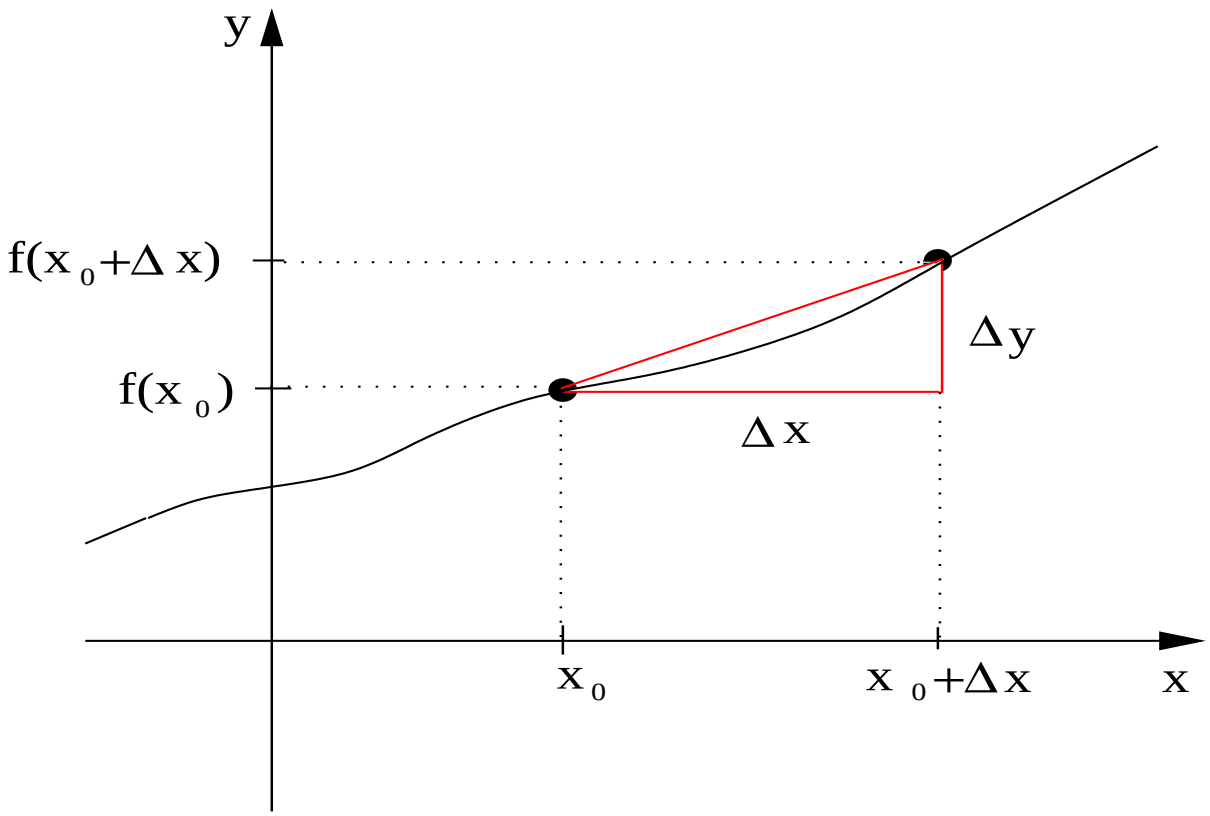
The University of Sussex – Department of Mathematics

G1103 – Mathematical Methods 1 (for Physicists)

Lecture Notes – Autumn Term 2009

Kerstin Hesse



Contents

Introduction

This course is an **introductory course to (higher) mathematics**, and it covers a wide range of material which is needed in **physics**. In contrast to mathematics lectures for mathematics students, where proof techniques play a crucial role, this course focusses on **understanding and applying mathematical techniques** that are often needed for the solution of problems in physics.

Mathematics and mathematical techniques **cannot be learnt** by just attending and following the lectures! The lecture can only give an introduction to the mathematical concepts and techniques. It is crucial that you **spend some time to familiarize yourself with the mathematical concepts and techniques and then apply them yourself to solve the problems on the weekly exercise sheets**!

In the weekly workshops, the workshop tutors will provide help with solving the problems on the exercise sheets, but ideally you should have tried the problems yourself beforehand and then ask the workshop tutors for help with those problems you could not solve. The more problems you solve, the better you will learn the material and the better you will be prepared. Therefore it is recommendable to solve **all** provided problems and not just those to which solutions have to be handed in.

Note that understanding the solution to a problem is **not** the same as solving a problem yourself. **Finding the solution yourself without help, or even with some help, provides a far better understanding and learning effect than understanding someone else solution!**

The key to solving most of the problems on the exercise sheets is **firstly to know which mathematical technique or concept you have to apply** and secondly to **have understood the needed technique/concept!** Once the needed mathematical technique is identified and understood, finding the solution to the problem is relatively easy. Therefore it is the wrong approach to focus only on the problems that you have to hand in every week. **Instead you should first study the**

worked examples and understand the new mathematical techniques and concepts that are needed for the problems on the current exercise sheet. Then you will also find that most problems are rather straight-forward.

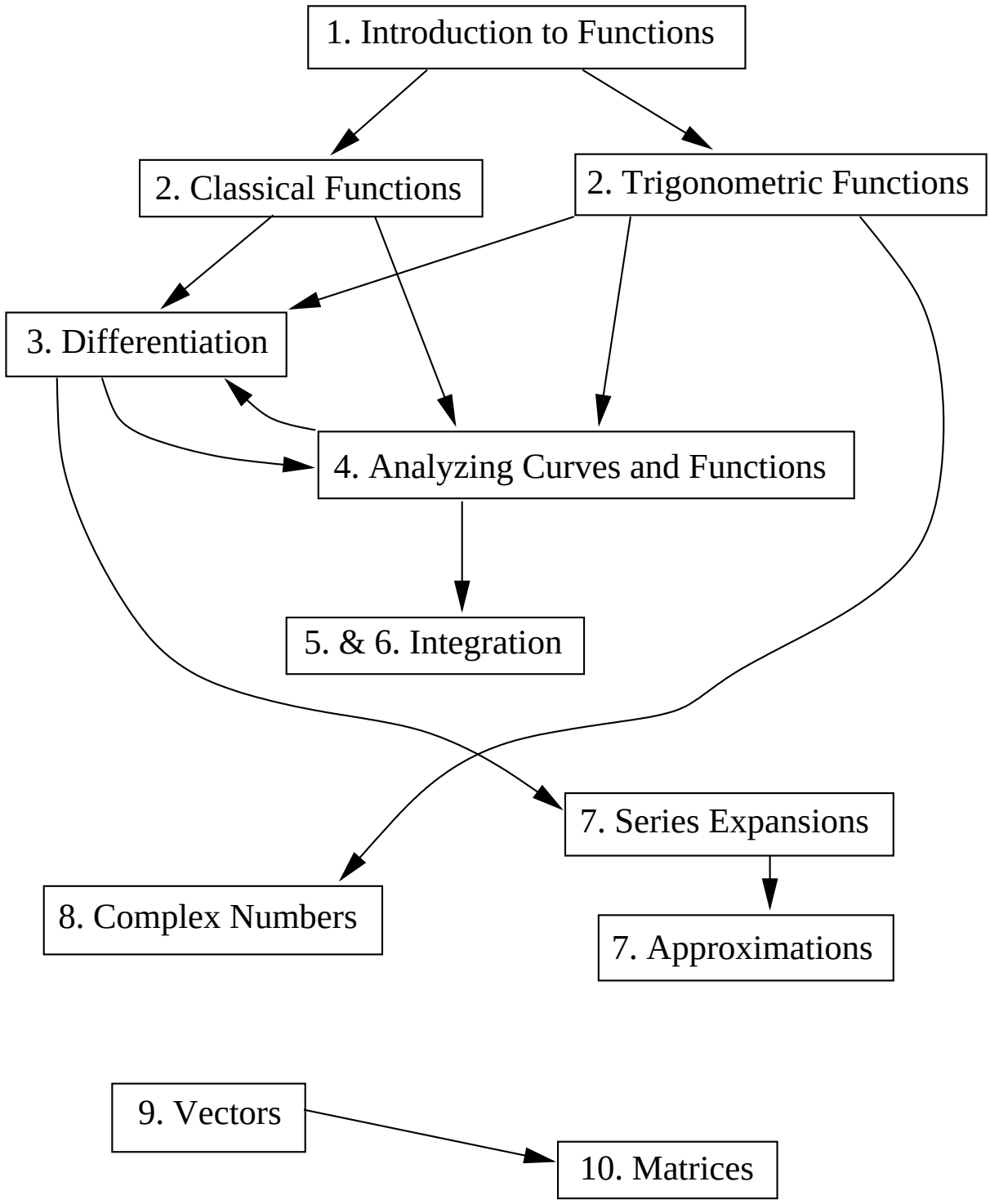


Figure 1: Map of the lecture material

The course covers mostly topics from **analysis** and **calculus**, but the last two chapters of the lecture give an introduction to the basics of **linear algebra**. The connection between the different topics is indicated in the diagram in Figure 1 below.

As a lecturer I am aware that your mathematical knowledge from school may vastly differ within the class. Therefore, I will try to assume as little as possible, but I have to assume that some basics are known. If you are not familiar with some notation or some basic concept, please let me know! Due to the differences in mathematical background knowledge within the class, some of you may be familiar with several of the topics and some of you may be familiar with only a few of the topics. I believe that most of you will find that you have not seen this arrangement of the material and the connection between the different topics in this form. You may also find that the presentation is different from how you were taught in school. I hope we will all enjoy the course.

Chapter 1

Introduction to Functions

Functions are of paramount importance in physics and engineering in order to describe physical phenomena such as movement and speed over a period of time, the change of the temperature over a period of time, as well as gravitation and electricity.

We start the chapter with a motivating example from physics, the so-called 'free fall', which describes the motion of an object that is dropped and falls downward, due to the gravitation of the earth.

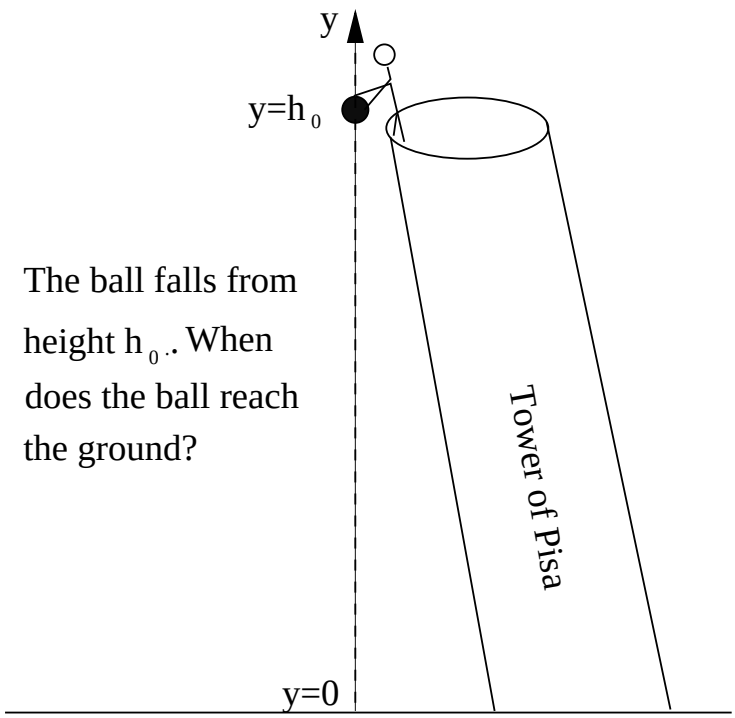


Figure 1.1: A ball is dropped from the Tower of Pisa.

Application 1.1 (free fall)
A man stands on top of the Tower of Pisa and drops a ball. Due to the gravitational acceleration, the ball falls down. Describe the motion of the ball and determine when it reaches the ground (see Figure 1.1)

Solution: *We assume for convenience that the* ***time*** t *is zero at the moment the ball is dropped. Thus* t *measures the time since the ball was dropped. We want to describe the* ***motion*** *of the ball as a function* $y = y(t)$*, where* $y(t)$ *is the* ***position of the ball above the ground at the time*** t*. The* ***(vertical) distance*** *traveled at the time* t *is known from physics to be*

$$s(t) = \frac{1}{2}\, g\, t^2,$$

where g *is the* ***gravitational acceleration*** $g \approx 9.8\,\mathrm{m/s^2}$ *(at the earth's surface). Since the ball is at height* $y(0) = h_0$ *at the time* $t = 0$ *and since the ball travels downward, we find that*

$$y(t) = h_0 - s(t) = h_0 - \frac{1}{2}\, g\, t^2. \tag{1.1}$$

At the ***time*** t_0 ***when the ball hits the ground*** *we have* $y(t_0) = 0$*, and thus*

$$y(t_0) = h_0 - \frac{1}{2}\, g\, t_0^2 = 0 \qquad \Rightarrow \qquad h_0 = \frac{1}{2}\, g\, t_0^2. \qquad \Rightarrow \qquad \frac{2\, h_0}{g} = t_0^2.$$

Since $t_0 > 0$*, we see that the ball hits the ground at the time*

$$t_0 = \sqrt{\frac{2\, h_0}{g}}.$$

At its higher side the Tower of Pisa has the height $h_0 = 56.7\,\mathrm{m}$*. Thus we find that*

$$t_0 \approx \sqrt{\frac{2 \times 56.7\,\mathrm{m}}{9.8\,\mathrm{m/s^2}}} = \sqrt{\frac{113.4}{9.8}}\,\sqrt{\frac{\mathrm{m\,s^2}}{\mathrm{m}}} \approx 11.6\,\mathrm{s},$$

that is, the ball falls 11.6 *seconds before it hits the ground. Of course, we have used a simplified scenario in this example and neglected other in influences on the ball like the aerodynamic resistance.* □

The function (1.1) that describes the **free fall** of the ball from the tower of Pisa, is a so-called **quadratic function**. The class of quadratic functions will be discussed in detail in Section 1.4 of this chapter.

