpha = \tan 45^o = 1 .$$

This gives

$$\begin{aligned} y - 4 &= 1 \cdot (x - 3) \\ y &= x + 1 . \end{aligned}$$

For writing the equation of this line in point-direction form, we can use e.g. the second point $Q = (4, 5)$, which gives the direction vector

$$\mathbf{a} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} .$$

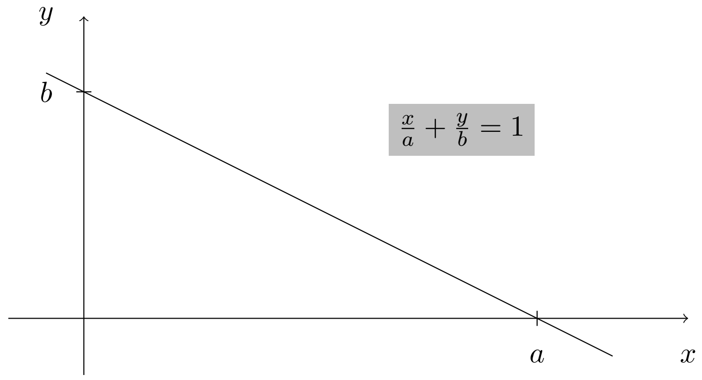


Figure 6.2: Intercept form of the equation of a line

Thus, the equation of this line is

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad t \in \mathbb{R}\,.$$

For instance, for $t = 10$*, we get the point* $R = (13, 14)$ *from the above equation and for* $t = -5$*, we get the point* $S = (-2, -1)$*.*

For writing the equation of this line in intercept form, we determine the intercepts with the x*- and* y*-axes. If* $y = 0$*, we have* $x = -1$ *and if* $x = 0$*, we have* $y = 1$*. This gives*

$$\frac{x}{-1} + \frac{y}{1} = 1\,.$$

Example 6.2 *Let the line*

$$5x - 2y + 6 = 0$$

be given. Solving for y*, we get the normal form*

$$\begin{aligned} y &= \frac{1}{2} \cdot (5x + 6) \\ y &= \frac{5}{2}x + 3\,. \end{aligned}$$

Writing this equation in the intercept form, we first find the intercepts with the axes at $y = 3$ *and* $x = -6/5$ *which gives*

$$\frac{x}{-\frac{6}{5}} + \frac{y}{3} = 1\,.$$

In the plane, two points $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ have the (Euclidean) distance

$$\boxed{d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}\,.}$$

If we consider a given point $P = (x_0, y_0)$ and a given line (in the plane), the distance of the point P from the line is determined by the smallest distance of this point P to some point Q belonging to the line. The **distance** d of point P from this line can be found according to the following formula:

$$d = \frac{Ax_0 + By_0 + C}{\sqrt{A^2 + B^2}} ,$$

where A, B, C are the parameters in the general form of the line.

Example 6.3 *Consider the point $P = (2, -4)$ and the line $y = 3x + 2$ or in the general form*

$$3x - y + 2 = 0 .$$

We determine the distance d of the point P from this line and obtain with $A = 3, B = -1$ and $C = 2$

$$d = \frac{3 \cdot 2 - 1 \cdot (-4) + 2}{\sqrt{3^2 + (-1)^2}} = \frac{12}{\sqrt{10}} \approx 3.7947 .$$

6.2 Curves of Second Order

A **curve of second order** has the equation

$$Ax^2 + By^2 + Cx + Dy + E = 0 .$$

Both variables x and y occur at most with the second power, and we assume that the product xy of both variables does not occur. Such a curve can be a circle, an ellipse, a parabola or a hyperbola. We discuss the individual cases separately.

6.2.1 Circles

Circle:
A **circle** is the set of all points having from a midpoint (or center) M the same distance r.

A circle is determined by the midpoint (x_0, y_0) and the radius r. The **equation of the circle** is given by

$$(x - x_0)^2 + (y - y_0)^2 = r^2 \; .$$

For the special case when the origin of the coordinate system is the midpoint the equation simplifies to

$$x^2 + y^2 = r^2 \; .$$

The equation of a circle is illustrated in Fig. 6.3.

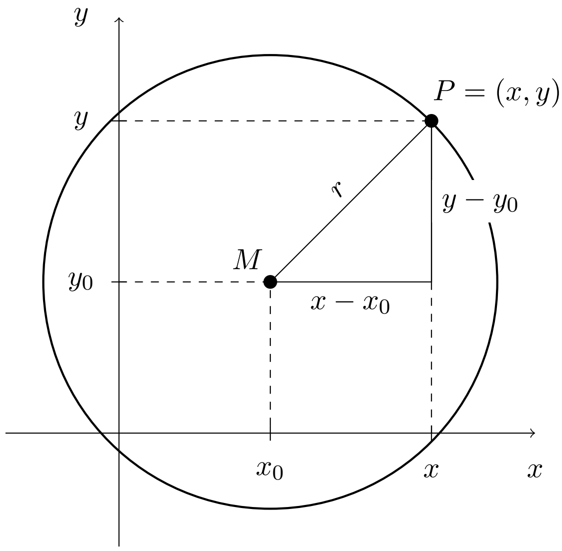


Figure 6.3: Definition of a circle

Example 6.4 *A circle has the midpoint* $M = (2, 3)$ *and it includes the point* $P = (5, 7)$*. We determine the equation of this circle. Since the point* P *must satisfy the equation of the circle, we obtain with* $x = 5$ *and* $y = 7$

$$\begin{aligned}(5-2)^2 + (7-3)^2 &= r^2 \\ 9 + 16 &= r^2\end{aligned}$$

Therefore, the equation of the circle is

$$(x-2)^2 + (y-3)^2 = 25,$$

i.e., the radius of the circle is $r = 5$*.*

6.2.2 Ellipses

Ellipse:
An **ellipse** is defined as the locus of points whose distances from two fixed points, called the foci or focal points, have a constant sum.

The **equation of an ellipse** with the midpoint $M = (x_0, y_0)$ and the half axes a and b is given by

$$\boxed{\frac{(x-x_0)^2}{a^2} + \frac{(y-y_0)^2}{b^2} = 1\,.}$$

The equation of an ellipse is illustrated in Fig. 6.4. The focal points are F_1 and F_2, and according to the definition of an ellipse we have

$$\overline{PF_1} + \overline{PF_2} = 2a\,.$$

If $a = b$, we have the special case of a circle.

Example 6.5 *We consider the equation*

$$7x^2 + 4y^2 - 28x + 8y + 4 = 0\,.$$

Transforming the terms depending on the variables x *and* y*, respectively, into complete squares, we obtain*

$$\begin{aligned}7x^2 - 28x &= 7(x^2 - 4x) = 7(x^2 - 4x + 4) - 7 \cdot 4 = 7(x-2)^2 - 28 \\ 4y^2 + 8y &= 4(y^2 + 2y) = 4(y^2 + 2y + 1) - 4 \cdot 1 = 4(y+1)^2 - 4\end{aligned}$$

Substituting this into the original equation, we obtain

$$7(x-2)^2 - 28 + 4(y+1)^2 - 4 + 4 = 0$$

which turns into

$$7(x-2)^2 + 4(y+1)^2 = 28$$

Dividing both sides by 28, we get the equation of an ellipse:

$$\frac{(x-2)^2}{4} + \frac{(y+1)^2}{7} = 1\,,$$

i.e., the midpoint of the ellipse is $(2, -1)$*, and the half axes are* $a = 2$ *and* $b = \sqrt{7}$*.*

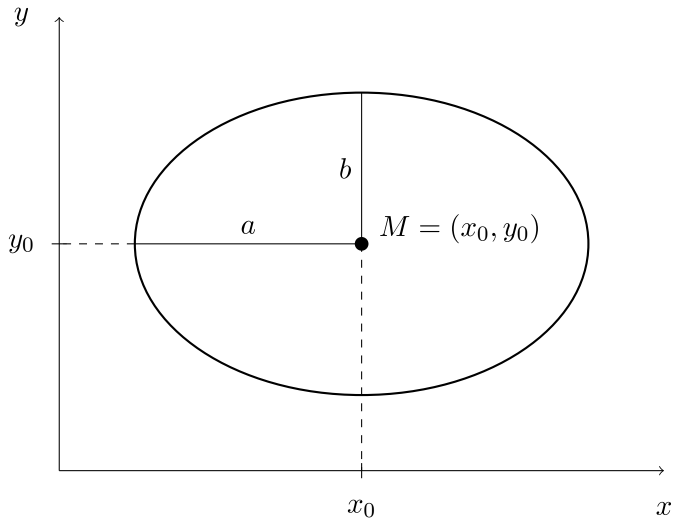


Figure 6.4: Definition of an ellipse

6.2.3 Parabolas

Parabola:
A **parabola** is the locus of points which are equidistant from a line (called the directrix) and from a fixed point (called the focus or focal point).

The definition of a parabola is illustrated in Fig. 6.5. For the **equation of a parabola** we have to distinguish two cases. If the parabola is parallel to the x-axis, the equation is given by

$$\boxed{(y - y_0)^2 = \pm 2p(x - x_0) \ .}$$

The point (x_0, y_0) is called the **apex** (or **vertex**) of the parabola. The value $2p$ denotes the parameter of the parabola. If a plus sign occurs in the equation, the parabola is open from the right (with $x \geq x_0$). If the minus sign occurs, the parabola is open from the left (with $x \leq x_0$). If the parabola is parallel to the y-axis, its equation is given by

$$\boxed{(x - x_0)^2 = \pm 2p(y - y_0) \ .}$$

If the plus sign occurs in the equation, the parabola is open from above (with $y \geq y_0$) while in the other case, the parabola is open from below (with $y \leq y_0$). A parabola open from the right and one open from above are illustrated in Fig. 6.6.

Accordingly, one can consider parabolas which are open from the left or from below.

Example 6.6 *Consider the equation*

$$y^2 + 4y - 2x + 10 = 0.$$

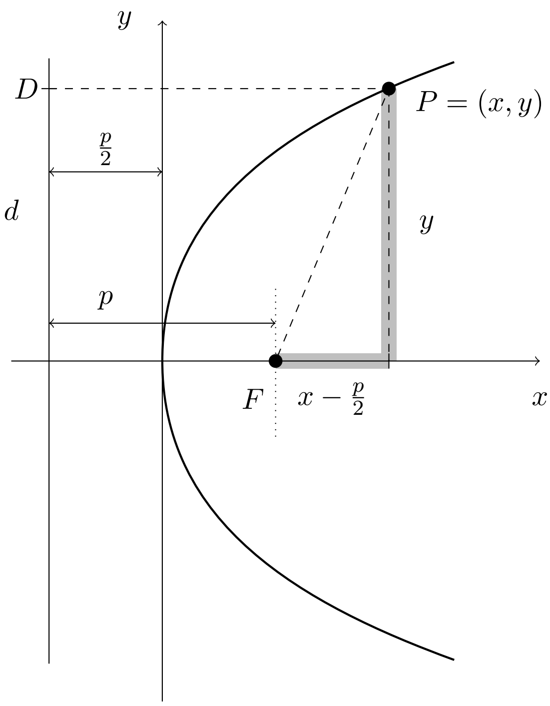


Figure 6.5: Definition of a parabola

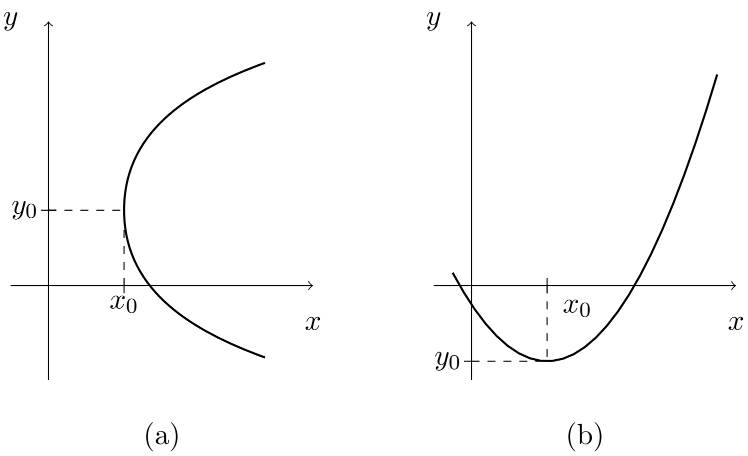


Figure 6.6: A parabola, which is (a) open from the right and (b) open from above

Putting all terms depending on y on the left-hand site and the rest on the right-hand side, we get

$$y^2 + 4y = 2x - 10.$$

Transforming now the term on the left-hand side into a complete square of the form $(y + a)^2$, we get

$$(y + 2)^2 - 4 = 2x - 10$$

which gives

$$(y + 2)^2 = 2x - 6 = 2(x - 3) .$$

Therefore, the above parabola is open from the right, and the apex A has the coordinates $(x_0, x_0) = (3, -2)$.

6.2.4 Hyperbolas

Hyperbola:
A **hyperbola** is defined as the locus of all points whose difference of the distances from two fixed points (call the foci or focal points) is constant.

A hyperbola consists of two branches. An illustration is given in Fig. 6.7 for the case, where the center M of the hyperbola is the origin of the coordinate system. F_1 and F_2 are the foci, and we have for the difference of the distances between the points P and F_1 as well as P and F_2

$$|PF_1| - |PF_2| = \pm 2a ,$$

where $2a$ denotes the principal axis of the hyperbola, and the minus sign holds for the left branch of the hyperbola.

The **equation of a hyperbola** with the center $M = (x_0, y_0)$ is given by

$$\frac{(y - y_0)^2}{a^2} - \frac{(x - x_0)^2}{b^2} = 1 .$$

or

$$\frac{(x - x_0)^2}{a^2} - \frac{(y - y_0)^2}{b^2} = 1 .$$

In the first case, the hyperbola is symmetric to the line $x = x_0$ (i.e., parallel to the y-axis), and we also say that it is a **vertical hyperbola** (see Fig. 6.8 (a)). In the second case, the hyperbola is symmetric to the line $y = y_0$ (i.e., parallel to the x-axis), and we say that it is a **horizontal hyperbola** (see Fig. 6.8 (b)). Here the parameters a and b denote the **half axes** of the hyperbola, and $2a$ denotes the principal axis of the hyperbola.

For graphing a hyperbola, it is useful to determine the asymptotes L_1 and L_2 of the hyperbola. From the equation of a hyperbola, they are obtained as follows:

$$L_1 : \quad y = \frac{b}{a}x ; \qquad\qquad L_2 : \quad y = -\frac{b}{a}x .$$

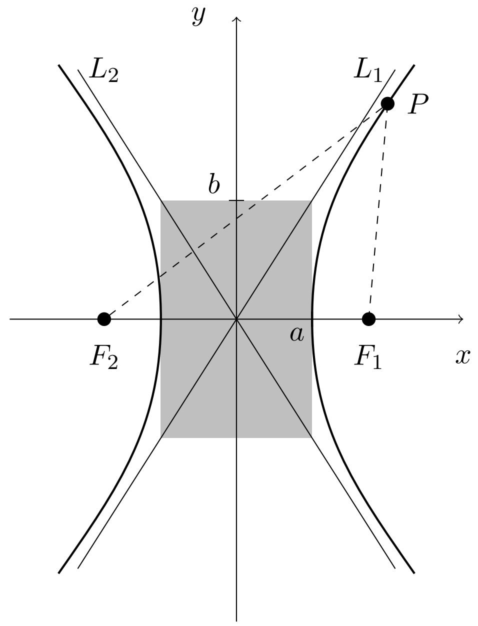


Figure 6.7: Definition of a hyperbola

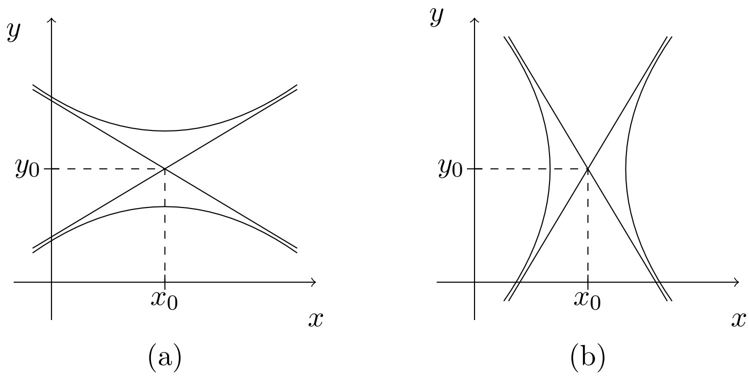


Figure 6.8: (a) Vertical and (b) horizontal hyperbolas

Example 6.7 *Let us consider the equation*

$$4x^2 - 9y^2 + 32x + 18y + 19 = 0$$

Transforming the terms depending on x *and* y *into a complete square, we obtain*

$$\begin{aligned} 4x^2 + 32x &= 4(x^2 + 8x) = 4(x^2 + 8x + 16) - 4 \cdot 16 = 4(x+4)^2 - 64 \\ -9y^2 + 18y &= -9(y^2 - 2y) = -9(y^2 - 2y + 1) + 9 \cdot 1 = -9(y-1)^2 + 9 \end{aligned}$$

Substituting this into the original equation, we get

$$4(x+4)^2 - 64 - 9(y-1)^2 + 9 + 19 = 0$$

which gives

$$4(x+4)^2 - 9(y-1)^2 - 36 = 0 .$$

Dividing this equation by 36, we obtain the equation of a hyperbola:

$$\frac{(x+4)^2}{9} - \frac{(y-1)^2}{4} = 1 .$$

This hyperbola has the center (or midpoint) $M = (-4, 1)$ *and the half axes* $a = 3$ *and* $b = 2$*. It is horizontally open and symmetric to the line* $x = -4$*, and the asymptotes are*

$$L_1 = \frac{4}{9}x \qquad \text{and} \qquad L_2 = -\frac{4}{9}x .$$

.

EXERCISES

6.1 For the following lines, determine the normal form and the intercept form (if possible).:

(a) $3x + 7y - 2 = 0$; (b) $-2x + 5y + 5 = 0$; (c) $3x + 11y = 0$.

6.2 Determine the normal form of the following lines going through point P and having the angle α with the positive x-axis:

(a) $P = (3, 1),\ \alpha = 60^o$; (b) $P = (-1, -1),\ \alpha = 45^o$; (c) $P = (4, -2),\ \alpha = 135^o$.

6.3 Determine the normal form of the following lines going through the point P and Q:

(a) $P = (-3, 4),\ Q = (-1, -3)$; (b) $P = (1, 7),\ Q = (4, 9)$.

6.4 Determine the equation of the circle having the midpoint $M = (1, 1)$ and going through the point $P = (5, 4)$.

6.5 Consider the parabola $8x - y^2 + 16x - 80 = 0$. Determine its apex and from which side it is open.

6.6 Characterize the location of the following parabolas and graph them:

(a) $x + y^2 + 4y - 24 = 0$; (b) $x^2 - 8y + 10 = 0$; (c) $x^2 + 6x + 5y + 8 = 0$.

6.7 An ellipse has the midpoint $(0, 0)$ and goes through the point $P = (3, 8)$. One of the half axes is 5. Determine the equation of this ellipse.

6.8 Determine the equation of the hyperbola with the center $M = (1, 1)$ and the half axes $a = 4$ and $b = 7$.

Chapter 7

Sequences and Partial Sums

Sequences play a role when listing infinitely many objects. For instance, sequences are used in connection with time series. They are also important when giving construction instructions for algorithms. We also discuss the fundamental notion of a limit of a sequence. In this chapter, we briefly review

- some basic notions about sequences;
- arithmetic and geometric sequences as important special sequences;
- properties of a sequence and
- partial sums of sequences.

7.1 Basic Notions

Sequence:
If a real number a_n is assigned to each natural number $n \in \mathbb{N}$, then

$$\{a_n\} = a_1, a_2, a_3, \dots, a_n, \dots$$

is called a **sequence**.

The numbers $a_1, a_2, a_3, \dots$ are called the **terms** of the sequence. In particular, the number a_n is denoted as the **nth term**. One can define a sequence in two different forms:

- **explicitly** by giving a mathematical term how the nth term of the sequence can be calculated;
- **recursively** by giving the first term(s) of the sequence and a recursive formula for calculating the remaining terms of the sequence.

We illustrate the calculations of the terms of a sequence by the following examples.

Example 7.1 *Consider the sequence*

$$\{a_n\} = \left\{ \frac{n-2}{3n+1} \right\} .$$

We determine the terms $a_1, a_2, \ldots, a_6$ *and obtain*

$$a_1 = \frac{1-2}{3\cdot 1+1} = -\frac{1}{4}\,, \qquad a_2 = \frac{2-2}{3\cdot 2+1} = 0\,, \qquad a_3 = \frac{3-2}{3\cdot 3+1} = \frac{1}{10}\,,$$
$$a_4 = \frac{4-2}{3\cdot 4+1} = \frac{2}{13}\,, \qquad a_5 = \frac{5-2}{3\cdot 5+1} = \frac{3}{16}\,, \qquad a_6 = \frac{6-2}{3\cdot 6+1} = \frac{4}{19}\,.$$

Example 7.2 *Consider the sequence*

$$\{b_n\} = \left\{ (-1)^{n+1} \cdot \frac{n^2}{n+1} \right\} .$$

Then we obtain for the first six terms:

$$b_1 = (-1)^{1+1} \cdot \frac{1^2}{1+1} = \frac{1}{2}\,; \qquad b_2 = (-1)^{2+1} \cdot \frac{2^2}{2+1} = -\frac{4}{3}\,; \qquad b_3 = (-1)^{3+1} \cdot \frac{3^2}{3+1} = \frac{9}{4}\,;$$
$$b_4 = (-1)^{4+1} \cdot \frac{4^2}{4+1} = -\frac{16}{5}\,; \quad b_5 = (-1)^{5+1} \cdot \frac{5^2}{5+1} = \frac{25}{6}\,; \qquad b_6 = (-1)^{6+1} \cdot \frac{6^2}{6+1} = -\frac{36}{7}\,.$$

We observe that any two successive terms of the above sequence have a different sign. Such a sequence is called an **alternating sequence**.

Example 7.3 *Let a sequence* $\{c_n\}$ *be given by*

$$c_1 = 3, \qquad c_n = c_{n-1}^2 - 5,\ n \geq 2\,.$$

In this case, the sequence is recursively defined and we can calculate a particular term if the preceding term is known. We obtain for the first six terms:

$$c_1 = 3; \qquad c_2 = c_1^2 - 5 = 3^2 - 5 = 4;$$
$$c_3 = c_2^2 - 5 = 4^2 - 5 = 11; \qquad c_4 = c_3^2 - 5 = 11^2 - 5 = 116;$$
$$c_5 = c_4^2 - 5 = 116^2 - 5 = 13,451; \qquad c_6 = c_5^2 - 5 = (13,451)^2 - 5 = 180,929,396\,.$$

Example 7.4 *Consider the sequence* $\{d_n\}$ *given by*

$$d_1 = 0, \qquad d_2 = 1, \qquad d_n = d_{n-2} + d_{n-1},\ n \geq 2\,.$$

This is a recursively defined sequence with two initial terms. Notice here that the calculation of a particular term requires to know the last two preceding terms. In particular, we get for the first six recursively calculated terms $d_3, d_4, \ldots, d_8$*:*

$$d_3 = d_1 + d_2 = 0 + 1 = 1; \qquad d_4 = d_2 + d_3 = 1 + 1 = 2;$$
$$d_5 = d_3 + d_4 = 1 + 2 = 3; \qquad d_6 = d_4 + d_5 = 2 + 3 = 5;$$
$$d_7 = d_5 + d_6 = 3 + 5 = 8; \qquad d_8 = d_6 + d_7 = 5 + 8 = 13;$$

This particular sequence is also known as **Fibonacci sequence**.

7.2 Arithmetic Sequences

Arithmetic sequence:
A sequence $\{a_n\}$, where the difference of any two successive terms is constant, is called an **arithmetic sequence**, i.e., the equality

$$a_{n+1} - a_n = d$$

holds for all $n \in \mathbb{N}$, where d is constant.

Thus, the terms of an arithmetic sequence with the first term a_1 are as follows:

$$a_1, \qquad a_2 = a_1 + d, \quad a_3 = a_1 + 2d, \quad a_4 = a_1 + 3d, \quad \ldots,$$

and we obtain the following explicit formula for the nth term:

$$a_n = a_1 + (n-1)\,d \qquad \text{for } n \in \mathbb{N}.$$

Example 7.5 *We consider the sequence*

$$\{a_n\} = 1,\ \frac{4}{3},\ \frac{5}{3},\ 2,\ \frac{7}{3},\ \ldots\ .$$

This is an arithmetic sequence with the first term $a_1 = 1$ and the difference $d = 1/3$. Therefore, the nth term is obtained by

$$a_n = a_1 + (n-1) \cdot d = 1 + \frac{1}{3} \cdot (n-1)\ .$$

Example 7.6 *We consider an arithmetic sequence with the third term* $a_3 = 6$ *and the 14th term* $a_{14} = 50$*. We get the difference* d *of this arithmetic sequence from the equality*

$$a_{14} - a_3 = 11 \cdot d$$

which gives

$$50 - 6 = 44 = 11 \cdot d$$

and thus, $d = 4$*. Since* $a_3 = a_1 + 2 \cdot d$*, we obtain from* $6 = a_1 + 2 \cdot d = a_1 + 8$ *the first term* $a_1 = -2$*.*

7.3 Geometric Sequences

Geometric sequence:
A sequence $\{a_n\}$, where the ratio of any two successive terms is the same number $q \neq 0$, is called a **geometric sequence**, i.e., the equality

$$\frac{a_{n+1}}{a_n} = q$$

holds for all $n \in \mathbb{N}$, where q is constant.

Thus, the terms of a geometric sequence with the first term a_1 are as follows:

$$a_1, \quad a_2 = a_1 \cdot q, \quad a_3 = a_1 \cdot q^2, \quad a_4 = a_1 \cdot q^3, \quad \ldots,$$

and we obtain the following explicit formula for the nth term:

$$a_n = a_1 \cdot q^{n-1} \qquad \text{for all } n \in \mathbb{N}.$$

Example 7.7 *Consider the sequence*

$$\{a_n\} = 1,\ 3,\ 9,\ 27,\ 81,\ 243,\ 729,\ \ldots.$$

This is a geometric sequence with the first term $a_1 = 1$ and the quotient $q = 3$.

Example 7.8 *A geometric sequence $\{b_n\}$ has the second term $b_2 = -8$ and the fifth term $b_5 = 512$. From $b_2 = b_1 \cdot q$ and $b_5 = b_1 \cdot q^4$, we get*

$$\frac{b_5}{b_2} = \frac{512}{-8} = -64 = q^3$$

which gives as the only real solution $q = -4$. Thus, for the first term, we have

$$b_2 = b_1 \cdot q$$

which yields

$$-8 = b_1 \cdot (-4)$$

from which we obtain the first term $b_1 = 2$.

Example 7.9 *Let the sequence $\{c_n\}$ be given by*

$$c_n = \frac{2}{3^{n+2}}\,.$$

Then we obtain

$$c_1 = \frac{2}{3^{1+2}} = \frac{2}{3^3} = \frac{2}{27}$$

and

$$\frac{c_{n+1}}{c_n} = \frac{2 \cdot 3^{n+2}}{3^{(n+1)+2} \cdot 2} = 3^{n+2-(n+3)} = 3^{-1} = \frac{1}{3}\,.$$

Thus, the sequence $\{c_n\}$ is a geometric sequence with the first term $c_1 = 2/27$ and the quotient $q = 1/3$.

7.4 Properties of Sequences

Often one wishes to find characteristic properties of a sequence, e.g. whether the terms of a sequence can become arbitrarily small or large or whether each subsequent term of a sequence is larger than the previous one. In this section, we discuss the monotonicity and the boundedness of a sequence.

Monotonicity of a sequence:
A sequence $\{a_n\}$ is called **increasing** (or equivalently, **non-decreasing**) if

$$a_n \leq a_{n+1} \qquad \text{for all } n \in \mathbb{N}.$$

A sequence $\{a_n\}$ is called **strictly increasing** if

$$a_n < a_{n+1} \qquad \text{for all } n \in \mathbb{N}.$$

A sequence $\{a_n\}$ is called **decreasing** (or equivalently, **non-increasing**) if

$$a_n \geq a_{n+1} \qquad \text{for all } n \in \mathbb{N}.$$

A sequence $\{a_n\}$ is called **strictly decreasing** if

$$a_n > a_{n+1} \qquad \text{for all } n \in \mathbb{N}.$$

A sequence $\{a_n\}$ which is (strictly) increasing or (strictly) decreasing is also denoted as (strictly) **monotonic** (or **monotone**).

When checking a sequence $\{a_n\}$ for monotonicity, we may investigate the difference

$$D_n = a_{n+1} - a_n$$

of two successive terms. Then:

- If $D_n \geq 0$ for all $n \in \mathbb{N}$, the sequence $\{a_n\}$ is increasing;
- If $D_n > 0$ for all $n \in \mathbb{N}$, the sequence $\{a_n\}$ is strictly increasing;
- If $D_n \leq 0$ for all $n \in \mathbb{N}$, the sequence $\{a_n\}$ is decreasing;
- If $D_n < 0$ for all $n \in \mathbb{N}$, the sequence $\{a_n\}$ is strictly decreasing.

Example 7.10 *We investigate the sequence $\{a_n\}$ with*

$$a_n = 3(n+1)^2 - 2n, \qquad n \in \mathbb{N},$$

for monotonicity, i.e., we investigate the difference of two successive terms and obtain

$$\begin{aligned} a_{n+1} - a_n &= 3(n+2)^2 - 2(n+1) - \Big[3(n+1)^2 - 2n\Big] \\ &= 3n^2 + 12n + 12 - 2n - 2 - \big(3n^2 + 6n + 3 - 2n\big) \\ &= 6n + 7. \end{aligned}$$

Since $6n + 7 > 0$ for all $n \in \mathbb{N}$, we get $a_{n+1} > a_n$ for all $n \in \mathbb{N}$. Therefore, the sequence $\{a_n\}$ is strictly increasing.

Boundedness of a sequence:
A sequence $\{a_n\}$ is called **bounded** if there exists a finite real constant C such that

$$|a_n| \leq C \qquad \text{for all } n \in \mathbb{N}.$$

Such a constant C is also denoted as a **bound**. According to the definition of the absolute value, this means that the inequalities $-C \leq a_n \leq C$ hold for all terms a_n.

Example 7.11 *Let us investigate the alternating sequence*

$$\{a_n\} = \left\{(-1)^n \cdot \frac{n-1}{n+2}\right\}$$

whether it is bounded. We estimate the absolute value of the nth term and obtain

$$|a_n| = \left|(-1)^n \cdot \frac{n-1}{n+2}\right| = |(-1)^n| \cdot \left|\frac{n-1}{n+2}\right| = \left|\frac{n-1}{n+2}\right| \leq 1.$$

Therefore, the sequence $\{a_n\}$ is bounded, e.g., $C = 1$ delivers such a bound.

7.5 Limit of a Sequence

We start with the introduction of a limit which is a central notion in mathematics. It is needed for many considerations in analysis. In general, this notion is necessary in mathematics to make the step from ‘finiteness’ to ‘infiniteness’. The correct introduction of irrational numbers is also only possible by means of limit considerations.

Limit of a sequence:
A finite number a is called the **limit** of a sequence $\{a_n\}$ if, for any given $\varepsilon > 0$, there exists an index $n(\varepsilon)$ such that

$$|a_n - a| < \varepsilon \qquad \text{for all } n \geq n(\varepsilon).$$

Note that the above inequality with the absolute value is used for a simpler presentation of the equivalent inequalities that

$$a - \epsilon < a_n < a_n + \varepsilon \qquad \text{for all } n \geq n(\varepsilon)$$

To indicate that the number a is the limit of the sequence $\{a_n\}$, we write

$$\lim_{n\to\infty} a_n = a.$$

The notion of the limit a is illustrated in Fig. 7.1. The sequence $\{a_n\}$ has the limit a, if there exists some index $n(\varepsilon)$ (depending on the positive number ε) such that the absolute value of the difference between the term a_n and the limit a becomes smaller than the given value ε for all terms a_n with $n \geq n(\varepsilon)$, i.e., from some number n on, all terms of the sequence $\{a_n\}$ are very close to the limit a. If ε becomes smaller, the corresponding value $n(\varepsilon)$ becomes larger.

We note that a limit must be a **finite real number**. If the terms of a sequence tend e.g. to ∞, then we also write as an abbreviation

$$\lim_{n\to\infty} a_n = \infty\ .$$

However, in this case the limit of the sequence $\{a_n\}$ does not exist.

Next, we give a few rules for working with limits of sequences. Assume that the limits

$$\lim_{n\to\infty} a_n = a \qquad \text{and} \qquad \lim_{n\to\infty} b_n = b$$

exist. Then the following limits exist, and we obtain:

Rules for working with limits of a sequence:

(1) $\lim\limits_{n\to\infty} (a_n \pm C) = \lim\limits_{n\to\infty} (a_n) \pm C = a \pm C, \quad (C \in \mathbb{R})\,;$

(2) $\lim\limits_{n\to\infty} (C \cdot a_n) = C \cdot a, \quad (C \in \mathbb{R})\,;$

(3) $\lim\limits_{n\to\infty} (a_n \pm b_n) = a \pm b\,;$

(4) $\lim\limits_{n\to\infty} (a_n \cdot b_n) = a \cdot b\,;$

(5) $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = \dfrac{a}{b} \qquad (b_n \neq 0 \text{ for all } n,\, b \neq 0)\,.$

In the formulas above, it is assumed that $C \in \mathbb{R}$ is constant. The use of the above formulas requires the knowledge of some limits. Next, we summarize some known limits.

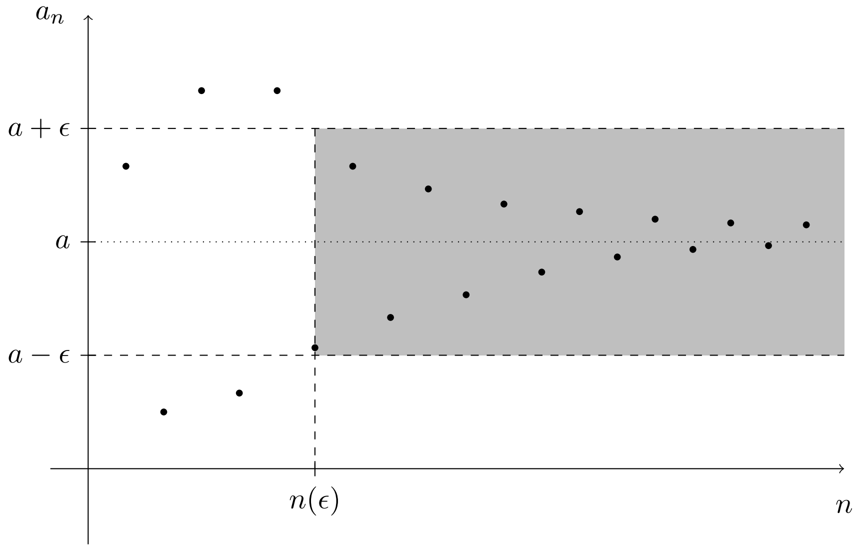


Figure 7.1: Limit of a sequence

Some limits:

(1) $\lim_{n\to\infty} \frac{1}{n^a} = 0$ for $a > 0$;

(2) $\lim_{n\to\infty} a^n = 0$ for $a \in \mathbb{R}$, $|a| < 1$;

(3) $\lim_{n\to\infty} \sqrt[n]{a} = 1$ for $a \in \mathbb{R}_{>0}$;

(4) $\lim_{n\to\infty} \frac{a^n}{n!} = 0$ for $a \in \mathbb{R}$;

(5) $\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e$.

We illustrate the calculation of limits by the following examples.

Example 7.12 *Consider the sequence* $\{a_n\}$ *with*

$$a_n = \frac{3}{n^2}.$$

We can rewrite the term a_n *as*

$$a_n = 3 \cdot \frac{1}{n^2}.$$

Using rule (2) and limit (1) above, we get

$$\lim_{n\to\infty} a_n = 3 \cdot \lim_{n\to\infty} \frac{1}{n^2} = 3 \cdot 0 = 0\,.$$

Example 7.13 *Let the sequences* $\{a_n\}$ *and* $\{b_n\}$ *with*

$$a_n = \frac{3n^2+n-4}{4n^2-2}, \qquad b_n = \frac{6n^2+2}{n^3+n}, \qquad n \in \mathbb{N},$$

be given. Using the above rules and known limits, we obtain

$$\begin{aligned}\lim_{n\to\infty} a_n &= \lim_{n\to\infty} \frac{3n^2+n-4}{4n^2-2} = \lim_{n\to\infty} \frac{n^2\left(3+\frac{1}{n}-\frac{4}{n^2}\right)}{n^2\left(4-\frac{2}{n^2}\right)} \\ &= \frac{\lim\limits_{n\to\infty} 3 + \lim\limits_{n\to\infty} \frac{1}{n} - 4\cdot \lim\limits_{n\to\infty} \frac{1}{n^2}}{\lim\limits_{n\to\infty} 4 - 2\cdot \lim\limits_{n\to\infty} \frac{1}{n^2}} = \frac{3+0-0}{4+0} = \frac{3}{4}.\end{aligned}$$

Similarly, we get

$$\begin{aligned}\lim_{n\to\infty} b_n &= \lim_{n\to\infty} \frac{6n^2+2}{n^3+n} = \lim_{n\to\infty} \frac{n^3\left(\frac{6}{n}+\frac{2}{n^3}\right)}{n^3\left(1+\frac{1}{n^2}\right)} \\ &= \frac{6\cdot \lim\limits_{n\to\infty} \frac{1}{n} + 2\cdot \lim\limits_{n\to\infty} \frac{1}{n^3}}{\lim\limits_{n\to\infty} 1 + \lim\limits_{n\to\infty} \frac{1}{n^2}} = \frac{0+0}{1+0} = 0.\end{aligned}$$

Consider now the sequence $\{c_n\}$ with $c_n = a_n + b_n$, $n \in \mathbb{N}$. Applying the given rules for limits, we immediately get

$$\lim_{n\to\infty} c_n = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n = \frac{3}{4} + 0 = \frac{3}{4}.$$

One can generalize the result from the above example. Let $p, q \in \mathbb{N}$, then we obtain for such special rational terms given as a fraction of two polynomial terms in n:

$$\lim_{n\to\infty} \frac{a_p n^p + a_{p-1} n^{p-1} + \ldots + a_1 n + a_0}{b_q n^q + b_{q-1} n^{q-1} + \ldots + b_1 n + b_0} = \begin{cases} 0 & \text{for } p < q \\ \dfrac{a_p}{b_q} & \text{for } p = q \\ \pm\infty & \text{for } p > q. \end{cases}$$

This means that, in order to find the above limit, we have to check only the terms with the largest exponent in the numerator and in the denominator.

Example 7.14 *Using the previous formula, we immediate get (without calculations):*

$$\begin{aligned}\lim_{n\to\infty} \frac{2n^4+n^3-2n}{n^4-n^2+5n+7} &= 2\,; \\ \lim_{n\to\infty} \frac{-4n^3+3n^2+n-3}{n^2-n-1} &= -\infty\,; \\ \lim_{n\to\infty} \frac{6n^3+n+7}{n^4+2n^3-n^2+2} &= 0\,;\end{aligned}$$

7.6 Partial Sums

Next, we consider partial sums of sequences. They are often required in several applications, e.g. in finance when one wishes to know the final amount in a saving account after some years when depositing a monthly amount and interest is paid or in connection with loan repayments.

nth partial sum:
Let $\{a_n\}$ be a sequence. Then the sum of the first n terms

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

is called the **nth partial sum** s_n.

Example 7.15 *Consider the sequence* $\{a_n\}$ *with*

$$a_n = 2 + (-1)^{n-1} \cdot \frac{1}{n}, \qquad n \in \mathbb{N}.$$

This gives

$$a_1 = 2 + (-1)^0 \cdot \frac{1}{1} = 3, \quad a_2 = 2 + (-1)^1 \cdot \frac{1}{2} = \frac{3}{2}, \quad a_3 = 2 + (-1)^2 \cdot \frac{1}{3} = \frac{7}{3},$$
$$a_4 = 2 + (-1)^3 \cdot \frac{1}{4} = \frac{7}{4}, \quad a_5 = 2 + (-1)^4 \cdot \frac{1}{5} = \frac{11}{5}, \quad \ldots.$$

Then we get the following partial sums:

$$s_1 = a_1 = 3, \quad s_2 = a_1 + a_2 = 3 + \frac{3}{2} = \frac{9}{2}, \quad s_3 = a_1 + a_2 + a_3 = 3 + \frac{3}{2} + \frac{7}{3} = \frac{41}{6},$$
$$s_4 = a_1 + a_2 + a_3 + a_4 = 3 + \frac{3}{2} + \frac{7}{3} + \frac{7}{4} = \frac{82 + 21}{12} = \frac{103}{12},$$
$$s_5 = a_1 + a_2 + a_3 + a_4 + a_5 = 3 + \frac{3}{2} + \frac{7}{3} + \frac{7}{4} + \frac{11}{5} = \frac{515 + 132}{60} = \frac{647}{60}.$$

Of course, we can also use that $s_i = s_{i-1} + a_i$ for $i = 2, 3, \ldots$, which simplifies the calculations.

For arithmetic and geometric sequences, one can easily calculate the corresponding partial sums as follows.