Before we start with the actual mathematics, we introduce some common notation that we will use throughout the course.

Notation 1.2 (sets of numbers)

(a) $\mathbb{N} = \{1, 2, 3, \ldots\}$ *denotes the set of all* ***positive integers****.*

(b) $\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$ *denotes the set of all* ***non-negative integers****.*

(c) $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ *denotes the set of all* ***integers****.*

(d) $\mathbb{Q}$ *denotes the set of all* ***rational numbers****, that is, all numbers that can be represented as fractions* n/m*, where* n *and* m *are both integers and* $m \neq 0$*.*

(e) $\mathbb{R}$ *denotes the set of all* ***real numbers****. The real numbers contain all integers and rational numbers, but they contain also numbers like* $\sqrt{2}$*.*

Notation 1.3 (for sets)
Let A *and* B *be sets.*

(a) If x *is in* A*, that is,* x ***is an element of*** A*, then we write* $x \in A$*.*

(b) If x ***is not an element of*** A*, then we write* $x \notin A$*.*

(c) If all elements of A *are also elements of* B*, that is,* A ***is a subset of*** B*, then we write* $A \subset B$ *(or equivalently,* $B \supset A$*).*

(d) The set B ***without*** A*, denoted by* $B \setminus A$*, is the* ***set of all those numbers from*** B ***that are not in*** A*. In formulas,*

$$B \setminus A = \{x \in B : x \notin A\}\,.$$

(e) The set $\emptyset$ *(sometimes also denoted by* $\{\}$*) is called the* ***empty set****. It contains no elements at all; it is empty.*

Definition 1.4 (odd and even integers)
An integer $n \in \mathbb{Z}$ *is* ***even*** *if we can write it as* $n = 2\,m$ *with some* $m \in \mathbb{Z}$*. An integer is* ***odd*** *if it is not even.*

The even integers are $\ldots, -8, -6, -4, -2, 0, 2, 4, 6, 8, \ldots$, and the odd integers are $\ldots, -7, -5, -3, -1, 1, 3, 5, 7, \ldots$. We can describe the **set off all even integers** by

$$\text{set of even integers} = \{2k \,:\, k \in \mathbb{Z}\},$$

and we can describe the **set off all odd integers** by

$$\text{set of odd integers} = \{2k + 1 \,:\, k \in \mathbb{Z}\}.$$

For **intervals** we use the following standard notation:

Notation 1.5 (intervals)
Let $a, b \in \mathbb{R}$ with $a < b$. Then we have the following types of ***intervals****:*

$$\begin{aligned}
[a,b] &= \{x \in \mathbb{R} : a \le x \le b\} && \textit{(closed interval)},\\
(a,b) &= \{x \in \mathbb{R} : a < x < b\} && \textit{(open interval)},\\
(a,b] &= \{x \in \mathbb{R} : a < x \le b\} && \textit{(half-open interval)},\\
[a,b) &= \{x \in \mathbb{R} : a \le x < b\} && \textit{(half-open interval)}.
\end{aligned}$$

We see that the round parentheses '(' and ')' mean that the corresponding endpoint is not included in the interval, whereas the angular parentheses '[' and ']' mean that the corresponding endpoint is included in the interval. We also have the following special types of ***unbounded intervals*** *which have ∞ (infinity) or $-\infty$ (minus infinity) as an endpoint that is not included in the interval.*

$$\begin{aligned}
[a,\infty) &= \{x \in \mathbb{R} : a \le x < \infty\} = \{x \in \mathbb{R} : x \ge a\} && \textit{(half-open interval)},\\
(a,\infty) &= \{x \in \mathbb{R} : a < x < \infty\} = \{x \in \mathbb{R} : x > a\} && \textit{(open interval)},\\
(-\infty,b] &= \{x \in \mathbb{R} : -\infty < x \le b\} = \{x \in \mathbb{R} : x \le b\} && \textit{(half-open interval)},\\
(-\infty,b) &= \{x \in \mathbb{R} : -\infty \le x < b\} = \{x \in \mathbb{R} : x < b\} && \textit{(open interval)},\\
(-\infty,\infty) &= \{x \in \mathbb{R} : -\infty < x < \infty\} = \mathbb{R} && \textit{(open interval)}.
\end{aligned}$$

We call them ***unbounded*** *intervals because the numbers in any of these intervals get either arbitrarily large (if the upper endpoint is ∞) or arbitrarily small (if the lower endpoint is $-\infty$).*

For example, with this new notation we have

$$[2,7] = \{x \in \mathbb{R} : 2 \le x \le 7\}$$

and

$$[1,2) \subset [1,5] \qquad \text{and} \qquad [1,5] \setminus [1,2) = [2,5].$$

1.1 Functions and Graphs

In this section we learn what a **function** is and discuss some examples. We also learn how to **plot** functions.

Definition 1.6 (function)
*A **function** $f : A \to B$ is a rule that associates with each element x of some set A **one** particular element $y = f(x)$ of some set B. A is called the **domain** of the function, and B is called the **codomain** of the function. We call x the **independent variable** and $y = f(x)$ the **dependent variable**.*

It is also common to write $f : x \mapsto f(x)$ to indicate that $x \in A$ is mapped onto $f(x) \in B$.

Usually A and B are subsets of the real numbers $\mathbb{R}$, that is, $A \subset \mathbb{R}$ and $B \subset \mathbb{R}$. Often A and B are $\mathbb{R}$ or an interval or $\mathbb{R}$ without some point.

Example 1.7 (function $f(x) = x$)
An example of a function is

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = x.$$

Here $\mathbb{R}$ is the domain, and $\mathbb{R}$ is also the codomain. □

Example 1.8 (function $f(x) = 1/(2 - x)$)
Another example of a function is

$$f : \mathbb{R} \setminus \{2\} \to \mathbb{R}, \qquad f(x) = \frac{1}{2 - x}.$$

Here $\mathbb{R} \setminus \{2\}$ (that is, the real numbers without 2) is the domain, and $\mathbb{R}$ is the codomain. □

To discuss functions it would be helpful to have a graphical representation of the function, a so-called **plot**.

Definition 1.9 (graph of a function)
*The **graph** $\mathcal{G}$ of a function $f : A \to B$ is the set of all pairs $(x, f(x))$ as x assumes all values in the domain A of f. In formulas,*

$$\mathcal{G} = \big\{ \big(x, f(x)\big) \,:\, x \in A \big\}.$$

*We can **plot/sketch** the graph by drawing a suitable collection of pairs $(x, f(x))$ from the graph in the (x, y)-plane and then connecting $(x_1, f(x_1))$ smoothly with $(x_2, f(x_2))$ if x_1 and x_2 are adjacent.*

Example 1.10 (plot of the graph of the function $f(x) = x$)
To plot the function $f : \mathbb{R} \to \mathbb{R}$, given by $f(x) = x$, from Example 1.7, we determine the graph of f. Since $f(x) = x$, the graph is given by

$$\mathcal{G} = \{(x, x) \,:\, x \in \mathbb{R}\},$$

and we see that the graph consists of all points on the diagonal of the (x, y)-plane. We are also interested in what happens with $f(x)$ if x gets larger and larger, that is, what happens with $f(x)$ for $x \to \infty$. Likewise we are interested in what happens with $f(x)$ if x gets smaller and smaller, that is, what happens with $f(x)$ for $x \to -\infty$. We have

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} x = \infty \qquad \text{and} \qquad \lim_{x\to-\infty} f(x) = \lim_{x\to-\infty} x = -\infty.$$

Thus $f(x) = x$ has the graph shown in Figure 1.2 below. □

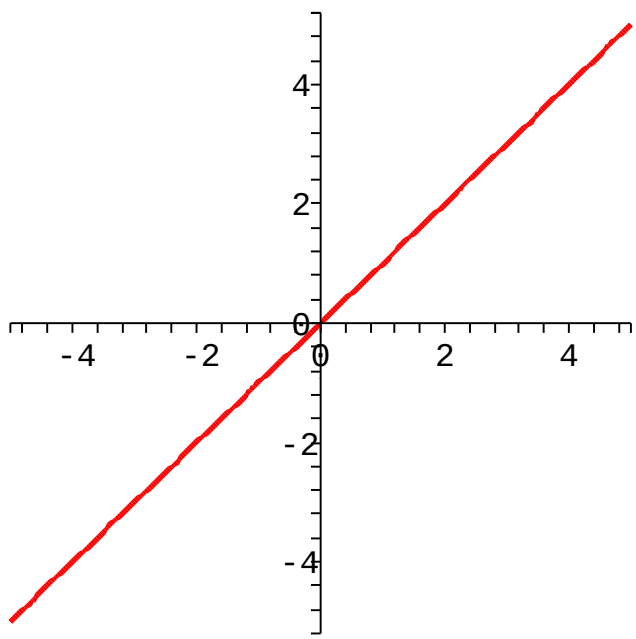


Figure 1.2: Graph of $f(x) = x$.

Before the next example we will introduce another definition.

Definition 1.11 (asymptote)
Let $f : A \to B$ be a function, and $A \subset \mathbb{R}$, $B \subset \mathbb{R}$. A ***straight line*** *is called an* ***asymptote*** *to the graph of the function f if* ***one*** *of the following two conditions is satisfied:*

(i) The straight line is vertical (to the x-axis) and goes through a point $(x_0, 0)$ and we have

$$\lim_{x \to x_0} |f(x)| = \infty.$$

(ii) The straight line can be described as an affine linear function, that is, as $g(x) = m\,x + c$, and we have either

$$\lim_{x \to \infty} \big(f(x) - g(x)\big) = 0 \qquad \textit{or} \qquad \lim_{x \to -\infty} \big(f(x) - g(x)\big) = 0.$$

Example 1.12 (plot of the graph of the function $f(x) = 1/(2-x)$)
The graph $\mathcal{G}$ of the function $f(x) = 1/(2-x)$ from Example 1.8 is given by

$$\mathcal{G} = \left\{ \left(x, \frac{1}{2-x} \right) \,:\, x \in \mathbb{R} \setminus \{2\} \right\},$$

and it is not so easy to visualize in the (x, y)-plane as the graph in the last example. First we observe that $f(x) > 0$ for $x < 2$ and that $f(x) < 0$ for $x > 2$. We observe that as $x < 2$ tends to 2, the function get arbitrary large because the denominator tends to zero as $x \to 2$, or, in other words, $f(x)$ tends to infinity as $x < 2$ tends to 2. In formulas, we describe this by

$$\lim_{x<2,\, x \to 2} \frac{1}{2-x} = \infty.$$

Likewise we find that

$$\lim_{x>2,\, x \to 2} \frac{1}{2-x} = -\infty.$$

At the point $x = 2$ the function f has a so-called singularity, and we see that the straight line that goes through $(2, 0)$ and is vertical (to the x-axis) is an **asymptote** to the graph of f. For large $|x|$, we see that $f(x)$ behaves like $1/(-x)$, that is,

$$f(x) = \frac{1}{2-x} \approx \frac{1}{-x} \qquad \text{for } |x| \text{ large.} \tag{1.2}$$

From (1.2), we also see that

$$\lim_{x \to \infty} \frac{1}{2-x} = 0 \qquad \text{and} \qquad \lim_{x \to -\infty} \frac{1}{2-x} = 0.$$

Thus the straight line given by $g(x) = 0$ is an **asymptote** for the graph of $1/(2-x)$. To get a sketch of the graph, we evaluate f at the points

$$x \in \{-2, -1, 0, 1, 1.5, 1.75, 2.25, 2.5, 3, 4, 5, 6, \}$$

and obtain

$$(-2, 1/4),\ (-1, 1/3),\ (0, 1/2),\ (1, 1),\ (1.5, 2),\ (1.75, 4),$$
$$(2.25, -4),\ (2.5, -2),\ (3, -1),\ (4, -1/2),\ (5, -1/3),\ (6, -1/4).$$

We have chosen the points closer together in the vicinity of $x = 2$ because the function values increase rapidly as we approach $x = 2$. Now we plot these points in the (x, y)-plane and draw a smooth curve through these points such that each point $(x, f(x))$ is connected via the graph with with those points with the next larger and next smaller x-coordinate. We also use our knowledge about the asymptotes of the graph of $1/(2-x)$ and that $f(x) > 0$ for $x < 2$ and that $f(x) < 0$ for $x > 2$. Thus we obtain the plot in Figure 1.3 below. □

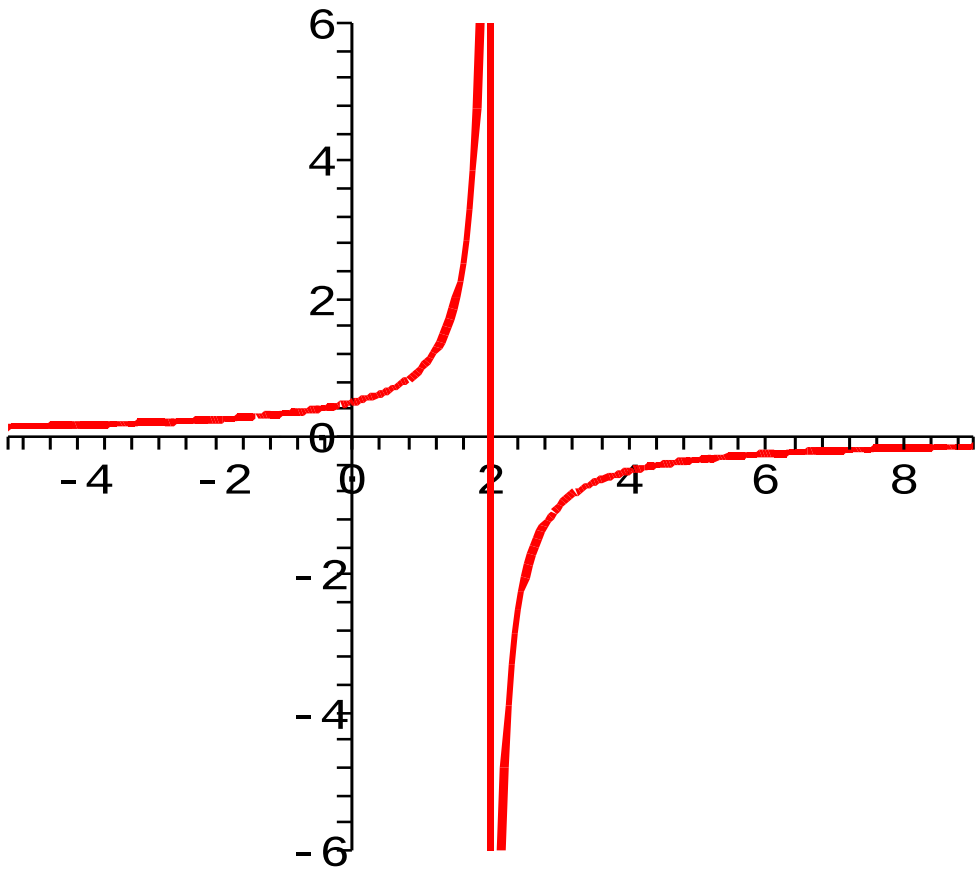


Figure 1.3: The graph of $f(x) = 1/(2-x)$.

1.2 Growth of Functions

In order to characterize functions it is helpful to know whether the **function values** $f(x)$ **increase or decline as** x **grows**.

Definition 1.13 ((strictly) monotonically increasing/decreasing)
Let $A \subset \mathbb{R}$, $B \subset \mathbb{R}$, and let $I \subset A$ be a subinterval of A. Let $f : A \to B$ be a function.

*(i) We say that f is **monotonically increasing on** I if for all $x_1, x_2 \in I$ with $x_1 < x_2$ we have $f(x_1) \leq f(x_2)$. If we have a strict inequality, that is, if for all $x_1, x_2 \in I$ with $x_1 < x_2$ we have $f(x_1) < f(x_2)$ then we call f **strictly monotonically increasing on** I.*

*(ii) We say that f is **monotonically decreasing on** I if for all $x_1, x_2 \in I$ with $x_1 < x_2$ we have $f(x_1) \geq f(x_2)$. If we have a strict inequality, that is, if for all $x_1, x_2 \in I$ with $x_1 < x_2$ we have $f(x_1) > f(x_2)$ then we call f **strictly monotonically decreasing on** I.*

Definition 1.13 is illustrated in Figure 1.4 below.

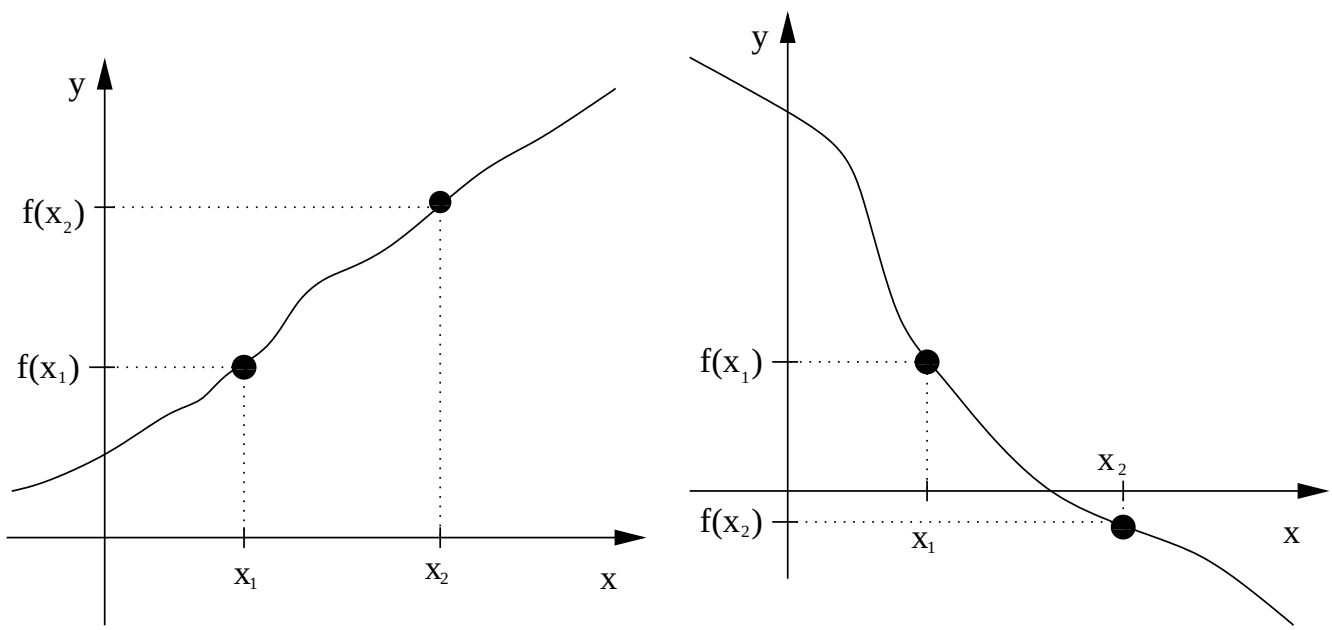


Figure 1.4: The function in the left picture is everywhere strictly monotonically increasing, and the function in the right picture is everywhere strictly monotonically decreasing.

Let us consider the functions from the examples in the previous section and also

some additional functions.

Example 1.14 ((strictly) monotonically increasing/decreasing)

(a) The function $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = x$, from Example 1.7 is strictly monotonically increasing on $\mathbb{R}$. Indeed, if $x_1, x_2 \in \mathbb{R}$ with $x_1 < x_2$, then

$$f(x_1) = x_1 < x_2 = f(x_2).$$

(b) The function $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = -x$, (see Figure 1.5) is strictly monotonically decreasing on $\mathbb{R}$. Indeed, if $x_1, x_2 \in \mathbb{R}$ with $x_1 < x_2$, then $-x_1 > -x_2$, and thus $f(x_1) > f(x_2)$.

(c) The function $f : \mathbb{R} \setminus \{2\} \to \mathbb{R}$, defined by $f(x) = 1/(2-x)$, from Example 1.8 is strictly monotonically increasing on $(-\infty, 2)$ and on $(2, \infty)$. To verify this let us first consider $x_1, x_2 \in (-\infty, 2)$ with $x_1 < x_2$. Then

$$f(x_1) = \frac{1}{2-x_1} < \frac{1}{2-x_2} = f(x_2) \qquad \text{because} \quad 0 < 2 - x_2 < 2 - x_1.$$

That $0 < 2 - x_2 < 2 - x_1$ can be seen as follows: from $x_1 < x_2 < 2$ we get (by subtracting 2) $x_1 - 2 < x_2 - 2 < 0$, and from multiplication by -1 we have $2 - x_1 > 2 - x_2 > 0$. Now let us consider $x_1, x_2 \in (2, \infty)$ with $x_1 < x_2$. Then

$$f(x_1) = \frac{1}{2-x_1} < \frac{1}{2-x_2} = f(x_2) \qquad \text{because} \quad 2 - x_2 < 2 - x_1 < 0.$$

That $2 - x_2 < 2 - x_1 < 0$ follows from $2 < x_1 < x_2$ and thus $0 < x_1 - 2 < x_2 - 2$ (from subtracting 2). Multiplying by (-1) yields $0 > 2 - x_1 > 2 - x_2$. □

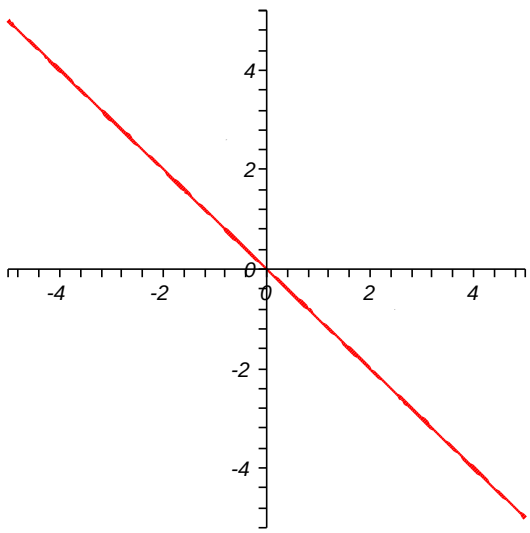


Figure 1.5: Graph of $f(x) = -x$.

We discuss some more examples.

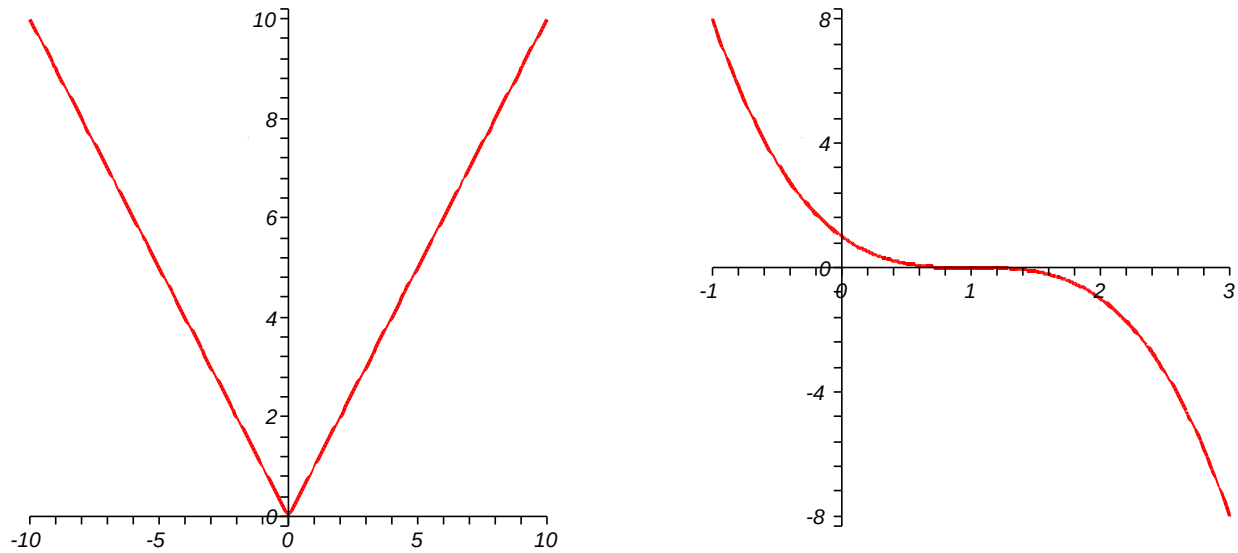


Figure 1.6: The graph of $f(x) = |x|$ on the left, and the graph of $f(x) = -(x-1)^3$ on the right.

Example 1.15 ((strictly) monotonically increasing/decreasing)

(a) The **absolute value function** $f : \mathbb{R} \to \mathbb{R}$, defined by (see picture on the left-hand side in Figure 1.6)

$$f(x) = |x| = \begin{cases} -x & \text{if} \quad x < 0, \\ 0 & \text{if} \quad x = 0, \\ x & \text{if} \quad x > 0, \end{cases}$$

is strictly monotonically decreasing on $(-\infty, 0]$ and strictly monotonically increasing on $[0, \infty)$. Indeed consider $x_1, x_2 \in (-\infty, 0]$, with $x_1 < x_2$. Then $-x_1 > -x_2$ and thus $f(x_1) > f(x_2)$. If $x_1, x_2 \in [0, \infty)$ with $x_1 < x_2$, then

$$f(x_1) = x_1 < x_2 = f(x_2).$$

(b) The function $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = -(x-1)^3$, (see picture on the right-hand side in Figure 1.6) is strict monotonically decreasing on $\mathbb{R}$.

To verify this we consider three cases:

(1) Let $x_1, x_2 \in \mathbb{R}$ with $x_1 < 1 < x_2$. Then $x_1 - 1 < 0$ and $x_2 - 1 > 0$, and thus $f(x_1) = -(x_1 - 1)^3 > 0$ and $f(x_2) = -(x_2 - 1)^2 < 0$. Therefore we clearly have $f(x_1) > f(x_2)$.