The **nth partial sum of an arithmetic sequence** $\{a_n\}$ with $a_n = a_1 + (n-1) \cdot d$ is given by

$$\boxed{s_n = \frac{n}{2} \cdot (a_1 + a_n) = \frac{n}{2} \cdot \Big[2a_1 + (n-1)d\Big].}$$

The **nth partial sum of a geometric sequence** $\{a_n\}$ with $a_n = a_1 \cdot q^{n-1}$ and $q \neq 1$ is given by

$$s_n = a_1 \cdot \frac{1-q^n}{1-q} .$$

Notice that for $q = 1$, we get a sequence having all the same term a_1 which gives the nth partial sum $s_n = n \cdot a_1$. Let us consider for illustration the following problems:

Example 7.16 *We determine the sum of the first 50 even natural numbers, i.e., we calculate*

$$S = 2 + 4 + 6 + \ldots + 98 + 100 .$$

This can be done by calculating the 50th partial sum of an arithmetic sequence with $a_1 = 2$ and $d = 2$, which gives

$$a_{50} = a_1 + 49 \cdot d = 2 + 49 \cdot 2 = 100.$$

Thus, we obtain

$$s_{50} = \frac{50}{2} \cdot (a_1 + a_{50}) = 25 \cdot (2 + 100) = 2,550 .$$

Example 7.17 *Let $\{a_n\}$ be an arithmetic sequence with the first term $a_1 = 3$, the difference $d = 3$ and the nth partial sum $s_n = 45$. We wish to determine the corresponding value n and the nth term a_n. From the formula for the nth partial sum of an arithmetic sequence, we obtain*

$$45 = s_n = \frac{n}{2} \cdot [2a_1 + (n-1) \cdot d] = \frac{n}{2} \cdot (6 + (n-1) \cdot 3) = \frac{n}{2} \cdot (3n + 3).$$

This gives the quadratic equation $90 = 3n^2 + 3n$ or

$$n^2 + n - 30 = 0$$

which has the two real solutions

$$n_1 = -\frac{1}{2} + \sqrt{\frac{1}{4} + 30} = -\frac{1}{2} + \frac{11}{2} = 5$$

and

$$n_2 = -\frac{1}{2} - \sqrt{\frac{1}{4} + 30} = -\frac{1}{2} - \frac{11}{2} = -6 .$$

Since $n_2 < 0$, the solution to our problem is $n = n_1 = 5$ which yields the 5th term

$$a_5 = a_1 + 4 \cdot d = 3 + 4 \cdot 3 = 15 .$$

Example 7.18 *Consider the sequence*

$$\{b_n\} = 2,\ 3,\ \frac{9}{2},\ \frac{27}{4},\ \frac{81}{8},\ \ldots .$$

This is a geometric sequence with the first term $b_1 = 2$ and the quotient $q = 3/2$. For the nth partial sum, we obtain

$$s_n = a_1 \cdot \frac{1-q^n}{1-q} = 2 \cdot \frac{1 - \left(\frac{3}{2}\right)^n}{1 - \frac{3}{2}} = 2 \cdot \frac{1 - \left(\frac{3}{2}\right)^n}{-\frac{1}{2}} = 2 \cdot \frac{\left(\frac{3}{2}\right)^n - 1}{\frac{1}{2}} = 4 \cdot \left[\left(\frac{3}{2}\right)^n - 1\right] .$$

Example 7.19 *We want to determine the sum*

$$S = \frac{9}{2} - \frac{9}{4} + \frac{9}{8} - \frac{9}{16} + \cdots - \frac{9}{256} + \frac{9}{512}.$$

This is the partial sum of a geometric sequence with the first term $a_1 = 9/2$ and the quotient $q = -1/2$. More precisely, it is the 9th partial sum since

$$a_9 = a_1 \cdot q^8 = \frac{9}{2} \cdot \left(-\frac{1}{2}\right)^8 = \frac{9}{2} \cdot \frac{1}{(-2)^8} = \frac{9}{2} \cdot \frac{1}{256} = \frac{9}{512}\,.$$

Therefore, we obtain

$$S = s_9 = a_1 \cdot \frac{1-q^9}{1-q} = \frac{9}{2} \cdot \frac{1-\left(-\frac{1}{2}\right)^9}{1-\left(-\frac{1}{2}\right)} = \frac{9}{2} \cdot \frac{2}{3} \cdot \left(1 - \frac{1}{2^9}\right) = 3 \cdot \left(1 - \frac{1}{512}\right) = \frac{1533}{512}\,.$$

Example 7.20 *Let $\{s_n\}$ be given by*

$$s_n = \sum_{k=1}^{n} \frac{3}{2^{k-2}}\,.$$

We determine the 20th partial sum s_{20}. We have

$$a_1 = \frac{3}{2^{1-2}} = \frac{3}{2^{-1}} = 6 \quad \text{and} \quad \frac{a_{k+1}}{a_k} = \frac{3 \cdot 2^{k-2}}{2^{k-1} \cdot 3} = 2^{k-2-(k-1)} = 2^{-1} = \frac{1}{2}\,.$$

Therefore, the sequence $\{s_n\}$ gives the partial sums of a geometric sequence with $a_1 = 6$ and $q = 1/2$, and we obtain

$$s_{20} = a_1 \cdot \frac{1-q^{20}}{1-q} = 6 \cdot \frac{1-\left(\frac{1}{2}\right)^{20}}{1-\frac{1}{2}} = 12 \cdot \left[1 - \left(\frac{1}{2}\right)^{20}\right]\,.$$

EXERCISES

7.1 Determine the first six terms of the following sequences:

(a) $\{a_n\} = \left\{2 + \frac{4}{n}\right\}$; (b) $\{b_n\} = \left\{(-1)^{n+1} \cdot \frac{n-1}{4+n}\right\}$; (c) $\{c_n\} = \{1 + 2n^2\}$.

7.2 Determine the nth term of the sequence $\{a_n\}$ given by

$$\{a_n\} = 4\,, \frac{3}{2}\,, \frac{2}{9}\,, \frac{5}{8}\,, \frac{12}{25}\,, \frac{7}{18}\,, \cdots\,.$$

7.3 Given is an arithmetic sequence $\{a_n\}$ with a first term $a_1 = 15$ and the difference $d = 8$. Find the term a_{101}.

7.4 For an arithmetic sequence, the terms $a_8 = 21$ and $a_{11} = 30$ are known. Find the difference d and the terms a_1 and a_n.

7.5 The first row in a football stadium has 376 seats and the last row has 1026 seats. Assuming that the number of seats increases from row to row by the same number, how many rows does the stadium have and what is the total number of seats?

7.6 Check the sequence

$$\{a_n\} = \left\{ \frac{n^2}{2^{n+1}} \right\}$$

for monotonicity! What is the largest term of the sequence? Is the sequence bounded?

7.7 Let $\{a_n\}$ be a geometric sequence with the ratio of successive terms $q = -2/3$ and the term $a_7 = 64/243$. Find the first term a_1. Which of the terms is the first with an absolute value less than 0.01?

7.8 How many terms of the arithmetic sequence $\{a_n\}$ with

$$a_n = 5 + 4(n-1)$$

are less than 700?

7.9 A geometric sequence has the terms $a_2 = 6$ and $a_7 = 2/81$. Find the first term a_1 and the ratio q of any successive terms.

7.10 Investigate the following sequences for monotonicity:

(a) $\{a_n\} = \left\{ \frac{n+5}{n+1} \right\}$; (b) $\{b_n\} = \left\{ \frac{2-n^2}{3n} \right\}$;

(c) $\{c_n\} = \left\{ \frac{5n}{2n-1} \right\}$.

7.11 Are the following sequences monotonic and bounded? Find the limits of the sequences if they exist:

(a) $\{a_n\} = \left\{ \frac{3}{n} - 6 \right\}$; (b) $\{b_n\} = \left\{ \frac{7-2n}{n^2} \right\}$; (c) $\{c_n\} = \left\{ \frac{2^n}{n!} \right\}$.

7.12 Find the limits of the following sequences if they exist ($n \in \mathbb{N}$):

(a) $\{a_n\} = \left\{ \frac{2n^n}{(n+1)^n} \right\}$; ($b$) $\{b_n\} = \left\{ \frac{an^4 + n^3}{3n^3 + 4n} \right\}$; $a \in \mathbb{R}$;

(c) $\{c_n\}$ with $c_n = c_{n-1}/2$. Check it for $c_1 = 1$ and $c_1 = 4$.

7.13 Determine the first six partial sums $s_1, s_2, \ldots, s_6$ for the following sequences:

(a) $\{a_n\} = \left\{ 3 + \frac{n}{2} \right\}$; ($b$) $\{b_n\} = \left\{ \frac{2n^3 - 2n^2 + 3n - 1}{2n^2 + 4n} \right\}$;

(c) $\{c_n\} = \left\{ (-1)^n \cdot \frac{n-1}{n+1} \right\}$.

7.14 Determine the partial sum s_{10} of the following arithmetic sequences:

(a) $\{a_n\} = 3, \frac{9}{2}, 6, \ldots$; (b) $\{b_n\} = 25,\ 13,\ -1,\ \ldots$;

(c) $\{c_n\} = -2,\ -14,\ -26,\ \ldots$.

7.15 Determine the partial sum s_{12} of the following geometric sequences:

(a) $\{a_n\} = \frac{1}{2}, -\frac{1}{3}; \frac{2}{9}, \ldots;$ (b) $\{b_n\} = \left\{\frac{2}{5} \cdot \left(\frac{1}{3}\right)^{n-1}\right\};$

(c) $\{c_n\} = \{(-3) \cdot 2^{n-1}\}$.

.

7.16 Determine the following sums:

(a) $2 + 6 + 10 + \ldots + 338\,;$ (b) $4 - \frac{4}{3} + \frac{4}{9} - \ldots + \ldots + \frac{4}{729}\,;$

(c) $32 + 16 + 8 + \ldots + \frac{1}{64}$.

Chapter 8

Functions

Functions are widely applied in many disciplines to express a relationship between an independent variable x and a dependent variable y, e.g. to express the production cost y in dependence on the quantity x produced, or the profit y in dependence on the quantity x produced, or the demand for a product y in dependence on its price x. The learning objectives of this chapter are to review

- the basic notion of a function (of a real variable) and some properties,
- the major types of a function, and
- composite and inverse functions.

8.1 Basic Notions and Properties

We start with the fundamental notion of a function.

Function of a real variable:
A **function** f **of a real variable** assigns to any real number $x \in D_f \subseteq \mathbb{R}$ a unique real number y. The set D_f is called the **domain** of function f and includes all numbers x for which function f is defined. The set

$$R_f = \{y \mid y = f(x),\ x \in D_f\}$$

is called the **range** of function f.

The real number $y \in R_f$ denotes the **function value** of x, i.e., the value of function f at the point x. The variable x is called the **independent variable** or the **argument**, and y is called the **dependent variable**. The domain and range of a function are illustrated in Fig. 8.1, where also the graph of function f is drawn in a coordinate system. Note that in mathematics, the horizontal axis is always the x-axis, and the vertical axis is always the y-axis.

In contemporary mathematics, one also writes

$$f : D_f \to \mathbb{R} \qquad \text{or} \qquad f : D_f \to R_f$$

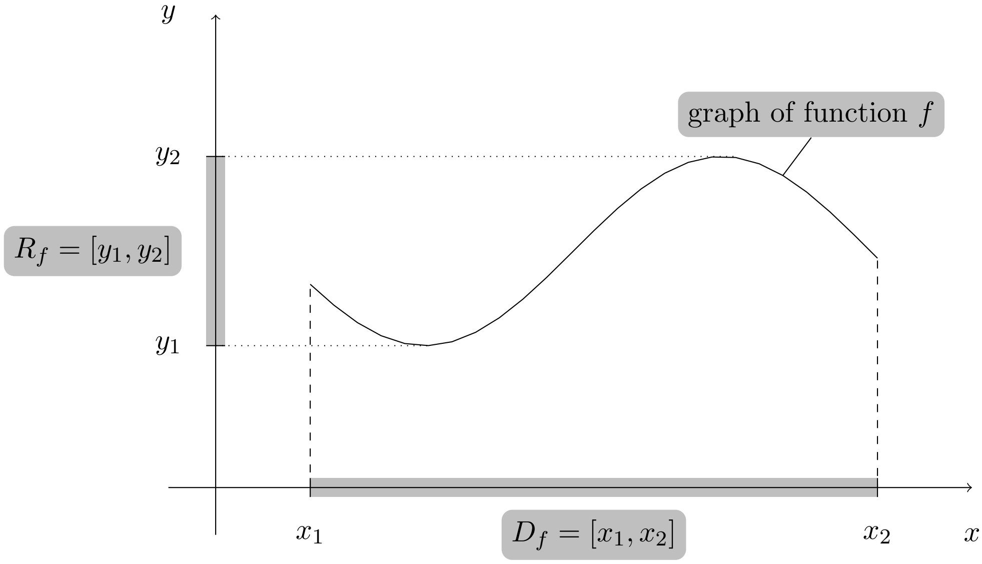


Figure 8.1: Domain, range and graph of a function

for a function of a real variable. We note that the second representation above is more precise because it gives the exact range of the function. However, often the first representation is used when it is not necessary to know the range explicitly. In this case, it simply expresses that the values of function f are real numbers. If the domain D_f is not stated explicitly, it is assumed that the **largest possible domain** is considered, i.e., the set of all real numbers x for which the function value can be calculated.

A function can be given in the following ways:

- analytically by an equation $y = f(x)$;
- by a table that gives for any $x \in D_f$ the resulting function value $f(x) \in R_f$;
- graphically or
- by a verbal description.

Example 8.1 *Consider the function* $f : D_f \to \mathbb{R}$ *with*

$$y = f(x) = \sqrt{x-1} + \frac{1}{x-2}.$$

We first determine the domain D_f *of function* f*. The term* $\sqrt{x-1}$ *is defined for non-negative numbers* $x-1$ *and therefore, for all real numbers* $x \geq 1$*. For the term* $1/(x-2)$*, we have to exclude the case when the denominator is equal to zero and thus, this term is defined for all real numbers* $x \neq 2$*. Both previous conditions have to be satisfied. This gives the largest possible domain*

$$D_f = \{x \in \mathbb{R} \mid x \geq 1 \text{ and } x \neq 2\},$$

or in interval notation

$$D_f = [1, 2) \cup (2, \infty).$$

So, the domain D_f *is the union of the above two intervals.*

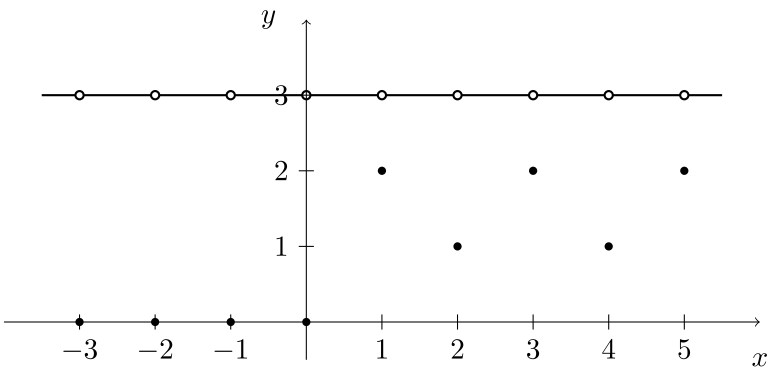


Figure 8.2: Graph of function f in Example 8.2

Example 8.2 *We consider the following function $f : \mathbb{R} \to \{0, 1, 2, 3\}$ described verbally:*

$$y = f(x) = \begin{cases} 0 & \text{if } x \text{ is a non-positive integer} \\ 1 & \text{if } x \text{ is an even positive integer} \\ 2 & \text{if } x \text{ is an odd positive integer} \\ 3 & \text{otherwise} \end{cases}$$

So, the function f is defined for all real numbers, and only four function values are possible. The graph of this function is given in Fig. 8.2.

The next four notions deal with **monotonicity** properties of a function.

Increasing function:
A function $f : D_f \to \mathbb{R}$ is said to be **increasing** (or equivalently, **non-decreasing**) on an interval $I \subseteq D_f$ if for arbitrary $x_1, x_2 \in I$ with $x_1 < x_2$, the inequality

$$f(x_1) \leq f(x_2)$$

holds.

Strictly increasing function:
A function $f : D_f \to \mathbb{R}$ is said to be **strictly increasing** on an interval $I \subseteq D_f$ if for arbitrary $x_1, x_2 \in I$ with $x_1 < x_2$, the inequality

$$f(x_1) < f(x_2)$$

holds.

Decreasing function:
A function $f : D_f \to \mathbb{R}$ is said to be **decreasing** (or equivalently, **strictly decreasing**) on an interval $I \subseteq D_f$ if for arbitrary $x_1, x_2 \in I$ with $x_1 < x_2$, the inequality

$$f(x_1) \geq f(x_2)$$

holds.

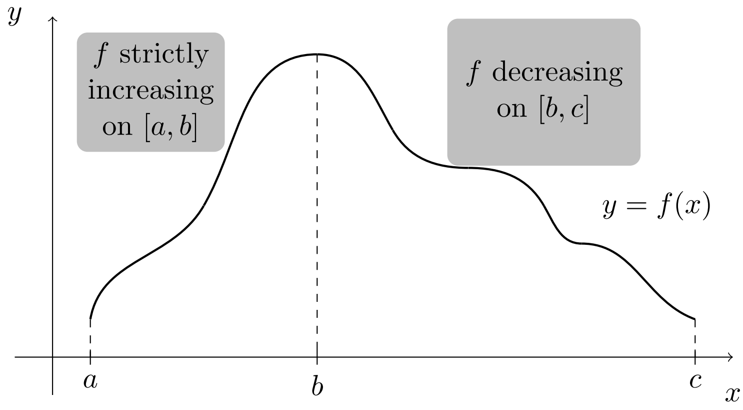


Figure 8.3: Monotonicity intervals of a function

Strictly decreasing function:
A function $f : D_f \to \mathbb{R}$ is said to be **strictly decreasing** on an interval $I \subseteq D_f$ if for arbitrary $x_1, x_2 \in I$ with $x_1 < x_2$, the inequality

$$f(x_1) > f(x_2)$$

holds.

If a function has one of the above four properties, we also say that it is **monotonic** (or **monotone**). An illustration is given in Fig. 8.3. In part (c), where function f is strictly increasing on the interval $[a, b]$, but decreasing on the interval $[b, c]$.

Figure 8.3: Monotonicity intervals of a function

We note that a function $f : D_f \to R_f$, which is strictly monotonic on the domain (i.e., either strictly increasing on D_f or strictly decreasing on D_f, is a so-called **one-to-one** (or equivalently, a **bijective**) function.

Example 8.3 *We investigate the function $f : D_f \to \mathbb{R}$ with*

$$f(x) = \frac{1}{x^3 + 8}$$

for monotonicity. This function is defined for all $x \neq -2$, i.e., $D_f = \mathbb{R} \setminus \{-2\}$. First, we consider the interval $I_1 = (-2, \infty)$. Let $x_1, x_2 \in I_1$ with $x_1 < x_2$. We get $0 < x_1^3 + 8 < x_2^3 + 8$ and thus,

$$f(x_1) = \frac{1}{x_1^3 + 8} > \frac{1}{x_2^3 + 8} = f(x_2)$$

(note that the inequality sign changes when considering the reciprocal terms of positive terms). Therefore, function f is strictly decreasing on the interval I_1.

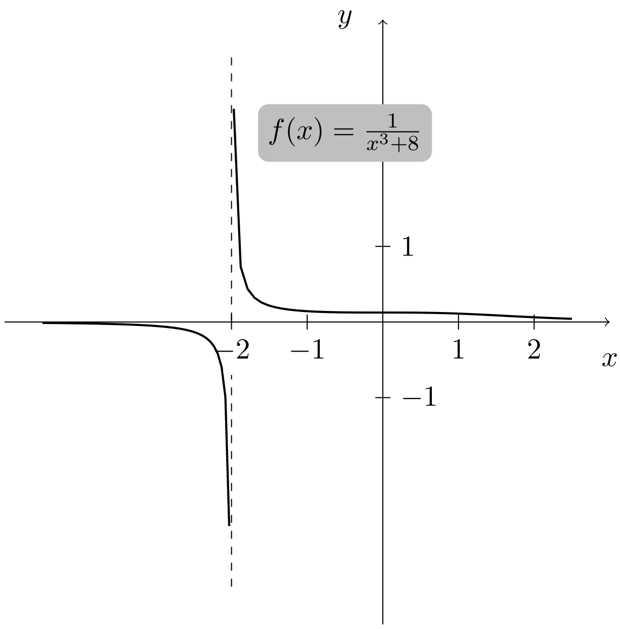


Figure 8.4: Graph of function f in Example 8.3

Consider now the interval $I_2 = (-\infty, -2)$ *and let* $x_1 < x_2 < -2$. *In this case, we first get* $x_1^3 + 8 < x_2^3 + 8 < 0$ *and then*

$$f(x_1) = \frac{1}{x_1^3 + 8} > \frac{1}{x_2^3 + 8} = f(x_2).$$

Therefore, function f *is also strictly decreasing on the interval* $(-\infty, -2)$. *The graph of function* f *is given in Fig. 8.4.*

Now we deal with the question whether the function value of a function can become arbitrarily small (large) or not.

Bounded function:
A function $f : D_f \to \mathbb{R}$ is said to be **bounded from below (from above)** if there exists a real constant C such that

$$f(x) \geq C \qquad \text{(resp. } f(x) \leq C\text{)}$$

for all $x \in D_f$.
A function $f : D_f \to \mathbb{R}$ is said to be **bounded from above** if there exists a constant C such that

$$f(x) \leq C)$$

for all $x \in D_f$.
Function f is said to be **bounded** if $f(x)$ is bounded from below and from above, i.e., we have

$$|f(x)| \leq C$$

for all $x \in D_f$.

According to the definition of the absolute values, a function f is bounded if there exists a constant C such that

$$-C \leq f(x) \leq C.$$

A bounded and an unbounded function are illustrated in Fig. 8.5.

Example 8.4 *We consider the function* $f : \mathbb{R} \to \mathbb{R}$ *with*

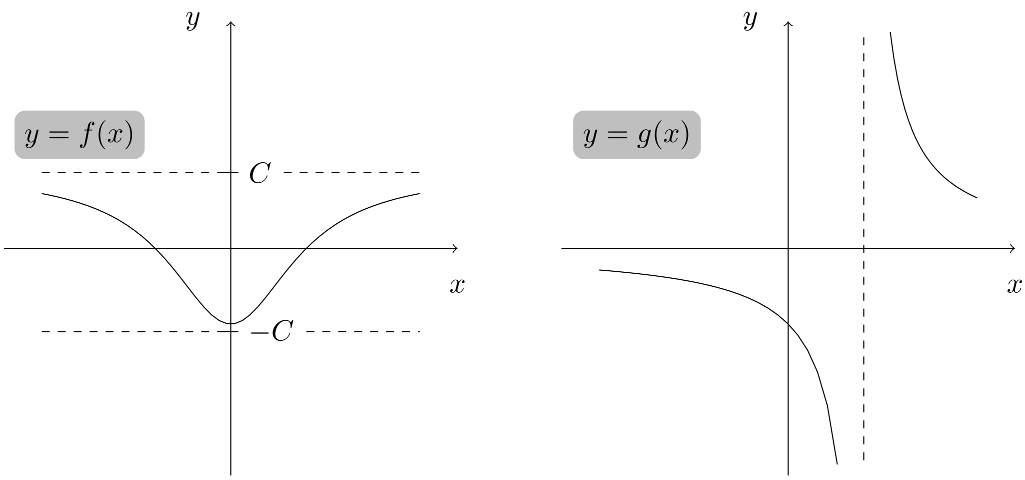
$$y = f(x) = e^x + \sin 2x + 3.$$

The value of the term e^x *is always greater than zero, the value of the term* $\sin 2x$ *is in the interval* $[-1, 1]$*. Thus, the value of function* f *is always greater than 0 + (-1) + 3 = 2. Therefore, function* f *is bounded from below since we can choose e.g. the constant* $C = 2$ *such that* $f(x) \geq C$ *for all* $x \in \mathbb{R}$*. However, function* f *is not bounded from above since the function value* $f(x)$ *can become arbitrarily large when* x *becomes large because the value of the term* e^x *can become arbitrarily large. Therefore, function* f *is also not bounded.*

Example 8.5 *We consider the function* $g : [3, \infty) \to \mathbb{R}$ *with*

$$y = g(x) = 5 - \sqrt{x-3}.$$

First, the values of the term $\sqrt{x-3}$ *are non-negative and therefore, the function values of* f *are smaller than or equal to 5. Consequently, for* $C = 5$ *we get* $g(x) \leq C$ *for all* $x \in \mathbb{R}$ *and thus, function* f *is bounded from above. However, since the value of the term* $\sqrt{x-3}$ *can become arbitrarily large, the function values of* g *can become arbitrarily small and function* g *is not bounded from below (and therefore also not bounded).*



(a) f bounded (b) g unbounded

Figure 8.5: Bounded and unbounded functions

Example 8.6 *We consider the function* $h : \mathbb{R} \to \mathbb{R}$ *with*

$$y = h(x) = 3\cos x + 4.$$

First, the possible values of the term $\cos x$ *are in the interval* $[-1, 1]$. *Thus, the values of the term* $3\cos x$ *are in the interval* $[-3, 3]$, *and hence the values of function* h *are in the interval* $[-3+4, 3+4] = [1, 7]$. *Consequently, function* h *is bounded, e.g. for* $C = 7$, *we have* $|h(x)| \le C$ *for all* $x \in \mathbb{R}$.

Next, we consider symmetry properties of a function.

Even/odd function:
A function $f : D_f \to \mathbb{R}$ is called **even** if

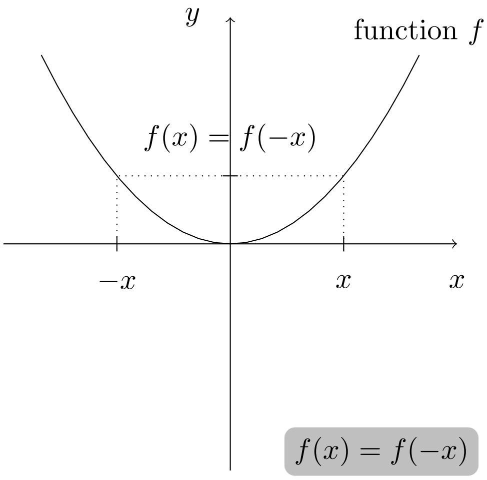
$$f(-x) = f(x) \text{ for all } x \in D_f.$$

A function f is called **odd** if

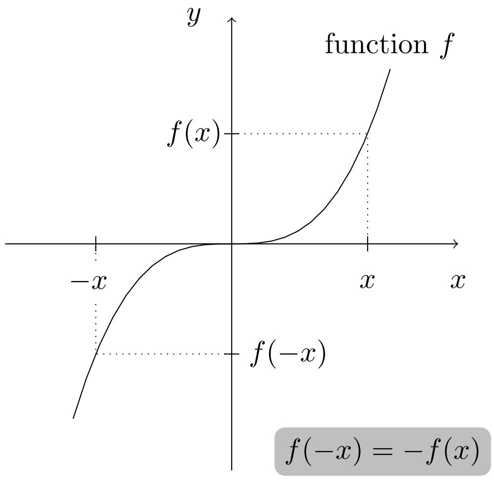
$$f(-x) = -f(x) \text{ for all } x \in D_f.$$

In both cases, the domain D_f has to be symmetric with respect to the origin of the coordinate system.

An even and an odd function are illustrated in Fig. 8.6.



(a) even function



(b) odd function

Figure 8.6: Even and odd functions

An even function is symmetric to the y-axis. An odd function is symmetric with respect to the origin of the coordinate system as a point. It is worth noting that a function f is not necessarily either even or odd. To be even or odd is a specific symmetry property of a function.

Example 8.7 *We consider the function $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ with*

$$f(x) = 3x^5 - x^3 + \frac{2}{x}.$$

We determine $f(-x)$ and obtain

$$\begin{aligned} f(-x) &= 3(-x)^5 - (-x)^3 + \frac{2}{-x} = -3x^5 + x^3 - \frac{2}{x} \\ &= -\left(3x^5 - x^3 + \frac{2}{x}\right) = -f(x). \end{aligned}$$

Thus, function f is an odd function.

Example 8.8 *Let the function $g : \mathbb{R} \to \mathbb{R}$ with*

$$g(x) = 4x^4 + x^2 + 2\,|x|$$

be given. Using the definition of the absolute value of a number, we obtain

$$g(-x) = 4(-x)^4 + (-x)^2 + 2|-x| = 4x^4 + x^2 + 2|x| = g(x).$$

Hence, function g is an even function.

Example 8.9 *Consider the function $h : \mathbb{R} \to \mathbb{R}$ with*

$$h(x) = x^2 + 2x + 5.$$

This function is neither even nor odd because we get e.g.

$$h(1) = 8 \qquad \textit{and} \qquad h(-1) = 4,$$

i.e., for the particular value $x = 1$, we get $h(1) \neq h(-1)$ and $h(1) \neq -h(1)$.

Periodic function:
A function $f : D_f \to \mathbb{R}$ is called **periodic** if there exists a real number T such that for all $x \in D_f$ with $x + T \in D_f$, the equality

$$f(x + T) = f(x)$$

holds. The smallest real number T with the above property is called the **period** of function f.

Two periodic functions are illustrated in Fig. 8.7.

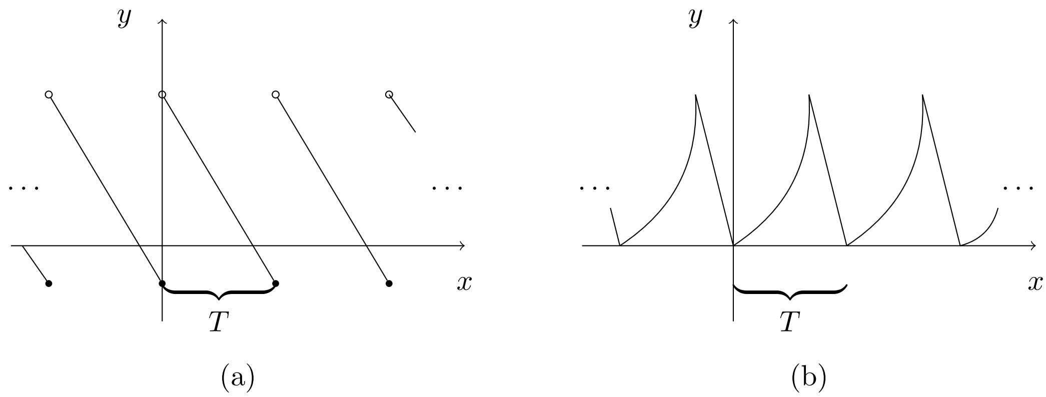


Figure 8.7: Periodic functions

Zero of a function:
A point x with

$$f(x) = 0$$

is called a **zero** (or **root**) of function f.

We will give the zeroes of some functions in the subsequent sections. However, in general, the problem of determining all zeroes of an arbitrary function is a difficult problem and often one has to apply numerical procedures for an approximate calculation.

Convex/concave function:
A function $f : D_f \to \mathbb{R}$ is called **convex** on an interval $I \subseteq D_f$, if for any choice of x_1 and x_2 in I and $0 \le \lambda \le 1$, inequality

$$f(\lambda x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda)f(x_2) \tag{8.1}$$

holds.

A function $f : D_f \to \mathbb{R}$ is called **concave** on an interval $I \subseteq D_f$, if for any choice of x_1 and x_2 in I and $0 \le \lambda \le 1$, inequality

$$f(\lambda x_1 + (1-\lambda)x_2) \ge \lambda f(x_1) + (1-\lambda)f(x_2)\Big) \tag{8.2}$$

holds.

If for $0 < \lambda < 1$ and $x_1 \ne x_2$, the sign $<$ holds in inequality (8.1) (and the sign $>$ holds in inequality (8.2), respectively), function f is called **strictly convex** (**strictly concave**).

The definition of a convex function is illustrated in Fig. 8.8. A function f is convex on an interval I if for any choice of two points x_1 and x_2 from the interval and for any intermediate point x from the interval $[x_1, x_2]$, which can be written as $x = \lambda x_1 + (1-\lambda)x_2$, the function value of this point (i.e., the value $f(\lambda x_1 + (1-\lambda)x_2)$) is not greater than the function value of the straight line through the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$ at the intermediate point x. The latter function value can be written as $\lambda f(x_1) + (1-\lambda)f(x_2)$.

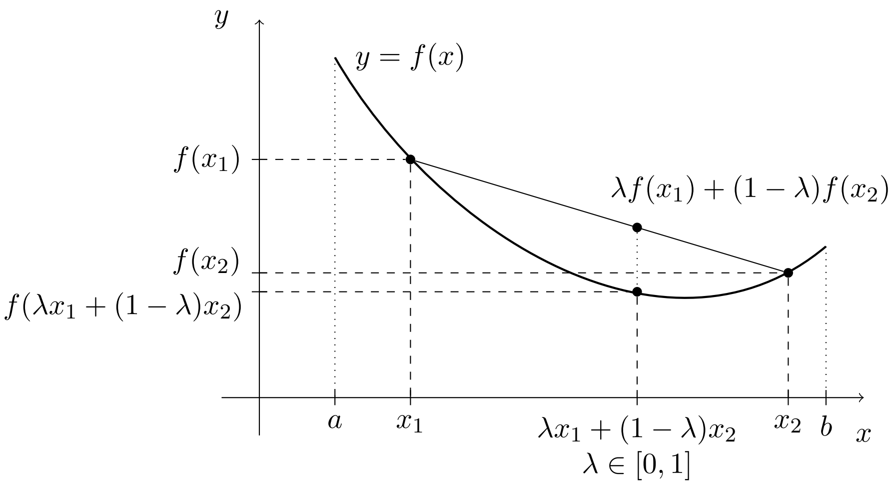


Figure 8.8: Definition of a convex function

Checking whether a function is convex or concave by applying the above definition can be rather difficult. By means of differential calculus, we give criteria in Chapter 9 which are easier to use.

In the next sections, we consider classes of functions for which their properties are well known.

8.2 Linear Functions

Linear function:
A function $f : \mathbb{R} \to \mathbb{R}$ given by

$$y = f(x) = a_1 x + a_0 \qquad (a_1 \neq 0)$$

is called a **linear function**.

In this case, the graph is a (straight) **line** which is uniquely defined by any two distinct points P_1 and P_2 belonging to this line. Assume that P_1 has the coordinates (x_1, y_1) and P_2 has the coordinates (x_2, y_2), then the above parameter a_1 is given by

$$a_1 = \frac{y_2 - y_1}{x_2 - x_1}$$

and a_1 is denoted as the **slope** of the line. The parameter a_0 gives the y-coordinate of the intersection of the line with the y-axis (see also Section 6.1). These considerations are illustrated in Fig. 8.9 (a) for $a_0 > 0$ and $a_1 > 0$ and in Fig. 8.9 (b) for $a_0 > 0$ and $a_1 < 0$. Different linear functions with the same parameter a_0 are illustrated in Fig. 8.10 (a), and different functions with the same parameter a_1 are illustrated in Fig. 8.10 (b). In case (a), the graphs are lines going through the same point $(0, a_0)$ while in case (b), the graphs are parallel lines.

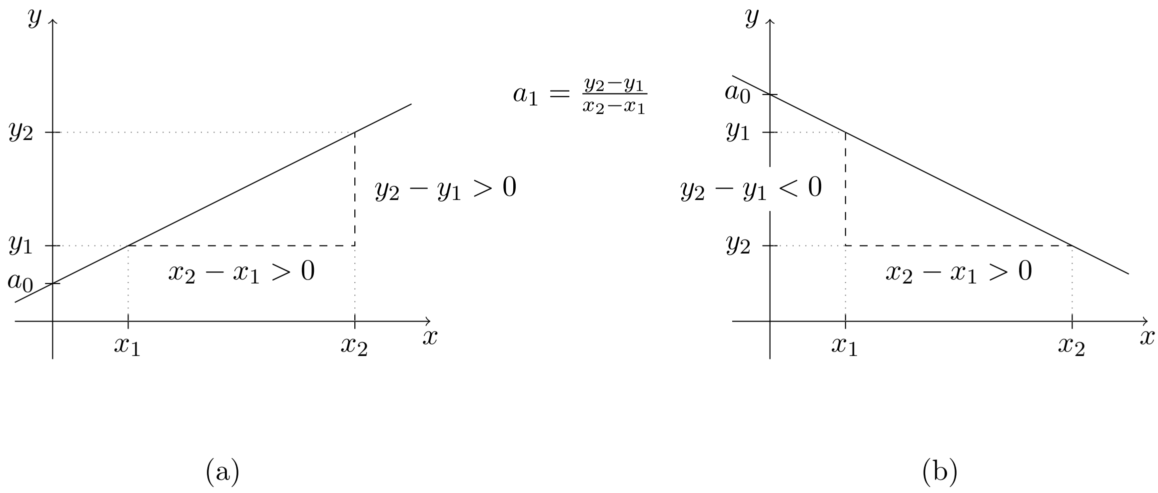


Figure 8.9: Linear functions with (a) $a_0 > 0, a_1 > 0$ and (b) $a_0 > 0, a_1 < 0$

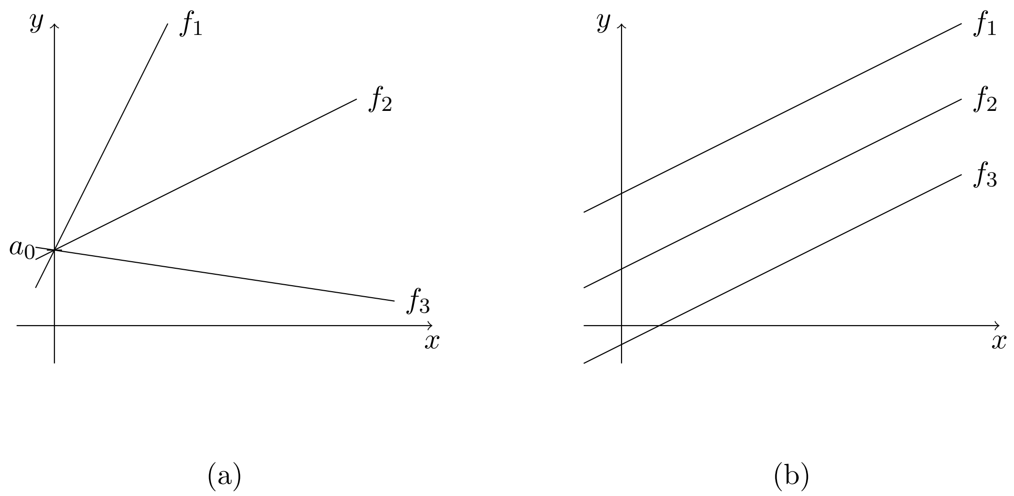


Figure 8.10: Linear functions with (a) identical a_0 and (b) identical a_1

The **point-slope formula of a line** passing through the point P_1 with the coordinates (x_1, y_1) and having the slope a_1 is given by

$$\boxed{y - y_1 = a_1 \cdot (x - x_1).}$$

A linear function $f(x) = a_1 x + a_0$ with $a_1 \neq 0$ has exactly one zero x_0 given by

$$\boxed{x_0 = -\frac{a_0}{a_1}\,.}$$

Example 8.10 *It is known that a line goes through the points* $(x_1, y_1) = (2, 3)$ *and* $(x_2, y_2) = (6, 11)$. *We find the equation of the corresponding linear function represented by this line. First, we determine the slope of this line and obtain*

$$a_1 = \frac{y_2 - y_1}{x_2 - x_1} = \frac{11 - 3}{6 - 2} = \frac{8}{4} = 2.$$

The equation of this line is given by

$$\begin{aligned} y - y_1 &= a_1 \cdot (x - x_1) \\ y - 3 &= 2 \cdot (x - 2) \\ y - 3 &= 2x - 4 \\ y &= 2x - 1, \end{aligned}$$

i.e., the slope of this linear function is $a_1 = 2$ *and the line intersects the* y*-axis at the point* $y = a_0 = -1$*, and the zero is given by*

$$x_0 = -\frac{a_0}{a_1} = -\frac{-1}{2} = \frac{1}{2}.$$

8.3 Quadratic Functions

Quadratic function:
A function $f : \mathbb{R} \to \mathbb{R}$ given by

$$y = f(x) = a_2x^2 + a_1x + a_0, \qquad a_2 \neq 0$$

is called a **quadratic equation**.

The graph of a quadratic function is called a **parabola**. If $a_2 > 0$, then the parabola is open from above (**upward parabola**, see Fig. 8.11 (a)) while, if $a_2 < 0$, then the parabola is open

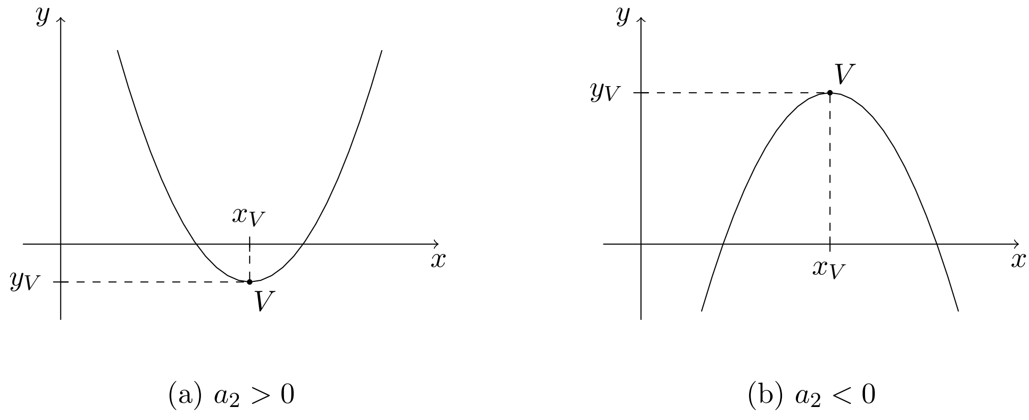


Figure 8.11: A quadratic function and the apex

from below (**downward parabola**, see Fig. 8.11 (b)). The point V in Fig. 8.11 is called the **apex** (or **vertex**), and its coordinates can be determined by rewriting the quadratic function in the following form:

$$y - y_V = a_2 \cdot (x - x_V)^2, \tag{8.3}$$

where (x_V, y_V) are the coordinates of the point V. The procedure of obtaining the vertex is illustrated in Example 8.11. In general, we get

$$\boxed{x_V = -\frac{a_1}{2a_2} \qquad \text{and} \qquad y_V = -\frac{a_1^2}{4a_2} + a_0.}$$

In the case of $a_1^2 \geq 4a_2a_0$, a quadratic function has the two real zeroes

$$x_{1,2} = \frac{-a_1 \pm \sqrt{a_1^2 - 4a_2a_0}}{2a_2}. \tag{8.4}$$

In the case of $a_1^2 = 4a_2a_0$, we get a zero $x_1 = x_2$ occurring twice. In the case of $a_1^2 < 4a_2a_0$, there does not exist a real zero (since there does not exist a real zero which is the square root of a negative number). If a quadratic function is given in **normal form** $y = x^2 + px + q$, formula (8.4) for finding the zeroes simplifies to

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

For the two zeroes x_1 and x_2 of the quadratic function $y = f(x) = x^2 + px + q$, we have

$$x_1 + x_2 = -p \qquad \text{and} \qquad x_1x_2 = q.$$

If one knows the two zeroes x_1 and x_2, one can easily determine the x-coordinate of the apex as follows:

$$x_V = \frac{x_1 + x_2}{2}.$$

Example 8.11 *Let the quadratic function $f : \mathbb{R} \to \mathbb{R}$ with*

$$y = f(x) = 2x^2 - 12x + 16.$$

be given. Since $a_2 = 2 > 0$, the parabola is open from above. For determining the apex, we rewrite function f as follows:

$$\begin{aligned} y &= 2(x^2 - 6x) + 16 \\ y &= 2(x^2 - 6x + \mathbf{9}) + 16 - \mathbf{18} \\ y &= 2(x-3)^2 - 2 \\ y + 2 &= 2(x-3)^2. \end{aligned}$$

From the latter representation we find the apex

$$V = (x_V, y_V) = (3, -2).$$

In the above transformation, we have written the right-hand side as a complete square of the form $a_2(x - x_V)^2$ by adding 2 times the number 9 in parentheses and then we subtracted outside 18 so that the value of the right-hand side does not change. To determine the zeroes, we divide the given equation by 2 and solve

$$x^2 - 6x + 8 = 0$$

which gives

$$x_1 = 3 + \sqrt{9-8} = 4 \qquad \textit{and} \qquad x_2 = 3 - \sqrt{9-8} = 2.$$

8.4 Polynomials

Linear and quadratic functions are a special case of a so-called **polynomial**. Polynomials are often used for approximating more complicated functions. They can be easily differentiated and integrated, and their properties are well investigated. Polynomials are defined as follows.

Polynomial of degree n:
The function $P_n : \mathbb{R} \to \mathbb{R}$ with

$$y = P_n(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0 \tag{8.5}$$

with $a_n \neq 0$ is called a **polynomial function** (or **polynomial) of degree n**. The numbers $a_0, a_1, \ldots, a_n$ are called the **coefficients** of the polynomial.