(2) Now we consider $x_1, x_2 \in \mathbb{R}$ with $x_1 < x_2 < 1$. Then $x_1 - 1 < 0$ and $x_2 - 1 < 0$ and therefore $f(x_1) = -(x_1 - 1)^3 > 0$, $f(x_2) = -(x_2 - 1)^3 > 0$. Thus $f(x_1) = |x_1 - 1|^3$ and $f(x_2) = |x_2 - 1|^3$. Since $x_1 < x_2 < 1$, we find that $x_1 - 1 < x_2 - 1 < 0$, and thus $|x_1 - 1| > |x_2 - 1|$. Therefore we finally get

$$f(x_1) = |x_1 - 1|^3 > |x_2 - 1|^3 = f(x_2).$$

(3) Finally we consider $x_1, x_2 \in \mathbb{R}$ with $1 < x_1 < x_2$. Then $0 < x_1 - 1 < x_2 - 1$, and thus $0 < (x_1 - 1)^3 < (x_2 - 1)^3$. Therefore

$$f(x_1) = -(x_1 - 1)^3 > -(x_2 - 1)^3 = f(x_2).$$

Thus we have verified that $f(x) = -(x-1)^3$ is indeed strictly monotonically decreasing on $\mathbb{R}$. □

1.3 One-to-one Functions and Inverse Functions

Let us consider $y = x + 2$. Then we can solve for x and obtain $x = y - 2$. If $y = x^2$, where $y > 0$, we have that $x = \pm\sqrt{y}$, that is, $y = x^2$ does not have a unique solution.

This leads us to the concept of an **inverse function**: The function $f(x) = x + 2$ has the inverse function $f^{-1}(y) = y - 2$. The function $g : \mathbb{R} \to \mathbb{R}$, $g(x) = x^2$ does **not** have an inverse function because we cannot uniquely solve $y = g(x) = x^2$ for $x \in \mathbb{R}$, since we get the two solutions $x = -\sqrt{y}$ and $x = \sqrt{y}$.

More formally, the question is whether we can **'invert'** a function: A function $f : A \to B$ associates with each $x \in A$ **exactly one** $y = f(x) \in B$; but for a given y in $\{f(x) : x \in A\}$, **how many** $x \in A$ satisfy $f(x) = y$? If for **each**

$$y \in \{f(x) : x \in A\},$$

there exists **only one** $x \in A$ with $f(x) = y$, then we can define the **inverse function** $f^{-1} : \{f(x) : x \in A\} \to A$ by $f^{-1}(y) = x$, where x is the **unique** element of A satisfying $f(x) = y$. A function f with this property are called **one-to-one**.

Definition 1.16 (one-to-one)
*Let $f : A \to B$ be a function from some set A into another set B. We say that f is **one-to-one** if $f(x_1) = f(x_2)$ for any $x_1, x_2 \in A$ implies that $x_1 = x_2$. This means the following: if for $y_0 \in B$, there exists some $x_0 \in A$ with $f(x_0) = y_0$, then this $x = x_0$ is the **only element** in A satisfying $f(x) = y_0$.*

Remark 1.17 (not one-to-one)
*If a function $f : A \to B$ is **not one-to-one**, then there exist $x_1, x_2 \in A$ with $x_1 \neq x_2$ and $f(x_1) = f(x_2)$. Thus in order to show that a function $f : A \to B$ is not one-to-one, it is sufficient to find $x_1, x_2 \in A$ with $x_1 \neq x_2$ and $f(x_1) = f(x_2)$.*

Example 1.18 (one-to-one)

(a) $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = x$, is one-to-one because $f(x_1) = f(x_2)$ means just $x_1 = x_2$.

(b) $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = x^2$, is **not** one-to-one, because $x^2 = (-x)^2$, and thus $f(x) = f(-x)$ but $x \neq -x$.

(c) $f : [0, \infty) \to \mathbb{R}$, defined by $f(x) = x^2$, is one-to-one , because $x^2 = y^2$ implies $x = y$, since $x \geq 0$ and $y \geq 0$.

(d) The constant function $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = c$ with some constant $c \in \mathbb{R}$, is **not** one-to-one. Indeed, for any two $x_1 \neq x_2$, we have $f(x_1) = c = f(x_2)$. $\square$

Only for a one-to-one function we can define its so-called 'inverse function'.

Definition 1.19 (inverse function of a one-to-one function)
*Let $f : A \to B$ be a **one-to-one** function. Then the function f has an **inverse function** $f^{-1} : f(A) \to A$, $y \mapsto f^{-1}(y)$, defined by the properties*

$$f^{-1}\big(f(x)\big) = x \quad \textit{for all } x \in A, \qquad \textit{and} \qquad f\big(f^{-1}(y)\big) = y \quad \textit{for all } y \in f(A),$$

*where $f(A)$ is the **range** of f, that is,*

$$f(A) = \{y \in B \,:\, y = f(x) \textit{ for some } x \in A\}.$$

To **find the inverse function** of a given one-to-one function $f : A \to B$, we **consider** $y = f(x)$ **and solve for** x. Then we will find $x = f^{-1}(y)$, with some algebraic expression $f^{-1}(y)$ and this expression defines the inverse function $f^{-1} : f(A) \to A$.

Example 1.20 (inverse functions)

(a) The one-to-one function $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = x$, has the range $f(\mathbb{R}) = \mathbb{R}$. By setting $y = f(x) = x$, we find that $x = f^{-1}(y) = y$, and the inverse function $f^{-1} : \mathbb{R} \to \mathbb{R}$ is defined by $f^{-1}(y) = y$. We see that indeed

$$f^{-1}\big(\underbrace{f(x)}_{=x}\big) = f^{-1}(x) = x \qquad \text{for all } x \in \mathbb{R},$$

and

$$f\big(\underbrace{f^{-1}(y)}_{=y}\big) = f(y) = y \qquad \text{for all } y \in \mathbb{R}.$$

(b) To find the inverse of the one-to-one function $f : [0, \infty) \to \mathbb{R}$, defined by $f(x) = x^2$, we set

$$x^2 = y \quad \Rightarrow \quad x = y^{1/2} = \sqrt{y}.$$

We have that the range of f is $f([0, \infty)) = [0, \infty)$. Thus the inverse function of $f(x) = x^2$ is given by

$$f^{-1} : [0, \infty) \to \mathbb{R}, \qquad f^{-1}(y) = \sqrt{y}.$$

We see that we have indeed

$$f^{-1}\big(\underbrace{f(x)}_{=x^2}\big) = \sqrt{x^2} = x \qquad \text{for all } x \geq 0,$$

and

$$f\big(\underbrace{f^{-1}(y)}_{=\sqrt{y}}\big) = (\sqrt{y})^2 = y \qquad \text{for all } y \geq 0. \qquad \square$$

1.4 Affine Linear Functions and Quadratic Functions

In this section we learn about two very important classes of functions: **affine linear functions** and **quadratic functions**.

Definition 1.21 (affine linear function)
A function of the form

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = m\,x + c,$$

where $m \neq 0$ and c are real constants, is called an ***affine linear function****.*

Example of an affine linear function: $f(x) = 3\,x - 5$

The graph of an affine linear function $f(x) = m\,x + c$ is given by

$$\mathcal{G} = \{(x,\, m\,x + c) \,:\, x \in \mathbb{R}\}\,.$$

Plotting the graph (see Figure 1.7), we see that it is a **straight line** and therefore has **constant slope**. Setting $x = 0$ yields

$$f(0) = m\,0 + c = c$$

and we see that the function f **intersects the y-axis at $y = c$**. The **constant slope** of the function is given by

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = \frac{(m\,x_2 + c) - (m\,x_1 + c)}{x_2 - x_1} = \frac{m\,(x_2 - x_1)}{x_2 - x_1} = m.$$

If we choose $x_2 - x_1 = 1$ then we find that $f(x_2) - f(x_1) = m$ as illustrated in Figure 1.7 below.

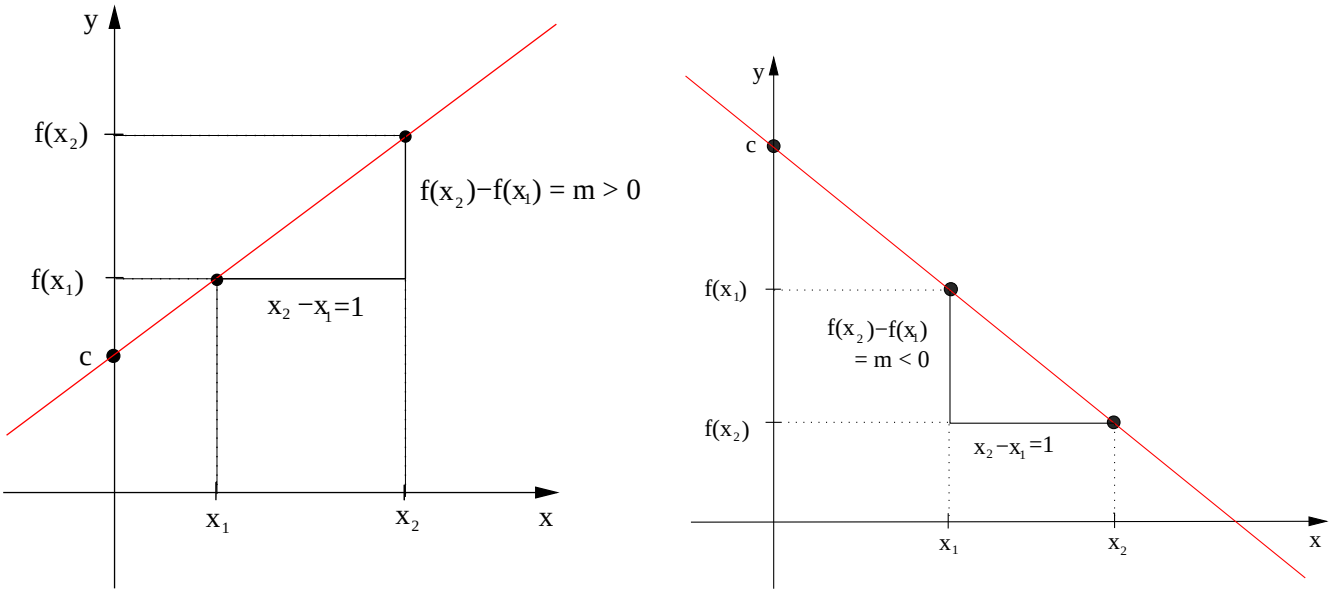


Figure 1.7: On the left the graph of $f(x) = m\,x + c$ with $m > 0$, and on the right the graph of $f(x) = m\,x + c$ with $m < 0$.

If $|x|$ gets large then we find the

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} \big(m\,x + c\big) = \begin{cases} \infty & \text{if} \quad m > 0, \\ -\infty & \text{if} \quad m < 0, \end{cases}$$

and

$$\lim_{x\to-\infty} f(x) = \lim_{x\to-\infty} \big(m\,x + c\big) = \begin{cases} -\infty & \text{if} \quad m > 0, \\ \infty & \text{if} \quad m < 0. \end{cases}$$

Finally we want to investigate whether an affine linear function is (strictly) monotonically increasing or decreasing. Let $x_1, x_2 \in \mathbb{R}$ with $x_1 < x_2$. If $m > 0$, then $m\,x_1 < m\,x_2$, and thus

$$f(x_1) = m\,x_1 + c < m\,x_2 + c = f(x_2),$$

and f is **strictly monotonically increasing on** $\mathbb{R}$. If $m < 0$, then multiplication of $x_1 < x_2$ with $m < 0$ yields $m\,x_1 > m\,x_2$, and thus

$$f(x_1) = m\,x_1 + c > m\,x_2 + c = f(x_2).$$

We see that f is **strictly monotonically decreasing on** $\mathbb{R}$.

Example 1.22 (inverse function of an affine linear function)
Investigate whether the affine linear function $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = m\,x + c$, with some constants c and m, where $m \neq 0$, is one-to-one. If yes, find the range $f(\mathbb{R})$ of f and find the inverse function $f^{-1} : f(\mathbb{R}) \to \mathbb{R}$ of f.

Solution: The affine linear function $f(x) = m\,x + c$, with $m \neq 0$ is **one-to-one**. Indeed, $f(x_1) = m\,x_1 + c = m\,x_2 + c = f(x_2)$ implies that $m\,x_1 = m\,x_2$ and thus $x_1 = x_2$.

The range $f(\mathbb{R})$ of f is $\mathbb{R}$, and this is the domain of f^{-1}. To find the inverse function $f^{-1} : \mathbb{R} \to \mathbb{R}$ of the one-to-one function $f(x) = m\,x + c$, we set

$$m\,x + c = y \quad \Rightarrow \quad m\,x = y - c \quad \Rightarrow \quad x = \frac{y - c}{m} = \frac{1}{m}\,y - \frac{c}{m}.$$

Thus the inverse function $f^{-1} : \mathbb{R} \to \mathbb{R}$ is given by

$$f^{-1}(y) = \frac{1}{m}\,y - \frac{c}{m},$$

and we observe that it is **an affine linear function** with slope $1/m$ that intersects the y-axis at $y = -c/m$. □

Summary 1.23 (facts about affine linear functions $f(x) = m\,x + c$)
*An **affine linear function** $f : \mathbb{R} \to \mathbb{R}$ is of the form*

$$f(x) = m\,x + c,$$

where $m \neq 0$.

- *Its **domain**, **codomain**, and **range** are $\mathbb{R}$.*
- *It is a **straight line** with **slope** m that **intersects the y-axis at** $y = c$.*
- *It is **one-to-one** and its **inverse function is also an affine linear function**.*

The next class of functions that we want to discuss are so-called **quadratic functions**. The simplest example of a quadratic function is $f(x) = x^2$.

Example 1.24 (quadratic function $f(x) = x^2$)
The function $f : \mathbb{R} \to \mathbb{R}$, defined by,

$$f(x) = x^2,$$

has the graph

$$\mathcal{G} = \left\{(x, x^2) \,:\, x \in \mathbb{R}\right\}.$$

We observe that $(-x)^2 = x^2$, that is, f will be **mirror-symmetric with respect to the y-axis**. The function has **non-negative values** since $f(x) = x^2 \geq 0$ for all $x \in \mathbb{R}$. We find that $\lim_{x\to\infty} x^2 = \infty$ and that $\lim_{x\to-\infty} x^2 = \infty$. The function f is **strictly monotonically increasing on** $[0, \infty)$ because for any $x_1, x_1 \in [0, \infty)$ we have

$$0 \leq x_1 < x_2 \qquad \Rightarrow \qquad f(x_1) = x_1^2 < x_1\, x_2 \leq x_2^2 = f(x_2).$$

The function f is **strictly monotonically decreasing on** $(-\infty, 0]$ because for any $x_1, x_1 \in (-\infty, 0]$ we have

$$x_1 < x_2 \leq 0 \qquad \Rightarrow \qquad -x_1 > -x_2 \geq 0 \qquad \Rightarrow \qquad |x_1| > |x_2| \geq 0$$
$$\Rightarrow \qquad f(x_1) = x_1^2 = |x_1|^2 > |x_2|^2 = x_2^2 = f(x_2).$$

We have for $x \in \{-3, -2, -1, -1/2, 0, 1/2, 1, 2, 3\}$ the points $(x, f(x))$ given by

$$(-3, 9),\ (-2, 4),\ (-1, 1),\ (-1/2, 1/4),\ (0, 0),\ (1/2, 1/4),\ (1, 1),\ (2, 4),\ (3, 9),$$

and we can now easily plot $f(x) = x^2$ (see Figure 1.8). The graph in Figure 1.8 is called a **parabola** and the point $(0, 0)$ is called the **vertex** of the parabola. □

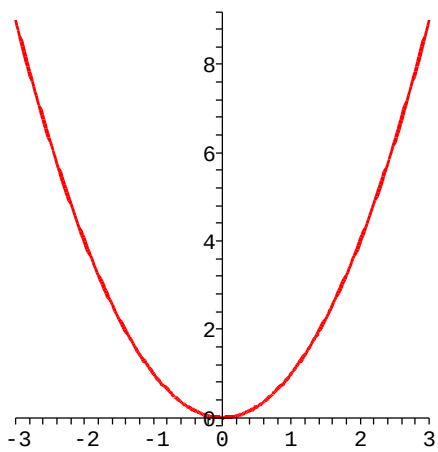


Figure 1.8: Graph of $f(x) = x^2$.

Now we want to discuss **quadratic functions** in their most general form.

Definition 1.25 (quadratic functions)
A function $f : \mathbb{R} \to \mathbb{R}$ of the form

$$f(x) = a\,x^2 + b\,x + c,$$

with real constants a, b, and c, where $a \neq 0$, is called a ***quadratic function****.*

We will show that the **graph of such a function is essentially of the form shown in Figure 1.8, but the graph may by upside-down, scaled, and shifted**. To analyze $f(x) = a\,x^2 + b\,x + c$, we try to bring it into the form

$$\widetilde{f}(x) = \widetilde{a}\,(x - \widetilde{b})^2 + \widetilde{c}, \tag{1.3}$$

with $\widetilde{a} \neq 0$.

Why is a representation of the form (1.3) useful?

- To see this, we first consider the case where $\widetilde{a} = 1$ and $\widetilde{c} = 0$. Then the function becomes $\widetilde{f}(x) = (x - \widetilde{b})^2$, and the graph of this function is obtained by **shifting the graph of $f(x) = x^2$ in the horizontal direction by $\widetilde{b}$** (to the left if $\widetilde{b} < 0$ and to the right if $\widetilde{b} > 0$).
- Now consider the case that $\widetilde{b} = 0$ and $\widetilde{a} = 1$. Then we obtain the function $\widetilde{f}(x) = x^2 + \widetilde{c}$, whose graph can be obtained by **shifting the graph of $f(x) = x^2$ in the vertical direction by $\widetilde{c}$** (upwards if $\widetilde{c} > 0$ and downwards if $\widetilde{c} < 0$).
- Finally we consider the case that $\widetilde{b} = 0$ and $\widetilde{c} = 0$. Then we have $\widetilde{f}(x) = \widetilde{a}\,x^2$, and we see that the graph can be obtained by **scaling the graph of $f(x) = x^2$ by the factor $\widetilde{a}$**, that is, all y-coordinates get multiplied by $\widetilde{a}$.

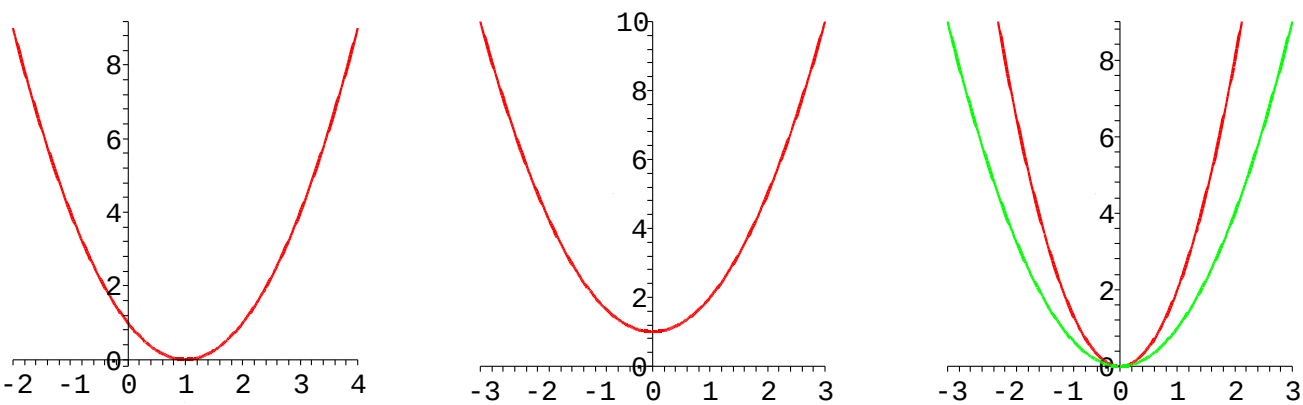


Figure 1.9: The graph of $f(x) = (x-1)^2$ on the left, the graph of $f(x) = x^2 + 1$ in the middle, and the graphs of $f(x) = 2\,x^2$ and $g(x) = x^2$ on the right.

From these observations, we see that the function (1.3) is a **parabola** with **vertex** in the point $(\widetilde{b}, \widetilde{c})$ and it is **scaled by a factor** $\widetilde{a}$ compared to the 'standard' parabola $f(x) = x^2$. This parabola may intersect the x-axis or not. This can be seen by setting

$$\widetilde{f}(x) = \widetilde{a}\,(x - \widetilde{b})^2 + \widetilde{c} = 0 \qquad \Rightarrow \qquad (x - \widetilde{b})^2 = \frac{-\widetilde{c}}{\widetilde{a}}.$$

Since $(x - \widetilde{b})^2 \geq 0$, this equality holds if and only if $-\widetilde{c}/\widetilde{a} \geq 0$. **If** $-\widetilde{c}/\widetilde{a} \geq 0$, then the function $\widetilde{f}$ **intersects the** x**-axis** at the point(s)

$$(x, 0) = \left(\widetilde{b} \pm \sqrt{-\frac{\widetilde{c}}{\widetilde{a}}}, 0\right).$$

If $-\widetilde{c}/\widetilde{a} > 0$ we have two points of intersection. If $-\widetilde{c}/\widetilde{a} = 0$, then we have $\widetilde{c} = 0$, and we get one point of intersection; this point is the vertex of the parabola.

Now we come back to our general quadratic function and transform it to the form (1.3). With the **binomial formulas**

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

we find that

$$\begin{aligned} \widetilde{f}(x) = a\,x^2 + b\,x + c &= a\left(x^2 + \frac{b}{a}\,x + \frac{c}{a}\right) \\ &= a\left[\left(x + \frac{b}{2\,a}\right)^2 + \left(\frac{c}{a} - \frac{b^2}{4\,a^2}\right)\right] \\ &= a\left(x + \frac{b}{2\,a}\right)^2 + \left(c - \frac{b^2}{4\,a}\right). \end{aligned} \tag{1.4}$$

Now we can apply our knowledge for the case (1.3). From (1.4), we see that the parabola $f(x) = a\,x^2 + b\,x + c$ has its **vertex** at the point

$$\left(-\frac{b}{2\,a}, c - \frac{b^2}{4\,a}\right),$$

and is scaled by the factor a. If

$$-\left(\frac{c}{a} - \frac{b^2}{4\,a^2}\right) = \frac{b^2 - 4\,a\,c}{4\,a^2} \geq 0 \qquad \Leftrightarrow \qquad b^2 - 4\,a\,c \geq 0,$$

then we have

$$f(x) = a\,x^2 + b\,x + c = \left(x + \frac{b}{2a} + \sqrt{\frac{b^2 - 4\,a\,c}{4\,a^2}}\right)\left(x + \frac{b}{2a} - \sqrt{\frac{b^2 - 4\,a\,c}{4\,a^2}}\right),$$

and the function $f(x) = a\,x^2 + b\,x + c$ **intersects the x-axis** at the point(s)

$$(x, 0) = \left(-\frac{b}{2a} \pm \sqrt{\frac{b^2 - 4\,a\,c}{4\,a^2}}, 0\right) = \left(\frac{-b \pm \sqrt{b^2 - 4\,a\,c}}{2\,a}, 0\right).$$

Summary 1.26 (how to plot a quadratic function)
*A **quadratic function** $f : \mathbb{R} \to \mathbb{R}$ is of the form*

$$f(x) = a\,x^2 + b\,x + c, \qquad \text{with } a \neq 0.$$

- *Use the binomial formulas to bring it into the standard form*

$$\widetilde{f}(x) = \widetilde{a}\,(x - \widetilde{b})^2 + \widetilde{c}. \tag{1.5}$$

- *The **vertex** of the **parabola** (1.5) is then at $(\widetilde{b}, \widetilde{c})$.*
- *The parabola (1.5) is **scaled** by a factor $\widetilde{a}$.*
- *If $-\widetilde{c}/\widetilde{a} \geq 0$, then the parabola (1.5) **intersects the x-axis** at the point(s)*

$$(x, 0) = \left(\widetilde{b} \pm \sqrt{-\frac{\widetilde{c}}{\widetilde{a}}}, 0\right).$$

If $-\widetilde{c}/\widetilde{a} > 0$ we have two points of intersection. If $-\widetilde{c}/\widetilde{a} = 0$, then we have $\widetilde{c} = 0$, and we get only one point of intersection; this point is the vertex of the parabola.

We close this chapter by summarizing what we have learned so far about plotting and analyzing functions. If we want to **analyze and plot a function** $f : A \to B$ then we are interested in the following:

- for which x then $f(x)$ is positive or negative or zero
- if $x \in A$ can get arbitrary large, what happens with $f(x)$ as $x \to \infty$; and if $x \in A$ can get arbitrary small, what happens with $f(x)$ as $x \to -\infty$
- if f is undefined at a single point, then what happens with $f(x)$ for $x \to x_0$ from below (that is, $x < 0$) and for $x \to x_0$ from above (that is, $x > 0$)
- to find the asymptotes of f, if it has any
- on which subintervals of A is f (strictly) monotonically increasing or (strictly) monotonically decreasing

Chapter 2

Classical Functions

In this chapter we discuss the classical functions. In Section 2.1, we discuss briefly polynomials, the easiest classical functions. In the last chapter we have already encountered polynomials of degree zero, one, and two, that is, constant functions, affine linear functions, and quadratic functions, respectively. Trigonometric functions, that is, the sine function, the cosine function, the tangent and cotangent functions, as well as their inverse functions, are investigated in Section 2.2. After that we introduce exponential functions and logarithmic functions (which are defined as the inverse functions of exponential functions) in Sections 2.3 and 2.4, respectively. We will discuss exponential and logarithmic functions in some detail. In Section 2.5 we finally encounter hyperbolic functions and their inverse functions.

2.1 Polynomials

We start by introducing an important class of functions, of which the affine linear functions and quadratic functions are special cases.

Definition 2.1 (polynomial)
A function $f : \mathbb{R} \to \mathbb{R}$ of the form

$$f(x) = a_0 + a_1\, x + a_2\, x^2 + \ldots + a_{n-1}\, x^{n-1} + a_n\, x^n = \sum_{k=0}^{n} a_k\, x^k, \tag{2.1}$$

*where the **coefficients** $a_0, a_1, \ldots, a_{n-1}, a_n$ are fixed real numbers and $a_n \neq 0$, is a **polynomial** of **degree** $n \in \mathbb{N}_0$.*

Example of polynomial of degree 5: $f(x) = 3\,x^5 + 7\,x^4 - 2\,x^2 + 3$.