Example 8.12 *The function $f : \mathbb{R} \to \mathbb{R}$ with*

$$f(x) = 2x^5 + x^3 - 4x + 2$$

is a polynomial of degree 5. The function $g : \mathbb{R} \to \mathbb{R}$ with

$$g(x) = -x^3 + 1$$

is a polynomial of degree 3.

Zero of multiplicity k**:**
Let $f = P_n : \mathbb{R} \to \mathbb{R}$ be a polynomial of degree n with

$$P_n(x) = (x - x_0)^k \cdot S_{n-k}(x) \qquad \text{and} \qquad S_{n-k}(x_0) \neq 0.$$

Then x_0 is called a **zero** (or **root) of multiplicity k** of the polynomial P_n.

If $x = x_0$ is a zero of multiplicity k of the polynomial, then the polynomial contains exactly k times the factor $x - x_0$. If a polynomial $P_n(x)$ according to (8.5) has the zeroes $x_1, x_2, \ldots, x_n$, it can be written as a product in the form

$$\boxed{P_n(x) = a_n(x - x_1)(x - x_2)\ldots(x - x_n).}$$

However, we have to give two remarks. First, not all n zeroes of a polynomial need to be real numbers. Second, finding the real zeroes of a polynomial is in general a difficult (but important) problem. Often one has to apply numerical procedures for finding the zeroes approximately such as Regula falsi (see Section 3.6) or Newton's method (see Section 9.8).

Example 8.13 *Let a polynomial* $P_5 : \mathbb{R} \to \mathbb{R}$ *with*

$$P_5(x) = x^2(x - 1)(x^2 - 2x - 3)$$

be given. From the first two factors we see that

$$x_1 = x_2 = 0, \quad x_3 = 1.$$

are zeroes (note that $x = 0$ is a zero of multiplicity two). The remaining two zeroes are obtained from $x^2 - 2x - 3 = 0$ which gives

$$x_4 = 1 + \sqrt{1+3} = 3 \qquad \textit{and} \qquad x_5 = 1 - \sqrt{1+3} = -1.$$

8.5 Rational Functions

Next, we consider rational functions T which can be written as a quotient of polynomials P and Q.

Rational function:
A function $T : D_T \to \mathbb{R}$ with

$$T(x) = P(x)/Q(x) = (P/Q)(x)$$

is called a **rational function.** The rational function T is called **proper** if $deg\ P < deg\ Q$ and **improper** if $deg\ P \geq deg\ Q$, where *deg* is an abbreviation for the degree.

So,

$$T_1(x) = \frac{x^3 + 4}{x^2 - 3x + 1}$$

is an improper rational function, while

$$T_2(x) = \frac{x^2 + 4x + 5}{x^4 - 2x^3 + 3}$$

is a proper rational function.

By means of **polynomial division**, any improper rational function $T = P/Q$ can be written as the sum of a polynomial and a proper rational function. For $\deg P \geq \deg Q$, the rational function $T(x)$ can be written as

$$\boxed{T(x) = \frac{P(x)}{Q(x)} = S(x) + \frac{R(x)}{Q(x)}}$$

with $D_T = \{x \in \mathbb{R} \mid Q(x) \neq 0\}$. To get the latter representation, we consider the terms in both polynomials P and Q in decreasing order of their powers and divide in each step the first term of P by the first term of Q. Then the result is multiplied by Q and the product is subtracted from P. This yields a new polynomial P_1 having a smaller degree than polynomial P_n. Now the first term (with largest exponent) of polynomial P_1 is divided by the first term of polynomial Q, the resulting term is multiplied by polynomial Q and subtracted from P_1 yielding a polynomial P_2 and so on. The procedure stops if some resulting polynomial P_i has a smaller degree than the polynomial Q. This procedure is illustrated by the following example.

Example 8.14 *Let the two polynomials P and Q with*

$$P(x) = x^4 + x^3 - 2x^2 + 3x - 1 \qquad \textit{and} \qquad Q(x) = x^2 + x + 4$$

be given. The function T with $T(x) = (P/Q)(x)$ is an improper rational function, and by polynomial division we obtain:

$$\begin{array}{rrrrrrrrr}
(x^4 & + & x^3 & - & 2x^2 & + & 3x & - & 1) \\
-(x^4 & + & x^3 & + & 4x^2) & & & & \\
\hline
 & & & & -6x^2 & + & 3x & - & 1 \\
 & & & - & (-6x^2 & - & 6x & - & 24) \\
\hline
 & & & & & & 9x & + & 23
\end{array}
\quad : (x^2 + x + 4) = x^2 - 6 + \frac{9x + 23}{x^2 + x + 4}$$

Recall that the terms in the polynomials have to be considered according to non-increasing exponents. Thus, the rational function T can be written as the sum of a polynomial of degree 2 and a proper rational function, where the polynomial in the numerator has the degree 1.

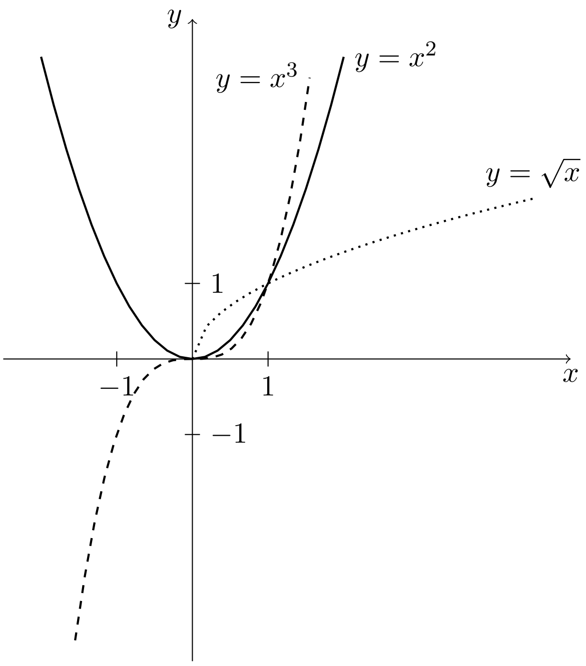


Figure 8.12: Graphs of power and root functions

8.6 Power and Root Functions

Power function:
The function $f : D_f \to \mathbb{R}$ with

$$f(x) = x^r, \qquad r \in \mathbb{R},$$

is called a **power function**. In dependence on the exponent r, the domain D_f and the range R_f are as follows:

1. $r \in \{1, 3, 5, \ldots\} \subseteq \mathbb{N} : D_f = (-\infty, \infty),\ R_f = (-\infty, \infty);$
2. $r \in \{2, 4, 6, \ldots\} \subseteq \mathbb{N} : D_f = (-\infty, \infty),\ R_f = [0, \infty);$
3. $r \in \mathbb{R} \setminus (\mathbb{N} \cup \{0\}) : D_f = (0, \infty),\ R_f = (0, \infty).$

In case (1), the function f is strictly increasing, unbounded and odd. In case (2), the function f is strictly decreasing on the interval $(-\infty, 0]$, strictly increasing on the interval $[0, \infty)$, bounded from below and even. If in case (3), we have $r > 0$, then the function is strictly increasing and bounded from below. If we have $r < 0$ in case (3), the function is strictly decreasing and also bounded from below. For all $r \in \mathbb{R}_{>0}$, the power function goes through the point $(x, y) = (1, 1)$. The graphs of some power and root functions are given in Fig. 8.12.

Case (3) includes **root functions** as a special case when r is a positive value with $r = 1/n$, i.e.,

$$f(x) = x^{1/n} = \sqrt[n]{x}.$$

In the case of a root function, the number zero belongs to the domain and also to the range,

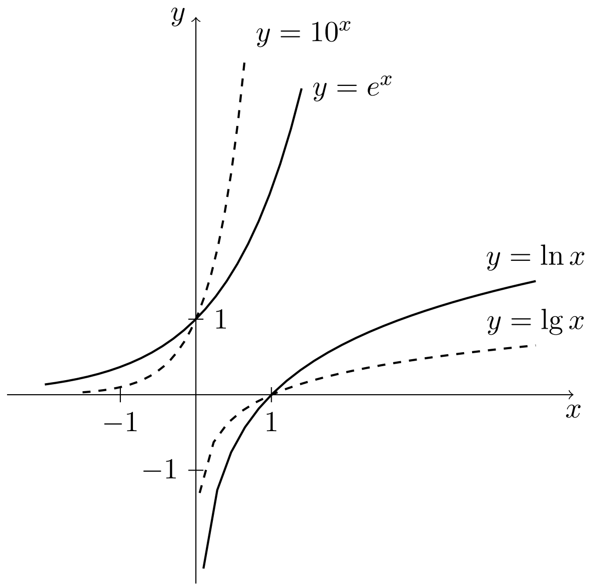


Figure 8.13: Graphs of exponential and logarithmic functions

i.e., $D_f = R_f = [0, \infty)$.

8.7 Exponential and Logarithmic Functions

Exponential function:
The function $f : \mathbb{R} \to (0, \infty)$ with

$$f(x) = a^x, \qquad a > 0,\ a \neq 1,$$

is called an **exponential function**.

All exponential functions go through the point $(0, 1)$. If $a > 1$, the exponential function is strictly increasing. If $0 < a < 1$, the corresponding exponential function is strictly decreasing.

Logarithmic function:
The function $f : (0, \infty) \to \mathbb{R}$ with

$$f(x) = \log_a x, \qquad a > 0,\ a \neq 1,$$

is called a **logarithmic function**.

The graphs of some exponential and logarithmic functions are given in Fig. 8.13.

If the base $a = e \approx 2.71828$ is chosen, we use the abbreviation ln which denotes the natural logarithm, i.e.,

$$\log_e x = \ln x.$$

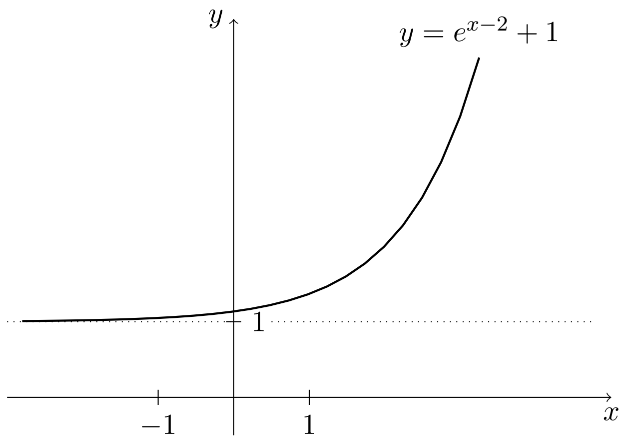


Figure 8.14: The graph of function h in Example 8.16

If the base $a = 10$ is chosen, often the abbreviation lg is used, i.e.,

$$\log_{10} x = \lg x.$$

All logarithmic functions go through the point $(1, 0)$. If $a > 1$, the logarithmic function is strictly increasing. If $0 < a < 1$, the corresponding logarithmic function is strictly decreasing.

Example 8.15 *Let the logarithmic function $g : D_g \to R_g$ with*

$$y = g(x) = \ln(2x + 5) - 3$$

be given. Since any logarithmic function is defined only for positive real numbers, we must have $2x + 5 > 0$ which gives the domain

$$D_g = \left\{ x \in \mathbb{R} \mid x > -\frac{5}{2} \right\}.$$

The range of function g corresponds to the range of an arbitrary logarithmic function (since the modification in the argument and the subtraction of 3 from the logarithmic value do not change the range). A logarithmic function has the only zero at the real number one. Thus the only zero of function h is obtained from $2x + 5 = 1$ which gives $x_0 = -2$.

Example 8.16 *Let us consider the exponential function $h : D_h \to R_h$ with*

$$y = h(x) = e^{x-2} + 1.$$

Since an exponential function is defined for all real numbers, function h has the domain $D_h = \mathbb{R}$ too. The range of any exponential function is the interval $(0, \infty)$. Therefore, the we get the range $D_h = (1, \infty)$ of function h. As a consequence, function h has no zeroes. The graph of function h is obtained from the graph of function $y = f(x) = e^x$ by shifting it two units to the right (since the argument is $x - 2$ instead of x) and then one unit above (since we have to add one to the function value). The graph of function h is given in Fig. 8.14.

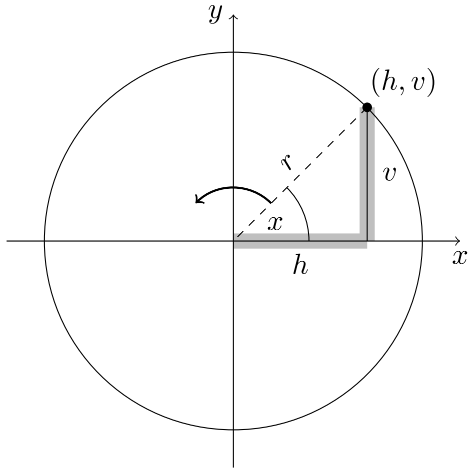


Figure 8.15: Definition of trigonometric functions

8.8 Trigonometric Functions

In this section, we repeat the four trigonometric functions and their properties. They can be defined on a circle with the angle x being the variable (see Fig. 8.15).

Sine and cosine functions:
The function $f : \mathbb{R} \to [-1, 1]$ with

$$f(x) = \sin x = \frac{v}{r}$$

is called the **sine function**. The function $f : \mathbb{R} \to [-1, 1]$ with

$$f(x) = \cos x = \frac{h}{r}$$

is called the **cosine function**.

The **zeroes of the sine function** are given by

$$x_k = k\pi \qquad \text{with} \qquad k \in \mathbb{Z}\,.$$

The **zeroes of the cosine function** are given by

$$x_k = \frac{\pi}{2} + k\pi \qquad \text{with} \qquad k \in \mathbb{Z}\,.$$

The graphs of the sine and cosine functions are given in Fig. 8.16.

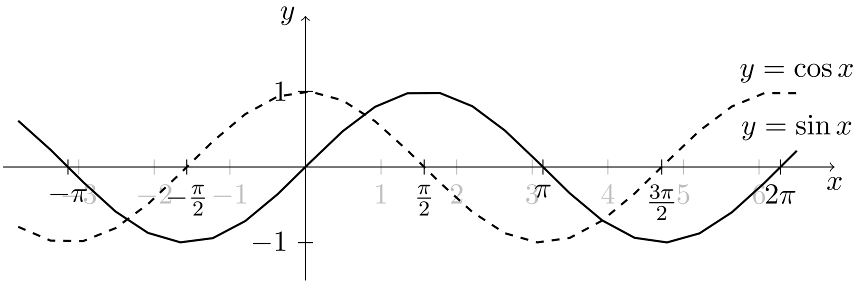


Figure 8.16: Graphs of the sine and the cosine functions

Tangent and cotangent functions:
The function $f : D_f \to \mathbb{R}$ with

$$f(x) = \tan x = \frac{\sin x}{\cos x} = \frac{v}{h} \quad \text{and} \quad D_f = \left\{x \in \mathbb{R} \mid x \neq \frac{\pi}{2} + k\pi,\ k \in \mathbb{Z}\right\}$$

is called the **tangent function**. The function $f : D_f \to \mathbb{R}$ with

$$f(x) = \cot x = \frac{\cos x}{\sin x} = \frac{h}{v} \quad \text{and} \quad D_f = \{x \in \mathbb{R} \mid x \neq k\pi,\ k \in \mathbb{Z}\}$$

is called the **cotangent function**.

The **zeroes of the tangent function** coincide with those of the sine function and thus are given by

$$\boxed{x_k = k\pi \qquad \text{with} \qquad k \in \mathbb{Z}\,.}$$

The **zeroes of the cotangent function** coincide with those of the cosine function and thus are given by

$$\boxed{x_k = \frac{\pi}{2} + k\pi \qquad \text{with} \qquad k \in \mathbb{Z}\,.}$$

The graphs of the tangent and cotangent functions are given in Fig. 8.17.

All trigonometric functions are periodic functions. In particular, the sine and cosine functions have a period of 2π, i.e., we have

$$\sin(x + 2k\pi) = \sin x, \quad \cos(x + 2k\pi) = \cos x, \qquad k \in \mathbb{Z}.$$

Moreover, the tangent and cotangent functions have the period π, i.e., we have

$$\tan(x + \pi) = \tan x, \quad \cot(x + k\pi) = \cot x, \qquad k \in \mathbb{Z}.$$

Example 8.17 *Consider the function g with*

$$y = g(x) = \sin\left(x + \frac{\pi}{2}\right).$$

The domain and range of function g do not change in comparison to $f(x) = \sin x$: $D_g = \mathbb{R}$ and $R_g = [-1, 1]$. To determine the zeroes, we use that function f has the zeroes $x_k = k\pi, \in \mathbb{Z}$. Therefore, for function g, we must have

$$x + \frac{\pi}{2} = k\pi, \qquad k \in \mathbb{Z}.$$

Hence, function g has the zeroes

$$x_k = k\pi - \frac{\pi}{2}, \qquad k \in \mathbb{Z}.$$

Function g has the same period of 2π as function f. The graph of function g is obtained from the graph of f by shifting it $\pi/2$ units left along the x axis.

Example 8.18 *Given is the trigonometric function $h : \mathbb{R} \to R_h$ with*

$$y = h(x) = 2\cos ax, \qquad a \in \mathbb{R}.$$

First, if $a = 0$, function h is a constant function with $h(x) = 2$. Consider now the case $a \neq 0$. Function h is defined for all real numbers x: $D_h = \mathbb{R}$. The range of the term $\cos ax$ for $a \neq 0$ is the interval $[-1, 1]$. Therefore, the range of function h is given by $R_h = [-2, 2]$. To determine the zeroes of function h, we remember that the cosine function $y = f(x) = \cos x$ has the zeroes

$$x_k = \frac{\pi}{2} + k\pi = \frac{\pi}{2} \cdot (2k + 1), \qquad k \in \mathbb{Z}.$$

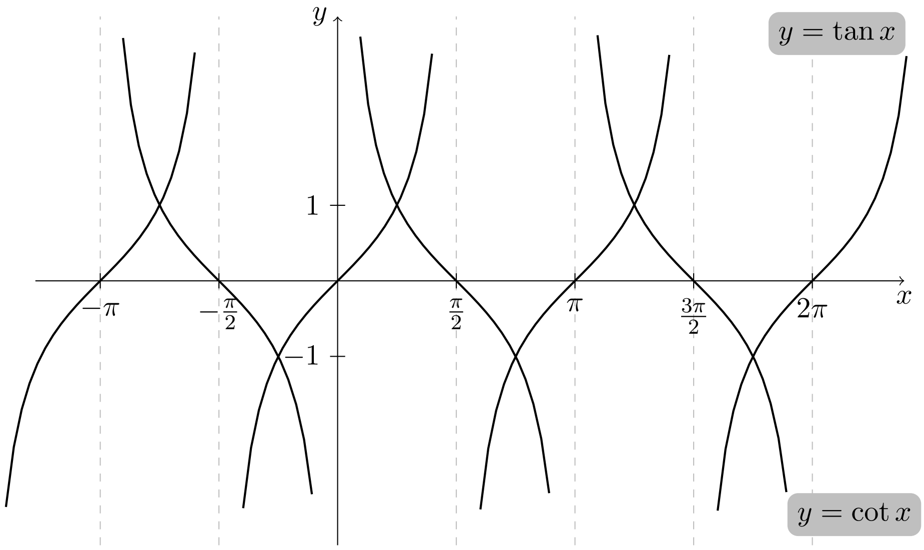


Figure 8.17: Graphs of the tangent and the cotangent functions

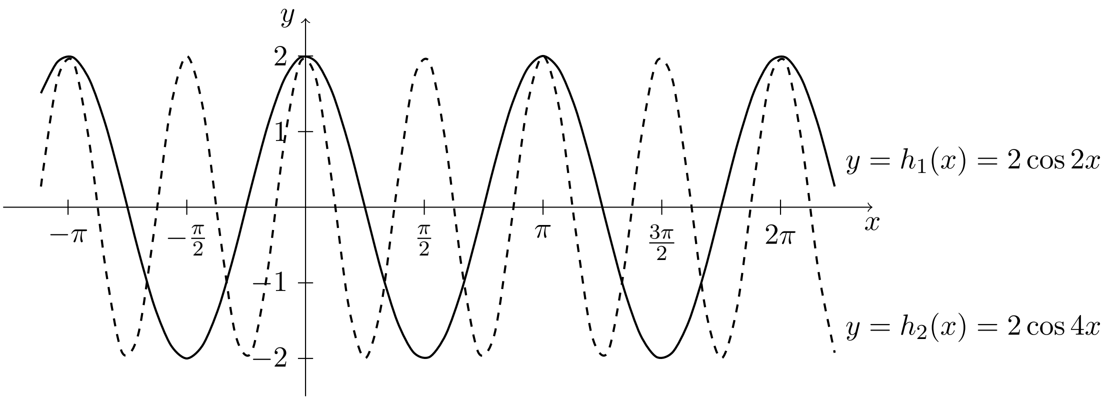


Figure 8.18: Graph of function h in Example 8.18 for $a = 2$ ($h_1(x)$) and $a = 4$ ($h_2(x)$)

For function h, we must therefore have

$$ax = \frac{\pi}{2} \cdot (2k+1)$$

which gives the zeroes

$$x_k = \frac{\pi}{2a} \cdot (2k+1), \qquad k \in \mathbb{Z}.$$

Moreover, function h has the period T with

$$T = \frac{2\pi}{a}$$

(since the term ax must be a multiple of 2π). The graph of function h is given in Fig. 8.18 for $a = 2$ and $a = 4$.

Example 8.19 *We consider the trigonometric function* $l : D_l \to R_l$ *with*

$$l(x) = 1 + 3\cot(x - 1).$$

Since the cotangent function is defined for all $x \neq k\pi, k \in \mathbb{Z}$, *we get the domain*

$$D_l = \{x \in \mathbb{R} \mid x \neq k\pi + 1,\ k \in \mathbb{Z}\}.$$

The range R_l *of function* l *is the whole set* $\mathbb{R}$. *Since the zeroes of function* $f(x) = \cot x$ *are given by*

$$x_k^* = \frac{\pi}{2} + k\pi = \frac{\pi}{2}(2k+1), \qquad k \in \mathbb{Z}$$

we get the zeroes of function l *as follows:*

$$x_k = \frac{\pi}{2} + k\pi + 1 = \frac{\pi}{2}(2k+1) + 1, \qquad k \in \mathbb{Z}.$$

The period of function l *is also* π *as for the function* $f(x) = \cot x$.

8.9 Composite and Inverse Functions

In this section, we review the definition of a composite and an inverse function and how to work with them.

Composite function:
Let the two functions $f : D_f \to R_f$ and $g : D_g \to R_g$ be given with $R_f \subseteq D_g$. Then the **composite function** $g \circ f$ of the functions f and g is defined as follows:

$$g \circ f : D_f \to \mathbb{R}$$

with

$$y = g(f(x)), \qquad x \in D_f.$$

The function f is called the **inside function**, and the function g is called the **outside function**. This means that first function f is applied to some number $x \in D_f$, this gives the function value $f(x)$ which must belong to the domain D_g of function g (otherwise the composite function would not be defined). Finally, the function value is $y = g(f(x)) \in R_g \subseteq \mathbb{R}$. The determination of the composite function is illustrated in Fig. 8.19.

If both compositions $g \circ f$ and $f \circ g$ are defined, this operation is not necessarily commutative, i.e., in general, we have

$$\boxed{(g \circ f)(x) \neq (f \circ g)(x).}$$

Thus, in general, the inside and outside functions cannot be interchanged. We illustrate the use of composite functions by the following examples.

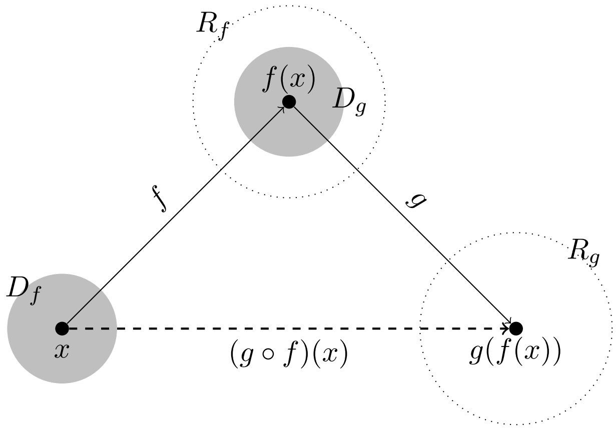


Figure 8.19: The composite function $g \circ f$

Example 8.20 *Let the functions $f : \mathbb{R} \to R_f$ and $g : \mathbb{R} \to R_g$ with*

$$f(x) = 2x + 5 \qquad \textit{and} \qquad g(x) = 3x^2 - x + 4$$

be given. We determine both composite functions $f \circ g$ and $g \circ f$ and obtain

$$\begin{aligned}(f \circ g)(x) = f(g(x)) &= f(3x^2 - x + 4) \\ &= 2(3x^2 - x + 4) + 5 \\ &= 6x^2 - 2x + 13\end{aligned}$$

and

$$\begin{aligned}(g \circ f)(x) = g(f(x)) &= g(2x + 5) \\ &= 3(2x + 5)^2 - (2x + 5) + 4 \\ &= 3(4x^2 + 20x + 25) - (2x + 5) + 4 \\ &= 12x^2 + 58x + 74.\end{aligned}$$

Both compositions are defined since the range of the inside function is either the set $\mathbb{R}$ (function f) or a subset (function g) while in both cases the outside function is defined for all real numbers: $D_f = D_g = \mathbb{R}$.

Example 8.21 *Given are the functions $f : \mathbb{R} \to R_f$ and $g : \mathbb{R} \to R_g$ with*

$$f(x) = x^2$$

and

$$g(x) = e^x + 1.$$

For the composite functions we obtain

$$(f \circ g)(x) = f(g(x)) = f(e^x + 1) = \left(e^x + 1\right)^2.$$

and

$$(g \circ f)(x) = g(f(x)) = g(x^2) = e^{x^2} + 1.$$

The composition $f \circ g$ is defined since the range $R_g = (1, \infty)$ is a subset of the domain $D_f = \mathbb{R}$, and the composition $g \circ f$ is defined since the range $R_f = \mathbb{R}_{\geq 0}$ is a subset of the domain $D_g = \mathbb{R}$.

Example 8.22 *Given are the functions* $f : D_f \to R_f$ *and* $g : D_g \to R_g$ *with*

$$f(x) = \sqrt{x+2}, \qquad D_f = [-2, \infty)$$

and

$$g(x) = 2\cos x + 1, \qquad D_g = \mathbb{R}.$$

For the composite functions we obtain

$$(f \circ g)(x) = f(g(x)) = f(2\cos x + 1) = \sqrt{(2\cos x + 1) + 2} = \sqrt{2\cos x + 3}.$$

and

$$(g \circ f)(x) = g(f(x)) = g(\sqrt{x+2}) = 2\cos\sqrt{x+2} + 1.$$

The composition $f \circ g$ *is defined since the range* $R_g = [-1, 3]$ *is a subset of the domain* $D_f = [-2, \infty)$*, and the composition* $g \circ f$ *is defined since the range* $R_f = [0, \infty)$ *is a subset of the domain* $D_g = \mathbb{R}$*.*

Example 8.23 *Given are the functions* $f : D_f \to R_f$ *and* $g : D_g \to R_g$ *with*

$$f(x) = 3x - 2, \qquad D_f = \mathbb{R}_{>0}$$

and

$$g(x) = \frac{1}{x} + 2, \qquad D_g = \mathbb{R}_{>0}.$$

For the composite function $f \circ g$*, we obtain*

$$(f \circ g)(x) = f(g(x)) = f\left(\frac{1}{x} + 2\right) = 3 \cdot \left(\frac{1}{x} + 2\right) - 2 = \frac{3}{x} + 4.$$

The composition $f \circ g$ *is defined since the range* $R_g = \mathbb{R}_{>2}$ *is a subset of the domain* D_f*. However, the composition* $g \circ f$ *is not defined since we get the range* $R_f = \mathbb{R}_{>-2}$ *and thus,* $R_f \not\subseteq D_g$*.*

Inverse function:
If a function $f : D_f \to R_f, D_f \subseteq \mathbb{R}$, is strictly monotonic (i.e., either strictly increasing or strictly decreasing) on the domain D_f, then there exists a function f^{-1} which assigns to each real number $y \in R_f$ a unique real value $x \in D_f$ with

$$x = f^{-1}(y),$$

which is called the **inverse function** f^{-1} of function f.

Here it should be mentioned that, if $R_f \subseteq \mathbb{R}$ but $R_f \neq \mathbb{R}$ (in this case we say that R_f is a proper subset of $\mathbb{R}$), the inverse function of $f : D_f \to \mathbb{R}$ would not exist since there exist values $y \in \mathbb{R}$ for which $x = f^{-1}(y)$ does not exist. Nevertheless, such a formulation is often used and, in order to determine f^{-1}, one must find the range R_f of function f (or, what is the same, the domain $D_{f^{-1}}$ of the inverse function f^{-1} of f).

We write $f^{-1} : R_f \to \mathbb{R}$ or more precisely, $f^{-1} : R_f \to D_f$. Since x usually denotes the independent variable, we can exchange variables x and y and write

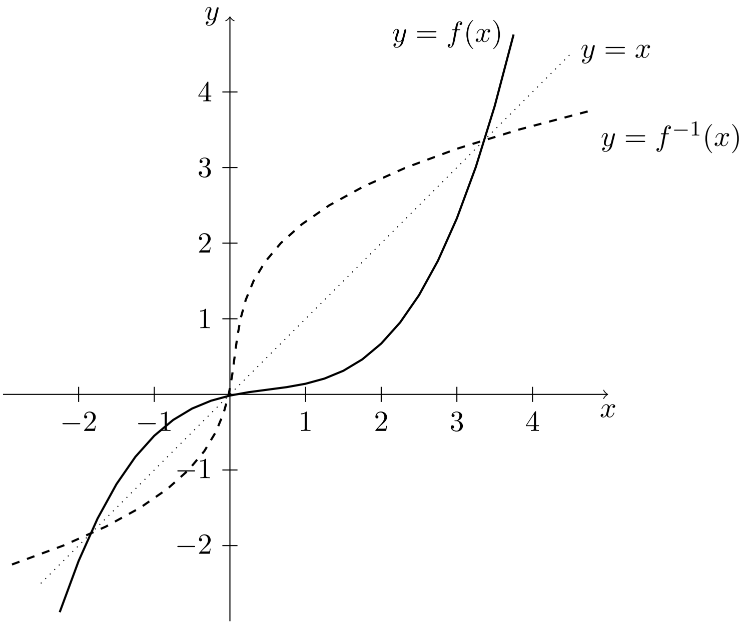


Figure 8.20: Function f and the inverse function f^{-1}

$$y = f^{-1}(x)$$

for the inverse function of function f with $y = f(x)$. The graph of the inverse function f^{-1} is given by the mirror image of the graph of function f with respect to the line $y = x$. We emphasize that in general, we have

$$f^{-1}(x) \neq \frac{1}{f(x)}.$$

If there exists the inverse function f^{-1} of a function $f : \mathbb{R} \to \mathbb{R}$, we have

$$f^{-1} \circ f = f \circ f^{-1},$$

i.e.,

$$f^{-1}(f(x)) = f(f^{-1})(x) = x \qquad \text{for all } x \in \mathbb{R}.$$

A function f and its inverse function f^{-1} are illustrated in Fig. 8.20.

We illustrate the determination of an inverse function by the following examples.

Example 8.24 *We consider the linear function $f : \mathbb{R} \to \mathbb{R}$ given by*

$$y = f(x) = 3x + 7.$$

Solving for x, we get

$$\begin{aligned} y - 7 &= 3x \\ x &= \frac{1}{3}\,(y - 7). \end{aligned}$$

Interchanging the variables x and y, we obtain the inverse function f^{-1} of function f with

$$y = f^{-1}(x) = \frac{1}{3}\,(x-7).$$

Hence, the inverse function of a linear function is a linear function as well, and we have $D_{f^{-1}} = R_{f^{-1}} = \mathbb{R}$.

Example 8.25 *Let the power function $f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ with*

$$y = x^{2.5}$$

be given. Using power rules when solving for x, we obtain

$$\begin{aligned} y^{1/2.5} &= \left[x^{2.5}\right]^{1/2.5} \\ x &= y^{1/2.5} = y^{0.4}. \end{aligned}$$

Interchanging the variables x and y, we get the inverse function f^{-1} of function f with

$$y = f^{-1}(x) = x^{0.4}\,. \tag{8.6}$$

For the inverse function, we have

$$D_{f^{-1}} = R_{f^{-1}} = \mathbb{R}_{>0}.$$

Note than we can also use the logarithmic function for determining the inverse function. By taking the natural logarithm, we obtain

$$\ln x = \frac{\ln y}{2.5} = 0.4 \ln y = \ln y^{0.4}$$

from which we also obtain (8.6).

Example 8.26 *Let the function $f : \mathbb{R} \to (5, \infty)$ with*

$$y = f(x) = \frac{1}{3}\,e^{2x} + 5$$

be given. Solving for x, we obtain

$$\begin{aligned} y - 5 &= \frac{1}{3}\,e^{2x} \\ 3(y-5) &= e^{2x} \\ \ln\Big[3(y-5)\Big] &= 2x \\ x &= \frac{1}{2}\,\ln\Big[3(y-5)\Big]. \end{aligned}$$

Note that $y > 5$ so that the argument of the logarithmic function is always positive. Interchanging the variables x and y, we get the inverse function $y = f^{-1}(x)$ of function f with

$$y = \frac{1}{2}\ln\Big[3(x-5)\Big],$$

where

$$D_{f^{-1}} = (5, \infty) \qquad \text{and} \qquad R_{f^{-1}} = \mathbb{R}\,.$$

Example 8.27 *Let the function* $f : \mathbb{R}_{\geq 0} \to [3, \infty)$ *with*

$$y = f(x) = 2x^3 + 3$$

be given. To determine the inverse function f^{-1}*, we solve the above equation for* x *and obtain*

$$\begin{aligned} 2x^3 &= y - 3 \\ x^3 &= \frac{y-3}{2} \\ x &= \sqrt[3]{\frac{y-3}{2}}. \end{aligned}$$

Note that $y \geq 3$ *so that the term under the root is always non-negative. Interchanging now both variables, we obtain the inverse function* f^{-1} *of function* f *with*

$$y = f^{-1}(x) = \sqrt[3]{\frac{x-3}{2}}.$$

The domain of the inverse function f^{-1} *is the range of function* f*, and the range of function* f^{-1} *is the domain of function* f*, i.e.,*

$$D_{f^{-1}} = [3, \infty) \qquad \text{and} \qquad R_{f^{-1}} = \mathbb{R}_{\geq 0}.$$

EXERCISES

8.1 Determine the domain D_f and the range R_f of the following functions:

(a) $f(x) = 2x + 5\,;$ (b) $f(x) - x^2 + 2\,;$ (c) $f(x) = \frac{1}{2}x^2 + \frac{1}{2}x - 3\,;$

(d) $z = \sqrt{x + 3}\,;$ (e) $f(x) = 2\ln x + 1\,;$ (f) $f(x) = \ln(x^2 + 1)\,;$

(g) $f(x) = \sqrt{x^2 - 9}\,;$ (h) $f(x) = \ln x^5\,;$ (i) $f(x) = \sqrt{|x| - x}\,.$

8.2 Graph the following functions:

(a) $f(x) = 2x + 5\,;$ (b) $f(x) = -x^2 + 2\,;$ (c) $f(x) = e^{-x} + 2\,;$

(d) $f(x) = \sqrt{x + 3}\,;$ (e) $f(x) = 2\ln x + 1\,;$ (f) $f(x) = \tan\left(x - \frac{\pi}{2}\right)\,.$

8.3 Give the following functions as the sum of a polynomial and a proper rational function:

(a) $f(x) = \dfrac{3x^5 + 2x^4 - x^3 + 2x}{x^2 + 2}\,;$ (b) $f(x) = \dfrac{x^5 - x^4 + 2x^2 + 1}{x^3 + 3x + 1}\,.$

8.4 Determine the domain D_f, the range R_f, zeroes and periodicity of the following trigonometric functions:

(a) $f(x) = 2\sin\left(2x + \frac{\pi}{2}\right)\,;$ (b) $f(x) = \frac{1}{2}\cos\left(\frac{x}{2} - \pi\right)\,;$

(c) $f(x) = \pi\tan\left(x + \sqrt{3}\right)\,;$ (d) $f(x) = s\cot(x + 1) \quad (s \in \mathbb{R})\,.$

8.5 Determine the composite functions $f \circ g$ and $g \circ f$ for the following functions f and g provided that they exist:

(a) $f(x) = x + 3\,, \quad D_f = \mathbb{R}\,; \qquad g(x) = e^x\,, \quad D_g = \mathbb{R}\,;$

(b) $f(x) = (x-1)(x^2+2)\,, \quad D_f = [1,\infty); \qquad g(x) = \sqrt{x}\,, \quad D_g = [0,\infty)\,;$

(c) $f(x) = (x+1)^2\,, \quad D_f = \mathbb{R}\,; \qquad g(x) = \ln x\,, \quad D_g = (0,\infty)\,;$

(d) $f(x) = 2x+1\,, \quad D_f = \mathbb{R}\,; \qquad g(x) = x^2+x-3\,, \quad D_f = \mathbb{R}\,;$

8.6 Determine the inverse functions f^{-1} of the following functions $f : D_f \to R_f$ provided that they exist:

(a) $f(x) = 2x-7\,, \quad D_f = \mathbb{R}\,;$

(b) $f(x) = x^2+3\,, \quad D_f = [0,\infty)\,;$

(c) $f(x) = x^{1/4}-1\,, \quad D_f = [0,\infty)\,;$

(d) $f(x) = e^{2x}-1\,, \quad D_f = \mathbb{R}\,;$

(e) $f(x) = \dfrac{\sqrt{x}-4}{\sqrt{x}+4}\,, \quad D_f = \mathbb{R}_{\geq 0}\,;$

(f) $f(x) = x^4 + 2x^2 + 1; \quad D_f = \mathbb{R}.$

Chapter 9

Differentiation

Differentiation can be used for describing and modeling movement and alteration processes in many disciplines. It can be applied e.g. in optimization for finding extreme points or as a base for differential equations when simulating dynamic systems. The learning goals of this chapter are as follows:

- refreshing the concepts of a limit, continuity and the derivative of a function;
- repeating the major rules of differentiation and getting practice in their application; and
- reviewing some applications of differential calculus like graphing functions, extremum problems and determining the zeroes of a function approximately.

9.1 Limit and Continuity of a Function

First, we formally introduce the notation of a limit and illustrate it then in Fig. 9.1.

Limit of a function:
The real number L is called the **limit** of the function $f : D_f \to \mathbb{R}$ as x tends to x_0 if for any sequence $\{x_n\}$ with $x_n \neq x_0,\ x_n \in D_f,\ n = 1, 2, \ldots$, which converges to x_0, the sequence of the function values $\{f(x_n)\}$ converges to L.

The limit of a function is closely related to the limit of a sequence. In the above description, the sequence $\{x_n\}$ related to the independent variable x and the sequence of $\{f(x_n)\}$ related to the dependent variable $y = f(x)$ are considered. A function has the limit L as x tends to some value x_0 (represented by a sequence, where the terms x_n are from some value n on very close to x_0), if the sequence of the function values $\{f(x_n)\}$ tends to the value L. Often, the δ-ϵ notation is used in mathematics. This means that a function f is continuous at x_0 if for any positive real number ϵ, there exists a positive $\delta = \delta(\epsilon)$ depending on ϵ such that

$$|x - x_0| < \delta(\epsilon) \qquad \Longrightarrow \qquad |f(x) - f(x_0| < \epsilon\,.$$

(see Fig. 9.1). Thus, for any positive (in particular, very small) ϵ, one can give a positive number δ such that the function values $f(x)$ are from the open interval $(f(x)-\epsilon, f(x)+\epsilon)$ provided that

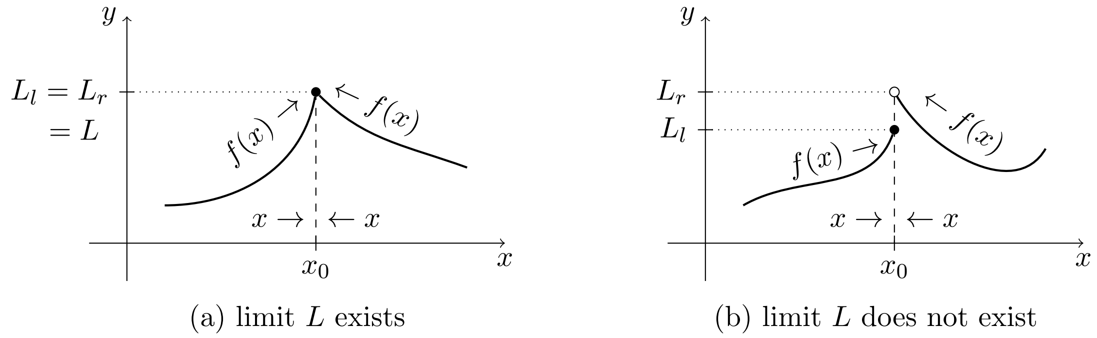


Figure 9.1: The limit of a function

x is from the open interval $(x_0 - \delta, x_0 + \delta)$. If ϵ becomes smaller, then usually also δ becomes smaller. Notice that for a limit of a function at a point x_0, it is not necessary that the function value $f(x_0)$ exists.

Similarly, we can consider **one-side limits**. If for the sequence $\{x_n\}$, we have $x_n > x_0$ for $n = 1, 2, \ldots,$ (i.e., one approaches from the right side to x_0), we obtain the **right-side limit** L_r and we indicate this by $x \to x_0 + 0$. Analogously, if $x_n < x_0$ for all $n = 1, 2, \ldots,$ (i.e., one approaches from the left side), we obtain the **left-side limit** L_l and we indicate this by $x \to x_0 - 0$. The limit L of a function f as x tends to x_0 exists if and only if both the right- and left side limits exist and coincide: $L_r = L_l = L$. Next, we review some rules for working with limits.

Assume that the limits

$$\lim_{x\to x_0} f(x) = L_1 \qquad \text{and} \qquad \lim_{x\to x_0} g(x) = L_2.$$

exist. Then we have the following rules for determining limits:

Rules for limits:

(1) $$\lim_{x\to x_0} \Big[f(x) + g(x)\Big] = \lim_{x\to x_0} f(x) + \lim_{x\to x_0} g(x) = L_1 + L_2\,;$$

(2) $$\lim_{x\to x_0} \Big[f(x) - g(x)\Big] = \lim_{x\to x_0} f(x) - \lim_{x\to x_0} g(x) = L_1 - L_2\,;$$

(3) $$\lim_{x\to x_0} \Big[f(x) \cdot g(x)\Big] = \lim_{x\to x_0} f(x) \cdot \lim_{x\to x_0} g(x) = L_1 \cdot L_2\,;$$

(4) $$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \frac{\lim\limits_{x\to x_0} f(x)}{\lim\limits_{x\to x_0} g(x)} = \frac{L_1}{L_2} \qquad \text{provided that } L_2 \neq 0\,;$$

(5) $$\lim_{x\to x_0} \sqrt{f(x)} = \sqrt{\lim_{x\to x_0} f(x)} = \sqrt{L_1} \qquad \text{provided that } L_1 \geq 0\,;$$

(6) $$\lim_{x\to x_0} [f(x)]^n = \left[\lim_{x\to x_0} f(x)\right]^n = L_1^n\,;$$

(7) $$\lim_{x\to x_0} \Big[a^{f(x)}\Big] = a^{\left[\lim\limits_{x\to x_0} f(x)\right]} = a^{L_1}\,.$$

We illustrate the calculation of limits by the following examples.

Example 9.1 *We calculate*

$$L = \lim_{x \to 0} \frac{|x|}{x}.$$

Using the definition of the absolute value, we obtain for the one-side limits

$$\begin{aligned} L_r &= \lim_{x \to 0+0} \frac{|x|}{x} = \lim_{x \to 0+0} \frac{x}{x} = 1 \\ L_l &= \lim_{x \to 0-0} \frac{|x|}{x} = \lim_{x \to 0-0} \frac{x}{x} = -1 \end{aligned}$$

Due to the definition of the absolute value, it is necessary to consider the right- and left-side limits separately. Since they are different, the limit L does not exist.

Example 9.2 *We calculate*

$$L = \lim_{x \to 1} \frac{x^2 + 3x - 2}{x + 7} = \frac{\lim\limits_{x \to 1} x^2 + 3 \lim\limits_{x \to 1} x - \lim\limits_{x \to 1} 2}{\lim\limits_{x \to 1} x + \lim\limits_{x \to 1} 7} = \frac{1 + 3 - 2}{1 + 7} = \frac{2}{8} = \frac{1}{4}.$$

Example 9.3 *We determine*

$$L = \lim_{x \to 4} \frac{(x+5)(x+2)}{\sqrt{x}}.$$

This yields

$$L = \frac{\lim_{x\to 4}(x+5)\cdot\lim_{x\to 4}(x+2)}{\sqrt{\lim_{x\to 4} x}} = \frac{9\cdot 6}{\sqrt{4}} = 27\,.$$

Example 9.4 *We calculate*

$$L = \lim_{x\to 2}\left(e^{3-x} + 2^x\right)\,.$$

We obtain

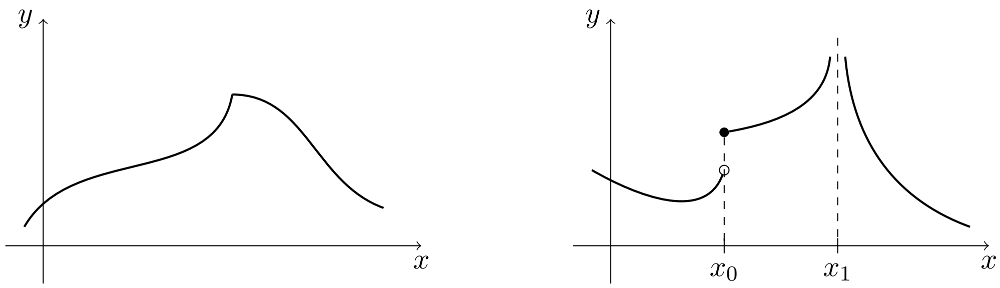
$$L = e^{\lim_{x\to 2}(3-x)} + 2^{\lim_{x\to 2} x} = e^1 + 2^2 = e + 4\,.$$

Example 9.5 *We determine*

$$L = \lim_{x\to 9}\frac{x-9}{\sqrt{x}-3}\,.$$

If we consider the limit of the numerator and the denominator as $x \to 4$, we observe that both terms tend to zero so that we cannot find a result in this way. Later we discuss a way how to treat such indeterminate forms (where both the numerator and denominator tend to zero) by means of differential calculus. However, here we can expand the fraction by $\sqrt{x}+3$ and obtain

$$L = \lim_{x\to 9}\frac{(x-9)(\sqrt{x}+3)}{(\sqrt{x}-3)(\sqrt{x}+3)} = \lim_{x\to 9}\frac{(x-9)(\sqrt{x}+3)}{x-9} = \lim_{x\to 9}(\sqrt{x}+3) = 6\,.$$



(a) f continuous (b) f discontinuous at x_0 and x_1

Figure 9.2: Continuous and discontinuous functions

Continuous function:
A function $f : D_f \to \mathbb{R}$ is said to be **continuous** at $x_0 \in D_f$ if the limit of function f as x tends to x_0 exists and if this limit coincides with the function value $f(x_0)$, i.e., we have

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

In Fig. 9.2, examples of a continuous and a discontinuous function are given. We review some types of discontinuities:

- **jump:** A function f has a (finite) jump at x_0 if both one-side limits as $x \to x_0 - 0$ and $x \to x_0 + 0$ exist but are different. The existence of the function value $f(x_0)$ does not play a role.

- **gap:** A function f has a gap at x_0 if

$$\lim_{x \to x_0} f(x)$$

exists but the function value $f(x_0)$ does not exist.

- **pole:** A rational function $f = P/Q$ of two polynomials P and Q has a pole at x_0 if $P(x_0) \neq 0$ and $Q(x_0) = 0$. Thus, the function value $f(x_0)$ does not exist and for $x \to x_0+0$ and $x \to x_0 - 0$, the function values tend to $\pm\infty$.

Example 9.6 *Given is the function $f : D_f \to \mathbb{R}$ with*

$$f(x) = \frac{x^2 + x - 6}{x - 2}\,.$$

Function f is continuous for all points $x \neq 2$ so that the only critical point is $x = 2$. The term in the numerator has the zeroes $x = 2$ and $x = -3$. Therefore, we get

$$\lim_{x \to 2} f(x) = \lim_{x \to 2} \frac{(x-2) \cdot (x+3)}{x-2} = \lim_{x \to 2} (x+3) = 5\,.$$

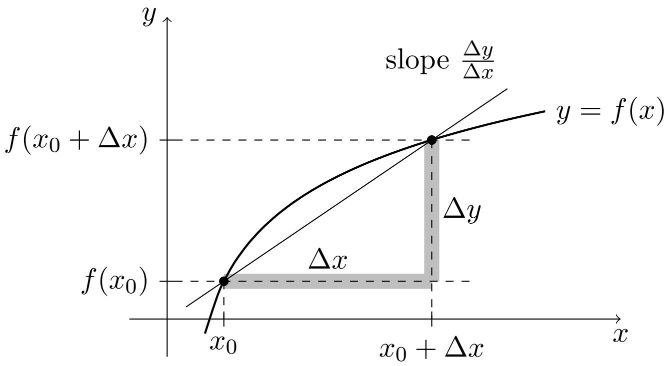


Figure 9.3: The difference quotient of a function

Note that for the calculation of the limit, we can assume $x \neq 2$ so that the above transformation is correct. However, the function value $f(2)$ does not exist and so function f is discontinuous at $x = 2$, i.e., it has a gap.

Example 9.7 *Given is the function $f : \mathbb{R} \to \mathbb{R}$ with*

$$f(x) = \begin{cases} 2x - 2 & \textit{for } x \in (-\infty, 2] \\ a + (x-2)^2 & \textit{for } x \in (2, \infty), \end{cases}$$

where a is a real parameter. We determine the parameter a such that function f is continuous at all points. Function f is obviously continuous at any point $x \neq 2$. For $x = 2$, we obtain

$$\lim_{x \to 2-0} f(x) = f(2) = 2$$

and

$$\lim_{x \to 2+0} \left(a + (x-2)^2\right) = a\,.$$

Thus, function f is continuous at $x = 2$ for $a = 2$.

9.2 The Derivative of a Function

Difference quotient of a function:
Let $f : D_f \to \mathbb{R}$ and $x_0, x_0 + \Delta x \in (a, b) \subseteq D_f$. The ratio

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

is called the **difference quotient** of function f with respect to points $x_0 + \Delta x$ and x_0.

The difference quotient characterizes the **average increase or decrease** of a function f over the interval $[x_0, x_0 + \Delta x]$. It gives the quotient of the change in the y variable over the change in the x variable which corresponds to the slope of the line going through the points $(x_0, f(x_0))$ and $(x_0 + \Delta x, f(x_0 + \Delta x))$. The difference quotient is illustrated in Fig. 9.3.

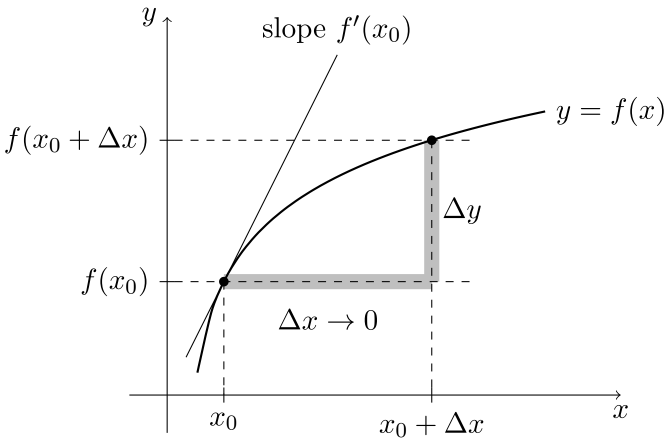


Figure 9.4: The derivative of a function

We are now interested in the change of the value of function f when only a small change in the independent variable x is allowed. This means that we look what happens if we determine the limit of the difference quotient as Δx tends to zero. This leads to the derivative of a function as follows.

Derivative of a function:
A function $f : D_f \to \mathbb{R}$ with $y = f(x)$ is said to be **differentiable** at the point $x_0 \in (a, b) \subseteq D_f$ if the limit

$$L = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

exists. This limit L is called the **derivative** (or the **differential quotient**) of the function f at point x_0 and denoted as

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

.

Geometrically, the value $f'(x_0)$ gives the slope of the tangent to the curve $y = f(x)$ at the point $(x_0, f(x_0))$. It describes the change in the function value of f when only a 'very small' change in the independent variable is considered. Roughly, we can interpret it as the number of units by which the variable y changes (i.e., increases if greater than zero or decreases if smaller than zero) when the independent variable x increases by one unit (i.e., from x_0 to $x_0 + 1$). The derivative $y' = f'(x)$ is illustrated in Fig. 9.4.

We consider the following example for finding the derivative of a function (at the moment still by applying the definition).

Example 9.8 *Let function $f : \mathbb{R} \to \mathbb{R}$ be given by*

$$y = f(x) = x^3\,.$$

We determine the first derivative of function f at the point x_0. Applying the above definition, *this gives*

$$\begin{aligned} f'(x_o) = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} &= \lim_{\Delta x \to 0} \frac{(x_0 + \Delta x)^3 - x_0^3}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{x_0^3 + 3x_0^2 \Delta x + 3x_0 (\Delta x)^2 + (\Delta x)^3 - x_0^3}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{\Delta x \cdot (3x_0^2 + 3x_0 \Delta x + (\Delta x)^2}{\Delta x} \\ &= 3x_0^2 \,. \end{aligned}$$

9.3 Elementary Rules

In this section, we review the derivatives of the major types of functions.

Table 9.1: Derivatives of elementary functions:

$y = f(x)$	$y' = f'(x)$	D_f	
C	0	$-\infty < x < \infty,$	C constant
x^n	nx^{n-1}	$-\infty < x < \infty,$	$n \in \mathbb{N}$
x^α	$\alpha x^{\alpha-1}$	$0 < x < \infty,$	$\alpha \in \mathbb{R}$
e^x	e^x	$-\infty < x < \infty$	
a^x	$a^x \ln a$	$-\infty < x < \infty,$	$a > 0$
$\ln x$	$\frac{1}{x}$	$0 < x < \infty$	
$\log_a x$	$\frac{1}{x \ln a}$	$0 < x < \infty,$	$a > 0$
$\sin x$	$\cos x$	$-\infty < x < \infty$	
$\cos x$	$-\sin x$	$-\infty < x < \infty$	
$\tan x$	$\frac{1}{\cos^2 x} = 1 + \tan^2 x$	$x \neq \frac{\pi}{2} + k\pi,$	$k \in \mathbb{Z}$
$\cot x$	$-\frac{1}{\sin^2 x} = -(1 + \cot^2 x)$	$x \neq k\pi,$	$k \in \mathbb{Z}$

Next, we review some more differentiation rules considering sums, differences, products and quotients of two functions f and g.

Derivative of a Sum, Difference, Product and Quotient:

(1) $(f+g)'(x) = f'(x) + g'(x)\,;$

(2) $(f-g)'(x) = f'(x) - g'(x)\,;$

(3) $(f \cdot g)'(x) = f'(x) \cdot g(x) + f(x) \cdot g'(x)\,;$

(4) $\left(\frac{f}{g}\right)'(x) = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{[g(x)]^2}\,.$

Example 9.9 *We illustrate the use of the above rules by the differentiation of the following functions:*

1. *Let* $f : D_f \to \mathbb{R}$ *with* $f(x) = \sqrt{x}$. *Using* $\sqrt{x} = x^{1/2}$, *we obtain*

$$f'(x) = \frac{1}{2} \cdot x^{-1/2} = \frac{1}{2\,x^{1/2}} = \frac{1}{2\,\sqrt{x}}\,.$$

2. *Let* $f : D_f \to \mathbb{R}$ *with* $f(x) = 2x^4 - 3x^3 + 5x - 7$. *Then*

$$f'(x) = 8x^3 - 9x^2 + 5\,.$$

3. *Let* $f : D_f \to \mathbb{R}$ *with* $f(x) = x^2 \sin x$. *Applying the product rule (3), we get*

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

4. *Let* $f : D_f \to \mathbb{R}$ *with* $f(x) = x^3 \cdot a^x$. *Using the product rule, we get*

$$f'(x) = 3x^2 \cdot a^x + x^3 \cdot a^x \cdot \ln a = x^2 \cdot a^x(3 + x \cdot \ln a) \ .$$

5. *Let* $f : D_f \to \mathbb{R}$ *with*

$$f(x) = \frac{x^2 + 2}{x - 5} \ .$$

Applying the quotient rule (4), we get

$$f'(x) = \frac{2x \cdot (x - 5) - (x^2 + 2) \cdot 1}{(x - 5)^2} = \frac{2x^2 - 10x - x^2 - 2}{(x - 5)^2} = \frac{x^2 - 10x - 2}{(x - 5)^2} \ .$$

6. *Let* $f : D_f \to \mathbb{R}$ *with*

$$f(x) = \frac{\ln x}{x^3} \ .$$

Then, using the quotient rule,

$$f'(x) = \frac{\frac{1}{x} \cdot x^3 - \ln x \cdot 3x^2}{(x^3)^2} = \frac{x^2 - 3x^2 \cdot \ln x}{x^6} = \frac{x^2(1 - 3 \ln x)}{x^6} = \frac{1 - 3 \ln x}{x^4} \ .$$

Let us now consider a composite function $f = g \circ h$. The rule for the differentiation of function f is given by the following formula, known as the **chain rule**:

Derivative of a composite function $f = g \circ h$**:**

$$[g\,[h(x)]]' = g'(z) \cdot h'(x) \qquad \text{with } z = h(x)$$

Example 9.10 *Let us consider the following examples for illustrating the use of the chain rule.*

1. *Let* $f : D_f \to \mathbb{R}$ *with*
$$f(x) = e^{-x^3} \; .$$
Then we have
$$g(z) = e^z \qquad \text{and} \qquad z = h(x) = -x^3 \; .$$
We obtain
$$g'(z) = e^z \qquad \text{and} \qquad h'(x) = -3x^2 \; .$$
Therefore, we obtain
$$f'(x) = g'(h(x)) \cdot h'(x) = e^{-x^3} \cdot (-3x^2) = -3x^2 e^{-x^3} \; .$$

2. *Let* $f : D_f \to \mathbb{R}$ *with*
$$f(x) = (2x^2 + 3)^4 \; .$$
We have
$$g(z) = z^4 \qquad \text{and} \qquad h(x) = 2x^2 + 3 \; .$$
Then we obtain
$$g'(z) = 4z^3 \qquad \text{and} \qquad h'(x) = 4x \; .$$
Therefore, we obtain
$$f'(x) = g'(h(x)) \cdot h'(x) = 4(2x^2 + 3)^3 \cdot 4x = 16x\,(2x^2 + 3)^3 \; .$$

3. *Let* $f : D_f \to \mathbb{R}$ *with*
$$f(x) = \sqrt{x^2 - 6x + 5} \; .$$
We have
$$g(z) = \sqrt{z} \qquad \text{and} \qquad z = h(x) = x^2 - 6x + 5 \; .$$
We obtain
$$g'(z) = \frac{1}{2}\, z^{-1/2} = \frac{1}{2\,\sqrt{z}} \qquad \text{and} \qquad h'(x) = 2x - 6 \; .$$
Thus, we get
$$f'(x) = g'(h(x)) \cdot h'(x) = \frac{1}{2\sqrt{x^2 - 6x + 5}} \cdot (2x - 6) = \frac{x - 3}{\sqrt{x^2 + 2x - 3}} \; .$$

An alternative formulation of the chain rule for $f = g \circ h$ with $f = g(y)$ and $y = h(x)$ is given by
$$\frac{df}{dx} = \frac{dg}{dy} \cdot \frac{dy}{dx} \, .$$

The chain rule can also be applied if a function is composed of more than two functions. We illustrate this by the following example.

Example 9.11 *Let $f : D_f \to \mathbb{R}$ with*

$$f(x) = e^{\sqrt{x^2+3x}}\,.$$

Since the term under the square root must be negative, the function is only defined for $x^2+3x \geq 0$. Solving this quadratic inequality, we have $x \leq -3$ or $x \geq 0$. Therefore, the domain D_f is given by $D_f = (-\infty, -3] \cup [0, \infty)$. Function f is of the type $f = g \circ h \circ k$, where

$$w = g(z) = e^z, \quad z = h(y) = \sqrt{y}, \quad y = k(x) = x^2 + 3x\,.$$

We get

$$g'(z) = e^z, \qquad h'(y) = \frac{1}{2\sqrt{y}}, \qquad k'(x) = 2x + 3\,.$$

Thus, we obtain

$$f'(x) = e^{\sqrt{x^2+3x}} \cdot \frac{2x+3}{2 \cdot \sqrt{x^2+3x}}\,.$$

Next, we review **higher-order derivatives.** If function f' with $y' = f'(x)$ is again differentiable, function

$$y'' = f''(x) = \frac{df'}{dx}(x) = \frac{d^2 f}{(dx)^2}(x)$$

is called the **second derivative** of function f at point x. We can continue with this procedure and obtain in general:

$$y^{(n)} = f^{(n)}(x) = \frac{df^{n-1}}{dx}(x) = \frac{d^n f}{(dx)^n}(x), \qquad n \geq 2,$$

which denotes the **nth derivative** of function f at the point $x \in D_f$. Notice that we use $f'(x), f''(x), f'''(x)$, and for $n \geq 4$, we use the notation $f^{(n)}(x)$. Higher-order derivatives are used for instance in the rest of this chapter when investigating specific properties of functions.

Example 9.12 *Let function $f : D_f \to \mathbb{R}$ with*

$$f(x) = 6x^3 + \frac{1}{x} + e^{2x+1}$$

be given. We determine all derivatives until fourth order and obtain

$$\begin{aligned}
f'(x) &= 12x^2 - \frac{1}{x^2} + 2e^{2x+1} \\
f''(x) &= 24x + \frac{2}{x^3} + 4e^{2x+1} \\
f'''(x) &= 24 - \frac{6}{x^4} + 8e^{2x+1} \\
f^{(4)}(x) &= \frac{24}{x^5} + 16e^{2x+1}\,.
\end{aligned}$$

Example 9.13 *Let function $f : D_f \to \mathbb{R}$ with*

$$f(x) = 2\sin(3x+1) + 4x^2$$

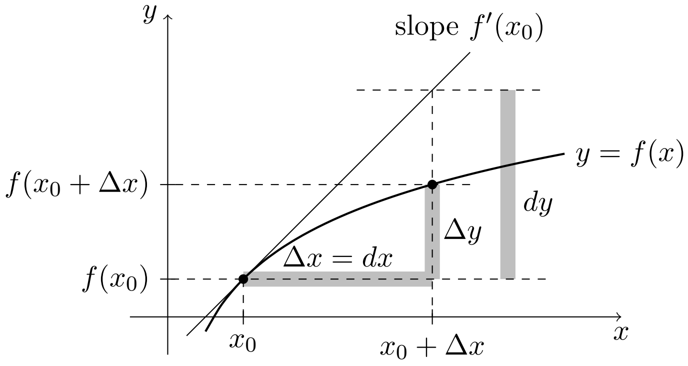


Figure 9.5: The differential of a function

be given. We determine all derivatives until the fourth order and obtain

$$\begin{aligned}
f'(x) &= 2\cos(3x+1)\cdot 3 + 8x = 6\cos(3x+1) + 8x \\
f''(x) &= 6(-\sin(3x+1)\cdot 3 + 8 = -18\sin(3x+1) + 8 \\
f'''(x) &= -18\cos(3x+1)\cdot 3 = -54\cos(3x+1) \\
f^{(4)}(x) &= 54\sin(3x+1)\cdot 3 = 162\sin(3x+1)\,.
\end{aligned}$$

9.4 The Differential

Next, we introduce the notion of the differential which can be used for estimating the error when calculating function values, i.e., what is the influence on the error in the y-value if the x-value is only known with some possible deviations.

Differential of a function:
Let function $f : D_f \to \mathbb{R}$ be differentiable at point $x_0 \in (a,b) \subseteq D_f$. The **differential** of function f at point x_0 is defined as

$$dy = f'(x_0)\cdot dx\ .$$

The differential is illustrated in Fig. 9.5.

The differential can be used e.g. for the estimation the change in the function value when a small change in the independent variable x is considered. For small changes in the variable x, we have

$$\Delta y \approx dy = f'(x_0)\cdot dx\ .$$

Example 9.14 *Let function* $f : D_f \to \mathbb{R}$ *with*

$$f(x) = \left(\frac{x}{2} - 1\right)^2$$

and $x_0 = 25$ *be given. We compare* Δy *and* dy *for the case when* $\Delta x = dx = 1$*. Then we obtain*

$$\begin{aligned} \Delta y &= f(x_0 + 1) - f(x_0) = \left(\frac{25+1}{2} - 1\right)^2 - \left(\frac{25}{2} - 1\right)^2 = \left(\frac{26}{2} - 1\right)^2 - \left(\frac{25}{2} - 1\right)^2 \\ &= 12^2 - 11.5^2 = 144 - 132.25 = 11.75 \end{aligned}$$

Using

$$f'(x) = 2 \cdot \left(\frac{x}{2} - 1\right) \cdot \frac{1}{2} = \frac{x}{2} - 1,$$

we get

$$f'(25) = \frac{25}{2} - 1 = 11.5$$

and therefore

$$dy = f'(25) \cdot dx = 11.5 \cdot 1 = 11.5\ ,$$

i.e., Δy *and* dy *differ by 0.25.*

Example 9.15 *The radius* R *of a sphere has been measured as* $R = (4.502 \pm 0.005)$ *cm, i.e., we have* $|\Delta R| = |dR| \leq 0.005$*. We estimate the absolute and relative error of the volume and the surface of the sphere. The formula for the volume of a sphere is given by*

$$V(R) = \frac{4}{3}\pi R^3\ .$$

Differentiation gives

$$V'(R) = 4\pi R^2$$

and thus

$$|dV| = 4\pi R^2 \, |dR| \, .$$

Using $R = 4.502$ *and* $|dR| \leq 0.005$*, we get for the absolute error*

$$|dV| \leq 4\pi(4.502)^2 \cdot 0.005 = 1.2735 \ cm^3 \, .$$

Moreover, we get

$$V(4.502) = \frac{4}{3}\pi \cdot (4.502)^3 = 382.2128 \ cm^3$$

and therefore, we have for the relative error

$$\left|\frac{dV}{V}\right| = \frac{|dV|}{V} \leq \frac{1.2735}{382.2128} = 0.0033 \, .$$

Thus, the relative error for the volume is estimated to be not larger than 0.33 %.

The formula for the surface of a sphere is given by

$$S(R) = 4\pi R^2 \, .$$

Differentiation gives

$$S'(R) = 8\pi R \, .$$

Using again $R = 4.502$ *and* $|\Delta R| = |dR| \leq 0.005$, *we get*

$$|dS| \leq 8\pi \cdot 4.502 \cdot 0.005 = 0.5657 \; cm^2 .$$

Moreover, the surface is obtained as

$$S(4.502) = 4\pi \cdot (4.502)^2 = 254.6952 \; cm^2 .$$

Therefore, we obtain

$$\left|\frac{dS}{S}\right| = \frac{|dS|}{S} \leq \frac{0.5657}{254.6952} = 0.0022 ,$$

i.e., the relative error for the surface can be estimated to be not larger than 0.22 %.

9.5 Graphing functions

To get a quantitative overview on a function $f : D_f \to \mathbb{R}$ and its graph, respectively, we determine and investigate:

- domain D_f (if not given) and possibly the range R_f;
- zeroes and discontinuities;
- monotonicity of the function;
- extreme points and values;

- convexity and concavity of the function;
- inflection points;
- limits and asymptotic behavior, i.e., how does the function behave when x tends to $\pm\infty$.

Having the detailed information listed above, we can draw the graph of function f. In the following, we discuss the above subproblems in detail.

In connection with functions of one variable, we have already discussed how to determine the domain D_f and we have classified the different types of discontinuities. As far as the determination of zeroes is concerned, we have already considered special cases such as zeroes of a quadratic function. For more complicated functions, where finding the zeroes is difficult or analytically impossible, we give in addition to Regula falsi from Chapter 3 another numerical procedure for the approximate determination of zeroes later in this chapter. We start with the investigation of the monotonicity of a function.

9.5.1 Monotonicity

By means of the first derivative f', we can determine intervals in which a function f is (strictly) increasing or decreasing. In particular, the following property can be formulated.

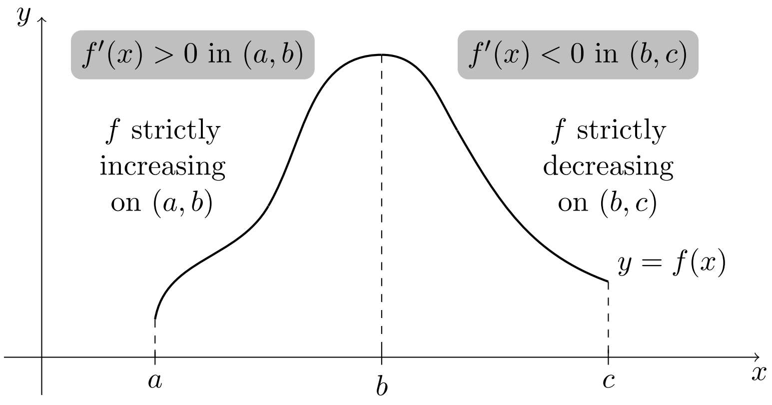


Figure 9.6: Monotonicity of a function

Check of a function for (strict) monotonicity:
Let function $f : D_f \to \mathbb{R}$ be differentiable on the open interval (a, b) and let $I = [a, b] \subseteq D_f$. Then:

(1) Function f is **increasing** on I if and only if $f'(x) \geq 0$ for all $x \in (a, b)$.

(2) Function f is **decreasing** on I if and only if $f'(x) \leq 0$ for all $x \in (a, b)$.

(3) If $f'(x) > 0$ for all $x \in (a, b)$, then function f is **strictly increasing** on I.

(4) If $f'(x) < 0$ for all $x \in (a, b)$, then function f is **strictly decreasing** on I.

We remind from Chapter 8 that, if a function is (strictly) increasing or (strictly) decreasing on an interval $I \subseteq D_f$, we say that function f is (strictly) monotonic on the interval I. Checking a function f for monotonicity requires to determine the intervals on which function f is monotonic and strictly monotonic, respectively. Fig. 9.6 illustrates the determination of monotonicity intervals by the first derivative.

9.5.2 Extreme Points

First we give the definition of a local and of a global extreme point which can be either a minimum or a maximum.

Local extreme point:
A function $f : D_f \to \mathbb{R}$ has a **local maximum (minimum)** at point $x_0 \in D_f$ if there is an interval $(a, b) \subseteq D_f$ containing x_0 such that

$$f(x) \leq f(x_0) \qquad (f(x) \geq f(x_0),\ \text{respectively}) \tag{9.1}$$

for all points $x \in (a, b)$. Point x_0 is called a **local maximum (minimum) point**.

If inequality (9.1) holds for all points $x \in D_f$, function f has at point x_0 a **global maximum (minimum)**, and x_0 is called a **global maximum (minimum) point**.

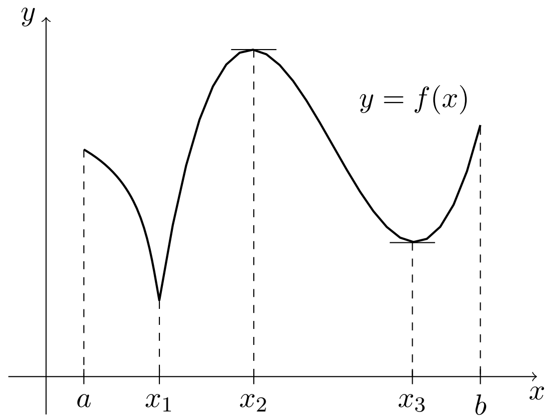


Figure 9.7: Local extreme points of a function

Necessary condition for local optimality:
Let function $f : D_f \to \mathbb{R}$ be differentiable on the open interval $(a, b) \subseteq D_f$. If function f has a local maximum or local minimum at point $x_0 \in (a, b)$, then $f'(x_0) = 0$.

Stationary point:
A point $x_0 \in (a, b)$ with $f'(x_0) = 0$ is called a **stationary point** (or **critical point**).

When searching global maximum and minimum points for a function f in a closed interval $I = [a, b] \subseteq D_f$, we have to search among the following types of points:

(1) points in the open interval (a, b), where $f'(x) = 0$ (stationary points);
(2) end points a and b of I;
(3) points in (a, b), where $f'(x)$ does not exist.

Local extreme points are illustrated in Fig. 9.7.

We note that only points according to (1) can be found by means of differential calculus. Points according to (2) and (3) have to be checked separately. Returning to Fig. 9.7, there are two stationary points x_2 and x_3. The local (and global) minimum point x_1 cannot be found by differential calculus since the function drawn in Fig. 9.7 is not differentiable at the point x_1.

The following two claims present sufficient conditions for so-called **isolated** local extreme points, for which in Inequality (9.1) the strict inequality holds for all $x \in (a, b)$ different from x_0.

First, we give a criterion for deciding whether a stationary point is a local extreme point in the case of a differentiable function which uses only the first derivative of function f.

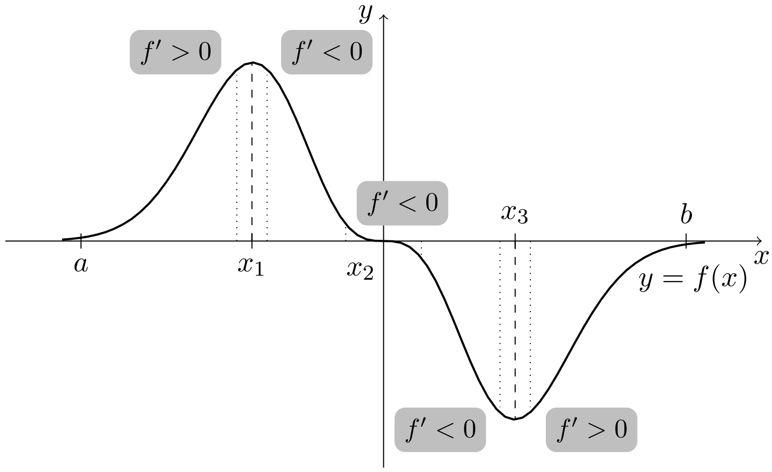


Figure 9.8: First-order derivative test for local extrema

Sufficient condition for local optimality (first-order derivative test):
Let function $f : D_f \to \mathbb{R}$ be differentiable on the open interval $(a, b) \in D_f$ and $x_0 \in (a, b)$ be a stationary point of function f. Then:

(1) If $f'(x) > 0$ for all $x \in (a^*, x_0) \subseteq (a, b)$ and $f'(x) < 0$ for all $x \in (x_0, b^*) \subseteq (a, b)$, then x_0 is a **local maximum point** of function f.

(2) If $f'(x) < 0$ for all $x \in (a^*, x_0) \subseteq (a, b)$ and $f'(x) > 0$ for all $x \in (x_0, b^*) \subseteq (a, b)$, then x_0 is a **local minimum point** of function f.

(3) If $f'(x) > 0$ for all $x \in (a^*, x_0) \subseteq (a, b)$ and for all $x \in (x_0, b^*) \subseteq (a, b)$, then x_0 is not a local extreme point of function f. The same conclusion holds if $f'(x) < 0$ on both sides of x_0.

The first-order derivative test for local extreme points is illustrated in Fig. 9.8 (x_1 is a local maximum point, x_2 is not a local extreme point, and x_3 is a local minimum point).

Alternatively, one can use the subsequent criterion which uses higher-order derivatives.

Sufficient condition for local optimality (higher-order derivative test):
Let $f : D_f \to \mathbb{R}$ be n times continuously differentiable on the open interval $(a, b) \subseteq D_f$ and $x_0 \in (a, b)$ be a stationary point. If

$$f'(x_0) = f''(x_0) = f'''(x_0) = \ldots = f^{(n-1)}(x_0) = 0 \quad \text{and} \quad f^{(n)}(x_0) \neq 0,$$

where number n is even, then point x_0 is a local extreme point of function f, in particular:

(1) If $f^{(n)}(x_0) < 0$, then function f has at point x_0 a **local maximum**.

(2) If $f^{(n)}(x_0) > 0$, then function f has at point x_0 a **local minimum**.

Often the above sufficient condition works already for $n = 2$ and one only needs to check the sign of the second derivative of all stationary points.

Example 9.16 *We determine all monotonicity intervals and local extreme points for the function* $f : [1, \infty) \to \mathbb{R}$ *with*

$$f(x) = \sqrt{4x-4} \cdot e^{-2x}\,.$$

Differentiation gives

$$\begin{aligned} f'(x) &= \frac{1}{2}(4x-4)^{-1/2} \cdot 4 \cdot e^{-2x} - 2e^{-2x}\sqrt{4x-4} \\ &= 2\,e^{-2x}\left[(4x-4)^{-1/2} - \sqrt{4x-4}\right] \\ &= 2\,e^{-2x} \cdot \frac{1-(4x-4)}{\sqrt{4x-4}} = 2\,e^{-2x} \cdot \frac{5-4x}{\sqrt{4x-4}}\,. \end{aligned}$$

From $f'(x) = 0$*, we obtain the stationary point* $x_0 = 5/4$*. To investigate the sign of* $f'(x)$*, we need to investigate the sign of* $5 - 4x$ *(notice that both terms* e^{-2x} *and* $\sqrt{4x-4}$ *are positive for* $x > 1$*). Thus, we have*

$$f'(x) > 0 \quad \text{for } 1 < x < \frac{5}{4}$$

and function f *is strictly increasing on the interval* $[1, 5/4]$*. Moreover, we have*

$$f'(x) < 0 \quad \text{for } x > \frac{5}{4}$$

and thus, function f *is strictly decreasing on the interval* $[5/4, \infty)$*. Due to the signs of* f' *around* x_0*, this point is a local maximum (due to the monotonicity properties it is also a global maximum). Alternatively, one could also check the sign of* f'' *at the point* x_0 *which would give* $f''(5/4) < 0$ *and thus, it confirms that* x_0 *is a local maximum point. Note also that the global minimum point of function* f *is* $x_1 = 1$ *with* $f(1) = 0$ *since all other points from the domain* $D_f = [1, \infty)$ *have a positive function value. Since* x_1 *is a boundary point, this minimum cannot be found by differential calculus, and one has to check such points separately.*

9.5.3 Convexity and Concavity

Criterion for convexity/concavity of a function:
Let function $f : D_f \to \mathbb{R}$ be twice differentiable on the open interval $(a, b) \subseteq D_f$ and let $I = [a, b]$. Then:

(1) Function f is **convex** on I if and only if $f''(x) \geq 0$ for all $x \in (a, b)$.

(2) Function f is **concave** on I if and only if $f''(x) \leq 0$ for all $x \in (a, b)$.

(3) If $f''(x) > 0$ for all $x \in (a, b)$, then function f is **strictly convex** on I.

(4) If $f''(x) < 0$ for all $x \in (a, b)$, then function f is **strictly concave** on I.

For deciding whether a function is convex or concave, the following notion of an inflection point can be helpful.

Inflection point:
The point $x_0 \in (a, b)$ is called an **inflection point** of function f when f changes at x_0 from being convex to being concave or vice versa.

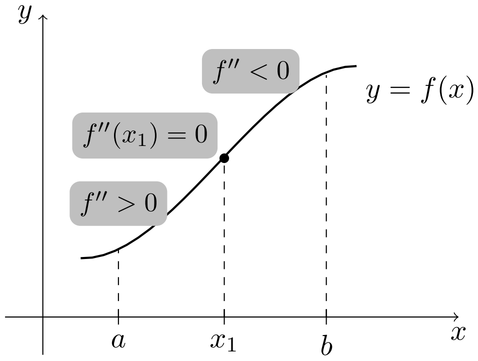


Figure 9.9: Convexity/concavity intervals of a function

Criterion for an inflection point:
Let function $f : D_f \to \mathbb{R}$ be n times continuously differentiable on the open interval $(a, b) \subseteq D_f$. Point $x_0 \in (a, b)$ is an **inflection point** of function f if and only if

$$f''(x_0) = f'''(x_0) = \ldots = f^{(n-1)}(x_0) = 0 \qquad \text{and} \qquad f^{(n)}(x_0) \neq 0,$$

where n is odd.

Often the above criterion works already with $n = 3$, i.e.: $f''(x_0) = 0$, $f'''(x_0) \neq 0$. One can also note that an inflection point can be interpreted as a local extreme point of the first derivative $y' = f'(x)$. The notions of convexity and concavity are illustrated in Fig. 9.9.

Example 9.17 *We discuss the properties of function $f : D_f \to \mathbb{R}$ given by*

$$f(x) = 6x^3 - 4x^2 - 10x \, .$$

Function f is defined for all real numbers and therefore, we have $D_f = \mathbb{R}$. In order to find all zeroes, we may factor out $2x$ which gives

$$f(x) = 2x \cdot (3x^2 - 2x - 5) = 0$$

The first zero is $x_1 = 0$ and the other zeroes can be obtained from

$$3x^2 - 2x - 5 = 0$$

or, after dividing both sides by 3, from

$$x^2 - \frac{2}{3}x - \frac{5}{3} = 0 \, .$$

This gives

$$x_{2,3} = \frac{1}{3} \pm \sqrt{\frac{1}{9} + \frac{5}{3}} = \frac{1}{3} \pm \sqrt{\frac{16}{9}}$$

which gives the solutions

$$x_2 = \frac{1}{3} + \frac{4}{3} = \frac{5}{3} \qquad \textit{and} \qquad x_3 = \frac{1}{3} - \frac{4}{3} = -1 \, .$$

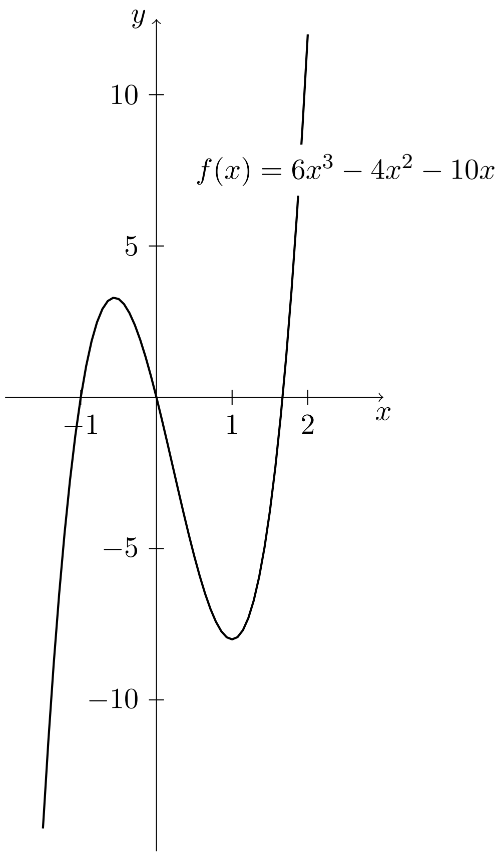


Figure 9.10: Graph of function f in Example 9.17

Function f has no discontinuities, and we get

$$\lim_{x\to-\infty} f(x) = -\infty \qquad \text{and} \qquad \lim_{x\to\infty} f(x) = \infty\,.$$

To find stationary points of function f, we determine zeroes of the first derivative:

$$f'(x) = 18x^2 - 8x - 10 = 0\,.$$

Dividing both sides by 18, we get a quadratic equation in normal form:

$$x^2 - \frac{4}{9}\,x - \frac{5}{9} = 0\,.$$

Thus, we get the solutions

$$x_{E1,E2} = \frac{2}{9} \pm \sqrt{\frac{4}{81} + \frac{5}{9}} = \frac{2}{9} \pm \sqrt{\frac{49}{81}},$$

i.e., we get

$$X_{E1} = \frac{2}{9} + \frac{7}{9} = 1 \qquad \text{and} \qquad x_{E2} = \frac{2}{9} - \frac{7}{9} = -\frac{5}{9}$$

as stationary points for function f. We check the sufficient condition by means of the second derivative and obtain

$$f''(x) = 36x - 8\,.$$

This yields

$$f''(1) = 36 \cdot 1 - 8 = 28 > 0 \qquad \text{and} \qquad f''\left(-\frac{5}{9}\right) = 36 \cdot \left(-\frac{5}{9}\right) - 8 = -28 < 0\,.$$

Therefore, $x_{E2} = 1$ is a local minimum point with $f(1) = -8$, and $x_{E2} = -5/9$ is a local maximum point with $f(-5/9) \approx 3.75$. For finding inflection points, we set $f''(x) = 0$ which gives

$$x_I = \frac{2}{9}$$

as a candidate for an inflection point. We check the sufficient condition by means of the derivative and obtain

$$f'''(x) = 36,$$

i.e., $f'''(2/9) \neq 0$, and thus x_I is an inflection point of function f. We determine the sign of the second derivation and find that for $x > 2/9$, we get $f''(x) > 0$, while for $x < 2/9$, we have $f''(x) < 0$. Therefore, function f is strictly convex for $x \geq 2/9$ and strictly concave for $x \leq 2/9$. The graph of function f is given in Fig. 9.10.