Example 2.2 (polynomials of degrees 0, 1, and 2)

(a) The polynomials of degree zero are the **constant functions**

$$f(x) = c \qquad \text{for all } x \in \mathbb{R},$$

where c is a real constant. In the representation (2.1) we have $n = 0$ and $a_0 = c$.

(b) The polynomials of degree one are the **affine linear functions**

$$f(x) = m\,x + c \qquad \text{for all } x \in \mathbb{R},$$

with real constants $m \neq 0$ and c. In the representation (2.1) we have $n = 1$ and $a_0 = c$ and $a_1 = m$.

(c) The polynomials of degree two are the **quadratic functions**

$$f(x) = a\,x^2 + b\,x + c \qquad \text{for all } x \in \mathbb{R},$$

with real constants $a \neq 0$, b, and c. In the representation (2.1) we have $n = 2$ and $a_0 = c$, $a_1 = b$, and $a_2 = a$.

2.2 Trigonometric Functions

From everyday life we are used to measuring angles in **degree**. For example, a very steep road may have a warning traffic sign announcing that the slope is 20°.

In science on the other hand, **angles** are often measured in **radians**.

Definition 2.3 (radians)
*For the **unit circle** with radius one, one **radians** is defined as the angle corresponding to an arc of length one. One radians is approximately 57.3 degree.*

Using this complicated definition is much simpler than it might appear at first glance. Since the length of the unit circle is 2π, we know that 360° are 2π radians. Observing that for an angle θ in degree and its angle ϕ in radians we have

$$\frac{\phi}{\theta} = \frac{2\pi}{360} \qquad \Leftrightarrow \qquad \phi = \frac{2\pi}{360}\,\theta$$

we can work out the following table.

degree	0	30	45	60	90	180	270	360	θ
radians	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$	π	$3\pi/2$	2π	$2\pi\,\theta/360$

Table 2.1: Conversion from degree to radians.

We see from the definition of radians that the arc of the unit circle corresponding to an angle ϕ measured in radians has just length ϕ.

After this preparation we can introduce the **trigonometric functions** whose argument is given in radians.

Definition 2.4 (sine function and cosine function on $[0, 2\pi]$)
The functions sine and cosine are geometrically defined as indicated in Figure 2.1 and as explained below. Consider the ***unit circle****, that is, the circle with radius one, and measure the angle ϕ starting at the origin $(0, 0)$ from the x-axis* ***anti-clockwise****. We draw a straight line with slope given by this angle ϕ up to the intersection of this line with the unit circle. From this intersection we draw a vertical straight line. The distance on this vertical straight line from its intersection with the unit circle to the intersection with the x-axis is $\sin(\phi)$. The distance along the x-axis from the origin to the intersection of the x-axis with the vertical straight line is $\cos(\phi)$.*

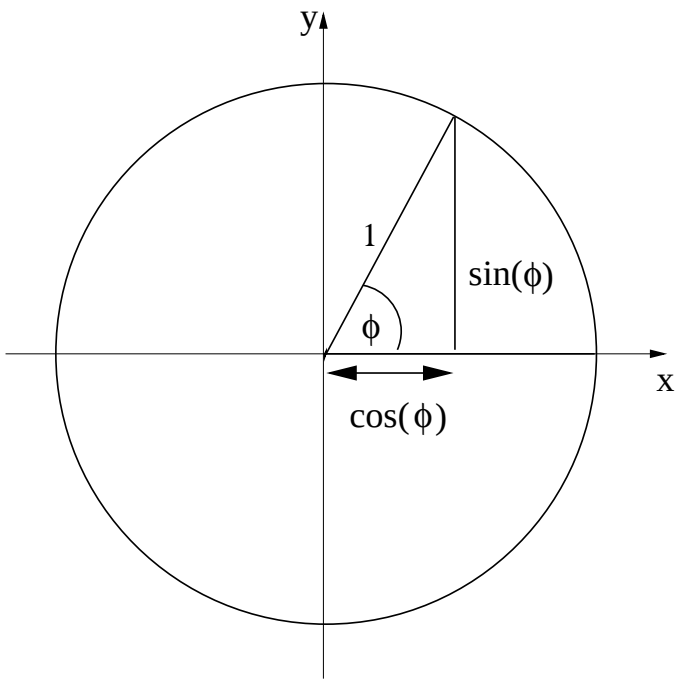


Figure 2.1: Definition of the sine function and the cosine function.

Since $\sin(0) = \sin(2\pi)$ and $\cos(0) = \cos(2\pi)$, we can easily extend the definition of $\sin(x)$ and $\cos(x)$ from $[0, 2\pi]$ to all $\mathbb{R}$.

Definition 2.5 (sine function and cosine function on $\mathbb{R}$)
Definition 2.4 defines $\cos(x)$ *and* $\sin(x)$ *for all* $x \in [0, 2\pi]$. *Since the angle* 2π *describes a full rotation, we have* $\cos(0) = \cos(2\pi)$ *and* $\sin(0) = \sin(2\pi)$. *Therefore we can define the cosine and the sine function for all* $x \in \mathbb{R}$ *in the following way: We can represent* $x \in \mathbb{R}$ *uniquely in the form* $x = 2\pi\, k + \phi$, *with* $\phi \in [0, 2\pi)$ *and* $k \in \mathbb{Z}$. *Then*

$$\cos(x) = \cos(2\pi\, k + \phi) = \cos(\phi), \qquad \sin(x) = \sin(2\pi\, k + \phi) = \sin(\phi), \tag{2.2}$$

and we have defined the cosine and the sine function on $\mathbb{R}$. *From Figure 2.4 and* ***Pythagoras' Theorem*** *follows the important formula*

$$\big(\sin(x)\big)^2 + \big(\cos(x)\big)^2 = 1 \qquad \textit{for all } x \in \mathbb{R}.$$

In order to plot the sine and the cosine function, we use Figure 2.4 to work out some of its important values.

x in radians	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$	$4\pi/6$	$3\pi/4$	$5\pi/6$	π	$3\pi/2$	2π
x in degree	0	30	45	60	90	120	135	150	180	270	360
$\cos(x)$	1	$\sqrt{3}/2$	$\sqrt{2}/2$	$1/2$	0	$-1/2$	$-\sqrt{2}/2$	$-\sqrt{3}/2$	-1	0	1
$\sin(x)$	0	$1/2$	$\sqrt{2}/2$	$\sqrt{3}/2$	1	$\sqrt{3}/2$	$\sqrt{2}/2$	$1/2$	0	-1	0

Table 2.2: Some important values of the cosine function and the sine function.

From the definition of sine and cosine in Figure 2.4, we can see the following two useful formulas

$$\cos(-x) = \cos(x) \qquad \text{and} \qquad \sin(-x) = -\sin(x) \qquad \text{for all } x \in \mathbb{R}. \tag{2.3}$$

This can be easily seen by considering for the angle $x > 0$ the angle $-x$ as a rotation about the angle x **clockwise**.

Now we can plot the cosine and the sine function.

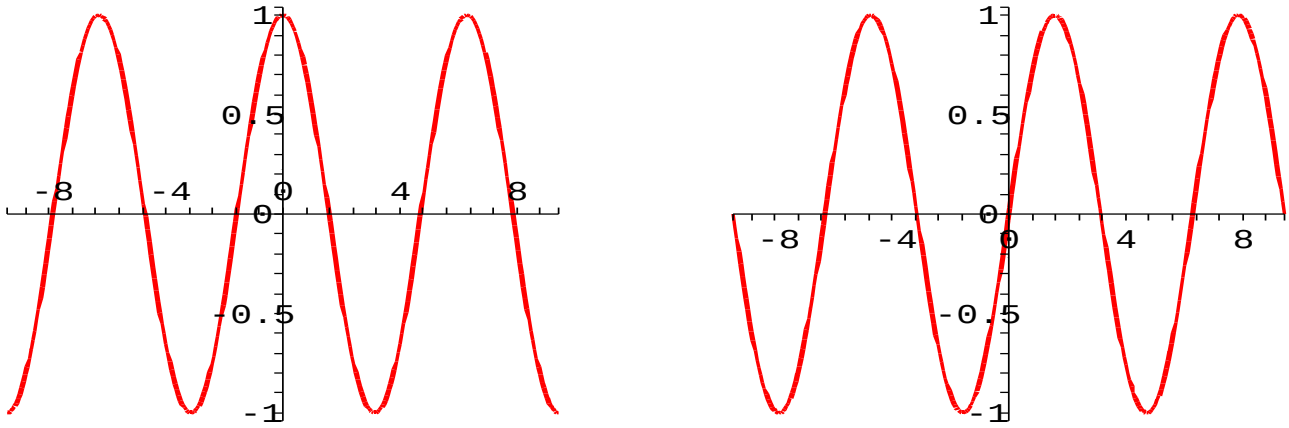


Figure 2.2: Graph of $\cos(x)$ on the left and the graph of $\sin(x)$ on the right.

Definition 2.6 (periodic function)
*A function $f : \mathbb{R} \to \mathbb{R}$ is called **periodic** with period $L > 0$ if we have*

$$f(x) = f(x + L) \qquad \text{for all } x \in \mathbb{R}. \tag{2.4}$$

Note that repeated application of (2.4) implies that

$$f(x) = f(x + k\,L) \qquad \text{for all } x \in \mathbb{R} \text{ and all } k \in \mathbb{Z}.$$

Example 2.7 (periodic functions)

(a) From (2.2), we see that $\cos(x)$ and $\sin(x)$ are periodic with period 2π.

(b) The **square wave** (see Figure 2.3), defined by

$$f(x) = \begin{cases} 1 & \text{if} \quad x \in [2k, 2k+1) \text{ with } k \in \mathbb{Z}, \\ -1 & \text{if} \quad x \in [2k-1, 2k) \text{ with } k \in \mathbb{Z}, \end{cases} \tag{2.5}$$

is periodic with period $L = 2$. Indeed, for any $x \in [2k-1, 2k)$ the point $x+2$ is in $[(2k-1)+2, 2k+2) = [2(k+1)-1, 2(k+1))$ and thus $f(x) = f(x+2) = -1$. Likewise if $x \in [2k, 2k+1)$, then we have $x + 2 \in [2(k+1), 2(k+1)+1)$, and thus $f(x) = f(x+2) = 1$. This shows the periodicity of the square wave with period $L = 2$.

(c) The functions $\sin(nx)$ and $\cos(nx)$ have period $2\pi/n$. This follows from the fact that $\sin(x)$ and $\cos(x)$ have period 2π, and thus

$$\sin(nx) = \sin(nx + 2\pi) = \sin\left(n\left[x + \frac{2\pi}{n}\right]\right),$$

$$\cos(nx) = \cos(nx + 2\pi) = \cos\left(n\left[x + \frac{2\pi}{n}\right]\right),$$

which proves the periodicity with period $2\pi/n$. □

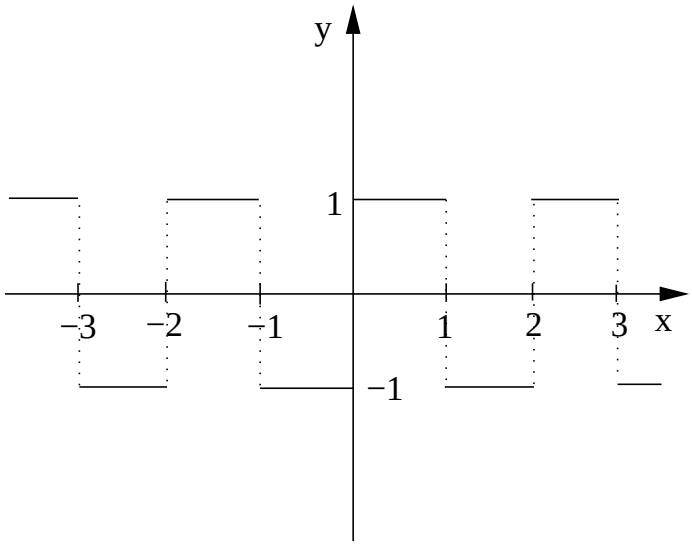


Figure 2.3: Graph of the square wave (2.5).

Definition 2.8 (tangent and cotangent function)
*The **tangent function** is defined by*

$$\tan(x) = \frac{\sin(x)}{\cos(x)}, \quad x \notin \left\{\ldots, -\frac{3\pi}{2}, -\frac{\pi}{2}, \frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}, \ldots\right\} = \left\{\frac{(2k+1)\pi}{2} \,:\, k \in \mathbb{Z}\right\}.$$

All real numbers of the form $(2k+1)\pi/2$, where $k \in \mathbb{Z}$, are excluded from the domain of the tangent function, because at these points the cosine function in the denominator is zero.
*The **cotangent function** is defined by*

$$\cot(x) = \frac{\cos(x)}{\sin(x)}, \qquad x \notin \{\ldots, -3\pi, -2\pi, -\pi, 0, \pi, 2\pi, 3\pi, \ldots\} = \{k\,\pi \,:\, k \in \mathbb{Z}\}.$$

All real numbers of the form $k\pi$, where $k \in \mathbb{Z}$, are excluded from the domain of the cotangent function, because at these points the sine function in the denominator is zero.