Example 9.18 *We investigate the function $f : D_f \to \mathbb{R}$ with*

$$f(x) = \ln(1 + x^2)\,.$$

Since $1 + x^2 \geq 1$, we get $D_f = \mathbb{R}$. In order to find zeroes, the argument of the logarithmic function must be equal to 1. From $1 + x^2 = 1$, we get the zero $x_0 = 0$. Function f has no discontinuities. We get

$$\lim_{x\to\infty} \ln(1 + x^2) = \lim_{x\to-\infty} \ln(1 + x^2) = \infty\,.$$

Moreover, function f is an even function:

$$f(-x) = \ln(1 + (-x)^2) = \ln(1 + x^2) = f(x)\,.$$

Determining stationary points, we obtain from

$$f'(x) = \frac{1}{1 + x^2} \cdot 2x = \frac{2x}{1 + x^2} = 0$$

the only stationary point $x_E = 0$. Since $f'(x) < 0$ for $x < 0$ and $f'(x) > 0$ for $x > 0$, the point $x_E = 0$ is a local minimum point with $f(0) = 0$. Alternatively, this can be confirmed by means of the second derivative. We obtain

$$f''(x) = \frac{2(1 + x^2) - 2x \cdot 2x}{(1 + x^2)^2} = \frac{2 + 2x^2 - 4x^2}{(1 + x^2)^2} = \frac{2 - 2x^2}{(1 + x^2)^2}\,.$$

For the stationary point $x_E = 0$, we get

$$f''(x_E) = f''(0) = \frac{2}{1^2} = 2 > 0\,.$$

This confirms again that $x_E = 0$ is a local minimum point. In order to find inflection points, we set $f''(x) = 0$ which gives $2 - 2x^2 = 0$. Thus, the candidates for an inflection point are

$$x_{I1} = 1 \qquad \text{and} \qquad x_{I2} = -1\,.$$

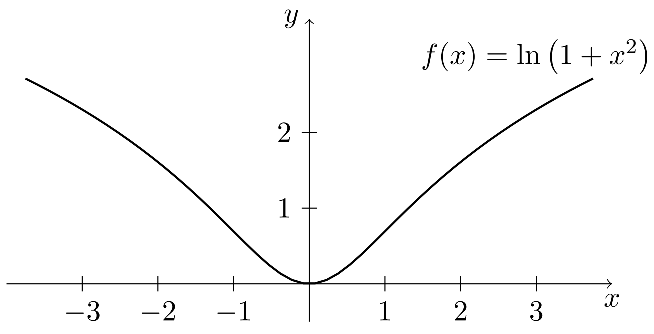


Figure 9.11: Graph of function f in Example 9.18

We observe that $f''(x) > 0$ for $x \in (-1, 1)$. Therefore, function f is strictly convex on the interval $[-1, 1]$. On the other hand, $f''(x) < 0$ for $x \in (-\infty, -1)$ and $x \in (1, \infty)$. Thus, the function f is strictly concave for all $x \in (-\infty, -1] \cup [1, \infty)$. Since the function f changes at the point $x_{I1} = -1$ from being concave to convex and at the point $x_{I2} = 1$ from being convex to being concave, both points are indeed inflection points. Alternatively, one may check that $f'''(-1) \neq 0$ and $f'''(1) \neq 0$. The graph of function f is given in Fig. 9.11.

Example 9.19 *We investigate the function $f : D_f \to \mathbb{R}$ with*

$$f(x) = \frac{x-1}{x^3} \, .$$

For finding the domain, we have to exclude points, where the denominator is equal to zero. Therefore, we have $D_f = \mathbb{R} \setminus \{0\}$. From $x - 1 = 0$ we get the zero $x_0 = 1$ (observe that for the denominator, we have $x_0^3 = 1 \neq 0$). Moreover, $x_0 = 0$ is the only discontinuity (namely a pole), and we get

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

Next, we determine stationary points. We get

$$f'(x) = \frac{x^3 - (x-1) \cdot 3x^2}{x^6} = \frac{x^2 \cdot (x - 3x + 3)}{x^6} = \frac{-2x + 3}{x^4} \, .$$

From $f'(x) = 0$, we get $x_1 = 3/2$ as the only stationary point. Since $f'(x) > 0$ for $x < 3/2$ and $f'(x) < 0$ for $x > 3/2$, point $x_E = 3/2$ is a local maximum point with $f(3/2) = 4/27$. Moreover, function f is strictly increasing on the interval $(-\infty, 3/2]$ and function f is strictly decreasing on the interval $[3/2, \infty)$.

To find inflection points, we determine the second derivative:

$$f''(x) = \frac{-2x^4 - (-2x + 3) \cdot 4x^3}{x^8} = \frac{x^3 \cdot (-2x + 8x - 12)}{x^8} = \frac{6x - 12}{x^5} \, .$$

From $f''(x) = 0$ we get $x_I = 2$ as the only candidate for a stationary point. Investigating the sign of the second derivative, we obtain $f''(x) > 0$ for $x < 0$, $f''(x) < 0$ for $0 < x < 2$, and $f''(x) > 0$ for $x > 2$. Therefore, function f is strictly convex for $x \in (-\infty, 0) \cup [2, \infty)$ and function f is strictly concave for $x \in [0, 2]$. Thus, function f changes it sign as $x_I = 2$ which is indeed an inflection point (note that $x = 0$ is not an inflection point by definition). The graph of function f is given in Fig. 9.12.

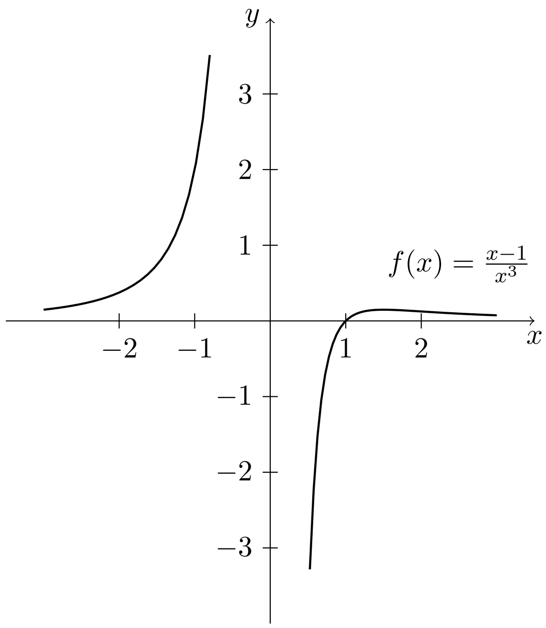


Figure 9.12: Graph of function f in Example 9.19

Example 9.20 *We investigate the function $f : D_f \to \mathbb{R}$ given by*

$$f(x) = e^{-2x^2+1}\,.$$

Since an exponential function is defined for all $x \in \mathbb{R}$, we have $D_f \to \mathbb{R}$. Moreover, any exponential function does not have zeroes and discontinuities. We get

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

Next, we determine stationary points and obtain:

$$f'(x) = -4x \cdot e^{-2x^2+1}\,.$$

Setting $f'(x) = 0$, we get $x_E = 0$ as the only stationary point. Moreover, we have $f'(x) < 0$ for all $x < 0$ and $f'(x) > 0$ for all $x > 0$. Therefore, $x_E = 0$ is a local maximum point with $f(0) = e$, and function f is strictly decreasing on the interval $(-\infty, 0]$ and strictly decreasing on the interval $[0, \infty)$. Next, we determine the second derivative and obtain:

$$f''(x) = (-4x) \cdot e^{-2x^2+1} \cdot (-4x) - 4e^{2x^2+1} = (16x^2 - 4) \cdot e^{-2x^2+1}\,.$$

Setting $f''(x) = 0$, we get $16x^2 - 4 = 0$ which has the two solutions

$$x_{I1} = \frac{1}{2} \qquad \text{and} \qquad x_{I2} = -\frac{1}{2}$$

as candidates for an inflection point. Investigating the sign of the second derivative, we see that $f''(x) > 0$ for $x < -1/2$ and for $x > 1/2$, while $f''(x) < 0$ for $-1/2 < x < 1/2$. Therefore, function f is strictly convex for $x \in (-\infty, -1/2] \cup [-1/2, \infty)$, and function f is strictly concave for $x \in [-1/2, 1/2]$. Since the sign of the second derivative is changing at each of the points $x_{I1} = 1/2$ and $x_{I2} = -1/2$, they are indeed both inflection points for function f. The graph of function f is given in Fig. 9.13.

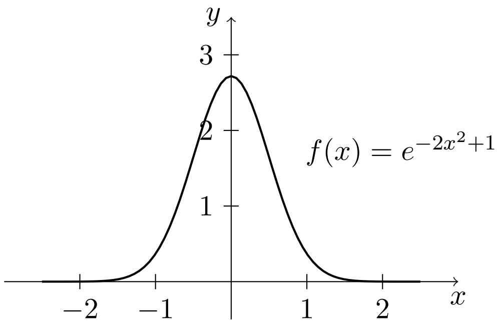


Figure 9.13: Graph of function f in Example 9.20

Example 9.21 *We investigate function $f : D_f \to \mathbb{R}$ given by*

$$f(x) = \sqrt{x+1} + \sqrt{x-1}\,.$$

For finding the domain, we have to take unto account that both terms under the roots must be non-negative. From $x+1 \geq 0$ and $x-1 \geq 0$, we get $x \geq -1$ and $x \geq 1$, which gives the domain $D_f = [1, \infty)$. This function has no zeroes because the range of function f is $R_f = [\sqrt{2}, \infty)$ (observe that for $x \geq 1$ the right square root is at least $\sqrt{2}$ and the first square root is non-negative). Moreover, function f has no discontinuities and

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

To find stationary points, we determine the first derivative:

$$f'(x) = \frac{1}{2\sqrt{x+1}} + \frac{1}{2\sqrt{x-1}} = \frac{\sqrt{x-1}+\sqrt{x+1}}{2\sqrt{x^2-1}}$$

Since both the numerator and the denominator are greater than zero for $x \geq 1$, function f has no stationary points and therefore no local extreme points. Looking for inflection points, we determine the second derivative and obtain:

$$f''(x) = -\frac{1}{4(x+1)^{3/2}} - \frac{1}{4(x-1)^{3/2}} = \frac{-(x-1)\sqrt{x-1} - (x+1)\sqrt{x+1}}{4(x^2-1)\sqrt{x^2+1}}$$

Since in the last term the numerator is negative but the denominator is positive, we obtain $f''(x) < 0$ for all $x > 1$. This means that function f is strictly concave on the domain D_f and that there are no inflection points. The graph of function f is drawn in Fig. 9.14.

9.5.4 Limits

We have already discussed some rules for computing limits of sums, differences, products or quotients of functions in Section 8.1. However, it was necessary that each of the limits exists, i.e., each of the limits was a finite number. However, what happens when we wish to determine the limit of a quotient of two functions, and both limits of the function in the numerator and the function in the denominator tend to ∞ as the variable x tends to a specific value x_0. Similarly, it is possible that both functions in the numerator and in the denominator tend to zero as x approaches some value x_0.

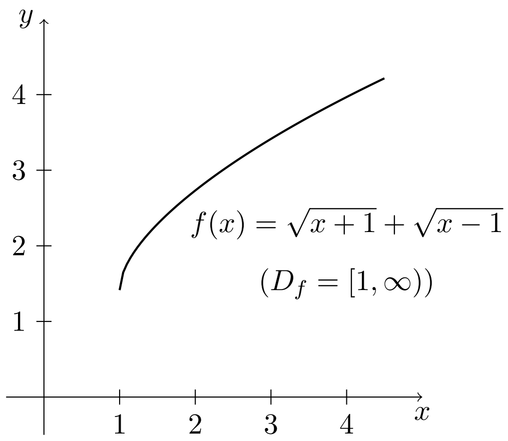


Figure 9.14: Graph of function f in Example 9.21

Indeterminate form:
If in a quotient both the numerator and the denominator tend to zero as x tends to x_0, we call such a limit an **indeterminate form** of type "0/0", and we write

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \text{" } 0/0\text{"}.$$

We only note that in addition to "0/0" and "∞/∞" there exist five other indeterminate forms of the type "$0 \cdot \infty$", "$\infty - \infty$", "0^0", "∞^0" and "1^∞" which all can be reduced to one of the forms considered in the theorem below.

(Bernoulli - de l'Hospital's rule)
Let functions $f : D_f \to \mathbb{R}$ and $g : D_g \to \mathbb{R}$ both tend to zero as x tends to x_0, or f and g both tend to ∞ as x tends to x_0. Moreover, let f and g be continuously differentiable on the open interval $(a, b) \in D_f \cap D_g$ containing x_0 and $g'(x) \neq 0$ for $x \in (a, b)$. Then

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{f'(x)}{g'(x)} = L.$$

Here either the limit exists (i.e., the value L is finite), or the limit does not exist.

Possibly, one needs to apply this rule repeatedly. We consider some examples.

Example 9.22 *We determine the limit*

$$L = \lim_{x \to 1} \frac{2x^2 - 3x + 1}{x^3 + x^2 + 2x - 4}.$$

This is an indeterminate form of the type "0/0". Applying Bernoulli- de l'Hospital's rule, we get

$$L = \lim_{x \to 1} \frac{4x - 3}{3x^2 + 2x + 2} = \frac{1}{7}.$$

Example 9.23 *We calculate the limit*

$$L = \lim_{x \to 0} \frac{e^{2x} + e^{-2x} - 2}{4x^2}\,.$$

This is in indeterminate form of the type "0/0". Using Bernoulli - de l'Hospital's rule, we get

$$L = \lim_{x \to 0} \frac{2e^{2x} - 2e^{-2x}}{8x}\,.$$

This is still an indeterminate form and we apply Bernoulli - de l'Hospital's rule again:

$$L = \lim_{x \to 0} \frac{4e^{2x} + 4e^{-2x}}{8} = 1\,.$$

Example 9.24 *We calculate the limit*

$$L = \lim_{x \to \infty} \frac{\sqrt{4x-4}}{e^{2x}}\,.$$

This is in indeterminate form of the type "∞/∞". Applying Bernoulli-de l'Hospital's rule, we get

$$L = \lim_{x \to \infty} \frac{\frac{4}{2\sqrt{4x-4}}}{2e^{2x}} = \lim_{x \to \infty} \frac{1}{\sqrt{4x-4} \cdot e^{2x}} = 0$$

Example 9.25 *We determine the limit*

$$L = \lim_{x \to 2} \frac{\ln(x-1)^3}{x^2 - 4}\,.$$

This is an indeterminate form of the type "0/0", and we obtain

$$L = \lim_{x \to 2} \frac{\frac{3(x-1)^2}{(x-1)^3}}{2x} = \lim_{x \to 2} \frac{3}{2x(x-1)} = \frac{3}{4}\,.$$

9.6 Extreme Points under Constraints

In this section, we discuss some problems of finding the extreme points of a function $f(x, y)$ subject to a constraint $g(x, y) = b$. Note that the function and constraint depend on two variables. If possible, one eliminates one of the variables in the constraint, e.g. $y = g^*(x)$ and then substitutes this term for y in the function $f(x, y)$ so that a function $f^*(x)$ results depending only on one variable (which we can treat). We illustrate this by the following examples.

Example 9.26 *Given is a rectangle with the fixed circumference C. Among these rectangles, determine the lengths of the sides a and b of that rectangle having the largest area. The area A of a rectangle is given by $A = a \cdot b$. Moreover, we have the constraint $2a + 2b = C$. Solving this constraint e.g. for b, we obtain*

$$b = \frac{1}{2}C - a\,.$$

Using the above term to determine the area A, we get a function depending only on the variable a:

$$A(a) = a \cdot \left(\frac{1}{2}C - a\right) = \frac{1}{2}Ca - a^2\,.$$

Differentiation gives

$$A'(a) = \frac{1}{2}\,C - 2a$$

Setting $A'(a) = 0$, *we get*

$$a = \frac{1}{4}C$$

and then from the constraint

$$b = \frac{1}{2}\,C - \frac{1}{4}\,C = \frac{1}{4}\,C\,.$$

Since $A''(a) = -2 < 0$, *a rectangle with the lengths*

$$a = b = \frac{1}{4}\,C$$

has the maximal area, i.e., the desired rectangle is a square with the area $A = \frac{1}{16}C^2$.

Example 9.27 *We determine the point* (x, y) *of the function* $f : (0, \infty) \to \infty)$ *with* $f(x) = 1/x$ *having the largest distance from the origin* $(0, 0)$. *According to the Pythagorean theorem, the distance* R *of a point* (x, y) *from the origin is given by*

$$R(x, y) = \sqrt{x^2 + y^2}\,.$$

Using $y = 1/x$ *and plugging in* $R(x, y)$, *we get*

$$R(x) = \sqrt{x^2 + \left(\frac{1}{x}\right)^2}\,.$$

Differentiating function R and looking for stationary points, we get

$$R'(x) = \frac{2x - 2 \cdot \frac{1}{x^3}}{2\sqrt{x^2 + \left(\frac{1}{x}\right)^2}} = 0$$

which corresponds to

$$x - \frac{1}{x^3} = 0\,.$$

After multiplying both sides by $x^3 > 0$, we obtain

$$x^4 - 1 = 0$$

which has the solution $x_{E1} = 1$ (note that $x_{E2} = -1 \notin D_f$). Since $R'(x) > 0$ for $x > 1$ and $R'(x) < 0$ for $0 < x < 1$, $x_{E1} = 1$ is a local and global minimum point with the distance $R(1,1) = \sqrt{2}$.

Example 9.28 *Given is an equilateral triangle with the length L of each side. It is desired to put an rectangle with the side lengths x and y into this triangle such that the area of the rectangle becomes maximal (see Fig. 9.15). The area of the rectangle is given by $A = x \cdot y$. Let H be the height of the triangle. According to the intercept theorems, we get the proportion*

$$H : \frac{L}{2} = (H - y) : \frac{x}{2}$$

which, after cross multiply, corresponds to

$$\frac{Hx}{2} = \frac{L}{2} \cdot (H - y)\,.$$

Solving for y, we obtain

$$y = H - \frac{H}{L} \cdot x\,.$$

Plugging the term for y in A, we get

$$A(x) = x \cdot \left(H - \frac{H}{L}\, x \right) = xH - \frac{H}{L}\, x^2$$

Differentiation gives

$$A'(x) = H - 2\,\frac{H}{L}\, x\,.$$

From $A'(x) = 0$, we obtain

$$2\,\frac{H}{L}\, x = H$$

which gives

$$x_E = \frac{L}{2}\,.$$

Using the above term for y and the known fact that in an equilateral triangle, we have $H = L\sqrt{3}/2$, we get

$$y_E = H\left(1 - \frac{x}{L}\right) = \frac{L}{2}\sqrt{3}\left(1 - \frac{1}{2}\right) = \frac{L}{4}\sqrt{3}\,.$$

Since $A''(x) = -2H/L < 0$, the point (x_E, y_E) is a local maximum. It is also a global maximum since $A(x)$ is a quadratic function and the graph is a parabola open from below.

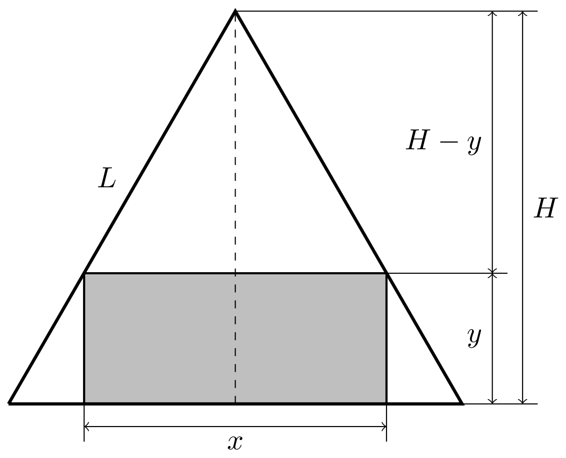


Figure 9.15: Rectangle inside an equilateral triangle

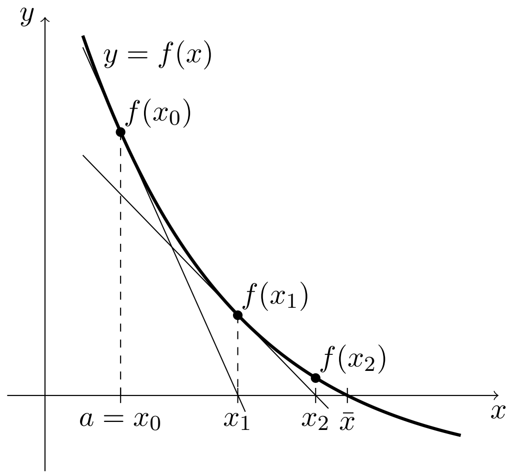


Figure 9.16: Illustration of Newton's method

9.7 Zero Determination by Newton's Method

In mathematics, it is often required to find the zeroes of a function. For instance, if we look for stationary points or inflection points, some function has to be set to be equal to zero and we need to determine the zeroes. However, for many functions this is a hard problem (although one can easily see that for a continuous function on the interval $[a, b]$, there must exist a zero in the open interval (a, b) if two points $x_1, x_2 \in [a, b]$ exist with $f(x_1) < 0$ and $f(x_2) > 0$). As an illustration, we have no formula for finding the zeroes of a polynomial of degree 4.

In this section, we refresh an algorithm for finding zeroes of a function approximately which is known as Newton's method. The idea is to approximate function f about x_0 by its tangent at x_0:

$$f(x) \approx P_1(x) = f(x_0) + f'(x_0) \cdot (x - x_0).$$

Let x_0 be an initial (approximate) value. Then we have $f(x) \approx P_1(x) = 0$, from which we obtain

$$f(x_0) + f'(x_0) \cdot (x - x_0) = 0.$$

We eliminate x and identify it with x_1 which yields:

$$\boxed{x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.}$$

Now we replace the value x_0 on the right-hand side by the new approximate value x_1, and we determine another approximate value x_2. The procedure can be stopped if two successive approximate values x_n and x_{n+1} are sufficiently close to each other which means that the function value is close to zero.

Here we only mention that some assumptions must be satisfied to guarantee that this procedure converges to a zero. We refer the reader e.g. to the textbook by Werner and Sotskov (Mathematics of Economics and Business, Routledge, 2006). Newton's method is illustrated in Fig. 9.16.

Example 9.29 *We look for zeroes of the function* $f : \mathbb{R} \to \mathbb{R}$ *with*

$$f(x) = x + e^{3x} .$$

We obtain

$$f'(x) = 1 + 3e^{3x} > 0 .$$

Since function f *is continuous,* $f(-1) = -1 + e^{-3} < 0$ *and* $f(0) = 1 > 0$*, there does exist a real zero in the interval* $(-1, 0)$*. According to Newton's method, we obtain*

$$x_{n+1} = x_n - \frac{x_n + e^{3x_n}}{1 + 3e^{3x_n}}, \qquad n = 0, 1, 2, \ldots$$

We obtain the values given in Table 9.2 (with four decimal places) and thus, $\overline{x} = -0.3500$ *is the approximate value for the only real zero of function* f*.*

Table 10.1: *Function values for Example 9.29*

n	x_n	$f(x_n)$	$f'(x_n)$
0	0.5	-0.2769	1.6694
1	-0.3341	0.0329	2.1011
2	- 0.3489	0.0086	2.0533
3	- 0.3531	-0.0064	2.0401
4	-0.3500	-0.0001	2.0500
5	- 0.3500		

EXERCISES

9.1 Find the left-side limit and the right-side limit of function $f : D_f \to \mathbb{R}$ as x approaches x_0. Can we conclude from these answers that function f has a limit as x approaches x_0?

(a) $f(x) = \begin{cases} s & \text{for} \quad x \neq x_0 \\ s+1 & \text{for} \quad x = x_0 \end{cases}$; (b) $f(x) = \begin{cases} \sqrt{x} & \text{for} \quad x \leq 4 \\ (x-2)^2 - 2 & \text{for} \quad x > 4 \end{cases}$, $x_0 = 4$;

(c) $f(x) = |x - 2|$, $x_0 = 2$; (d) $f(x) = \begin{cases} x+2 & \text{for} \quad x < 3 \\ 2x - 2 & \text{for} \quad x \geq 3 \end{cases}$, $x_0 = 3$.

9.2 Find the following limits provided that they exist:

(a) $\lim\limits_{x \to -1} \dfrac{x^2 - 2x - 3}{x + 1}$; (b) $\lim\limits_{x \to -1} \dfrac{x^3 - 3x^2}{x + 1}$; (c) $\lim\limits_{x \to -1} \dfrac{x^3 - x}{(x+1)^2}$.

9.3 Check the continuity of function $f : D_f \to \mathbb{R}$ at point $x = x_0$ and in case of a discontinuity, give its type:

(a) $f(x) = \dfrac{\sqrt{x+5} + 1}{x - 4}$, $x_0 = 4$;

(b) $f(x) = |x + 3|$, $x_0 = -3$;

(c) $f(x) = \begin{cases} \ln(x-1) & \text{for} \quad x < 1 \\ 4 - 2x & \text{for} \quad x \geq 1 \end{cases}$, $x_0 = 1$;

9.4 Find the derivative of the following functions $f : D_f \to \mathbb{R}$ given by:

(a) $f(x) = x^2 - 5x - 3\cos x + \sin(\pi/2)$;
(b) $f(x) = (x^3 - x)\cos x$;
(c) $f(x) = \dfrac{x - \sin x}{2 + \cos x}$;
(d) $f(x) = (2x^3 - 3x + \ln x)^4$;
(e) $f(x) = \sin(x^2 + 4x + 1)^3$;
(f) $f(x) = \cos^3(x^2 + 4x + 1)$;
(g) $f(x) = \sqrt{\sin(e^x)}$;
(h) $f(x) = \ln(2x^2 - 1)$;
(i) $f(x) = \dfrac{x + 1}{x^2 + 1}$;

9.5 Find and simplify the derivative of the following functions $f : D_f \to \mathbb{R}$ given by:

(a) $f(x) = \dfrac{\ln^2(3x)}{x}$;
(b) $f(x) = (\tan x + 2)\cos x$;
(c) $f(x) = \ln\sqrt{\dfrac{1 + \cos x}{1 - \cos x}}$;
(d) $f(x) = (1 - \sqrt[4]{x})^3$;
(e) $f(x) = x^2\, e^{x^2}$;
(f) $f(x) = \dfrac{e^x + \ln x}{\sin x}$;
(g) $f(x) = \sqrt{(x^2 + 4x)^3}$.

9.6 Find the third derivatives of the following functions $f : D_f \to \mathbb{R}$ with:

(a) $f(x) = x^2 \cos 2x$;
b) $f(x) = \ln(2x^2)$;
(c) $f(x) = \dfrac{3(x - 1)^3}{(x + 1)^2}$;
d) $f(x) = (x - 1)e^{2x}$.

9.7 Find all local extreme points of the following functions $f : D_f \to \mathbb{R}$ given by:

(a) $f(x) = x^4 - 3x^3 + x^2 - 5$;
(b) $f(x) = 3 - |x + 1|$;
(c) $f(x) = e^{-x^2/2}$;
(d) $f(x) = \dfrac{3x^2}{x - 2}$;
(e) $f(x) = \dfrac{1}{3}\, x^3\, \sqrt{4 - x^2}$;
(f) $f(x) = \dfrac{x - 1}{x^3}$;
(g) $f(x) = (x - 1)\, \ln^2(x - 1)$;
(h) $f(x) = \lg(15 - x^2 - 2x) - 1$.

9.8 Determine the following limits by Bernoulli-de l'Hospital's rule:

(a) $\lim\limits_{x \to 0} \dfrac{\sin 3x}{2x}$;
(b) $\lim\limits_{x \to 1} \dfrac{e^{3(x-1)} - x}{(x^2 - 1)^2}$;
(c) $\lim\limits_{x \to 2} \dfrac{x^3 + 2x^2 - 16}{x^2 + x - 6}$;
(d) $\lim\limits_{x \to \infty} \dfrac{x^2 - \sin x}{2x + \cos x}$;
(e) $\lim\limits_{x \to \infty} \dfrac{3 + 4e^{3x}}{5x + 1000}$;
(f) $\lim\limits_{x \to 0} \dfrac{x}{\tan x}$;
(g) $\lim\limits_{x \to 0} \dfrac{e^{x^2} - 1}{x^2}$;
(h) $\lim\limits_{x \to \pi + 0} \dfrac{\sin x}{\sqrt{x - \pi}}$;

9.9 For the following functions $f : D_f \to \mathbb{R}$, determine and investigate domains, zeros, discontinuities, monotonicity, extreme points and extreme values, convexity and concavity, inflection points and limits as x tends to $\pm\infty$. Graph the functions f with $f(x)$ given as follows:

(a) $f(x) = (x+2)\,(x-2)^3\,;$ (b) $f(x) = \dfrac{x^3}{(x-1)^2}\,;$

(c) $f(x) = \dfrac{3x-4}{(x-2)^3}\,;$ (d) $f(x) = e^{(x-1)^2/2}\,;$

(e) $f(x) = x^2\, e^{-x}\,;$ (f) $f(x) = \sqrt[3]{2x^2 - x^3}\,;$

(g) $f(x) = \frac{1}{2}\,(e^{-x} - e^{-4x})\,;$ (h) $f(x) = \dfrac{10x - 100}{100 + x^2}\,;$

(i) $f(x) = 8x^2 + \frac{1}{x}\,.$

9.10 A 400 m path in a sports stadium consists of two parallel sections of a length L and two added half spheres with a radius R. How must the values L and R be chosen such that the area of the playing field (i.e., $A = L \cdot 2R$) becomes maximal.

9.11 Among all cylinders with a given volume V^* determine that cylinder having the smallest surface.

9.12 Determine the zero $\overline{x}$ of function $f : D_f \to \mathbb{R}$ with

$$f(x) = x + e^{2x} \qquad \text{and} \qquad -1 \le \overline{x} \le 0$$

exactly to four decimal places. Use Newton's method and compare the results with Regula falsi.

9.13 Find one zero $\overline{x}$ of the function $f : D_f \to \mathbb{R}$ with

$$f(x) = x^3 - 3x - 4 \qquad \text{and} \qquad 2 \le \overline{x} \le 3\,.$$

Use Newton's method to find the value with an error less than 10^{-5} and compare the results with Regula falsi.

Chapter 10

Integration

One can find many applications of integration, e.g. in physics, chemistry, engineering or economical sciences. For instance, the area between curves or volumes can be determined by means of integration. The learning objectives of this chapter are to refresh

- indefinite integrals and the basic integration methods and
- definite integrals including their approximate calculation.

10.1 Indefinite Integrals

We start with the introduction of an antiderivative of a function f.

Antiderivative:
A function $F : D_F \to \mathbb{R}$ differentiable on an open interval $I \subseteq D_F$ is called an **antiderivative** of the function $f : D_f \to \mathbb{R}$ on $I \subseteq D_f$ if

$$F'(x) = f(x) \qquad \text{for all } x \in I.$$

Looking for an antiderivative of a function f means that we are looking for a function F whose derivative is equal to the given function f. If function $F : D_F \to \mathbb{R}$ is any antiderivative of function $f : D_f \to \mathbb{R}$, then all the antiderivatives F^* of function f are of the form

$$F^*(x) = F(x) + C,$$

where $C \in \mathbb{R}$ is an arbitrary constant. As a consequence, there exist infinitely many antideratives to a function f. Using the notion of an antiderivative, the indefinite integral can be introduced as follows.

Indefinite Integral:
Let function $F : D_F \to \mathbb{R}$ be an antiderivative of function f. The **indefinite integral** of function f, denoted by $\int f(x)\,dx$, is defined as

$$\int f(x)\,dx = F(x) + C,$$

where $C \in \mathbb{R}$ is any constant.

Function f is also called the **integrand**, and as we see from the above definition, the indefinite integral of function f gives the infinitely many antiderivatives of the integrand f. The notation dx in the integral indicates that x is the **variable of integration**, and C denotes the **integration constant**.

10.2 Basic Integrals

In general, it is a hard problem to find an antiderivative for an arbitrary function, often this is even impossible. In contrast to differentiation, we do not have complete rules but only methods. Integration as the reverse process of differentiation is much more difficult.

Table 10.1 gives some indefinite integrals which follow immediately from the differentiation rules. One can easily check their validity by differentiating the right-hand side, where we must obtain the integrand of the corresponding left-hand side integral.

Table 10.1: Some elementary indefinite integrals

$\int x^n\,dx$	$=$	$\frac{x^{n+1}}{n+1} + C$	$(n \in \mathbb{Z},\ n \neq -1)$
$\int x^r\,dx$	$=$	$\frac{x^{r+1}}{r+1} + C$	$(r \in \mathbb{R},\ r \neq -1,\ x > 0)$
$\int \frac{1}{x}\,dx$	$=$	$\ln\lvert x\rvert + C$	$(x \neq 0)$
$\int e^x\,dx$	$=$	$e^x + C$	
$\int a^x\,dx$	$=$	$\frac{a^x}{\ln a} + C$	$(a > 0,\ a \neq 1)$
$\int \sin x\,dx$	$=$	$-\cos x + C$	
$\int \cos x\,dx$	$=$	$\sin x + C$	
$\int \frac{dx}{\cos^2 x}$	$=$	$\tan x + C$	$\left(x \neq \frac{\pi}{2} + k\pi,\ k \in \mathbb{Z}\right)$
$\int \frac{dx}{\sin^2 x}$	$=$	$-\cot x + C$	$(x \neq k\pi,\ k \in \mathbb{Z})$

Next, we give two basic rules for indefinite integrals concerning the treatment of a constant factor in the integral and the integral of a sum or difference of two functions:

1. $\int C \cdot f(x)\, dx = C \cdot \int f(x)\, dx$ **(constant-factor rule)**;

2. $\int \Big[f(x) \pm g(x)\Big]\, dx = \int f(x)\, dx \pm \int g(x)\, dx$ **(sum-difference rule)**.

Rule (1) says that we can write a constant factor in front of the integral, and rule (2) says that, if the integrand is the sum (or difference) of two functions, we can determine the indefinite integral as the sum (or difference) of the corresponding two integrals. Using the above list of elementary definite integrals and the two rules mentioned above, we are now able to find the indefinite integral for some simple functions.

Example 10.1 *We find the integral*

$$\int \left(4x^3 + 5^x - 3\cos x\right)\, dx.$$

Applying the rules given above, we can split the integral into three integrals and solve each of them by using the list of indefinite integrals given in Table 10.1. We obtain:

$$\begin{aligned}\int \left(4x^3 + 5^x - 3\cos x\right)\, dx &= 4\int x^3 dx + \int 5^x dx - 3\int \cos x\, dx \\ &= 4 \cdot \frac{x^4}{4} + \frac{5^x}{\ln 5} - 3\sin x + C = x^4 + \frac{5^x}{\ln 5} - 3\sin x + C.\end{aligned}$$

Example 10.2 *We consider the integral*

$$\int \left(x^{2.5} + \frac{1}{x^3} + \frac{2}{x} - \frac{1}{3\cos^2 x}\right) dx.$$

Again, we can split the integral and obtain

$$\begin{aligned}\int \left(x^{2.5} + \frac{1}{x^3} + \frac{2}{x} - \frac{1}{3\cos^2 x}\right) dx &= \int x^{2.5}\, dx + \int x^{-3}\, dx + 2\int \frac{1}{x}\, dx - \frac{1}{3}\int \frac{1}{\cos^2 x}\, dx \\ &= \frac{x^{3.5}}{3.5} - 3x^{-2} + 2\ln|x| - \frac{1}{3}\tan x + C \\ &= \frac{2}{7}x^{3.5} - \frac{3}{x^2} + 2\ln|x| - \frac{1}{3}\tan x + C.\end{aligned}$$

Example 10.3 *We wish to find the integral*

$$\int \left(\frac{\sqrt{x} + 3x^3}{\sqrt[4]{x}}\right) dx.$$

Using power rules and the indefinite integral for a power function as integrand (see Table 10.1), we can transform the given integral as follows:

$$\begin{aligned}\int \left(\frac{\sqrt{x} + 3x^3}{\sqrt[4]{x}}\right) dx &= \int \left(\frac{\sqrt{x}}{\sqrt[4]{x}} + 3 \cdot \frac{x^3}{\sqrt[4]{x}}\right) dx = \int \left(\frac{x^{1/2}}{x^{1/4}} + 3 \cdot \frac{x^3}{x^{1/4}}\right) dx \\ &= \int \left(x^{1/4} + 3x^{11/4}\right) dx = \frac{x^{5/4}}{\frac{5}{4}} + 3 \cdot \frac{x^{15/4}}{\frac{15}{4}} + C \\ &= \frac{4}{5} \cdot \left(\sqrt[4]{x^5} + \sqrt[4]{x^{15}}\right) + C.\end{aligned}$$

10.3 Integration Methods

In this section, we discuss two integration methods: **integration by substitution** and **integration by parts**.

10.3.1 Integration by Substitution

Applying this method, one tries to transform the integral in such a way that an antiderivative of the resulting integrand can be easily found. To this end, one introduces a new variable t by means of an appropriate substitution $t = g(x)$ (or accordingly, $x = g^{-1}(t)$). We have the following property.

Integration by substitution:
Suppose that function $f : D_f \to \mathbb{R}$ has an antiderivative F and function $g : D_g \to \mathbb{R}$ with $R_g \subseteq D_f$ is continuously differentiable on an open interval $(a, b) \in D_g$. Then function $z = f \circ g$ exists with $z = (f \circ g)(x) = f(g(x))$ and using the substitution $t = g(x)$, we obtain

$$\int f\left(g(x)\right) \cdot g'(x)\, dx = \int f(t)\, dt = F(t) + C = F(g(x)) + C.$$

The symbol $\circ$ stands for the composition of two functions as introduced in Chapter 8. The above rule states that, if the integrand is the product of a composite function $f \circ g$ and the derivative

g' of the inside function, then the antiderivative is given by the composite function $F \circ g$, where function F is an antiderivative of function f. The validity of the above rule can be easily proven by differentiating the composite function $F \circ g$ (using the chain rule).

We illustrate the above method by the following examples.

Example 10.4 *Let us find*

$$\int 2\,(2x+5)^3\;dx.$$

Using the substitution $t = g(x) = 2x+5$, *we get*

$$\frac{dt}{dx} = 2\,, \qquad i.e: \quad 2\,dx = dt$$

and therefore, the derivative of the inside function is two (note that the outside function is $f(t) = t^3$*). Thus, using the formula for integration by substitution, we obtain*

$$\int 2\,(2x+5)^3\;dx = \int t^3\;dt = \frac{1}{4}\,t^4 + C = \frac{1}{4}\,(2x+5)^4 + C.$$

Example 10.5 *Consider the integral*

$$\int 3x^2 e^{x^3}\,dx.$$

Setting $t = g(x) = x^3$, *we obtain*

$$\frac{dt}{dx} = 3x^2\,, \quad i.e.: \quad 3x^2 dx = dt\,.$$

The application of the above formula for integration yields

$$\int 3x^2 e^{x^3}\;dx = \int e^t dt = e^t + C = e^{x^3} + C.$$

Example 10.6 *Consider the integral*

$$\int \cos x \sin^3 x\,dx\,.$$

Introducing the substitution $t = \sin x$, *we obtain by differentiation*

$$\frac{dt}{dx} = \cos x\,, \quad i.e.: \quad \cos x\,dx = dt\,.$$

Using the above substitution and replacing $\cos x\,dx$ *by* dt, *this yields now*

$$\int \cos x \sin^3 x\,dx = \int t^3 dt = \frac{1}{4}\,t^4 + C = \frac{1}{4}\,\sin^4 x + C\,.$$

Sometimes the integrand is not of the type $f(g(x)) \cdot g'(x)$ (or cannot be easily transformed into this form). However, integration by substitution can often also be applied in a more general form. If we try to use some substitution $t = g(x)$ and if, by using it and the differential $dt = g'(x)\,dx$, it is possible to replace all terms with x and dx in the original integrand by some terms depending on the new variable t and dt, then we can successfully apply integration by substitution (provided one can find an antiderivative of the resulting function depending on t). However, it is important that in the resulting integral only t and dt occur (but no longer x or dx). To illustrate, let us consider the following examples.

Example 10.7 *We find the integral*

$$\int \cos(4x - 2)\, dx\,.$$

Applying the substitution $t = 4x - 2$*, we get by differentiation*

$$\frac{dt}{dx} = 4\,, \qquad i.e.: \quad dx = \frac{dt}{4}\,.$$

Using the above substitution and the term obtained for dx*, we obtain*

$$\int \cos(4x - 2)\, dx = \int \cos t \cdot \frac{dt}{4} = \frac{1}{4}\int \cos t\, dt = \frac{1}{4}\sin t + C = \frac{1}{4}\sin(4x - 2) + C\,.$$

Example 10.8 *Consider the integral*

$$\int \sqrt{2 - 3x}\, dx\,.$$

We use the substitution $t = 2 - 3x$*. Differentiation gives*

$$\frac{dt}{dx} = -3\,, \qquad i.e.: \quad dx = -\frac{dt}{3}\,.$$

Introducing the variable t *and* dt*, we obtain:*

$$\int \sqrt{2 - 3x}\, dx = -\frac{1}{3}\int \sqrt{t}\, dt = -\frac{1}{3}\int t^{1/2} dt = -\frac{1}{3}\cdot\frac{2}{3}\, t^{3/2} + C = -\frac{2}{9}\sqrt{(2 - 3x)^3} + C\,.$$

Example 10.9 *We find the integral*

$$= \int \frac{5}{x} \ln(2x)\, dx\,.$$

Trying the substitution $t = \ln(2x)$, we obtain by differentiation

$$\frac{dt}{dx} = \frac{1}{2x} \cdot 2 = \frac{1}{x}\,, \quad i.e.: \quad \frac{dx}{x} = dt\,.$$

Using the above terms and writing the factor 5 outside the integral, we get

$$\int \frac{5}{x} \ln(2x)\, dx = 5 \int t\, dt = \frac{5}{2} t^2 + C = \frac{5}{2} \ln^2(2x) + C\,.$$

Example 10.10 *Let us consider the integral*

$$\int \frac{\sqrt{x}}{\sqrt{x}+2}\, dx\,.$$

We try the substitution $t = \sqrt{x} + 2$. Differentiation gives

$$\frac{dt}{dx} = \frac{1}{2\sqrt{x}}\,, \quad i.e.: \quad dx = 2\sqrt{x}\, dt = 2(t-2)\, dt\,.$$

Here we used $\sqrt{x} = t - 2$. Now we are able to replace the integrand $f(x)$ and dx and obtain an integral only depending on t and dt:

$$\begin{aligned} \int \frac{\sqrt{x}}{\sqrt{x}+2}\, dx &= \int \frac{t-2}{t} \cdot 2(t-2)\, dt \\ &= 2 \int \frac{t^2 - 4t + 4}{t}\, dt \\ &= 2 \int \left(t - 4 + \frac{4}{t} \right) dt \\ &= 2 \left(\frac{t^2}{2} - 4t + 4 \ln|t| \right) + C\,. \end{aligned}$$

Substituting back, we get the final result:

$$I = (\sqrt{x}+2)^2 - 8(\sqrt{x}+2) + 8\ln(\sqrt{x}+2) + C\,.$$

Note than we can skip the absolute values in the last term since $\sqrt{x} + 2$ is always greater than zero.

Example 10.11 *We consider the integral*

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

We try the substitution $x = \tan t$ and obtain by differentiation

$$\frac{dx}{dt} = 1 + \tan^2 t\,, \quad i.e.: \quad dx = 1 + \tan^2 t\, dt\,.$$

This gives

$$\int \frac{1}{1+x^2}\,dx = \int \frac{1}{1+\tan^2 t}\cdot(1+\tan^2 t)\,dt = \int dt = t + C.$$

From the substitution we get $t = \tan^{-1}\,x$ *and therefore,*

$$\int \frac{1}{1+x^2}\,dx = \tan^{-1}\,x + C\,.$$

We only note that the inverse function $\tan^{-1}\,x$ *of the tangent function is also often denoted as* $\arctan x$.

Example 10.12 *We consider the integral*

$$\int x\,\sqrt{x-1}\,dx\,.$$

We apply the substitution $x = t^2 + 1$ *which corresponds to* $t = \sqrt{x-1}$*. Differentiation gives*

$$\frac{dx}{dt} = 2t\,, \quad \textit{i.e.:} \quad dx = 2t\,dt\,.$$

Now we obtain

$$\begin{aligned}\int x\,\sqrt{x-1}\,dx &= \int (t^2+1)\,\sqrt{t^2+1-1}\cdot 2t\,dt\\ &= 2\int t^2\,(t^2+1)\,dt = 2\int (t^4+t^2)\,dt\\ &= 2\left(\frac{1}{5}\,t^5 + \frac{1}{3}\,t^3\right) + C = \frac{2}{5}(\sqrt{x-1})^5 + \frac{2}{3}(\sqrt{x-1})^3 + C\,.\end{aligned}$$

10.3.2 Integration by Parts

The formula for this method is obtained from the formula for the differentiation of a product of two functions u and v:

$$\Big[u(x)\cdot v(x)\Big]' = u'(x)\cdot v(x) + u(x)\cdot v'(x).$$

Integrating now both sides of the above equation, we obtain the formula for integration by parts as follows:

Integration by parts:
Let $u : D_u \to \mathbb{R}$ and $v : D_v \to \mathbb{R}$ be two functions differentiable on some open interval $I = (a,b) \subseteq D_u \cap D_v$. Then:

$$\int u(x)\cdot v'(x)\,dx = u(x)\cdot v(x) - \int u'(x)\cdot v(x)\,dx.$$

To apply integration by parts, one must be able to find an antiderivative of function v' and an antiderivative of function $u'\cdot v$. If we are looking for an antiderivative of a product of two functions, the successful use of integration by parts depends on an appropriate choice of the functions u and v'. Integration by parts can, for instance, be applied when one of the two factors is a polynomial $P_n(x)$ and the other one is a logarithmic, trigonometric (sine / cosine) or an exponential function. In most cases above, the polynomial P_n is taken as function u which has to be differentiated within the application of integration by parts (as a consequence, the derivative u' is a polynomial of smaller degree). We illustrate integration by parts by some examples.

Example 10.13 *Let us find*

$$\int x^2 \cos x \, dx.$$

We apply integration by parts with

$$u(x) = x^2 \qquad \text{and} \qquad v'(x) = \cos x.$$

Now we obtain

$$u'(x) = 2x \qquad \text{and} \qquad v(x) = \sin x.$$

Hence,

$$\int x^2 \cos x \, dx = x^2 \sin x - 2 \int x \sin x \, dx.$$

Applying now again integration by parts to the integral $\int x \sin x \, dx$ *with*

$$u(x) = x \qquad \text{and} \qquad v'(x) = \sin x,$$

we get

$$u'(x) = 1 \qquad \text{and} \qquad v(x) = -\cos x.$$

This yields

$$\begin{aligned} \int x^2 \cos x \, dx &= x^2 \sin x - 2\Big(- x \cos x - \int - \cos x \, dx \Big) \\ &= x^2 \sin x + 2x \cos x + 2 \sin x + C \\ &= (x^2 + 2) \sin x + 2x \cos x + C. \end{aligned}$$

Notice that integration constant C has to be written as soon as no further integral appears on the right-hand side.

Example 10.14 *Let us now consider the integral*

$$\int (x + 4)e^{2x} \, dx \, .$$

Again, we apply integration by parts and use

$$u(x) = x + 4 \qquad \text{and} \qquad v'(x) = e^{2x}.$$

Differentiation gives

$$u'(x) = 1 \qquad \text{and} \qquad v(x) = \frac{1}{2} e^{2x}.$$

According to the formula of integration by parts, this gives

$$\begin{aligned}\int (x+4)e^{2x}\,dx &= \frac{1}{2}\,e^{2x}(x+4) - \frac{1}{2}\int e^{2x}\,dx \\ &= \frac{1}{2}\,e^{2x}(x+4) - \frac{1}{4}\,e^{2x} + C\,.\end{aligned}$$

Example 10.15 *Let us determine*

$$\int \ln\frac{x}{2}\,dx.$$

Although the integrand $f(x) = \ln(x/2)$ is here not written as a product of two functions, we can nevertheless apply integration by parts by introducing factor one, i.e., we set

$$u(x) = \ln\frac{x}{2} \qquad \text{and} \qquad v'(x) = 1.$$

Then we obtain

$$u'(x) = \frac{2}{x}\cdot\frac{1}{2} = \frac{1}{x} \qquad \text{and} \qquad v(x) = x$$

which leads to

$$\int \ln\frac{x}{2}\,dx = x\ln\frac{x}{2} - \int dx = x\ln\frac{x}{2} - x + C = x\left(\ln\frac{x}{2} - 1\right) + C.$$

Example 10.16 *We find*

$$\int x^2 \ln x\,dx.$$

Setting

$$u(x) = \ln x \qquad \text{and} \qquad v'(x) = x^2\,,$$

we obtain

$$u'(x) = \frac{1}{x} \qquad \text{and} \qquad v(x) = \frac{1}{3}x^3\,.$$

It is worth to note that mostly the polynomial is taken as $u(x)$ (which has to be differentiated) while here it is taken as $v'(x)$ (so that it has to be integrated). Then we get

$$\int x^2\ln x\,dx = \frac{1}{3}x^3\ln x - \int \frac{1}{x}\cdot\frac{1}{3}\,x^3\,dx = \frac{1}{3}x^3\ln x - \frac{1}{3}\int x^2\,dx = \frac{1}{3}\,x^3\ln x - \frac{1}{9}\,x^3 + C\,.$$

Example 10.17 *We determine*

$$\int \cos^2 x\,dx.$$

Here we can set

$$u(x) = \cos x \qquad \text{and} \qquad v'(x) = \cos x.$$

This yields

$$u'(x) = -\sin x \qquad \text{and} \qquad v(x) = \sin x$$

Applying now integration by parts we get

$$\int \cos^2 x\,dx = \sin x\cos x + \int \sin^2 x\,dx.$$

Now one could try to apply again integration by parts to the integral on the right-hand side. However, this does not yield a solution to the problem. Instead of this, we use the equality (see Chapter 5)

$$\sin^2 x = 1 - \cos^2 x,$$

and then the above integral can be rewritten as follows:

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

Adding now $\int \cos^2 x \, dx$ *to both sides, dividing the resulting equation by two and introducing the integration constant* C*, we finally obtain*

$$\int \cos^2 x \, dx = \frac{1}{2}\Big(\sin x \cos x + x \Big) + C.$$

We finish this section with an example, where both integration methods have to applied.

Example 10.18 *We find*

$$\int \cos \sqrt{2x} \, dx.$$

First, we apply integration by substitution using $t = \sqrt{2x}$*. Differentiation gives*

$$\frac{dt}{dx} = \frac{2}{2\sqrt{2x}} = \frac{1}{t}$$

which can be rewritten as

$$dx = t \, dt \, .$$

Replacing now $\sqrt{2x}$ *and* dx *in the integral, we get*

$$\int \cos \sqrt{2x} \, dx = \int t \cos t \, dt.$$

So, the above substitution was useful and the new integral depends only on t *and* dt*. To solve this integral, we now apply integration by parts. We set*

$$u(t) = t \qquad \textit{and} \qquad v'(t) = \cos t$$

which gives

$$u'(t) = 1 \qquad \textit{and} \qquad v(t) = \sin t.$$

Then we get

$$\begin{aligned} \int t \cos t \, dt &= t \sin t - \int \sin t \, dt \\ &= t \sin t + \cos t + C. \end{aligned}$$

After substituting back, we finally obtain

$$\int \cos \sqrt{2x} \, dx = \sqrt{2x} \cdot \sin \sqrt{2x} + \cos \sqrt{2x} + C.$$

10.4 The Definite Integral

To introduce the definite integral, let us start with the following problem: Given is a function f with $y = f(x) \geq 0$ for all $x \in [a, b] \subseteq D_f$. How can we compute the area A under the graph of function f from a to b assuming that function f is continuous on the closed interval $[a, b]$. First, we split the interval $[a, b]$ into n subintervals of equal length. Thus, we choose the points

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b.$$

The idea is now to approximate the real area under the graph of function f by a sum of the areas of rectangles. Denote by u_i (and v_i, respectively) the point of the i-th closed interval $[x_{i-1}, x_i]$, where the function f takes the minimum (maximum) value, i.e., we have

$$f(u_i) = \min\{f(x) \mid x \in [x_{i-1}, x_i]\},$$

$$f(v_i) = \max\{f(x) \mid x \in [x_{i-1}, x_i]\},$$

and let $h = \Delta x_i = x_i - x_{i-1}$. For any interval, we consider the area of the rectangles with the heights $f(u_i)$ and $f(v_i)$, respectively, all having the same width h. Then we can give a lower bound A^n_{min} and an upper bound A^n_{max} on the area A in dependence on the number n of intervals for the area A as follows (see Fig. 10.1):

$$A^n_{min} = \sum_{i=1}^{n} f(u_i) \cdot \Delta x_i,$$

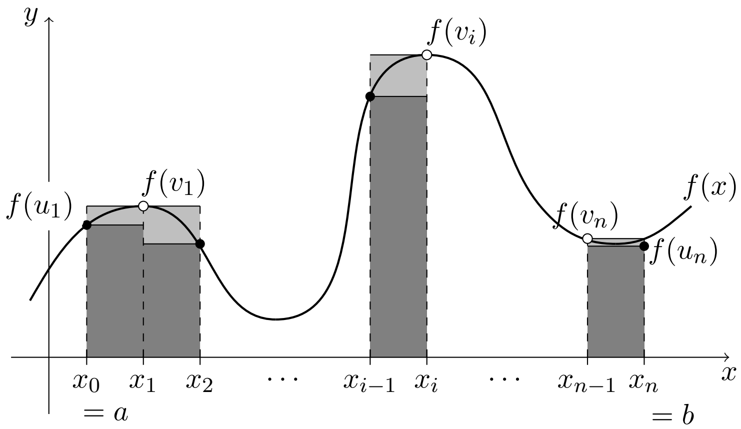


Figure 10.1: The definite integral

$$A^n_{max} = \sum_{i=1}^{n} f(v_i) \cdot \Delta x_i.$$

Since we have $f(l_i) \leq f(x) \leq f(u_i)$ for each $x \in [x_{i-1}, x_i]$ $i \in \{1, 2, \ldots, n\}$, the inequalities $A^n_{min} \leq A \leq A^n_{max}$ hold, i.e., the real area A is between the values A^n_{min} and A^n_{max}. Now one considers the limits of both values A^n_{min} and A^n_{max} as the number of intervals tends to ∞ (or equivalently, the widths h of the intervals tend to zero). If both limits of the sequences $\{A^n_{min}\}$ and $\{A^n_{max}\}$ as n tends to ∞ exist and are equal, we say that the definite integral of function f over the interval $[a, b]$ exists. Formally, we can write the following.

Definite integral:
Let function $f : D_f \to \mathbb{R}$ be continuous on the closed interval $[a, b] \subseteq D_f$. If the limits of the sequences $\{A^n_{min}\}$ and $\{A^n_{max}\}$ as n tends to ∞ exist and coincide, i.e.

$$\lim_{n\to\infty} A^n_{min} = \lim_{n\to\infty} A^n_{max} = I,$$

then I is called **definite integral** of function f over the closed interval $[a, b] \subseteq D_f$.

We write

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

for the definite integral of function f over the interval $[a, b]$. The numbers a and b, respectively, are denoted as **lower and upper limits of integration**. Of course, the above definition is not practical for evaluating definite integrals. For calculating the definite integral, the following formula by Newton and Leibniz is useful.

Evaluation of the definite integral (Newton-Leibniz's formula):
Let function $f : D_f \to \mathbb{R}$ be continuous on the closed interval $[a, b] \subseteq D_f$ and function F be an antiderivative of f. Then the **definite integral** of function f over $[a, b]$ is given by the change in the antiderivative between $x = a$ and $x = b$:

$$\int_a^b f(x)dx = \Big[F(x)\Big]\Big|_a^b = F(b) - F(a).$$

According to Newton-Leibniz's formula, the main difficulty is to find an antiderivative of the integrand f. Therefore, we again have to apply one of the methods discussed before for finding an antiderivative. The following properties of the definite integral are obvious.

Properties of the definite integral:

(1) $\int\limits_a^a f(x)\,dx = 0;$

(2) $\int\limits_a^b f(x)\,dx = -\int\limits_b^a f(x)\,dx;$

(3) $\int\limits_a^b C \cdot f(x)\,dx = C \cdot \int\limits_a^b f(x)\,dx \qquad (C \in \mathbb{R});$

(4) $\int\limits_a^b f(x)\,dx = \int\limits_a^c f(x)\,dx + \int\limits_c^b f(x)\,dx \qquad (a \leq c \leq b);$

When defining the definite integral, we assumed that function f is non-negative. However, for the evaluation of the definite integral this assumption is not required. If function f is non-positive, the value of the definite integral is negative: it corresponds to the area between the x-axis and function f with negative sign. If a function f has one or several zeroes and we wish to compute the area between the function f and the x-axis, we have to split the integral into several sub-integrals: the first integration is from the left boundary point up to the smallest zero, then from the smallest zero to the second smallest zero, and so on.

We continue with some examples for evaluating definite integrals.

Example 10.19 *We evaluate*

$$I = \int_1^9 (x^2 + \sqrt{x})\,dx$$

and obtain immediately by using the list of elementary integrals

$$I = \left[\frac{1}{3}x^3 + \frac{x^{3/2}}{\frac{3}{2}}\right]\Big|_1^9 = \left[\frac{1}{3}x^3 + \frac{2}{3}\sqrt{x^3}\right]\Big|_1^9 = \left(\frac{729}{3} + \frac{2}{3}\cdot 27\right) - \left(\frac{1}{3} + \frac{2}{3}\cdot\sqrt{1}\right) = (243+18) - 1 = 260\,.$$

Example 10.20 *We determine the lower limit of integration a of the following integral I such that it has the value* $112/3$*:*

$$I = \int_a^4 2x^2\,dx = \frac{112}{3}\,.$$

Then we obtain

$$I = \left[\frac{2}{3}\,x^3\right]\Big|_a^4 = \frac{2}{3}\,(64 - a^3) = \frac{112}{3}\,.$$

Solving for a^3, this gives $a^3 = 8$ from which we obtain the unique solution $a = 2$.

Example 10.21 *We evaluate the integral*

$$I = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} sin^2 x\, cos x\,dx\,.$$

We use the substitution $t = \sin x$. *By differentiation, we get*

$$\frac{dt}{dx} = \cos x \qquad i.e.: \quad \cos x dx = dt\,.$$

When writing the definite integral now in dependence on $f(t)$ *and* dt, *one has to transform the limits of integration. Since* $\sin(-\pi/2) = -1$ *and* $\sin(\pi/2) = 1$, *we now have* -1 *and* 1 *as lower and upper limits of integration, respectively. This gives*

$$I = \int_{-1}^{1} t^2\, dt = \left[\frac{1}{3}\, t^3\right]\Big|_{-1}^{1} = \frac{1}{3} - \left(-\frac{1}{3}\right) = \frac{2}{3}\,.$$

Example 10.22 *We evaluate the definite integral*

$$\int\limits_1^e \frac{dx}{x(2+\ln x)}\,.$$

We use the substitution

$$t = 2 + \ln x$$

and obtain by differentiation

$$\frac{dt}{dx} = \frac{1}{x}, \quad i.e.: \quad \frac{dx}{x} = dt.$$

Using the two previous representations, we get an integral only depending on t *and* dt:

$$\int\limits_1^e \frac{dx}{x(2+\ln x)} = \int\limits_{t(1)}^{t(e)} \frac{dt}{t}\,.$$

Now we obtain

$$\int\limits_{t(1)}^{t(e)} \frac{dt}{t} = \Big[\ln|t|\Big]\Big|_{t(1)}^{t(e)} = \Big[\ln|2+\ln x|\Big]\Big|_1^e = \ln(2+\ln e) - \ln(2+\ln 1) = \ln 3 - \ln 2 = \ln\frac{3}{2}.$$

In the above computations, we did not transform the limits of the definite integral into the corresponding t*-values but we have used the above substitution again (in the opposite direction) to the antiderivative in terms of* x. *In this case, it is not necessary to give the limits of integration with respect to the variable* t. *Of course, we can also transform the limits of integration (in this case, we get* $t(1) = 2 + \ln 1 = 2$ *and* $t(e) = 2 + \ln e = 3$*) and insert the obtained values directly into the obtained antiderivative* $\ln|t|$.

Example 10.23 *We evaluate the definite integral*

$$I = \int_0^{\frac{\pi}{2}} e^x \cos x\, dx\,.$$

First, we consider the indefinite integral and find an antiderivative using integration by parts. Setting

$$u(x) = e^x \qquad and \qquad v'(x) = \cos x\,,$$

we obtain

$$u'(x) = e^x \quad \text{and} \quad v(x) = \sin x\,.$$

According to the formula of integration by parts, this gives

$$\int e^x \cos x\, dx = e^x \sin x - \int e^x \sin x\, dx\,.$$

Using again integration by parts for the integral $\int e^x \sin x\, dx$ *with*

$$u(x) = e^x \quad \text{and} \quad v'(x) = \sin x\,,$$

we obtain

$$u'(x) = e^x \quad \text{and} \quad v(x) = -\cos x$$

and then

$$\int e^x \cos x\, dx = e^x \sin x - \left(-e^x \cos x + \int e^x \cos x\, dx\right)$$

which, after putting the integrals on one side and introducing the integration constant, finally gives

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

Evaluating now the definitive integral, we obtain

$$I = \frac{1}{2}\left[e^x(\sin x + \cos x)\right]\Big|_0^{\frac{\pi}{2}} = \frac{1}{2}\left[e^{\frac{\pi}{2}}(1+0) - e^0(0+1)\right] = \frac{1}{2}\left(e^{\frac{\pi}{2}} - 1\right)\,.$$

Example 10.24 *We determine the area bounded by the curve* $f(x) = -x^2 + 5x - 4$ *and the* x*-axis. First we determine the zeroes of function* f *and obtain (after multiplying* f *by* -1 *and applying the solution formula for a quadratic equation):*

$$x_1 = \frac{5}{2} - \sqrt{\frac{25}{4} - 4} = 1 \quad \text{and} \quad x_2 = \frac{5}{2} + \sqrt{\frac{25}{4} - 4} = 4\,.$$

Note that $f(x) \geq 0$ *for all* $x \in [1, 4]$. *Thus, the required area* A *is obtained as follows:*

$$A = \int_1^4 (-x^2 + 5x - 4)\, dx = \left[-\frac{1}{3}x^3 + \frac{5}{2}x^2 - 4x\right]\Big|_1^4 = \left(-\frac{64}{3} + 40 - 16\right) - \left(-\frac{1}{3} + \frac{5}{2} - 4\right) = \frac{9}{2}\,.$$

Example 10.25 *We want to compute the area enclosed by the graphs of the two functions* $f : \mathbb{R} \to \mathbb{R}$ *and* $g : \mathbb{R} \to \mathbb{R}$ *given by*

$$f(x) = x^2 - 16 \quad \text{and} \quad g(x) = -4x - x^2.$$

We first determine the points of intercept of both functions and obtain from

$$x^2 - 16 = -4x - x^2$$

the quadratic equation

$$2x^2 + 4x - 16 = 0$$

or, after dividing by 2,

$$x^2 + 2x - 8 = 0$$

This quadratic equation has the two real solutions

$$x_1 = -1 - \sqrt{1+8} = -4 \qquad \textit{and} \qquad x_2 = -1 + \sqrt{1+8} = 2.$$

The graphs of both functions are parabolas which intersect only in these two points $(-4, 0)$ *and* $(2, -12)$. *To compute the enclosed area* A, *we therefore have to evaluate the definite integral*

$$A = \int_{-2}^{4} \Big[g(x) - f(x)\Big]\, dx\,.$$

Note that the graph of function f *is a parabola open from above while the graph of function* g *is a parabola open from below. Therefore, the we have* $g(x) \geq f(x)$ *for all* $x \in [-2, 4]$ *and thus, the integrand is the difference* $g(x) - f(x)$*:*

$$\begin{aligned}
A &= \int_{-4}^{2} \Big[(-4x - x^2) - (x^2 - 16)\Big]\, dx = \int_{-4}^{2} (-2x^2 - 4x + 16)\, dx \\
&= 2\int_{-4}^{2} (-x^2 - 2x + 8)\, dx = 2\left(-\frac{x^3}{3} - x^2 + 8x\right)\Big|_{-4}^{2} \\
&= 2\left[\left(-\frac{8}{3} - 4 + 16\right) - \left(\frac{64}{3} - 16 - 32\right)\right] = 2 \cdot (60 - 24) = 2 \cdot 36 = 72\,.
\end{aligned}$$

Thus, the area enclosed by the graphs of the given functions is equal to 216/3 squared units.

Example 10.26 *We wish to determine the area* A *enclosed by the function* f *with* $f(x) = \sin x$ *with the* x*-axis between* $a = -\pi/2$ *and* $b = \pi/2$. *Since the function* f *is non-positive in the closed interval* $[-\pi/2, 0]$ *and non-negative in the closed interval* $[0, \pi/2]$ *(note that* $x = 0$ *is a zero of the sine function), we have to split the interval into two sub-intervals (from* $-\pi/2$ *to 0 and from 0 to* $\pi/2$ *and to find the two integrals separately. We obtain the area* A *as follows:*

$$\begin{aligned}
A &= -\int_{-\frac{\pi}{2}}^{0} \sin x\, dx + \int_{0}^{\frac{\pi}{2}} \sin x\, dx \\
&= \cos x\Big|_{-\frac{\pi}{2}}^{0} - \cos x\Big|_{0}^{\frac{\pi}{2}} \\
&= (1 - 0) + (1 - 0) = 2.
\end{aligned}$$

If we would not take into account that function f *has a zero and would integrate in one integral from* $-\pi/2$ *until* $\pi/2$, *the resulting value of the definite integral would be equal to zero (since the area between the* x*-axis and the sine function in the first subinterval is equal to the area between the sine function and the* x*-axis in the second subinterval, however, both the signs would be opposite).*

10.5 Approximation of Definite Integrals

Often, one cannot evaluate a definite integral due to several reasons. For some functions, there does not exist an antiderivative that can be determined analytically. As an example, we can mention here e.g. the function f with $f(x) = e^{-x^2}$, which is often applied in probability theory and statistics, or function g with $g(x) = (\sin x)/x$. Moreover, a function f may be given only as a set of points (x, y) which have been experimentally determined.

In such cases, we must be satisfied with determining the definite integral approximately by applying numerical methods. Similar to the definition of the definite integral, approximate methods divide the closed interval $[a, b]$ into n subintervals of equal width $h = (b-a)/n$, and so we get again the subintervals $[x_0, x_1], [x_1, x_2], \ldots, [x_{n-1}, x_n]$, where $a = x_0$ and $b = x_n$. Within each interval, we now replace function f by some other function which is 'close to the original one' and for which the integration can be easily performed. In all approximation methods, one replaces the function f by a polynomial of a small degree (mostly not greater than two). Here we discuss only two methods, namely an **approximation by trapeziums** and **Simpson's rule**.

Approximation by trapeziums

In this case, in each closed interval $[x_{i-1}, x_i]$, the function f is replaced by a line segment (i.e., a linear function) through the points $(x_{i-1}, f(x_{i-1}))$ and $(x_i, f(x_i))$. This gives the area

$$\frac{b-a}{n} \cdot \left[\frac{1}{2} f(x_{i-1}) + \frac{1}{2} f(x_i) \right].$$

If $n = 1$, this gives an approximation by one trapezium:

$$\int_a^b f(x)\,dx \approx (b-a) \cdot \frac{f(a) + f(b)}{2}.$$

The quality of the approximation improves when we use $n > 1$ which is also known as the **composite trapezoidal rule**. In this way, we get the following approximation formula for the definite integral:

$$\int_a^b f(x)\,dx \approx \frac{b-a}{n} \cdot \left[\frac{f(a) + f(b)}{2} + f(x_1) + f(x_2) + \ldots + f(x_{n-1}) \right].$$

Thus, by the above formula, we approximate the definite integral by the sum of the areas of n trapeziums, each with width $(b-a)/n$. Note also that any function value $f(x_1), f(x_2), \ldots, f(x_{n-1})$ is taken with factor $1/2$ as the right boundary point of an interval and with factor $1/2$ as a left boundary point of the succeeding interval. Approximation by trapeziums is illustrated in Fig. 10.2.

Simpson's rule

Here we consider the special case when the closed interval $[a, b]$ is divided only into $n = 2$ subintervals $[a, (a+b)/2]$ and $[(a+b)/2, b]$ of equal length. Now we approximate function f by a quadratic function which is uniquely defined by three points: $\big(a, f(a)\big)$, $\big((a+b)/2, f((a+b)/2)\big)$ and $\big(b, f(b)\big)$. This leads to Simpson's formula:

$$\int_a^b f(x)\,dx \approx \frac{b-a}{6} \cdot \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right].$$

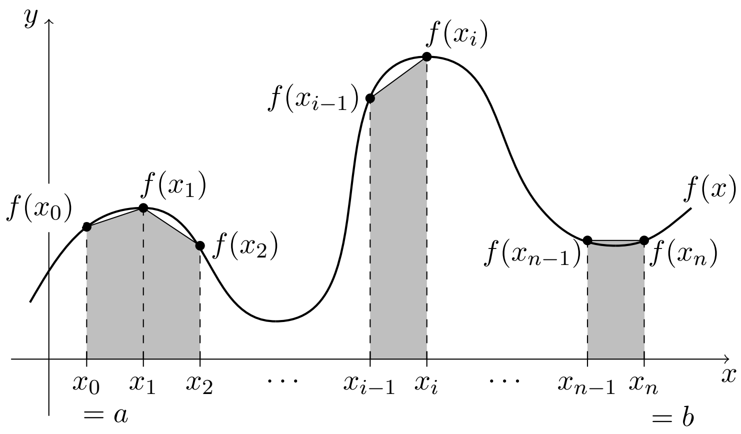


Figure 10.2: Approximation of definite integrals by trapeziums

We note that this formula is also known as **Kepler's rule**.

Example 10.27 *We illustrate the composite trapezoidal rule with* $n = 6$ *subintervals and Simpson's rule by evaluating numerically the integral*

$$I = \int_1^4 \frac{e^{2x}}{x}\, dx,$$

We get the function values given in Table 10.2.

Table 10.2: *Function values for Example 10.27*

i	x_i	$f(x_i)$
0	1	7.3891
1	1.5	13.3904
2	2	27.2991
3	2.5	59.3653
4	3	134.4763
5	3.5	313.3238
6	4	745.2395

Using the composite trapezoidal rule with $n = 6$*, we obtain*

$$\begin{aligned}\int_1^4 f(x)\, dx &\approx \frac{b-a}{n} \cdot \left[\frac{f(x_0) + f(x_6)}{2} + f(x_1) + f(x_2) + \ldots + f(x_5)\right] \\ &\approx \frac{3}{6} \cdot 1300.4835 = 650.2418.\end{aligned}$$

Using Simpson's rule with $a = 1$ *and* $b = 4$*, we obtain*

$$\begin{aligned}\int_1^4 f(x)\, dx &\approx \frac{b-a}{6} \cdot \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right] = \frac{1}{2} \cdot \left[f(1) + 4f(2.5) + f(4)\right] \\ &= \frac{1}{2} \cdot 990.0898 = 445.0449.\end{aligned}$$

We can see that the approximations still differ rather much. This is caused by the use of only two subintervals in the case of Simpson's rule: One can refine the calculations by considering subsequently an even number of intervals and applying Simpson's rule to the first and second intervals, then to the third and fourth intervals and so on, which is known as **composite Simpson's rule**.

EXERCISES

10.1 Find the following indefinite integrals

(a) $\int \left(3x^2 + \frac{1}{2x}\right) dx$; (b) $\int \frac{x^4 - \sqrt[3]{x}}{\sqrt[5]{x}} dx$; (c) $\int \left(2^x + x^{-3.5}\right) dx$.

10.2 Use the substitution rule for finding the following indefinite integrals:

(a) $\int \frac{\ln x}{3x} dx$; (b) $\int e^{\cos x} \sin x \, dx$; (c) $\int \frac{2}{3x-1} dx$;

(d) $\int \frac{dx}{e^{4+3x}}$; (e) $\int \sqrt{2+5x} \, dx$; (f) $\int \frac{x^3 \, dx}{\sqrt{1+x^2}}$;

(g) $\int \frac{x^3}{1+x^2} dx$; (h) $\int x \sqrt{2-x^2} \, dx$; (i) $\int \frac{\cos^3 x}{\sin^2 x} dx$;

(j) $\int \sin x \cos x \, dx$; (k) $\int \frac{e^x - 1}{e^x + 1} dx$.

10.3 Use integration by parts to find the following indefinite integrals:

(a) $\int x e^{4x} \, dx$; (b) $\int e^x \sin x \, dx$; (c) $\int \frac{x}{\cos^2 x} dx$;

(d) $\int \sin^2 x \, dx$; (e) $\int x \ln x \, dx$; (f) $\int x^2 \sin x \, dx$.

10.4 Evaluate the following definite integrals:

(a) $\int_{-2}^{4} (x^2 + 4x + 1) \, dx$; (b) $\int_{1}^{e} \frac{dx}{x}$; (c) $\int_{0}^{\frac{\pi}{2}} \sin^3 x \, dx$;

(d) $\int_{0}^{4} \frac{x \, dx}{\sqrt{x^2+4}}$; (e) $\int_{1}^{3} x \ln(x^2) \, dx$; (f) $\int_{0}^{\pi} \cos^2 x \sin x \, dx$;

(g) $\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \frac{2 - 3\cos^2 x}{\cos^2 x} dx$.

10.5 (a) Evaluate

$$\int_{0}^{2\pi} \cos x \, dx$$

and compute the area enclosed by function $f : \mathbb{R} \to \mathbb{R}$ with $f(x) = \cos x$ and the x-axis within the interval $[0, 2\pi]$.

(b) Compute the area of the triangle formed by the x-axis and the functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = x + 2 \qquad \text{and} \qquad g(x) = -\frac{x}{2} + 8\,.$$

(c) Compute the area enclosed by the two functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = -x^2 - 4x - 1 \qquad \text{and} \qquad g(x) = 4 - 10x.$$

10.6 Determine the following definite integral numerically:

$$\int_0^4 \frac{dx}{2 + 3x^2}.$$

(a) Use approximation by trapeziums with $n = 4$.

(b) Use approximation by trapeziums with $n = 8$.

(c) Use Simpson's rule.

Chapter 11

Vectors

Vectors are used for denoting points in the plane or space (e.g. n-tuples of numbers) and they describe (parallel) displacements. They are often applied for denoting e.g. physical variables. The learning objectives of this chapter are

- to review the basics about vectors with at most three components and
- to summarize the major operations with vectors.

11.1 Definition and Representation of Vectors

Vector:
A 3-dimensional **vector**

$$\mathbf{a} = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}$$

is an ordered sequence of three real numbers a_x, a_y, a_z. The numbers a_x, a_y, a_z are called the **components** (or **coordinates**) of the vector $\mathbf{a}$.

a_x, a_y, a_z are also called the **scalar components** of vector $\mathbf{a}$. If we consider only vectors with the z-coordinate equal to zero, we have a 2-dimensional vector, and we write

$$\mathbf{a} = \begin{pmatrix} a_x \\ a_y \end{pmatrix}.$$

In this chapter, we consider exclusively 2- or 3-dimensional vectors. In the following, we always assume that the components of a vector are indexed by x, y and possibly z. We describe all operations for 3-dimensional vectors, but they can also be applied if all vectors are 2-dimensional.

We can graph a vector as an arrow in the 3-dimensional space which can be interpreted as a **displacement** of a starting point P resulting in a terminal point Q. For instance, if point P has the coordinates (p_1, p_2, p_3) and point Q has the coordinates (q_1, q_2, q_3), then

$$\mathbf{a} = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} = \overrightarrow{PQ} = \begin{pmatrix} q_1 - p_1 \\ q_2 - p_2 \\ q_3 - p_3 \end{pmatrix}.$$

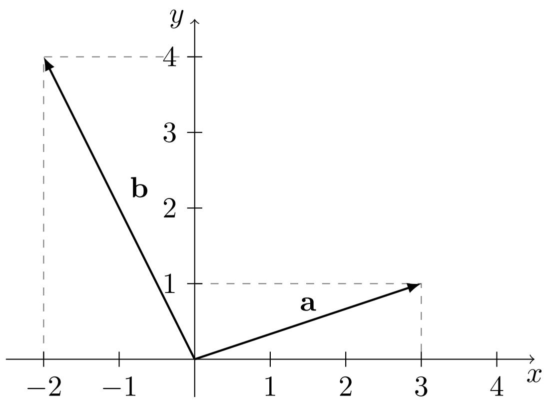


Figure 11.1: 2-dimensional vectors in the plane

It is often assumed that the starting point P is the origin of the coordinate system. In this case, the components of vector $\mathbf{a}$ are simply the coordinates of point Q and, therefore, a row vector $\mathbf{a}^T$ can be interpreted as a **point** (i.e., a **location**) in the 3-dimensional space. In the case of 2-dimensional vectors, we can illustrate them in the plane, e.g., the vectors

$$\mathbf{a} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \qquad \text{and} \qquad \mathbf{b} = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$$

are illustrated in Fig. 11.1. Finding the terminal point of vector $\mathbf{a}$ means that we are going three units to the right from the origin and one unit up. Similarly, to find the terminal point of $\mathbf{b}$, we are going two units to the left from the origin and four units up. We can see that a vector is characterized by the **length**, the **direction** and the **orientation**. The length is obtained by the length of the arrow representing the vector, the direction is determined by its location in the space, and the orientation is characterized by the direction of its arrowhead.

The vectors $\mathbf{a}$ and $\mathbf{b}$ are said to be **equal** if all their corresponding components are equal, i.e., we have

$$a_x = b_x, \qquad a_y = b_y, \qquad a_z = b_z\,.$$

We have $\mathbf{a} \le \mathbf{b}$ if the inequalities $a_x \le b_x$, $a_y \le b_y$, $a_z \le b_z$ hold for the components. In a similar way, we can introduce the inequalities $\mathbf{a} \ge \mathbf{b}$, $\mathbf{a} < \mathbf{b}$ and $\mathbf{a} > \mathbf{b}$.

Finally, we introduce some special vectors. A **unit vector** is a vector with a length equal to one. The special unit vectors pointing in the direction of one of the axes are denoted as $\mathbf{e^x}, \mathbf{e^y}$ and $\mathbf{e^z}$:

$$\mathbf{e^x} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{e^y} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \mathbf{e^z} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

The **zero vector** is a vector containing only zeroes as components:

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

11.2 Operations with Vectors

We start with the operations of adding and subtracting two vectors of the same dimension.

Sum of two vectors:
The **sum** of the two vectors $\mathbf{a}$ and $\mathbf{b}$ is the vector $\mathbf{a}+\mathbf{b}$ obtained by adding each component of vector $\mathbf{a}$ to the corresponding (i.e., at the same position) component of vector $\mathbf{b}$:

$$\mathbf{a}+\mathbf{b} = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} + \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} = \begin{pmatrix} a_x + b_x \\ a_y + b_y \\ a_z + b_z \end{pmatrix}$$

Difference of two vectors:
The **difference** between the vectors $\mathbf{a}$ and $\mathbf{b}$ is the vector $\mathbf{a}-\mathbf{b}$ defined by

$$\mathbf{a}-\mathbf{b} = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} - \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} = \begin{pmatrix} a_x - b_x \\ a_y - b_y \\ a_z - b_z \end{pmatrix}$$

Thus, the difference vector is obtained by subtracting the components of the vector $\mathbf{b}$ from the

corresponding components of the vector **a**.

If we multiply the vector **a** by a real number λ, also denoted as a scalar, the vector $\mathbf{b} = \lambda\mathbf{a}$ whose components are λ times the corresponding components of vector **a** is obtained :

$$\mathbf{b} = \lambda\,\mathbf{a} = \lambda \cdot \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} = \begin{pmatrix} \lambda a_x \\ \lambda a_y \\ \lambda a_z \end{pmatrix}.$$

If $\lambda > 0$, the vector **b** has the same orientation as the vector **a**. If $\lambda < 0$, the vector **b** has the opposite orientation than the vector **a**. The operation of multiplying a vector by a scalar is denoted as **scalar multiplication**.

Example 11.1 *Let*

$$\mathbf{a} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} \qquad \textit{and} \qquad \mathbf{b} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}.$$

Then we obtain

$$\mathbf{a}+\mathbf{b} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 6 \\ 7 \end{pmatrix} \qquad \textit{and} \qquad \mathbf{a}-\mathbf{b} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} - \begin{pmatrix} 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}.$$

Moreover, we obtain e.g.

$$2\,\mathbf{a} = 2 \cdot \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \end{pmatrix} \qquad \textit{and} \qquad (-1)\,\mathbf{b} = (-1) \cdot \begin{pmatrix} 2 \\ 5 \end{pmatrix} = \begin{pmatrix} -2 \\ -5 \end{pmatrix}.$$

The sum and difference of the two vectors as well as the scalar multiplication are geometrically illustrated in Fig. 11.2. The sum of the two vectors **a** *and* **b** *is the vector obtained when adding vector* **b** *to the terminal point of vector* **a**. *The resulting vector from the origin to the terminal point of vector* **b** *gives the sum* $\mathbf{a}+\mathbf{b}$. *Accordingly, the difference* $\mathbf{a}-\mathbf{b}$ *is obtained by adding to the terminal point of vector* **a** *the vector* $(-1)\cdot\mathbf{b}$, *i.e., the vector* **b** *with opposite orientation. If we multiply a vector by a scalar, the direction of the vector does not change (but the orientation may change).*

Example 11.2 *Let*

$$\mathbf{a}=\begin{pmatrix}2\\-1\\0\end{pmatrix} \quad \textit{and} \quad \mathbf{b}=\begin{pmatrix}4\\1\\-3\end{pmatrix}.$$

Then we obtain

$$\mathbf{a}+\mathbf{b}=\begin{pmatrix}2\\-1\\0\end{pmatrix}+\begin{pmatrix}4\\1\\-3\end{pmatrix}=\begin{pmatrix}6\\0\\-3\end{pmatrix} \quad \textit{and} \quad \mathbf{a}-\mathbf{b}=\begin{pmatrix}2\\-1\\0\end{pmatrix}-\begin{pmatrix}4\\1\\-3\end{pmatrix}=\begin{pmatrix}-2\\-2\\3\end{pmatrix}.$$

Moreover, we obtain e.g.

$$(-5)\,\mathbf{a}+3\,\mathbf{b}=(-5)\cdot\begin{pmatrix}2\\-1\\0\end{pmatrix}+3\begin{pmatrix}4\\1\\-3\end{pmatrix}=\begin{pmatrix}(-5)\cdot 2+3\cdot 4\\(-5)\cdot(-1)+3\cdot 1\\(-5)\cdot 0+3\cdot(-3)\end{pmatrix}=\begin{pmatrix}2\\8\\-9\end{pmatrix}.$$

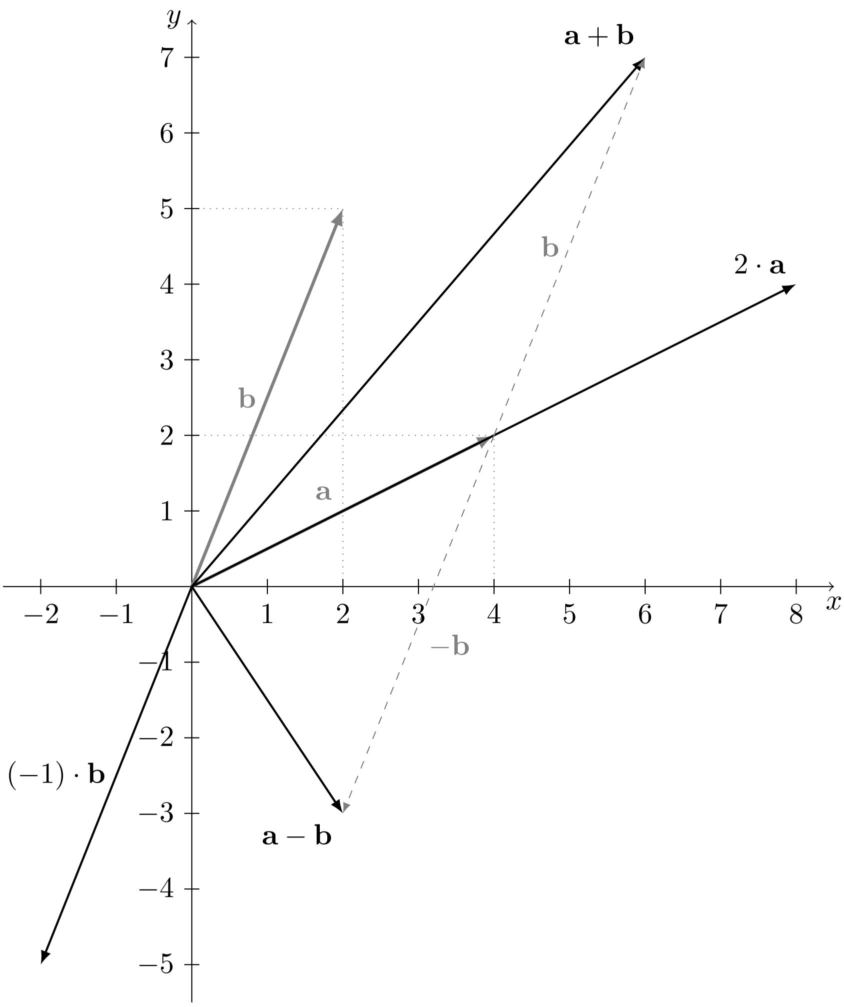


Figure 11.2: Vector operations: Sum, difference and scalar multiplication

The operation of adding two vectors is

- **commutative**, i.e., we have $\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$, and
- **associative**, i.e., we have $(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c})$.

There exist further rules for the sum of two vectors and the multiplication by a number so that one can interpret a vector in mathematics in a more general form as an element of a so-called **vector space**.