Remark 2.9 (geometric interpretation of the tangent function)
From the similarities of the triangles in Figure 2.4, we see that in the larger triangle the ***side parallel to the side with length*** $\sin(\phi)$ ***has length*** $\tan(\phi)$*, because its length a satisfies*

$$a = \frac{a}{1} = \frac{\sin(\phi)}{\cos(\phi)} = \tan(\phi).$$

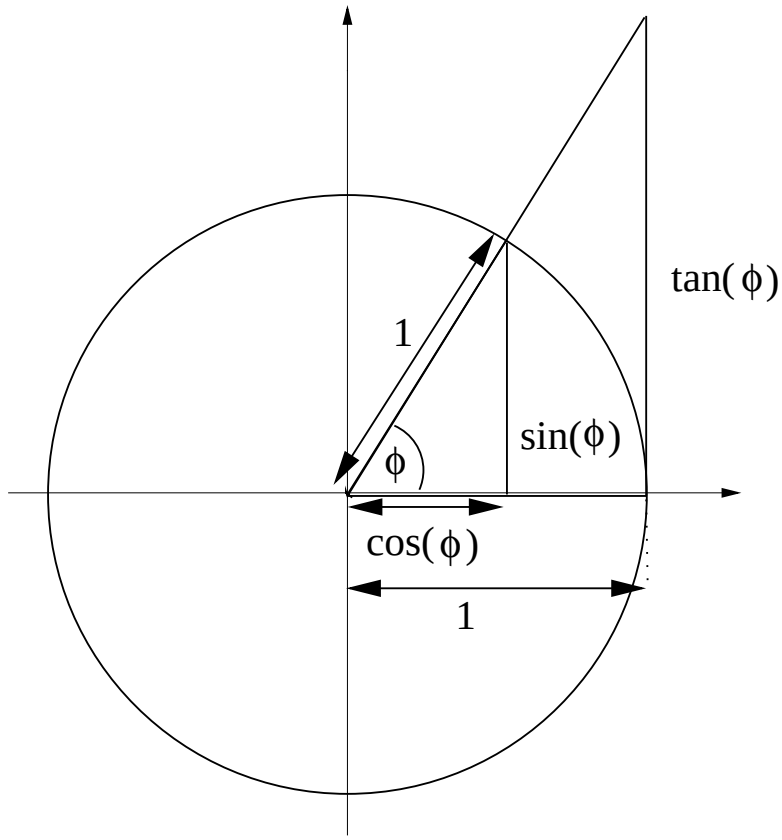


Figure 2.4: Geometric relation of $\tan(\phi)$ to $\sin(\phi)$ and $\cos(\phi)$.

Based on the values of $\sin(x)$ and $\cos(x)$, we can now plot the tangent function. Because

$$\lim_{x\to(2k+1)\pi/2} \cos(x) = 0, \qquad k \in \mathbb{Z},$$

and taking into account the signs of $\sin(x)$ and $\cos(x)$ as we approach $(2k+1)\pi/2$ from below and from above, we find that

$$\lim_{\substack{x<(2k+1)\pi/2,\\ x\to(2k+1)\pi/2}} \tan(x) = \infty, \qquad \lim_{\substack{x>(2k+1)\pi/2,\\ x\to(2k+1)\pi/2}} \tan(x) = -\infty, \qquad k \in \mathbb{Z},$$

Thus the vertical straight lines through $\big((2k+1)\pi/2, 0\big)$ are asymptotes of $\tan(x)$.

Furthermore we have

$$\tan(0) = \frac{\sin(0)}{\cos(0)} = \frac{0}{1} = 0, \qquad \text{and} \qquad \tan(\pi/4) = \frac{\sin(\pi/4)}{\cos(\pi/4)} = \frac{1/\sqrt{2}}{1/\sqrt{2}} = 1.$$

Thus we obtain the plot on the left-hand side in Figure 2.5. Analogous observations lead to the the plot of the cotangent function on the right-hand side in Figure 2.5.

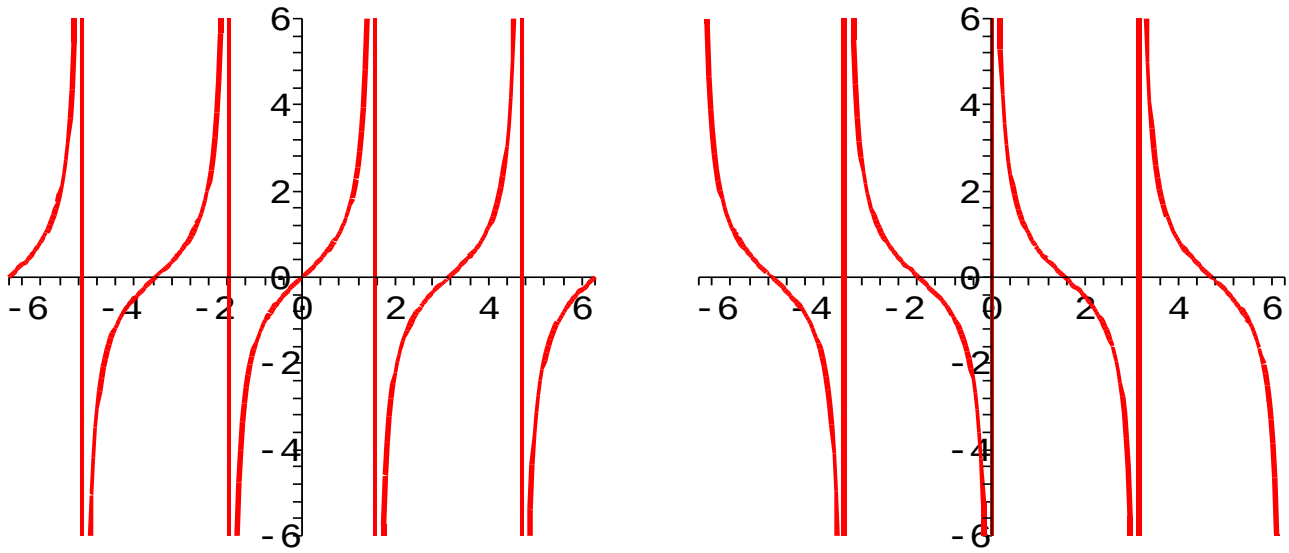


Figure 2.5: The graph of $\tan(x)$ on the left, and the graph of $\cot(x)$ on the right.

It is easily seen that the functions $\tan(x)$ and $\cot(x)$ are **periodic with period** π.

Now we list some important and useful properties of the cosine and sine function.

Lemma 2.10 (addition theorems for $\cos(x)$ and $\sin(x)$)

$$\begin{aligned} \sin(x+y) &= \sin(x)\,\cos(y) + \sin(y)\,\cos(x), \\ \sin(x-y) &= \sin(x)\,\cos(y) - \sin(y)\,\cos(x), \end{aligned}$$

and

$$\begin{aligned} \cos(x+y) &= \cos(x)\,\cos(y) - \sin(x)\,\sin(y), \\ \cos(x-y) &= \cos(x)\,\cos(y) + \sin(x)\,\sin(y). \end{aligned}$$

Lemma 2.11 (half angle formulas for $\cos(x)$ and $\sin(x)$)

$$\begin{aligned}
\sin(x)+\sin(y) &= 2\,\sin\left(\frac{1}{2}(x+y)\right)\cos\left(\frac{1}{2}(x-y)\right),\\
\sin(x)-\sin(y) &= 2\,\cos\left(\frac{1}{2}(x+y)\right)\sin\left(\frac{1}{2}(x-y)\right),\\
\cos(x)+\cos(y) &= 2\,\cos\left(\frac{1}{2}(x+y)\right)\cos\left(\frac{1}{2}(x-y)\right),\\
\cos(x)-\cos(y) &= 2\,\sin\left(\frac{1}{2}(x+y)\right)\sin\left(\frac{1}{2}(y-x)\right).
\end{aligned}$$

In the next definition, we learn some more properties to classify functions.

Definition 2.12 (even and odd functions)
*Let $B \subset \mathbb{R}$, and let $A \subset \mathbb{R}$ be a **'symmetric set'**, that is, if $x \in A$ then also $-x \in A$.*

*(i) A function $f : A \to B$ is called **even** if $f(-x) = f(x)$ for all $x \in A$.*

*(ii) A function $f : A \to B$ is called **odd** if $f(-x) = -f(x)$ for all $x \in A$.*

We note the 'symmetric sets' in $\mathbb{R}$ are just mirror symmetric at the $x = 0$. For example, the sets $\mathbb{R}$, $[-a, a]$ with $a > 0$, as well as $\mathbb{R} \setminus \mathbb{Z}$ and $\mathbb{R} \setminus \{k\pi : k \in \mathbb{Z}\}$ are all symmetric sets.

It is actually quite simple to see whether a function is even or odd, once we have plotted/sketched it. This is explained in the next remark.

Remark 2.13 (deduce from the plot whether a function is even/odd)

- *A function f is **even** if $f(x) = f(-x)$ for all x. This means that the function has at x and $-x$ the same value. In the plot this means the $f(x)$ is **mirror-symmetric on the y-axis**!*
- *A function f is **odd** if $f(-x) = -f(x)$ for all x. This means that at $-x$ the function has the value $(-1)f(x)$. This means that we obtain $f(x)$ for $x < 0$ by **rotating the graph of $f(x)$ for $x \geq 0$ by 180 degree about the origin!***

Example 2.14 (even and odd functions)

(a) We see from (2.3), that $\sin(x)$ is odd and that $\cos(x)$ is even.

(b) Since the product of an odd function and an even function is odd (see Exercise Sheet 2), (a) implies that $\tan(x)$ and $\cot(x)$ are odd. □

The functions $\sin(x)$, $\cos(x)$, $\tan(x)$, and $\cot(x)$ are called **trigonometric functions**.

Some particular functions defined with the help of the sine function and the cosine function are given their own name: $\operatorname{cosec} : \mathbb{R} \setminus \{(k\pi : k \in \mathbb{Z}\} \to \mathbb{R}$ and $\sec : \mathbb{R} \setminus \{(2k+1)\pi/2 : k \in \mathbb{Z}\} \to \mathbb{R}$ are defined by

$$\operatorname{cosec}(x) = \frac{1}{\sin(x)}, \qquad \sec(x) = \frac{1}{\cos(x)}.$$

Finally, we want to consider the **inverse trigonometric functions**. None of the trigonometric functions that we have encountered is one-to-one on $\mathbb{R}$, but all are one-to-one restricted to suitable subintervals of $\mathbb{R}$.

- $\cos(x)$ is one-to-one on any interval $[k\pi, (k+1)\pi]$, $k \in \mathbb{Z}$
- $\sin(x)$ is one-to-one on any interval $\left[\frac{(2k-1)\pi}{2}, \frac{(2k+1)\pi}{2}\right]$, $k \in \mathbb{Z}$.
- $\tan(x)$ is one-to-one on any interval $\left(\frac{(2k-1)\pi}{2}, \frac{(2k+1)\pi}{2}\right)$, $k \in \mathbb{Z}$.
- $\cot(x)$ is one-to-one on any interval $(k\pi, (k+1)\pi)$, $k \in \mathbb{Z}$.

Definition 2.15 (inverse trigonometric functions)

(i) The inverse function $\cos^{-1} : [-1,1] \to [0,\pi]$ *of* $\cos : [0,\pi] \to [-1,1]$ *is denoted by* $\cos^{-1}(x) = \arccos(x)$.

(ii) The inverse function $\sin^{-1} : [-1,1] \to [-\pi/2, \pi/2]$ *of* $\sin : [-\pi/2, \pi/2] \to [-1,1]$ *is denoted by* $\sin^{-1}(x) = \arcsin(x)$.

(iii) The inverse function $\tan^{-1} : \mathbb{R} \to (-\pi/2, \pi/2)$ *of* $\tan : (-\pi/2, \pi/2) \to \mathbb{R}$ *is denoted by* $\tan^{-1}(x) = \arctan(x)$.

(iv) The inverse function $\cot^{-1} : \mathbb{R} \to (0,\pi)$ *of* $\cot : (0,\pi) \to \mathbb{R}$ *is denoted by* $\cot^{-1}(x) = \operatorname{arcot}(x)$.