Scalar product:
The **scalar product** of the vectors **a** and **b** is defined as follows:

$$\mathbf{a}\cdot\mathbf{b}=\begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}\cdot\begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix}=a_xb_x+a_yb_y+a_zb_z.$$

Note also that the scalar product of two vectors is not a vector, but a **number** (i.e., a **scalar**) and that $\mathbf{a}\cdot\mathbf{b}$ is only defined if **a** and **b** are both of the **same dimension**, i.e., we cannot determine the scalar product of a 2- and a 3-dimensional vector. The commutative and distributive laws are valid for the scalar product, i.e., we have

$$\mathbf{a}\cdot\mathbf{b}=\mathbf{b}\cdot\mathbf{a} \qquad \text{and} \qquad \mathbf{a}\cdot(\mathbf{b}+\mathbf{c})=\mathbf{a}\cdot\mathbf{b}+\mathbf{a}\cdot\mathbf{c}$$

for all vectors $\mathbf{a},\mathbf{b},\mathbf{c}$ of the same dimension. It is worth noting that the associative law does not necessarily hold for the scalar product, i.e., in general we have

$$\mathbf{a}\cdot(\mathbf{b}\cdot\mathbf{c})\neq(\mathbf{a}\cdot\mathbf{b})\cdot\mathbf{c}.$$

Notice that the vector on the left-hand side is parallel to vector **a**, while the vector on the right-hand side is parallel to vector **c**. We emphasize that there is **no** operation that divides a vector by another one. However, we can divide a vector by a non-zero number λ, which corresponds to a scalar multiplication with

$$\mu=\frac{1}{\lambda}\,.$$

Example 11.3 *Let the vectors*

$$\mathbf{a}=\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \qquad \textit{and} \qquad \mathbf{b}=\begin{pmatrix} 4 \\ -1 \\ -6 \end{pmatrix}$$

be given. Then we obtain the scalar product $\mathbf{a}\cdot\mathbf{b}$ *as follows:*

$$\mathbf{a}\cdot\mathbf{b}=2\cdot4+3\cdot(-1)+(-1)\cdot(-6)=8-3+6=11.$$

Example 11.4 *For the scalar products composed of two of the three unit vectors parallel to the coordinate axes we obtain:*

$$\mathbf{e^x}\cdot\mathbf{e^x}=1\,; \qquad \mathbf{e^y}\cdot\mathbf{e^y}=1\,; \qquad \mathbf{e^z}\cdot\mathbf{e^z}=1\,;$$
$$\mathbf{e^x}\cdot\mathbf{e^y}=0\,; \qquad \mathbf{e^x}\cdot\mathbf{e^z}=0\,, \qquad \mathbf{e^y}\cdot\mathbf{e^z}=0\,.$$

Length of a vector:
The length of vector **a**, denoted by $|\mathbf{a}|$, is defined as

$$|\mathbf{a}|=\sqrt{a_x^2+a_y^2+a_z^2}.$$

A vector with length one is denoted as **unit vector** (remind that we have already introduced the specific unit vectors $\mathbf{e^x},\mathbf{e^y},\mathbf{e^z}$ which have obviously the length one). Each nonzero vector **a** can be written as the product of its length $|\mathbf{a}|$ and a unit vector $\mathbf{e^{(a)}}$ pointing in the same direction as the vector **a** itself, i.e.,

$$\mathbf{a}=|\mathbf{a}|\cdot\mathbf{e^{(a)}}.$$

By means of the unit vectors $\mathbf{e^x},\mathbf{e^y}$ and $\mathbf{e^z}$, we can present any vector in **vectorial components** as follows:

$$\mathbf{a}=a_x\mathbf{e^x}+a_y\mathbf{e^y}+a_z\mathbf{e^z}\,,$$

i.e., formally as the sum of three vectors.

Example 11.5 *Let the vector*

$$\mathbf{a}=\begin{pmatrix}-2\\3\\6\end{pmatrix}$$

be given. We are looking for a unit vector pointing in the same direction as the vector **a**. *Using*

$$|\mathbf{a}|=\sqrt{(-2)^2+3^2+6^2}=\sqrt{49}=7,$$

we find the corresponding unit vector

$$\mathbf{e^{(a)}}=\frac{1}{|\mathbf{a}|}\cdot\mathbf{a}=\frac{1}{7}\cdot\begin{pmatrix}-2\\3\\6\end{pmatrix}=\begin{pmatrix}-2/7\\3/7\\6/7\end{pmatrix}.$$

The distance between the two vectors **a** and **b** is obtained as follows:

$$|\mathbf{a}-\mathbf{b}|=\sqrt{(a_x-b_x)^2+(a_y-b_y)^2+(a_z-b_z)^2}\,.$$

Example 11.6 *Let the vectors*

$$\mathbf{a}=\begin{pmatrix}3\\2\\-3\end{pmatrix} \qquad \textit{and} \qquad \mathbf{b}=\begin{pmatrix}-1\\1\\5\end{pmatrix}$$

be given. The distance between both vectors is obtained as

$$|\mathbf{a}-\mathbf{b}|=\sqrt{[3-(-1)]^2+(2-1)^2+(-3-5)^2}=\sqrt{16+1+64}=\sqrt{81}=9.$$

Next, we present an alternative possibility to calculate the scalar product:

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}| \cdot |\mathbf{b}| \cdot \cos(\angle(\mathbf{a}, \mathbf{b})); \tag{11.1}$$

In Equation (11.1), $\cos(\angle(\mathbf{a}, \mathbf{b}))$ denotes the cosine value of the angle between the oriented vectors $\mathbf{a}$ and $\mathbf{b}$.

Example 11.7 *Using Equation (11.1) above and the earlier definition of the scalar product, one can easily determine the angle between two vectors* $\mathbf{a}$ *and* $\mathbf{b}$ *of the same dimension. Let*

$$\mathbf{a} = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} \qquad \textit{and} \qquad \mathbf{b} = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}.$$

Then we obtain

$$\begin{aligned} \cos(\angle(\mathbf{a}, \mathbf{b})) &= \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}| \cdot |\mathbf{b}|} \\ &= \frac{3 \cdot 2 + (-1) \cdot 1 + 2 \cdot 2}{\sqrt{3^2 + (-1)^2 + 2^2} \cdot \sqrt{2^2 + 1^2 + 2^2}} = \frac{9}{\sqrt{14} \cdot \sqrt{9}} = \frac{3}{\sqrt{14}} \approx 0.80178. \end{aligned}$$

We have to find the smallest positive argument of the cosine function which gives the value 0.80178. Therefore, the angle between vectors $\mathbf{a}$ *and* $\mathbf{b}$ *is approximately equal to* 36.7^o.

Next, we consider **orthogonal** vectors. Consider the triangle given in Fig. 11.3 formed by the three 2-dimensional vectors $\mathbf{a}, \mathbf{b}$ and $\mathbf{c} = \mathbf{a} - \mathbf{b}$. Denote the angle between vectors $\mathbf{a}$ and $\mathbf{b}$ by γ. From the Pythagorean theorem we know that the angle γ is equal to 90^o if and only if the sum of the squared lengths of vectors $\mathbf{a}$ and $\mathbf{b}$ is equal to the squared length of vector $\mathbf{c}$:

$$|\mathbf{a}|^2 + |\mathbf{b}|^2 = |\mathbf{a} - \mathbf{b}|^2 . \tag{11.2}$$

Using the rules for working with vectors, Equation (11.2) is satisfied if and only if the scalar product of the vectors $\mathbf{a}$ and $\mathbf{b}$ is equal to zero. We say in this case that vectors $\mathbf{a}$ and $\mathbf{b}$ are **orthogonal** (or **perpendicular**) and write $\mathbf{a} \perp \mathbf{b}$. Accordingly, two 3-dimensional vectors are orthogonal if their scalar product is equal to zero.

Example 11.8 *The 3-dimensional vectors*

$$\mathbf{a} = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} \qquad \textit{and} \qquad \mathbf{b} = \begin{pmatrix} 4 \\ 6 \\ -3 \end{pmatrix}$$

are orthogonal since

$$\mathbf{a} \cdot \mathbf{b} = 3 \cdot 4 + (-1) \cdot 6 + 2 \cdot (-3) = 0 .$$

Vector product:
The **vector product** (or **cross product**) $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ of the 3-dimensional vectors $\mathbf{a}$ and $\mathbf{b}$ is defined as follows:

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} = \begin{pmatrix} a_y b_z - a_z b_y \\ a_z b_x - a_x b_z \\ a_x b_y - a_y b_x \end{pmatrix},$$

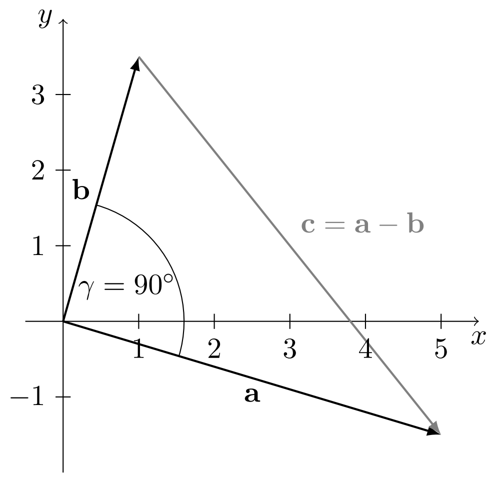


Figure 11.3: Orthogonality of 2-dimensional vectors

or equivalently,

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} = (a_y b_z - a_z b_y)\mathbf{e}^{\mathbf{x}} + (a_z b_x - a_x b_z)\mathbf{e}^{\mathbf{y}} + (a_x b_y - a_y b_x)\mathbf{e}^{\mathbf{z}}.$$

The vector product of two vectors is defined as a **vector** having the following properties:

1. The vector $\mathbf{c}$ is orthogonal to both vectors $\mathbf{a}$ and $\mathbf{b}$.
2. The vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ form a right-handed coordinate system, i.e., the thumb points into the direction of the vector $\mathbf{a} \times \mathbf{b}$.
3. The length of the vector c is equal to

$$|\mathbf{c}| = |\mathbf{a} \times \mathbf{b}| = |\mathbf{a}| \cdot |\mathbf{b}| \cdot \sin(\angle(\mathbf{a}, \mathbf{b})).$$

The length of the resulting vector $\mathbf{c}$ is equal to the area of the parallelogram spanned by the two vectors $\mathbf{a}$ and $\mathbf{b}$. As a consequence, the vector product of two parallel vectors $\mathbf{a}$ and $\mathbf{b}$ is equal to the zero vector.

We obtain

$$\begin{array}{lll} \mathbf{e}^{\mathbf{x}} \times \mathbf{e}^{\mathbf{x}} = \mathbf{0}\,; & \mathbf{e}^{\mathbf{x}} \times \mathbf{e}^{\mathbf{y}} = \mathbf{e}^{\mathbf{z}}\,; & \mathbf{e}^{\mathbf{x}} \times \mathbf{e}^{\mathbf{z}} = -\mathbf{e}^{\mathbf{y}}\,; \\ \mathbf{e}^{\mathbf{y}} \times \mathbf{e}^{\mathbf{x}} = -\mathbf{e}^{\mathbf{z}}\,; & \mathbf{e}^{\mathbf{y}} \times \mathbf{e}^{\mathbf{y}} = \mathbf{0}\,; & \mathbf{e}^{\mathbf{y}} \times \mathbf{e}^{\mathbf{z}} = \mathbf{e}^{\mathbf{x}}\,; \\ \mathbf{e}^{\mathbf{z}} \times \mathbf{e}^{\mathbf{x}} = \mathbf{e}^{\mathbf{y}}\,; & \mathbf{e}^{\mathbf{z}} \times \mathbf{e}^{\mathbf{y}} = -\mathbf{e}^{\mathbf{x}}\,; & \mathbf{e}^{\mathbf{z}} \times \mathbf{e}^{\mathbf{z}} = \mathbf{0}\,. \end{array}$$

The vector product is **not** commutative. In particular, we have

$$\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$$

Moreover, the vector product is **not** associative:

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} \neq \mathbf{a} \times (\mathbf{b} \times \mathbf{c})$$

Example 11.9 *Let*

$$\mathbf{a} = \begin{pmatrix} 4 \\ 1 \\ -2 \end{pmatrix} \qquad \textit{and} \qquad \mathbf{b} = \begin{pmatrix} 1 \\ -3 \\ 7 \end{pmatrix}$$

be given. We compute the vector product and obtain:

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} = \begin{pmatrix} 1 \cdot 7 - (-2) \cdot (-3) \\ (-2) \cdot 1 - 4 \cdot 7 \\ 4 \cdot (-3) - 1 \cdot 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -30 \\ -13 \end{pmatrix}.$$

Next, we compute the area of the parallelogram spanned by the vectors **a** *and* **b**. *This area is obtained as the length of the vector* **c***:*

$$|\mathbf{c}| = \sqrt{1^2 + (-30)^2 + (-13)^2} = \sqrt{1 + 900 + 169} = \sqrt{1070} \approx 32.71.$$

Example 11.10 *We determine the area of the triangle spanned by the three points*

$$P = (2, 0, 3), \quad Q = (5, 1, 4) \quad \textit{and} \quad R = (2, 6, 5).$$

Taking the vectors

$$\mathbf{a} = \overrightarrow{PQ} = \begin{pmatrix} 5-2 \\ 1-0 \\ 4-3 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}$$

and

$$\mathbf{b} = \overrightarrow{PR} = \begin{pmatrix} 2-2 \\ 6-0 \\ 5-3 \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \\ 2 \end{pmatrix},$$

we first determine the vector product of both vectors and obtain:

$$\mathbf{a} \times \mathbf{b} = \begin{pmatrix} 1 \cdot 2 - 1 \cdot 6 \\ 1 \cdot 0 - 3 \cdot 2 \\ 3 \cdot 6 - 1 \cdot 0 \end{pmatrix} = \begin{pmatrix} -4 \\ -6 \\ 18 \end{pmatrix}.$$

Finally, we determine

$$A = \frac{1}{2} \cdot |\mathbf{a} \times \mathbf{b}| = \frac{1}{2} \cdot \sqrt{(-4)^2 + (-6)^2 + 18^2} = \frac{1}{2} \cdot \sqrt{16 + 36 + 324} = \frac{1}{2} \cdot \sqrt{376} \approx 9.70,$$

i.e., the area of the triangle formed by the given three points is approximately equal to 9.70 squared units.

EXERCISES

11.1 Given are the vectors

$$\mathbf{a} = \begin{pmatrix} 4 \\ -3 \\ 2 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} -2 \\ 5 \\ 7 \end{pmatrix} \qquad \text{and} \qquad \mathbf{c} = \begin{pmatrix} 4 \\ 0 \\ -6 \end{pmatrix}.$$

(a) Determine for all pairs of two of the above vectors the sum and the difference of these vectors.
(b) Determine the vectors $6\mathbf{a} - 3\mathbf{b} + 2\mathbf{c}$ and $5\mathbf{b} - 3\mathbf{c} - \mathbf{a}$.
(c) Determine the angle between any two of the above vectors.

11.2 Determine the length of the vector

$$\mathbf{a} = \begin{pmatrix} 3 \\ \sqrt{6} \\ 1 \end{pmatrix}$$

and a unit vector pointing in the same direction as the vector **a**.

11.3 Given are the vectors

$$\mathbf{a} = \begin{pmatrix} -2 \\ -3 \\ z \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}.$$

For which values of z are the vectors **a** and **b** orthogonal?

11.4 Given are the vectors

$$\mathbf{a} = \begin{pmatrix} -2 \\ -11 \\ 3 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 4 \\ -3 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{c} = \begin{pmatrix} 6 \\ -2 \\ 6 \end{pmatrix}.$$

For which values of λ is the vector **a** orthogonal to the vector $\lambda\mathbf{b} + \mathbf{c}$?

11.5 Determine the vector product of the vectors

$$\mathbf{a} = \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} -7 \\ 1 \\ 3 \end{pmatrix}.$$

11.6 Given are the points $P = (-1, 4, 2), Q = (1, 1, 1)$ and $R = (5, 6, 0)$. Determine the area of the triangle formed by the points P, Q and R.

Chapter 12

Combinatorics, Probability Theory and Statistics

Probability theory and mathematical statistics, which are both a part of stochastic theory, deal with the analysis of random phenomena which play an important role practically everywhere in daily life. Probability theory analyzes abstract models of random events. In statistics, models are used for forecasts and further analyses and planning. Statistical methods transform big data sets into useful compact information. Statistical methods are not only used in mathematics, engineering or economics, but also to a large extent in empirical sciences. Some basic combinatorial relationships are needed e.g. for counting possible variants that may happen in an experiment or investigation.

The learning objectives of this chapter are

- to review some basic formulas for combinatorial problems,
- to survey the main probability distributions and
- to review some basics about a statistical test.

12.1 Combinatorics

In this section, we summarize some basics about combinatorics. In particular, we investigate two basic questions:

- How many possibilities do exist to sequence the elements of a given set?
- How many possibilities do exist to select a certain number of elements from a set?

Let us start with the determination of the number of possible sequences formed with a set of elements. To this end, we first introduce the notion of a permutation.

Permutation:
Let $M = \{a_1, a_2, \ldots, a_n\}$ be a set of n elements. Any sequence $(a_{p_1}, a_{p_2}, \ldots, a_{p_n})$ of all elements of the set M is called a **permutation**.

For instance, if $M = \{1, 2, 3, 4, 5\}$, one of the possible permutations is the sequence $(2, 1, 4, 3, 5)$. Similarly, for the set $M = \{a, b, s, t, u, v\}$, one of the possible permutations is (b, s, t, a, v, u).

In order to determine the number of possible permutations, we introduce the number $n!$ (read: n factorial) which is defined as follows:

$$n! = 1 \cdot 2 \cdot \ldots \cdot (n-1) \cdot n \quad \text{for } n \geq 1 .$$

For $n = 0$, we define $0! = 1$.

Number of permutations:
Let a set M consisting of $n \geq 1$ distinct elements be given. Then there exist $P(n) = n!$ permutations.

Example 12.1 *We enumerate all permutations of the elements of set* $M = \{1, 2, a\}$*. We can form* $P(3) = 3! = 6$ *sequences of the elements from the set* M*:*

$$(1, 2, a),\ (1, a, 2),\ (2, 1, a),\ (2, a, 1),\ (a, 1, 2),\ (a, 2, 1) .$$

Example 12.2 *Assume that nine jobs have to be processed on a single machine and that all job sequences are feasible. Then there exist*

$$P(9) = 9! = 1 \cdot 2 \cdot \ldots \cdot 9 = 362,880$$

feasible job sequences.

If there are some identical elements (which cannot be distinguished), the number of possible permutations of such elements reduces in comparison with $P(n)$. The number of such permutations with allowed repetition of elements can be determined as follows.

Number of permutations with allowed repetition:
Let n elements consisting of k groups of $n_1, n_2, \ldots, n_k$ identical elements with $n = n_1 + n_2 + \cdots + n_k$ be given. Then there exist

$$P(n;\ n_1, n_2, \ldots, n_k) = \frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_k!} \qquad (12.1)$$

permutations.

Example 12.3 *How many distinct numbers with 10 digits can one form which contain two times digit 1, three times digit 4, one time digit 5 and four times digit 9? We use formula (12.1) with* $n = 10, n_1 = 2, n_2 = 3, n_4 = 1$ *and* $n_4 = 4$ *and obtain*

$$P(10;\ 2, 3, 1, 4) = \frac{10!}{2! \cdot 3! \cdot 1! \cdot 4!} = \frac{5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot 10}{2 \cdot 6} = 12,600$$

distinct numbers with these properties.

Example 12.4 *In a wardrobe, there are put 11 pants: five black ones, three blue ones, two white ones and one brown one. How many different sequences do exist for arranging these 11 pants in the wardrobe when only the colors are distinguished (i.e., pants of the same color are not distinguished). The problem is to find the number of possible permutations with identical elements. We have* $n = 11, n_1 = 5, n_2 = 3, n_3 = 2$ *and* $n_4 = 1$. *Thus, due to formula (12.1), there are*

$$P(11;\ 5,3,2,1) = \frac{11!}{5! \cdot 3! \cdot 2! \cdot 1!} = \frac{7 \cdot 8 \cdot 9 \cdot 10 \cdot 11}{2} = 27,720$$

possibilities for arranging the pants in the wardrobe when they are only distinguished by their color.

For the subsequent considerations, the notion of a binomial coefficient is useful.

Binomial coefficient:
Let k, n be integers with $0 \leq k \leq n$. The term

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$$

is called the **binomial coefficient** (read: from n choose k). For $k > n$, we define

$$\binom{n}{k} = 0.$$

For instance, we get

$$\binom{8}{3} = \frac{8!}{3! \cdot (8-3)!} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8}{(1 \cdot 2 \cdot 3) \cdot (1 \cdot 2 \cdot 3 \cdot 4 \cdot 5)} = \frac{6 \cdot 7 \cdot 8}{1 \cdot 2 \cdot 3} = 56.$$

In order to determine the binomial coefficient $\binom{n}{k}$ with $n > k$, we have to compute the quotient of the product of the k largest integers not greater than n, i.e., $(n-k+1) \cdot (n-k+2) \cdot \ldots \cdot n$, divided by the product of the first k integers (i.e., $1 \cdot 2 \cdot \ldots \cdot k$). Then we have the following properties for working with binomial coefficients.

Rules for binomial coefficients:
Let k and n be integers with $0 \le k \le n$. Then we have:

1. $\binom{n}{0} = \binom{n}{n} = 1\,; \qquad \binom{n}{k} = \binom{n}{n-k}\,;$
2. $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}\,.$

Using the binomial coefficients, one can formulate the following equality for the computation of the n-th power of a binomial term $a+b$:

$$(a+b)^n = a^n + \binom{n}{1} a^{n-1} b + \binom{n}{2} a^{n-2} b^2 + \ldots + \binom{n}{n-1} a b^{n-1} + b^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k. \quad (12.2)$$

where a, b are real numbers (or more general, mathematical terms) and n is a natural number. The coefficients in formula (12.2) can be easily determined by using Pascal's triangle which is as follows:

Pascal's triangle

$$\begin{array}{lccccccc}
(a+b)^0 & & & & 1 & & & \\
(a+b)^1 & & & 1 & & 1 & & \\
(a+b)^2 & & 1 & & 2 & & 1 & \\
(a+b)^3 & 1 & & 3 & & \mathbf{3} & & \mathbf{1} \\
(a+b)^4 & 1 \quad 4 & & 6 & & \mathbf{4} & & 1 \\
\vdots & & & & \vdots & & & \\
(a+b)^n & \binom{n}{0} \quad \binom{n}{1} & \cdots & \binom{n}{k} & \binom{n}{k+1} & \cdots & \binom{n}{n-1} & \binom{n}{n} \\
(a+b)^{n+1} & \binom{n+1}{0} \quad \binom{n+1}{1} & \cdots & \binom{n+1}{k+1} & & \cdots & \binom{n+1}{n} & \binom{n+1}{n+1}
\end{array}$$

In Pascal's triangle, each inner number is obtained by adding the two numbers in the row above which are standing immediately to the left and to the right of the number considered (see rule (2) for binomial coefficients). For instance, the number 4 in the row for $(a+b)^4$ is obtained by adding the number 3 and the number 1 both in the row of $(a+b)^3$ (see the numbers in bold face in rows four and five of Pascal's triangle above). Moreover, the numbers in each row are symmetric. In the case of $n = 2$, we obtain the well-known binomial formula

$$(a+b)^2 = a^2 + 2ab + b^2.$$

Example 12.5 *We determine the term*

$$T = (2x + 3y)^4 \ .$$

Applying formula (12.2), we get

$$\begin{aligned} T &= (2x)^4 + \binom{4}{1} \cdot (2x)^3 \cdot 3y + \binom{4}{2} \cdot (2x)^2 \cdot (3y)^2 + \binom{4}{3} \cdot 2x \cdot (3y)^3 + (3y)^4 \\ &= 16x^4 + 4 \cdot 8x^3 \cdot 3y + 6 \cdot 4x^2 \cdot 9y^2 + 4 \cdot 2x \cdot 27y^3 + 81y^4 \\ &= 16x^4 + 96x^3y + 216x^2y^2 + 216xy^3 + 81y^4 \ . \end{aligned}$$

Next, we investigate how many possibilities do exist for selecting a certain number of elements from some given set when the order in which the elements are chosen is **important**. We distinguish the cases where a repeated selection of elements is allowed and forbidden, respectively.

Number of variations without / with repeated selection:
The number of possible selections of k elements from n elements with consideration of the sequence (i.e., the order in which the elements are selected), each of them denoted as a **variation**, is equal to

1. $V(k,n) = \dfrac{n!}{(n-k)!}$ if repeated selection is forbidden;
2. $\overline{V}(k,n) = n^k$ if repeated selection is allowed.

In the first case, every element may occur only once in each selection while in the second case, an element may occur arbitrarily often in a selection.

Example 12.6 *A dual number consists only of the digits 0 and 1. How many different dual numbers can one form with eight digits? The answer is obtained by the number of variations of 8 elements when repeated selection is allowed. We get*

$$\overline{V}(2,8) = 2^8 = 256\,.$$

Next, we consider the special case when the order in which the elements are chosen is **unimportant**, i.e., it does not matter which element is selected first.

Number of combinations without / with repeated selection:
The number of possible selections of k elements from n elements without consideration of the sequence (i.e., the order in which the elements are selected), each of them denoted as a **combination**, is equal to

1. $C(k,n) = \dbinom{n}{k}$ if repeated selection of the same element is forbidden;
2. $\overline{C}(k,n) = \dbinom{n+k-1}{k}$ if repeated selection of the same element is allowed.

Example 12.7 *In the German game Skat, 32 cards are distributed among three players. Each of them get 10 cards (and two are left for the skat, which one player can exchange later with two own cards). How many possibilities do exist for one player to get different collections of the initial 10 cards? This number is equal to the number of combinations $C(10,32)$ without repetitions:*

$$C(10,32) = \binom{32}{10} = \frac{32\cdot 31\cdot 30\cdot 29\cdot 28\cdot 27\cdot 26\cdot 25\cdot 24\cdot 23}{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7\cdot 8\cdot 9\cdot 10} = 64,512,240\,.$$

If we determine the number of overall distributions of the cards to the players (i.e., how many different distributions are possible), we have to take into account that the first player gets 10 cards out of 32 ones, the second player 10 out of 22 cards, and the third player gets 10 out of 12 cards. This gives on overall

$$N = \binom{32}{10}\cdot\binom{22}{10}\cdot\binom{12}{10} = \frac{32!}{(10!)^3\cdot 2!} = 688,323,602,126,160$$

possible distributions of the cards to the players.

Example 12.8 *In a small country, there are 10 major touristic attractions. One person can visit only three attractions on one day. How many possibilities do exist to select three attractions? This corresponds to the number of combinations without repeated selection:*

$$C(3,10) = \binom{10}{3} = \frac{10 \cdot 9 \cdot 8}{1 \cdot 2 \cdot 3} = \frac{720}{6} = 120\,.$$

If the tourist has chosen three attractions, he can visit them in different orders. Assume that every possible variant is considered as a different sub-tour, how many sub-tours including three different attractions are possible? This is the problem of determining the number of variations when repeated selection is forbidden. Thus, we obtain

$$V(3,10) = \frac{10!}{(10-3)!} = \frac{10!}{7!} = 8 \cdot 9 \cdot 10 = 720$$

different sub-tours with three attractions included.

Example 12.9 *In a German lottery, there have to be chosen six numbers out of 49 numbers. How many possibilities do exist to have exactly three out of the six right numbers chosen? There are* $\binom{6}{3}$ *possibilities to select from the 6 'right' numbers exactly three numbers, and there are* $\binom{43}{3}$ *possibilities to choose from the* $49 - 6 = 43$ *'false' numbers exactly three ones. Since each selection of three 'right' numbers can be combined with each selection of three 'false' numbers, we get the number*

$$N = \binom{6}{3} \cdot \binom{43}{3} = \frac{6 \cdot 5 \cdot 4}{1 \cdot 2 \cdot 3} \cdot \frac{43 \cdot 42 \cdot 41}{1 \cdot 2 \cdot 3} = 20 \cdot 12,341 = 246,820$$

of possible variants with exactly three right numbers in the lottery game.

Finally, we discuss some additional variants of combinations. Let $C_m(k,n)$ be the number of combinations without repeated selection which contain m given elements and $C_{\overline{m}}(k,n)$ be the number of combinations without repeated selection which do not contain m given elements $(m \le k)$. Then we have

$$C_m(k,n) = C(k-m, n-m) = \binom{n-m}{k-m}$$

and

$$C_{\overline{m}}(k,n) = C(k, n-m) = \binom{n-m}{k}\,.$$

Example 12.10 *How many 5-digit numbers without repeated selection can be formed by means of 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 containing the digits 3 and 6. We have* $n = 10, k = 5$ *and* $m = 2$. *This gives*

$$C_2(5,10) = C(3,8) = \binom{8}{3} = \frac{8 \cdot 7 \cdot 6}{1 \cdot 2 \cdot 3} = 8 \cdot 7 = 56\,.$$

Next, we determine how many 5-digit numbers can be formed by means of 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 without repeated selection containing at least one of the digits 3 or 6. This is the number

$C(5,10)$ *reduced by the number of possible 5-digit numbers which do not contain one of the digits 3 or 6. This gives*

$$\begin{aligned} C(5,10) - C_{\overline{2}}(5,10) &= C(5,10) - C(5,8) = \binom{10}{5} - \binom{8}{5} \\ &= \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} - \frac{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} \\ &= 252 - 56 = 196 \end{aligned}$$

of such numbers.

12.2 Events

In this section, we introduce the notion of an event which is of fundamental importance in probability theory and modeling aspects. An event is simply a collection of outcomes of a random experiment. One deals with the occurrence and non-occurrence of events. If one tosses a coin, there are two possible outcomes possible: either the obverse of a coin is visible or the reverse. However, it cannot be said in advance what will result. So, when tossing a coin, there are two **random events** possible, namely 'obverse' and 'reverse' of the coin. Similarly, if one tosses a dice, there are six random events possible: each of the numbers 1, 2, 3, 4, 5, or 6 can appear.

Impossible and certain event:
The event $\emptyset$ that never happens is called the **impossible event**.
The event S that always occurs is called the **certain event**.

One can work with events similar to sets. In particular, we can introduce the following events.

Union and intersection of events:
The **union** $A \cup B$ of the events A and B occurs if at least one of the events A or B happens.
The **intersection** $A \cap B$ of the events A and B occurs if both events A and B happen.

We emphasize that the commutative, associative and distributive laws are valid when working with events.

Exclusive events:
Two events A and B are called **exclusive** (or **disjoint**) if they cannot happen both:

$$A \cap B = \emptyset \, .$$

Complementary event:
The **complementary event** $\overline{A}$ occurs if and only if the event A does not happen:

$$A \cap \overline{A} = \emptyset, \qquad A \cup \overline{A} = S\,.$$

Example 12.11 *If one tosses a coin, one can consider the events:*

O*: The obverse of a coin appears.*
R*: The reverse of a coin appears.*

The union $O \cup R$ *is the certain event while* O *and* R *are exclusive: the event* R *is the complementary event of* $O : R = \overline{O}$.

Example 12.12 *An autonomous region of a country has three power stations. Let* A_i *be the event that power station* i *works without a failure (*$i = 1, 2, 3$*). Consequently,* $\overline{A_i}$ *is the event that power station* i *works with a failure. Moreover, introduce the following events:*

B*: All three power stations work without a failure;*
C*: Exactly one of the power stations works without a failure;*
D*: At least one of the power stations works without a failure.*

Then we get:

$$\begin{aligned} B &= A_1 \cap A_2 \cap A_3\,; \\ C &= (A_1 \cap \overline{A_2} \cap \overline{A_3}) \cup (\overline{A_1} \cap A_2 \cap \overline{A_3}) \cup (\overline{A_1} \cap \overline{A_2} \cap A_3)\,; \\ D &= A_1 \cup A_2 \cup A_3\,. \end{aligned}$$

Example 12.13 *Consider the events:*

A*: A person is younger than 30 years;*

B*: A person is between 30 and 50 years old;*

C*: A person is older than 50 years;*

M*: A person is male.*

Then we can consider e.g. the following events:

$(A \cup B) \cap M$*: A person is not older than 50 years and male.*

$A \cap \overline{M}$*: A person is younger than 30 years and female (i.e., not male).*

$(B \cap M) \cup (A \cap \overline{M})$*: A person is between 30 and 50 years old and male, or a person is younger than 30 years old and female.*

$(A \cup B \cup C) \cap M$*: This event corresponds to event* M.

$(A \cap B) \cup M$*: This is the impossible event* $\emptyset$ *since a person cannot be simultaneously younger than 30 and between 30 and 50 years old.*

12.3 Relative Frequencies and Probability

Absolute and relative frequency of an event:
The **absolute frequency** $H(A)$ of an event A is the number of times the event occurs in a particular investigation.
The **relative frequency** $h(A)$ of an event A is defined as the quotient of the absolute frequency $H(A)$ of an event and the total number n of events observed:

$$h(A) = \frac{H(A)}{n} .$$

The relative frequency is also known as **empirical probability**. If relationships between two events are considered, often a **four-array table** (which is a special case of a **contingency table**) is used. We illustrate this for the case of relative frequencies (for absolute frequencies, it can be done in a similar manner). We consider two events A and B and well as their complementary events $\overline{A}$ and $\overline{B}$.

Table 12.1: Four-array table for relative frequencies

	B	$\overline{B}$	Σ
A	$h(A \cap B)$	$h(A \cap \overline{B})$	$h(A)$
$\overline{A}$	$h(\overline{A} \cap B)$	$h(\overline{A} \cap \overline{B})$	$h(\overline{A})$
Σ	$h(B)$	$h(\overline{B})$	1

If a problem is given, one first fills the table with the known values and then computes the remaining entries by using the calculation rules. For instance, we have

$$h(A \cap B) + h(A \cap \overline{B}) = h(A) \qquad \text{and} \qquad h(A \cap B) + h(\overline{A} \cap B) = h(B) .$$

Moreover, we have

$$h(A) + h(\overline{A}) = h(B) + h(\overline{B}) = 1 .$$

Example 12.14 *200 persons took part in a qualifying test for entering a university. Among them, there were 107 women. Moreover, it is known that 153 persons passed the test and that 24 men failed. Using the events P (a person passed the test), $\overline{P}$ (a person failed, i.e., did not pass the test), M (a person is male) and $\overline{M}$ (a person is female, i.e., not male), we get the following four-array table, where the known numbers from the problem are given in bold face.*

Table 12.2: *Four-array table for Example 12.14.*

	M	$\overline{M}$	Σ
P	$\frac{69}{200}$	$\frac{84}{200}$	$\mathbf{\frac{153}{200}}$
$\overline{P}$	$\mathbf{\frac{24}{200}}$	$\frac{23}{200}$	$\frac{47}{200}$
Σ	$\frac{93}{200}$	$\mathbf{\frac{107}{200}}$	1

From the problem, we know

$$h(\overline{M}) = \frac{107}{200}; \qquad h(P) = \frac{153}{200}; \qquad h(\overline{P} \cap M) = \frac{24}{200} \,.$$

To find the other values in Table 12.2, we calculate

$$h(M) = 1 - h(\overline{M}) = 1 - \frac{107}{200} = \frac{93}{200}; \quad h(\overline{P}) = 1 - h(P) = 1 - \frac{153}{200} = \frac{47}{200} \,.$$

Using these values, we obtain further

$$\begin{aligned} h(\overline{P} \cap \overline{M}) &= h(\overline{P}) - h(\overline{P} \cap M) = \frac{47}{200} - \frac{24}{200} = \frac{23}{200}; \\ h(P \cap \overline{M}) &= h(\overline{M}) - h(\overline{P} \cap \overline{M}) = \frac{107}{200} - \frac{23}{200} = \frac{84}{200}; \\ h(P \cap M) &= h(P) - h(P \cap \overline{M}) = \frac{153}{200} - \frac{84}{200} = \frac{69}{200} \,. \end{aligned}$$

As a consequence, we found that 23 women did not pass the test, while 84 women and 69 men passed the test.

In a simplified manner, one can define the probability of an event A as follows.

Probability:
The **probability** $P(A)$ of an event A is given by

$$P(A) = \frac{\text{number of the favorable outcomes for the event}}{\text{number of the possible outcomes}} \,.$$

Example 12.15 *Let A be the event that there appears at least number 3 when tossing a dice. There are four favorable outcomes for the event (one of the numbers 3, 4, 5, or 6 appears) among the possible six outcomes. Therefore, we get*

$$P(A) = \frac{4}{6} = \frac{2}{3} \,.$$

Example 12.16 *Somebody tosses simultaneously two dice. Let A denote the event that both dice display at least number 5. There are 2^2 favorable outcomes (namely each of the dice displays one of the numbers 5 or 6) and 6^2 possible outcomes. Therefore, we obtain*

$$P(A) = \frac{4}{36} = \frac{1}{9} \,.$$

12.4 Basic Probability Theorems

In this short section, we review the calculation of the probabilities of the complementary event and the union of two events.

Complementary event:
The **probability of the complementary event** $\overline{A}$ of an event A is given by

$$P(\overline{A}) = 1 - P(A) \ .$$

Addition rule:
The **probability of the union** of two events A and B is given by

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \ .$$

In the case when A and B are **exclusive**, the addition rule simplifies to

$$P(A \cup B) = P(A) + P(B) \ .$$

Example 12.17 *Consider the events*
A: A person is less than 18 years old;
B: A person is at least 18 years old but younger than 65;
C: A person is at least 65 years old.
Let for some country these probabilities given as $P(A) = 0.21, P(B) = 0.56$ *and* $P(C) = 0.23$.

Any two events are mutually exclusive, and we have $A \cup B \cup C = S$. *We obtain*

$$\begin{aligned} P(\overline{A}) &= 1 - P(A) = 1 - 0.21 = 0.79; \\ P(A \cup B) &= P(A) + P(B) = 0.21 + 0.56 = 0.77; \\ P(A \cup C) &= P(A) + P(C) = 0.21 + 0.23 = 0.44; \\ P(B \cup C) &= P(B) + P(C) = 0.56 + 0.23 = 0.79\,. \end{aligned}$$

If two events are not exclusive, the calculation of the probability of the intersection of two elements requires in the general case the consideration of conditional probabilities which are introduced in the next section.

12.5 Conditional Probabilities and Independence of Events

In this section, we continue with some further rules for calculating probabilities of events.

Multiplication rule:
Let A and B be two events and $P(B) \neq 0$. Then the **probability of the intersection** of the two events A and B is

$$P(A \cap B) = P(A|B) \cdot P(B)$$

where $P(A|B)$ denotes the **conditional probability** of event A, given event B.

Note that we can interchange the events A and B, i.e., we have

$$P(A|B) \cdot P(B) = P(B|A) \cdot P(A) .$$

If the events A and B are **independent**, we have $P(A|B) = P(A)$ (i.e, the condition has no influence), and we get the special case

$$P(A \cap B) = P(A) \cdot P(B) .$$

Example 12.18 *12.19 Assume that one tosses simultaneously two distinguishable dice. Consider the events*
A_i : *The i-th dice displays an even number* ($i = 1, 2$)*.*
Hence, the complementary events are
$\overline{A_i}$: *The i-th dice displays an odd number* ($i = 1, 2$)*.*
We have

$$P(A_1) = P(A_2) = P(\overline{A_1}) = P(\overline{A_2}) = \frac{1}{2} .$$

Since A_1 and A_2 are independent, we obtain

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2) = \frac{1}{4}$$

and

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) = \frac{1}{2} + \frac{1}{2} - \frac{1}{4} = \frac{3}{4} .$$

Similarly, we get

$$P(\overline{A_1} \cap \overline{A_2}) = P(\overline{A_1}) \cdot P(\overline{A_2}) = \frac{1}{4} \quad \text{and} \quad P(\overline{A_1} \cup \overline{A_2}) = P(\overline{A_1}) + P(\overline{A_2}) - P(\overline{A_1} \cap \overline{A_2}) = \frac{3}{4} .$$

For calculating with probabilities in the case of considering only two events, one can also use a **four-array table** as introduced before.

Example 12.19 *It is known that in a big company 60 % of the employees have an salary less than 4000 EUR per month, 35 % of the employees have a university qualification, and 25 % of the workers both have a university qualification and earn at least 4000 EUR per month,*

Selecting a worker randomly, we can consider the above percentages as probabilities. Introducing the events
U *- the worker has a university qualification;*
G *- the worker has a salary of at least 4000 EUR per month;*
as well as the complementary events $\overline{U}$ *and* $\overline{G}$*, we have the following information (given in bold face in Table 12.2):*

$$P(\overline{G}) = 0.6, \qquad P(U) = 0.35, \qquad P(G \cap U) = 0.25 .$$

Table 12.2: *Four-array table for Example 12.19.*

	U	$\overline{U}$	Σ
G	**0.25**	0.15	0.4
$\overline{G}$	0.1	0.5	**0.6**
Σ	**0.35**	0.65	1

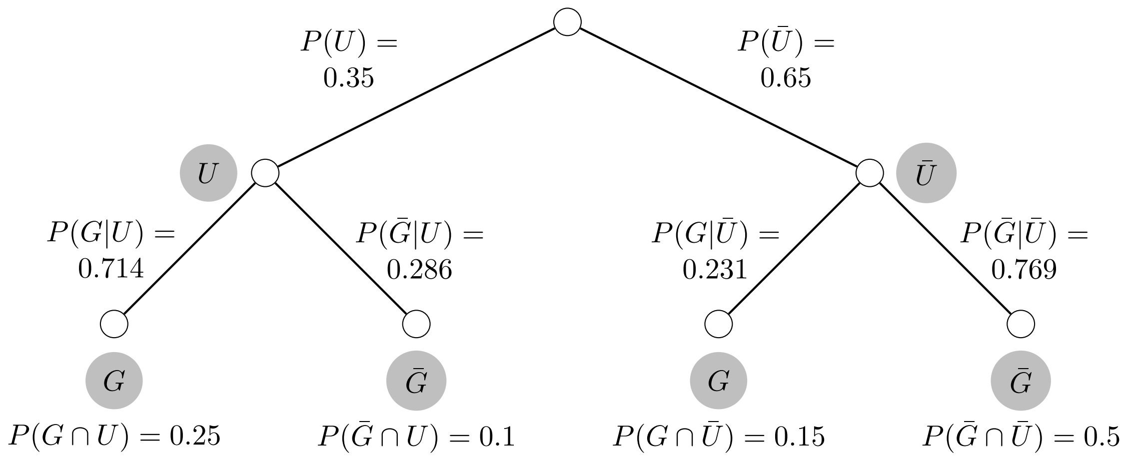


Figure 12.1: Tree diagram for example 8.19

Using the known probabilities, we first get

$$P(G) = 1 - P(\overline{G}) = 1 - 0.6 = 0.4 \qquad \textit{and} \qquad P(\overline{U}) = 1 - P(U) = 1 - 0.35 = 0.65\,.$$

Then we obtain further

$$\begin{aligned} P(\overline{G} \cap U) &= P(U) - P(G \cap U) = 0.35 - 0.25 = 0.1; \\ P(\overline{G} \cap \overline{U}) &= P(\overline{G}) - P(\overline{G} \cap U) = 0.6 - 0.1 = 0.5; \\ P(G \cap \overline{U}) &= P(G) - P(G \cap U) = 0.4 - 0.25 = 0.15 \end{aligned}$$

Using the above four-array table, one can also easily compute the conditional properties. We obtain:

$$\begin{aligned} P(G|U) &= \frac{P(G \cap U)}{P(U)} = \frac{0.25}{0.35} = 0.714 \\ P(\overline{G}|U) &= \frac{P(\overline{G} \cap U)}{P(\overline{U})} = \frac{0.1}{0.35} = 0.286 \\ P(G|\overline{U}) &= \frac{P(G \cap \overline{U})}{P(\overline{U})} = \frac{0.15}{0.65} = 0.231 \\ P(\overline{G}|\overline{U}) &= \frac{P(\overline{G} \cap \overline{U})}{P(\overline{U})} = \frac{0.5}{0.65} = 0.769\,. \end{aligned}$$

For illustrating the above calculations, often a **tree diagram** *is used which is illustrated in Fig. 12.1. At the second branching level, the conditional probabilities are given and at the end of the tree, the probabilities of the intersections of two possible events are given: The leftmost node in the second level displays the probability* $P(G \cap U) = 0.25$ *which is obtained as the product of the probabilities at the arcs leading to this node.*

12.6 Total Probability and Bayes' Theorem

In this section, we review two results using conditional probabilities.

Rule of total probability:
Let $B_1, B_2, \ldots, B_k$ be mutually exclusive events and $A = B_1 \cup B_2 \cup \ldots \cup B_k$. Then:

$$P(A) = \sum_{i=1}^{k} P(A|B_i) \cdot P(B_i) \,.$$

Example 12.20 *In a mass production, a company uses the three machines M_1, M_2 and M_3 which differ in their efficiency and precision. Machine M_1 produces 40 % of the items with an error rate of 8 %, machine M_2 produces 35 % of the items with an error rate of 6 %, and machine M_3 produces 25 % of the items with an error rate of 3 %. Let A be the event that a chosen item is faulty and B_i be the event that it is produced on machine M_i ($i = 1, 2, 3$). Thus, we have*

$$\begin{array}{ccc} P(B_1) = 0.4, & P(B_2) = 0.35, & P(B_3) = 0.25, \\ P(A|B_1) = 0.08, & P(A|B_2) = 0.06, & P(A|B_3) = 0.03 \,. \end{array}$$

We determine the probability that an arbitrarily selected item is faulty. Using the rule of total probability, we obtain

$$\begin{aligned} P(A) &= P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2) + P(A|B_3) \cdot P(B_3) \\ &= 0.08 \cdot 0.4 + 0.06 \cdot 0.35 + 0.03 \cdot 0.25 \\ &= 0.0605 \,. \end{aligned}$$

Thus, the probability of a randomly selected item being faulty is 6.05 %.

The following property gives a relationship between conditional probabilities of the form $P(A|B)$ and $P(B|A)$.

Bayes' theorem:
Let $B_1, B_2, \ldots, B_k$ be mutually exclusive events and $A = B_1 \cup B_2 \cup \ldots \cup B_k$. Then:

$$\begin{aligned} P(B_i|A) &= \frac{P(A|B_i) \cdot P(B_i)}{P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2) + \ldots + P(A|B_k) \cdot P(B_k)} \\ &= \frac{P(A|B_i) \cdot P(B_i)}{\sum_{i=1}^{k} P(A|B_i) \cdot P(B_i)} \end{aligned}$$

for $i = 1, 2, \ldots, k$.

Example 12.21 *We consider again the data given in Example 12.20 and assume that a randomly selected item is faulty. We determine the probability that this item was produced on machine M_3, i.e., we look for $P(B_3|A)$. Applying Bayes' theorem, we get*

$$\begin{aligned} P(B_3|A) &= \frac{P(A|B_3) \cdot P(B_3)}{P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2) + P(A|B_3) \cdot P(B_3)} \\ &= \frac{0.03 \cdot 0.25}{0.08 \cdot 0.4 + 0.06 \cdot 0.35 + 0.03 \cdot 0.25} = \frac{0.075}{0.452} \\ &= 0.166 \,. \end{aligned}$$

12.7 Random Variables and Specific Distributions

12.7.1 Random Variables and Probability Distributions

In this section, we deal with **random variables** and their distribution. A random variable is a variable whose value depends on random influences. We do not present here a measure-theoretic definition of a random variable but mention only that some other formal properties are required for a random variable. We denote random variables by capital letters and possible realizations (i.e., values) by small letters.

Distribution function:
Let X be a random variable. The function $F : \mathbb{R} \to [0,1]$ defined by

$$F(x) = P(X \leq x)$$

is denoted as the **distribution function** of the random variable X.

Any distribution function has the properties

$$\lim_{x \to -\infty} F(x) = 0 \qquad \text{and} \qquad \lim_{x \to \infty} F(x) = 1 \,.$$

Moreover, any distribution function is an increasing, but not necessarily strictly increasing function, and all function values of a distribution function are from the interval $[0,1]$.

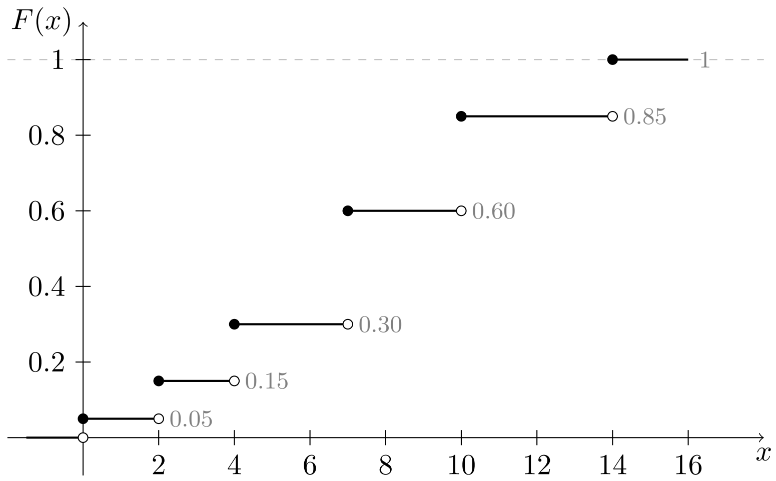


Figure 12.2: Distribution function $F(x)$ for Example 12.22

By means of the distribution function F, one can determine the probability that a particular random variable takes on a value from a specific interval $(a, b]$ as follows:

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)\ .$$

Among the random variables, we distinguish discrete and continuous random variables. A **discrete random variable** is characterized by the values $x_i, x_2, \ldots$, which it may take on and the corresponding probabilities $P(X = x_i) = p_i$. The possible realizations of a discrete random variable may include a finite or an infinite number of values (in the latter case, they need to be 'countable'). For the distribution function of a discrete random variable, we have

$$F(x) = \sum_{x_i \leq x} P(X = x_i) = \sum_{x_i \leq x} p_i\ .$$

Example 12.22 *We consider a discrete random variable X which can take on the values 0, 2, 4, 7, 10 and 14 with the following probabilities:*

$$P(X = 0) = 0.05; \qquad P(X = 2) = 0.1; \qquad P(X = 4) = 0.15;$$

$$P(X = 7) = 0.3; \qquad P(X = 10) = 0.25; \qquad P(X = 14) = 0.15.$$

The distribution function F is a so-called step function given in Fig. 12.2. The discontinuities of the function F are the possible value 0, 2, 4, 7, 10, and 14.

For instance, we have

$$\begin{aligned}
P(X \leq 5) &= F(5) = 0.3\,; \\
P(X > 7) &= 1 - P(X \leq 7) = 1 - F(7) = 0.4 = \sum_{x_i > 7} P(X = x_i)\,; \\
P(4 \leq X \leq 10) &= P(X \leq 10) - P(X < 4) \\
&= P(X \leq 10) - P(X \leq 4) + P(X = 4) \\
&= 0.85 - 0.30 + 0.15 = 0.7 = \sum_{4 \leq x_i \leq 10} P(X = x_i)\ .
\end{aligned}$$

Note that for a discrete random variable, one must carefully distinguish between $P(X \leq x)$ and $P(X < x)$. We have $P(X \leq x) = P(X < x) + P(X = x)$.

A random variable is called **continuous** if it can take on any value from an interval (a, b) (where also $\mathbb{R} = (-\infty, \infty)$ is possible, note also that the boundary points a and b might be included or not). Formally, we can define the following.

Density function:
A random variable X is said to be **continuous** if its distribution function $F(x)$ can be written in the following form:

$$F(x) = \int\limits_{-\infty}^{x} f(t)\, dt \,,$$

where f is denoted as **density function** of the continuous random variable X.

We only note that the last integral with an infinite limit of integration is a so-called **improper integral** which we do not consider in this book. However, we mention that it can be reduced by limit considerations to a usual definite integral:

$$\int_{-\infty}^{x} f(t)\, dt = \lim_{t \to -\infty} \int_{t}^{x} f(t)\, dt \,.$$

Due to the properties of the distribution function F, we always have

$$\int_{-\infty}^{\infty} f(x)\, dx = 1 \,.$$

Note also that for a continuous random variable X, we have

$$P(a < X < b) = P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b)\,,$$

i.e., it is not important, whether one or both of the boundary points are included or not.

12.7.2 Expected Value and Variance

In this section, we deal with the expectation value and the variance which are both fundamental notions of stochastic theory. They describe the average value of a random variable and the mean deviation of a value from this average value, respectively. Let us first consider discrete random variables.

Expected value of a discrete variable:
The **expected value** $E(X)$ of a discrete random variable X is defined as

$$E(X) = \sum_{x_i} x_i \cdot P(X = x_i) \,.$$

Here the summation is made over all possible values x_i which the random variable X may take on.

Example 12.23 *Let X be the discrete random variable giving the grade of a student in the examination in mathematics. From previous years, one knows the probabilities for the grades 1, 2, ..., 5 which are as follows:*

$$P(X=1)=0.1;\ P(X=2)=0.15;\ P(X=3)=0.35;\ P(X=4)=0.25; P(X=5)=0.15\,.$$

We determine the expected average grade $E(X)$:

$$\begin{aligned} E(X) &= \sum_{i=1}^{5} i\cdot P(X=i) \\ &= 1\cdot 0.1+2\cdot 0.15+3\cdot 0.35+4\cdot 0.25+5\cdot 0.15=3.2\,. \end{aligned}$$

Variance of a discrete variable:
The **variance** $\sigma^2(X)$ of a discrete variable is defined as

$$\sigma^2(X)=\sum_{x_i}\left[x_i-E(X)\right]^2\cdot P(X=x_i)\,.$$

Example 12.24 *A discrete random variable X may take on the values $x_1=1, x_2=2, x_3=3$ and $x_4=4$. The probabilities are given as follows:*

$$p_1=P(X=x_1)=0.2;\ p_2=P(X=x_2)=0.3;\ p_3=P(X=x_3)=0.4;\ p_4=P(X=x_4)=0.1.$$

For the expected value $E(X)$, we obtain

$$E(X)=\sum_{i=1}^{4} x_i p_i = 1\cdot 0.2+2\cdot 0.3+3\cdot 0.4+4\cdot 0.1=2.4\,.$$

For the variance $\sigma^2(X)$, we obtain

$$\begin{aligned} \sigma^2(X) &= \sum_{i=1}^{4}\left[x_i-E(X)\right]^2\cdot p_i \\ &= (1-2.4)^2\cdot 0.2+(2-2.4)^2\cdot 0.3+(3-2.4)^2\cdot 0.4+(4-2.4)^2\cdot 0.1 \\ &= 0.84\,. \end{aligned}$$

Let us now consider a continuous random variable X with the density function f. In this case, the expected value $E(X)$ and the variance $\sigma^2(X)$ of variable X are defined as follows:

Expected value of a continuous random variable X:

$$E(X)=\int_{-\infty}^{\infty} x\ f(x)dx\,.$$

Variance of a continuous random variable X:

$$\sigma^2(X) = \int_{-\infty}^{\infty} [x - E(X)]^2 \cdot f(x)\, dx.$$

Often one works with the so-called **standard deviation** σ which is defined as follows: $\sigma = +\sqrt{\sigma^2}$.

An equivalent but often easier way to compute the variance is as follows:

$$\sigma^2(X) = \int_{-\infty}^{\infty} x^2 \cdot f(x)\, dx - [E(X)]^2.$$

In an analogous way, such a transformation can be also made for a discrete random variable. If only values from some interval $[a, b]$ are possible for a continuous random variable, we get a usual definite integral with the limits of integration a and b as we demonstrate in the following example.

Example 12.25 *Consider a random variable with the density function*

$$f(x) = \frac{1}{b-a}, \qquad a \le x \le b$$

(for all other values, we have $f(x) = 0$). We determine $E(X)$ and obtain

$$\begin{aligned} E(X) &= \int_a^b x\, f(x)\, dx = \int_a^b x \cdot \frac{1}{b-a}\, dx \\ &= \frac{1}{b-a} \cdot \frac{x^2}{2}\Big|_a^b \\ &= \frac{1}{b-a} \cdot \frac{b^2 - a^2}{2} = \frac{(b-a)(b+a)}{(b-a)\cdot 2} = \frac{b+a}{2}\ . \end{aligned}$$

For the variance $\sigma^2(X)$, we obtain

$$\begin{aligned} \sigma^2(X) &= \int_a^b x^2\, f(x)\, dx - [E(X)]^2 \\ &= \int_a^b x^2\, \frac{1}{b-a}\, dx - \left(\frac{b+a}{2}\right)^2 \\ &= \frac{1}{b-a} \cdot \frac{x^3}{3}\Big|_a^b - \frac{(b+a)^2}{4} \\ &= \frac{a^2 + ab + b^2}{3} - \frac{(b+a)^2}{4} = \frac{a^2 - 2ab + b^2}{12} = \frac{(a-b)^2}{12}\ . \end{aligned}$$

Here we used that $(b^3 - a^3) = (a^2 + ab + b^2) \cdot (b - a)$.

Example 12.26 *Consider a non-negative random variable X with the density function*

$$f(x) = \lambda\, e^{\lambda x} \qquad (x \ge 0),$$

where $\lambda > 0$ is a given parameter. We only note that this distribution is called **exponential distribution**. *We determine $P(0 \leq X \leq 1)$ and obtain:*

$$\begin{aligned} P(0 \leq X \leq 1) &= \int_0^1 f(t)\,dt = \int_0^1 \lambda\, e^{-\lambda t}\,dt \\ &= \lambda \cdot \frac{1}{-\lambda} \cdot e^{-\lambda t}\Big|_0^1 = 1 - e^{-\lambda}\,. \end{aligned}$$

Note that to find the above integral, we applied integration by substitution.

12.7.3 Binomial Distribution

Bernoulli variable:
A random variable X is called a **Bernoulli variable**, if X can take on only the two values 0 and 1, and the corresponding probabilities are

$$P(X = 0) = p \qquad \text{and} \qquad P(X = 1) = 1 - p\,.$$

Binomial random variable:
Let $X_1, X_2, \ldots, X_n$ be n independent and identically distributed Bernoulli random variables. The variable

$$X = X_1 + X_2 + \ldots X_n$$

is called a **binomial random variable**.

We say that the variable $X = X_{n,p}$ is **binomially distributed** with the parameters n and p. The probability that the variable X takes on the value k can be calculated as follows:

$$P(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}\,.$$

The cumulative probabilities can be determined as follows:

$$P(X \leq l) = \sum_{k=0}^{l} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}\,.$$

For the mean and the variance of a binomial random variable, we have

$$E(X) = np \qquad \text{and} \qquad \sigma^2(X) = np(1-p)\,.$$

Example 12.27 *In a mass production of a product, it is known that 1.5 % of the items are faulty. We determine the probability that among $n = 10$ randomly selected items, there are $k = 2$ faulty ones.*

The discrete random variable $X = X_{n,p}$ describes the number of faulty items if one considers a sample of n items. For $n = 10$ and $k = 2$, we obtain

$$P(X_{10,0.015} = 2) = \binom{10}{2} \cdot 0.015^2 \cdot 0.985^{10-2} = 0.0091\,,$$

i.e., the probability of selecting exactly two faulty items is 0.91 %. Moreover, the probability that among that 10 selected items, there are at most two faulty ones is obtained as follows:

$$\begin{aligned}\sum_{k=0}^{2} P(X_{10,0.015} = k) &= \binom{10}{0} \cdot 0.015^0 \cdot 0.985^{10-0} + \binom{10}{1} \cdot 0.015^1 \cdot 0.985^{10-1} \\ &\quad + \binom{10}{2} \cdot 0.015^2 \cdot 0.985^{10-2} \\ &= 0.8597 + 0.1309 + 0.0091 = 0.9997\,,\end{aligned}$$

i.e., the probability of selecting at most two faulty items is 99.97 %.

12.7.4 Normal Distribution

Among the continuous distributions, the **normal distribution** is the most important one. Its importance results from the central limit theorem, according to which random variables resulting from the sum of a large number of independent variables are approximately normally distributed.

Normally distributed random variable X:
The density function of a **normally distributed random variable** X with the expected value μ and the variance σ^2 is given by

$$f(x) = \frac{1}{\sqrt{2\pi\,\sigma^2}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Here the expectation value μ and the variance σ^2 are the parameters of the normal distribution. As an abbreviation, we write that the random variable X is $N(\mu, \sigma^2)$-distributed. The density function f is symmetric with respect to the point $x = \mu$. It has the maximum at $x = \mu$ and two inflection points at $x_1 = \mu - \sigma$ and $x_2 = \mu + \sigma$.

For the probability $P(X \leq x)$, we get

$$F(x) = P(X \leq x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} e^{-\frac{(z-\mu)^2}{2\sigma^2}}\, dz\,.$$

If $\mu = 0$ and $\sigma = 1$, we have the **standard normal distribution**. In this case, the density function simplifies to

$$f(z) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}z^2}\,.$$

For transforming an $N(\mu, \sigma^2)$-distributed random variable X into an $N(0,1)$-distributed random variable Z, one can use the substitution

$$z = \frac{x-\mu}{\sigma}\,.$$

The distribution function of an $N(0,1)$-distributed random variable Z is tabulated and often denoted as $\Phi(z)$. We have

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) = \Phi(z)\,.$$

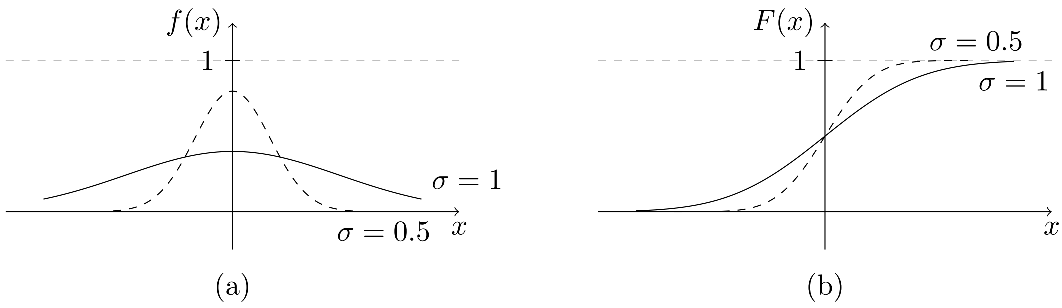


Figure 12.3: (a) Density function and (b) distribution function of an $N(0, \sigma^2)$-distributed random variable X

Often, the following relationship for the standard normal distribution is useful in the computations which holds due to symmetry:

$$\Phi(-z) = 1 - \Phi(z) \ .$$

In Fig. 12.3, the density and the distribution functions of an $N(0, \sigma^2)$-distributed random variable X are given for $\sigma = 0.5$ and $\sigma = 1$.

Example 12.28 *It is known that the length X of a workpiece is $N(\mu, \sigma^2)$-distributed with $\mu = 60$ cm and $\sigma = 2$ cm. We determine the probability that the length of the workpiece is*

(a) not greater than 61 cm;
(b) less than 58 cm;
(c) in the interval [57 cm, 62 cm] and
(d) does not deviate by more than 3 cm from the expectation value μ.

(a) We obtain

$$P(X \leq 61) = \Phi\left(\frac{61-60}{2}\right) = \Phi(0.5) = 0.69146$$

which corresponds (approximately) to 69.15 %.
(b) We get

$$P(X < 58) = \Phi\left(\frac{58-60}{2}\right) = \Phi(-1) = 1 - \Phi(1) = 1 - 0.84134 = 0.15866$$

which corresponds to 15.87 %.
(c) In this case, we obtain

$$\begin{aligned} P(57 \leq X \leq 62) &= \Phi\left(\frac{62-60}{2}\right) - \Phi\left(\frac{57-60}{2}\right) = \Phi(1) - \Phi(-1.5) \\ &= \Phi(1) + \Phi(1.5) - 1 = 0.84134 + 0.93319 - 1 = 0.77453 \end{aligned}$$

corresponding to 77.45 %.
(d) Here we obtain

$$\begin{aligned} P(|X-60| < 3) &= P(57 \leq x \leq 63) = \Phi\left(\frac{63-60}{2}\right) - \Phi\left(\frac{57-60}{2}\right) \\ &= \Phi(1.5) - \Phi(-1.5) = 2\Phi(1.5) - 1 = 0.86638 \end{aligned}$$

corresponding to 86.64 %.

Example 12.29 *For an $N(\mu, \sigma^2)$-distributed random variable X with $\mu = 90$ cm and $\sigma = 5$ cm, we determine the following probabilities:*

(a) $P(80 \le X < 97)$;
(b) $P(X > 100)$;
(c) $P(|X - \mu| > 10$.

(a) We obtain

$$\begin{aligned} P(80 \le X < 97) &= \Phi\left(\frac{97-90}{5}\right) - \Phi\left(\frac{80-90}{5}\right) = \Phi(1.4) - \Phi(-2) \\ &= \Phi(1.4) + \Phi(2) - 1 = 0.91924 + 0.97725 - 1 = 0.89649\,. \end{aligned}$$

(b) We get

$$P(X > 100) = 1 - P(X \le 100) = 1 - \Phi\left(\frac{100-90}{5}\right) = 1 - \Phi(2) = 1 - 0.97725 = 0.02275\,.$$

(c) Here we get

$$\begin{aligned} P(|X - \mu| > 10 &= P(X > 100) + P(X < 80) = 1 - P(X \le 100) + P(X < 80) \\ &= \left[1 - \Phi\left(\frac{100-90}{5}\right)\right] + \Phi\left(\frac{80-90}{5}\right) \\ &= [1 - \Phi(2)] + \Phi(-2) = 2 \cdot [1 - \Phi(2)] = 0.0455\,. \end{aligned}$$

Example 12.30 *Determine the value a such that for an $N(\mu, \sigma^2)$-distributed random variable X the interval $[\mu - a, \mu + a]$ contains 80 % of all values. For a $N(0, 1)$-distributed random variable Z, we get*

$$P(-z \le Z \le z) = 2\Phi(z) - 1 = 0.80$$

from which we obtain

$$\Phi(z) = \frac{1.80}{2} = 0.90\,.$$

Using the inverse function Φ^{-1} (or directly the table of the standard normal distribution), we get $z = 1.28$. Consequently, for an $N(0, 1)$-distributed random variable Z, 80 % of the realizations are from the interval $[-1.28, 1.28]$ and therefore, for an $N(\mu, \sigma^2)$-distributed random variable, 80 % of the realizations are from the interval $[\mu - 1.28\sigma, \mu + 1.28\sigma]$.

12.8 Statistical Tests

Finally, we briefly review some knowledge about statistical tests. The objective of such a test is to check a particular hypothesis H_0, called the **null hypothesis**. The opposite assumption is called the **alternative hypothesis** H_1.

In general, such a statistical hypothesis test includes the following steps:

1. Formulate the null hypothesis H_0 and the alternative hypothesis H_1;
2. Fix the significance level α;
3. Determine the region of acceptance and the region of rejection (critical region);

4. Take a sample;

5. Make a decision: either reject the null hypothesis or 'accept' it (i.e., H_0 is not rejected).

In the following, we explain one statistical **significance test** concerning the parameter (probability) p in a binomial distribution. The first question arises in connection with the formulation of the null and alternative hypotheses H_0 and H_1. There exist different variants:

1. $H_0 : p = p_0$; $\qquad H_1 : p \neq p_0$;

2. $H_0 : p \leq p_0$; $\qquad H_1 : p > p_0$;

3. $H_0 : p \geq p_0$; $\qquad H_1 : p < p_0$.

In the first case, we have a **two-side test** because deviations in both directions are considered. In the other two cases, we have **one-side tests** because only deviations in one direction are relevant. The selection of a particular variant depends on the specific interest of different groups of people interested in the results of the test. To illustrate, consider the following example. A wine store claims that a particular type of wine contains $p_0 = 12\%$ alcohol. A consumer protection organization is doubtful and conjectures that it contains less than 12% alcohol. In this case, it would be a test of the null hypotheses $p \geq p_0$ against the alternative hypothesis $p < p_0$. On the other side, an organization for legal protection for young people might conjecture that the alcoholic level is higher. In this case, one would test the null hypothesis $p \leq p_0$ against the alternative hypothesis $p > p_0$. In a simplified manner, one can see that the null hypothesis describes the 'current state' and in principle, one wishes to know whether there are reasons for the alternative hypothesis. Independently of the chosen variant, the calculations in the test are made for $p = p_0$.

Since statistical tests are based on probabilities, one has no guarantee to find in any case the right decision, i.e., is the null hypothesis indeed true or not. However, one wants to limit the risk of a wrong decision. There are two types of a possible error. First, one may reject the null hypothesis although it is true. This is know as a **type-1 error**. The other error arises when the null hypothesis is false but it is accepted. This kind of error is known as a **type-2 error**. This is summarized in Table 12.1.

Table 12.1: Type-1 and type-2 error in statistical tests.

	H_0 is true	H_1 is true
H_0 accepted	right decision	type 2 error (wrong decision)
H_0 rejected	type 1 error (wrong decision)	right decision

The simplest case is when only two possible decisions exist. For instance, consider a criminal trial where at the beginning the following two hypotheses are possible:

H_0: the defendant is innocent.
H_1: the defendant is guilty.

If the null hypothesis is rejected by the judge but the defendant is truly innocent, a type-1 error arises. However, if the null hypothesis is accepted by the judge but the defendant is truly guilty, a type-2 error occurs.

Typically, the major focus is on the type-1 error, i.e., one wishes to avoid to reject a null hypothesis although it is true. One introduces the probability

$$P(H_0 \text{ rejected} | H_0 \text{ is true }) = \alpha \,.$$

This probability should be small, and often one settles $\alpha = 0.05$ or $\alpha = 0.01$. The value of α is denoted as **significance level**, i.e., often the chosen significance level is 5 % or 1 %.

We next illustrate the determination of the **region of acceptance A** and the **region of rejection R** first for a one-side and then for a two-side test. For the hypothesis $H_0 : p \le p_0$, large values of the random variable do not confirm the null hypothesis. Therefore, for fixing the regions of acceptance and rejection, we consider

$$P(X \ge k) \le \alpha$$

and determine the smallest integer k satisfying the above inequality i.e., we determine the smallest number k such that the probability that an integer of at least k is observed is at most equal to the significance level α. In this case, we apply a **right-side test** and select

$$\mathbf{A} = \{0, 1, \ldots, k-1\} \qquad \text{and} \qquad \mathbf{R} = \{k, k+1, k+2, \ldots, n\}\,.$$

Accordingly, for a **left-side test**, we consider

$$P(X \le k) \le \alpha$$

and determine the largest integer k satisfying the above inequality i.e., we determine the largest number k such that the probability that an integer of at most k is observed is at most equal to the significance level α. In this case, we fix the regions of acceptance ($\mathbf{A}$) and rejection ($\mathbf{R}$) as follows:

$$\mathbf{A} = \{k+1, k+2, \dots, n\} \qquad \text{and} \qquad \mathbf{R} = \{0, 1, \dots, k\}\,.$$

When applying a two-side test, both small and large values do not confirm the null hypothesis. So we determine $\mathbf{A}$ and $\mathbf{R}$ such that the following two inequalities are satisfied:

$$P(X \le k_L) \le \frac{a}{2} \qquad \text{and} \qquad P(X \ge k_R) \le \frac{\alpha}{2}\,,$$

i.e., the significance level is split into equal parts for very small and large values. Again, we choose the largest integer satisfying the first inequality as k_L and the smallest integer satisfying the second inequality as k_R. Then we get the region of acceptance

$$\mathbf{A} = \{k_L+1, k_L+2, \dots, k_R-1\}$$

and the region of rejection

$$\mathbf{R} = \{0, 1, \dots, k_L, k_R, k_R+1, \dots, n\}\,.$$

The region of acceptance $\mathbf{A}$ and the region of rejection $\mathbf{R}$ are illustrated in Fig. 12.4 for a left-side and a two-side significance test. In Fig. 12.4, we used the binomial distribution with the parameters $n = 10$ and $p = 0.4$.

We illustrate the above variants of a significance test on the following three examples.

Example 12.31 *A car company gets parts from a supplier, where it is known that approximately 20 % of the parts are faulty (and must be reworked). Now a new supplier contacts the company and claims that the company can produce the same parts such that only 10 % of the parts are faulty. However, the car company would like to check and takes a sample of $n = 50$ parts among which 7 are faulty. Here one applies a right-side test with the hypotheses*

$$H_0 : p \le 0.1; \qquad H_1 : p > 0.1\,.$$

The random variable X denotes the number of faulty parts. The sample size is $n = 50$ and the chosen significance level is 5 % ($\alpha = 0.05$). If the null hypothesis H_0 is true, the random variable is binomially distributed with the parameters $n = 50$ and $p = 0.1$. We use

$$P(X \ge k) \le 0.05\,.$$

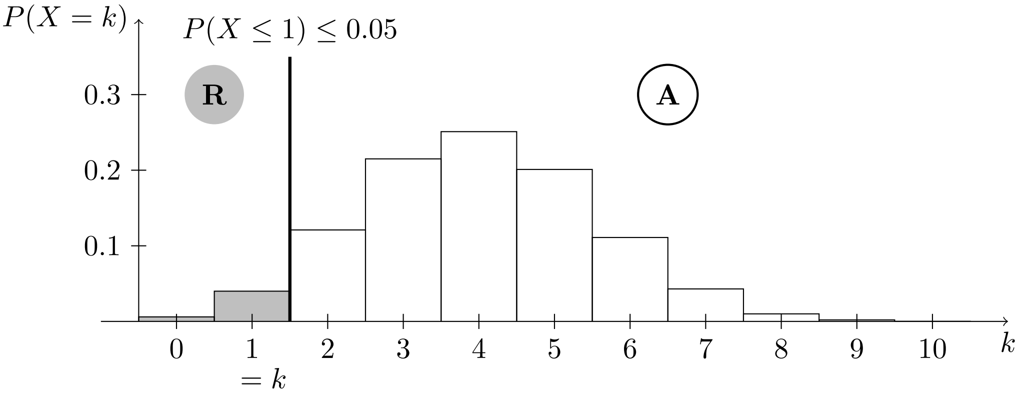
From the table of the cumulative binomial distribution, we get the smallest integer $k = 10$ satisfying the above inequality. Thus, we get the regions of acceptance ($\mathbf{A}$) and rejections ($\mathbf{R}$) as follows:

$$\mathbf{A} = \{0, 1, 2, \dots, 9\} \qquad \textit{and} \qquad \mathbf{R} = \{10, 11, \dots 50\}\,.$$

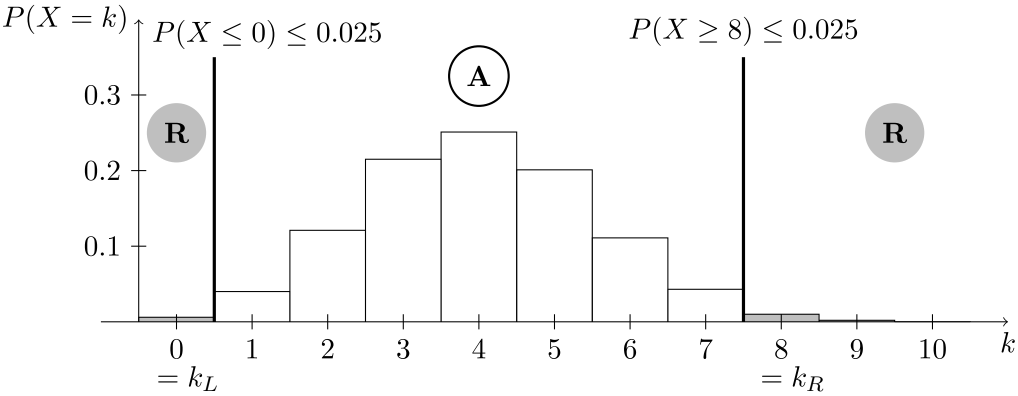
(we find $P(X \ge 10) = 0.025 < 0.05$ while $P(X \ge 9) = 0.058 > 0.05$). The null hypothesis is accepted since $7 \in \mathbf{A}$.

Example 12.32 *At the last elections, a particular party reached 30 % of the votes. However, in a recent survey, only 21 out of 100 people confirmed that they will vote next time again for this party. Can one conclude with a significance level of 5 % that the votes for this party will decrease? We use the hypotheses*

$$H_0 : p \ge 0.3; \qquad H_1 : p < 0.3\,.$$



(a) Left-side test



(b) Two-side test

Figure 12.4: Region of acceptance (**A**) and region of rejection (**R**) of a significance test

The sample size is $n = 100$ *and* $\alpha = 0.05$. *The random variable* X *denotes the number of votes for the party. If the null hypothesis is true, the random variable* X *is binomially distributed with the parameters* $n = 100$ *and* $p = 0.3$. *For this left-side test, we use*

$$P(X \leq k) \leq 0.05 \, .$$

From the table of the cumulative binomial distribution, we obtain the largest integer $k = 22$ *satisfying this inequality (we find* $P(X \leq 22) = 0.048 < 0.05$ *while* $P(X \leq 23) = 0.076 > 0.05$*). Therefore, we have*

$$\mathbf{A} = \{23, 24, \ldots, 50\} \qquad \text{and} \qquad \mathbf{R} = \{0, 1, \ldots 22\} \, .$$

The null hypothesis is rejected since $21 \in \mathbf{R}$.

Example 12.33 *It is conjectured that in an urn, there are 40 % white spheres (while the other 60 % are assumed to be black spheres). A sample of 20 spheres included 6 white spheres. If one chooses a significance level of* 5%*, will the hypothesis be rejected or not? Here we apply a two-side test. We use*

$$H_0 : p = 0.4 \qquad \text{and} \qquad H_1 : p \neq 0.4 \, .$$

The random variable X *denotes the number of white spheres. If the null hypothesis is correct, the random variable* X *is binomially distributed with the parameters* $n = 20$ *and* $p = 0.4$. *For this two-side test, we find from*

$$P(X \leq k_L) \leq \frac{\alpha}{2} = 0.025 \qquad \text{and} \qquad P(X \geq k_R) \leq \frac{\alpha}{2} = 0.025$$

$k_L = 3$ *(*$P(X \leq 3) = 0.016$*) and* $k_R = 13$ *(*$P(X \geq 13 = 0.021$*). Then we obtain the regions of acceptance*

$$\mathbf{A} = \{4, 5, \ldots, 12\} \, .$$

and the region of rejection

$$\mathbf{R} = \{0, 1, 2, 3, 13, 14, 15, 16, 17, 18, 19, 20\} \, .$$

The null hypothesis is accepted since $6 \in \mathbf{A}$.

EXERCISES

12.1 Four jobs J_1, J_2, J_3, J_4 have to be processed on a machine 1, and three jobs J_5, J_6, J_7 have to be processed on machine 2. How many different processing sequences of the jobs on the two machines are possible?

12.2 How many numbers containing each of the digits $1, 2, \ldots, 9$ exactly one can be formed which start with the digits 35 and end with digit 8?

12.3 How many distinct numbers of the eight digits $\{1, 1, 2, 2, 2, 2, 2, 3, 4, 4\}$ can be formed?

12.4 In an election of a commission in a school, 10 teachers, 25 pupils and 5 persons of the remaining staff take part. How many possibilities of different commissions do exist if from the 10 teachers exactly 5, from the 25 pupils exactly 3 and from the remaining staff exactly 2 persons have to be included into the commission?

12.5 In a questionnaire, one has to answer eight questions either with 'I agree', 'I partly agree' or 'I disagree'. How many different fillings of the questionnaire are possible?

12.6 A tourist has 12 exciting sights in the neighborhood of his resort. Unfortunately he can only visit four of them on one day. How many possibilities do exist
(a) to select four sights out of the 12 possible ones;
(b) to form a 4-sight subtour (where the sequence in which the sights are visited is important).

12.7 Consider a skat game with 32 cards. If one randomly selects one card, what is the probability
(a) to select a green card;
(b) to select a card with a number (i.e., 7, 8, 9, or 10);
(c) to select a green card with a number;
(d) to select a green card or a card with a number.

12.8 A big city has three soccer teams denoted by 1, 2, and 3 in the first league who play all at a particular weekend. Consider the events
A_i: Team i wins ($i = 1, 2, 3$).
Describe the events:
B: At least one of the three teams wins;
C: At most two of the three teams win;
D: None of the three teams wins.

12.9 A discrete random variable may take on the values $x_1 = -2, x_2 = -1, x_3 = 0, x_4 = 1, x_5 = 2, x_6 = 3$ with the probabilities

$$P(X = x_1) = 0.05; \quad P(X = x_2) = 0.08; \quad P(X = x_3) = 0.19;$$

$$P(X = x_4) = 0.22; \quad P(X = x_5) = 0.35; \quad P(X = x_6) = 0.11.$$

Determine the expectation value $E(X)$ and the variance $\sigma^2(X)$. Graph the distribution function $F(x)$.

12.10 Consider a continuous random variable X with the density function

$$f(x) = 0.5\, e^{-0.5x} \qquad (x \geq 0)\,.$$

Determine the probability $P(0 \leq X < 2)$.

12.11 It is known that in a mass production, 80 % of the items are of good quality while the other 20 % must be reworked. If 20 items are randomly selected, determine the probability that
(a) at least 15 items are of good quality;
(b) at least 2 items are not of good quality;
(c) exactly 3 items are not of good quality.

12.12 Is is known that a sportsmen in biathlon has a probability of 90 % of hitting the target with one shot. Determine the probability that with 10 shots
(a) at most eight times the target is hit;
(b) exactly seven times the target is hit;
(c) no more than five times the target is hit.

12.13 It is known that the length X of a workpiece is $N(\mu, \sigma^2)$-distributed with $\mu = 32$ cm and $\sigma = 1.5$ cm. Determine the probability that the length of the workpiece

(a) is not greater than 35 cm;
(b) is less than 26 cm;
(c) is in the interval [28 cm, 35 cm] and
(d) deviates by more than 6 cm from the expectation value μ.

12.14 Determine the probability that a $N(0,1)$-distributed random variable X is in the interval $[-2, 2]$.

12.15 A machine engineering company gets parts from a supplier, where it is known that approximately 90 % of the parts are of good quality. Now an alternative supplier contacts the company that the same parts with 95 % of the parts of good quality. The company takes a sample of $n = 30$ parts among which 26 are of good quality. Apply a left-side test with the hypotheses

$$H_0 : p \geq 0.95\,; \qquad H_1 : p < 0.95\,.$$

Will the null hypothesis be accepted or rejected?

12.16 It is conjectured that in an urn, there are 40 % blue (and 60 % red) spheres. A sample of 30 spheres included 10 blue spheres. Apply a two-side test with a significance level of 5% and the hypotheses

$$H_0 : p = 0.4 \qquad \text{and} \qquad H_1 : p \neq 0.4\,.$$

Will the null hypothesis be accepted or rejected?

List of Notations

$a \in A$	a is element of set A
$b \notin A$	b is not an element of set A
$\varnothing$	empty set (also used for impossible event in probability theory)
(a, b)	open interval between a and b
$[a, b]$	closed interval between a and b
$A \subseteq B$	set A is a subset of the set B
$A \cup B$	union of the sets (or events) A and B
$A \cap B$	intersection of the sets (or events) A and B
$A \setminus B$	difference of sets A and B
$\sum$	summation sign: $\sum_{i=1}^{n} a_i = a_1 + a_2 + \ldots + a_n$
$\prod$	product sign: $\prod_{i=1}^{n} a_i = a_1 \cdot a_2 \cdot \ldots \cdot a_n$
$\pm$	denotes two cases of a mathematical term: the first one with sign $+$ and the second one with sign $-$
$\mathbb{N}$	set of all natural numbers: $\mathbb{N} = \{1, 2, 3, \ldots\}$
$\mathbb{Z}$	union of the set $\mathbb{N}$ with the set of all negative integers and number 0
$\mathbb{Q}$	set of all rational numbers, i.e., set of all fractions $\frac{p}{q}$ with $p \in \mathbb{Z}$ and $q \in \mathbb{N}$
$\mathbb{R}$	set of all real numbers
$\mathbb{R}_{>a}$	set of all real numbers greater than a
$\mathbb{R}_{\geq a}$	set of all real numbers greater than or equal to a
$\neq$	sign 'not equal'
$\approx$	sign of approximate equality, e.g. $\sqrt{2} \approx 1.41$
π	irrational number: $\pi \approx 3.14159...$
e	Euler's number: $e \approx 2.71828...$
∞	infinity
$\lvert a \rvert$	absolute value of the number $a \in \mathbb{R}$
$n = 1, 2, \ldots, k$	equalities $n = 1, n = 2, \ldots, n = k$
$\sqrt{a}$	square root of a

log	notation used for the logarithm: if $y = \log_a x$, then $a^y = x$
lg	notation used for the logarithm with base 10: $\lg x = \log_{10} x$
ln	notation used for the logarithm with base e: $\ln a = \log_e x$
$\{a_n\}$	sequence: $\{a_n\} = a_1, a_2, a_3, \ldots, a_n, \ldots$
$\{s_n\}$	nth partial sum of a sequence $\{a_n\}$
lim	limit sign
$y = f(x)$	$y \in \mathbb{R}$ is the function value of $x \in \mathbb{R}$
D_f	domain of a function f of a real variable
R_f	range of a function f of a real variable
f^{-1}	inverse mapping or function of f
$g \circ f$	composite function of f and g
$x \to x_0$	x tends to x_0
$x \to x_0 + 0$	x tends to x_0 from the right-side
$x \to x_0 - 0$	x tends to x_0 from the left-side
f'	derivative of the function f
$f'(x)$, $y'(x)$	derivative of the function f with $y = f(x)$ at the point x
$f''(x)$, $y''(x)$	second derivative of the function f with $y = f(x)$ at the point x
$f^{(n)}(x)\, y^{(n)}(x)$	nth derivative of the function f with $y = f(x)$ at the point x
$dy, \quad df$	differential of function f with $y = f(x)$
$\int$	integral sign
$\mathbf{a}$	vector: ordered n-tuple of real numbers
$\mathbf{a}^T$	transposed vector of the vector $\mathbf{a}$
$\lvert\mathbf{a}\rvert$	length (or norm) of the vector $\mathbf{a}$
$\lvert\mathbf{a} - \mathbf{b}\rvert$	distance between the vectors $\mathbf{a}$ and $\mathbf{b}$
$\mathbf{a} \perp \mathbf{b}$	means that the vectors $\mathbf{a}$ and $\mathbf{b}$ are orthogonal
$\mathbf{a} \cdot \mathbf{b}$	scalar product of the vectors $\mathbf{a}$ and $\mathbf{b}$
$\mathbf{a} \times \mathbf{b}$	vector product of the vectors $\mathbf{a}$ and $\mathbf{b}$
$n!$	n factorial: $n! = 1 \cdot 2 \cdot \ldots \cdot (n-1) \cdot n$
$P(n)$	number of permutations of n elements
$V(k, n)$	number of variations (k out of n) with repeated selection
$\overline{V}(k, n)$	number of variations (k out of n)
$\binom{n}{k}$	binomial coefficient: $\binom{n}{k} = \dfrac{n!}{k! \cdot (n-k)!}$
$C(k, n)$	number of combinations (k out of n) without repeated selection

$\overline{C}(k,n)$	number of combinations (k out of n) with repeated selection
A	event
S	certain event
$\overline{A}$	complementary event
$H(A)$	absolute frequency of the event A
$h(A)$	relative frequency of the event A
$P(A)$	probability of the event A
$P(A\|B)$	conditional probability of the event A, given event B
X	random variable
$P(X = x_i)$	probability that the random variable X takes a value x
$P(X \leq x)$	probability that the random variable X takes a value at most equal to x
$E(X)$	expectation value of the random variable X
$\sigma^2(X)$	variance of the random variable X
$X_{n,p}$	the random variable X is binomially distributed with the parameters n and p
$N(\mu, \sigma^2)$	indicated that a random variable X is normally distributed with the parameters μ and σ^2
$\Phi(z)$	distribution function of the normal standard distribution
H_0	null hypothesis of a statistical test
H_1	alternative hypothesis of a statistical test

List of figures

Index