Now we plot the graphs of the inverse trigonometric functions. Rather than evaluating the functions above at a number of points to plot them, we can plot the

inverse function f^{-1} by **reflecting the plot of f on the diagonal axis though the origin given by $x = y$.**

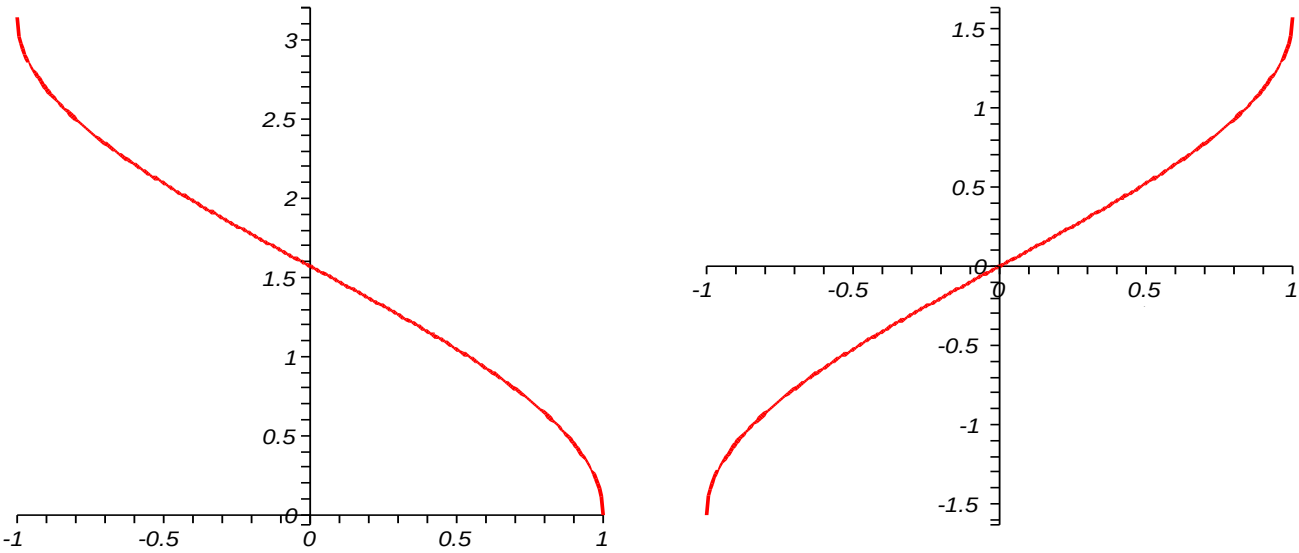


Figure 2.6: The graph of $\arccos(x)$ on the left, and the graph of $\arcsin(x)$ on the right.

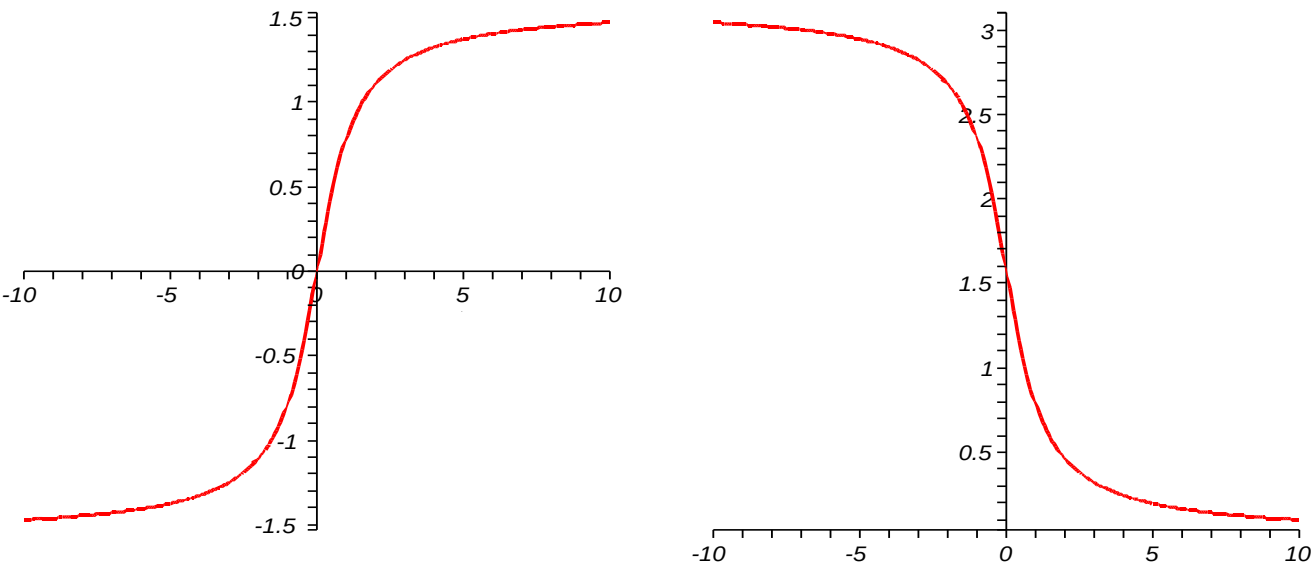


Figure 2.7: The graph of $\arctan(x)$ on the left and the graph of $\operatorname{arccot}(x)$ on the right.

Of course, it is also possible to find an inverse function to $\sin(x)$ and $\cos(x)$ on any interval where they are one-to-one, but the intervals chosen in Definition 2.15 are the ones usually considered.

2.3 Exponential Functions

In this section we will introduce **exponential functions**. In an expression of the form

$$a^x, \qquad \text{where } a \in \mathbb{R}, \text{ and } x \in \mathbb{R} \text{ with } x > 0,$$

as defined below, we call x the **exponent** (or **power**) and a the **base**.

Definition 2.16 (integer powers of a real number)
Let $a \in \mathbb{R}$ and let n be an integer. Then we define

$$a^0 = 1 \qquad \textit{for all } a \in \mathbb{R},$$

and for a positive integer n, a^n is defined by

$$a^n = \underbrace{a \times a \times \ldots \times a}_{n\text{-times}} \qquad \textit{for all } a \in \mathbb{R}.$$

If n is a negative integer $n = -m$, with $m \in \mathbb{N}$, we define

$$a^n = a^{-m} = \frac{1}{a^m} = \frac{1}{\underbrace{a \times a \times \ldots \times a}_{m\text{-times}}} \qquad \textit{for all } a \in \mathbb{R} \setminus \{0\}.$$

In particular, we have

$$a^{-1} = \frac{1}{a}.$$

Example 2.17 (integer powers of a real number)

(a) $2^3 = 2 \times 2 \times 2 = 8$.

(b) $10^4 = 10 \times 10 \times 10 \times 10 = 10000$.

(c) $2^{-1} = 1/2 = 0.5$.

(d) $10^{-2} = 1/(10 \times 10) = 1/100 = 0.01$.

(e) $3^{-3} = 1/(3 \times 3 \times 3) = 1/27$.

(f) $(-2)^3 = (-2) \times (-2) \times (-2) = -8$.

(g) $(-4)^{-2} = 1/(-4)^2 = 1/[(-4) \times (-4)] = 1/16 = 0.0625$. □

Lemma 2.18 (rules for integer powers)
Let $a, b \in \mathbb{R}$, and let n and m be non-zero integers. Then

$$a^{nm} = (a^n)^m = (a^m)^n, \tag{2.6}$$

and

$$a^{n+m} = a^n\, a^m, \qquad a^{n-m} = a^n\, a^{-m}. \tag{2.7}$$

Furthermore, we have

$$(a\, b)^n = a^n\, b^n.$$

Proof: We give the proofs only for $n > 0$ and $m > 0$. If $n \leq 0$ or $m \leq 0$, the proofs are analogously.

$$(a^n)^m = \underbrace{a^n \times a^n \times \ldots \times a^n}_{m\text{-times}} = \underbrace{a \times a \times \ldots \times a}_{nm\text{-times}} = \underbrace{a^m \times a^m \times \ldots \times a^m}_{n\text{-times}} = (a^m)^n,$$

and

$$a^{n+m} = \underbrace{a \times a \times \ldots \times a}_{(n+m)\text{-times}} = \big(\underbrace{a \times a \times \ldots \times a}_{n\text{-times}}\big) \times \big(\underbrace{a \times a \times \ldots \times a}_{m\text{-times}}\big) = a^n\, a^m$$

$$a^{n-m} = \underbrace{a \times a \times \ldots \times a}_{(n-m)\text{-times}} = \big(\underbrace{a \times a \times \ldots \times a}_{n\text{-times}}\big) \times \frac{1}{\underbrace{a \times a \times \ldots \times a}_{m\text{-times}}} = a^n\, a^{-m},$$

which verifies the formulas (2.6) and (2.7) for $m > 0$ and $n > 0$. Finally, we have for $n > 0$

$$(a\, b)^n = \underbrace{(a\, b) \times (a\, b) \times \ldots \times (a\, b)}_{n\text{-times}} = \big(\underbrace{a \times a \times \ldots \times a}_{n\text{-times}}\big) \times \big(\underbrace{b \times b \times \ldots \times b}_{n\text{-times}}\big) = a^n\, b^n,$$

as claimed. □

This definition is quite intuitive and natural, if we accept that $a^{-n} = 1/a^n$ for $n \in \mathbb{N}$ and $a > 0$. Now we want to go away from the assumption that the exponent is an integer, that is, we consider a^x with $x \in \mathbb{R}$ and $a \in \mathbb{R}$ with $a > 0$. First we consider the case that x is a rational number, that is, $x \in \mathbb{Q}$.

Definition 2.19 (nth roots of a non-negative real numbers)
Let $a \in \mathbb{R}$ be a non-negative real number, and let n be a positive integer. Then $a^{1/n}$ is defined to be the non-negative real number b such that

$$b^n = a,$$

*that is, the number $a^{1/n}$ is the n**th root of** a. We sometimes find the notation $\sqrt[n]{a}$ instead of $a^{1/n}$.*

Example 2.20 (nth roots of $a > 0$)

(a) $1000^{1/3} = 10$, because $10^3 = 1000$.

(b) $2^{1/2} = \sqrt{2}$, since $(\sqrt{2})^2 = \sqrt{2} \times \sqrt{2} = 2$.

(c) $81^{1/4} = 3$, because $3^4 = 81$.

(d) $8^{1/3} = 2$, since $2^3 = 8$.

(e) $a^{1/2} = \sqrt{a}$, because $(\sqrt{a})^2 = \sqrt{a} \times \sqrt{a} = a$.

(f) $0^{1/7} = 0$, because $0^7 = 0$. □

Lemma 2.21 (rules for nth roots)
Let $a, b \in \mathbb{R}$ be non-negative real numbers and let n and m be a positive integers. Then

$$a^{1/(nm)} = (a^{1/n})^{1/m} = (a^{1/m})^{1/n},$$

and

$$(a\,b)^{1/n} = a^{1/n}\,b^{1/n}.$$

Lemma 2.21 is a useful tool for working out roots and simplifying their representation.

Example 2.22 (illustration of Lemma 2.21)

(a) $8^{1/6} = 8^{1/(2\times 3)} = (8^{1/3})^{1/2}$, and since $2^3 = 8$, we have

$$8^{1/6} = (8^{1/3})^{1/2} = 2^{1/2} = \sqrt{2}.$$

(b) $6561^{1/8} = 6561^{1/(2\times 4)} = (6561^{1/2})^{1/4}$, and since $81^2 = 6561$, we have

$$6561^{1/8} = (6561^{1/2})^{1/4} = 81^{1/4} = (81^{1/2})^{1/2} = 9^{1/2} = 3,$$

where we have use that $9^2 = 81$ and $3^2 = 9$.

(c) $24^{1/3} = (3 \times 8)^{1/3} = 3^{1/3}\, 8^{1/3} = 3^{1/3} \times 2 = 2 \times 3^{1/3}$, where we have used that $2^3 = 8$. □

Definition 2.23 (fractional powers of a positive real number)
Let a be a positive real number, and let m be any integer and n be a positive integer. Then $a^{m/n}$ is defined by

$$a^{m/n} = \left(a^{1/n}\right)^m .$$

Note that we also have

$$a^{m/n} = (a^m)^{1/n},$$

and thus

$$(a^{1/n})^m = (a^m)^{1/n}.$$

Example 2.24 (fractional powers of positive real numbers)

(a) $2^{-1/2} = (2^{1/2})^{-1} = (\sqrt{2})^{-1} = 1/\sqrt{2}$.

(b) $9^{3/2} = (9^{1/2})^3 = (\sqrt{9})^3 = 3^3 = 27$, where we have used $3^2 = 9$.

(c) $1000^{-4/3} = (1000^{1/3})^{-4} = 10^{-4} = 1/10^4 = 0.0001$, where we have used the fact that $10^3 = 1000$.

(d) $(\sqrt{8})^{-2/3} = (\sqrt{8}^{1/3})^{-2} = \left((8^{1/2})^{1/3}\right)^{-2} = (8^{1/2\times 1/3})^{-2} = (8^{1/6})^{-2} = \left((8^{1/3})^{1/2}\right)^{-2} = (2^{1/2})^{-2} = 2^{-2/2} = 2^{-1} = 1/2$, where we have used $2^3 = 8$. □

Lemma 2.18 and Lemma 2.21 imply the following lemma.

Lemma 2.25 (properties of fractional powers of positive real numbers)
Let $a, b \in \mathbb{R}$ be positive real numbers, and let m and k be any integers, and let n and ℓ be positive integers. Then

$$a^{(mk)/(n\ell)} = a^{(m/n)(k/\ell)} = (a^{m/n})^{k/\ell} = (a^{k/\ell})^{m/n}.$$

We also have

$$a^{m/n+k/\ell} = a^{m/n}\, a^{k/\ell}, \qquad a^{m/n-k/\ell} = a^{m/n}\, a^{-k/\ell} = \frac{a^{m/n}}{a^{k/\ell}},$$

and

$$(a\, b)^{m/n} = a^{m/n}\, b^{m/n}.$$

Example 2.26 (fractional powers of positive real numbers)
In this example we will apply the properties stated in Lemma 2.25. to simplify expressions:

(a) $2^{1/3}\,2^{2/3} = 2^{1/3+2/3} = 2^1 = 2$.

(b) $50^{3/2} = (2\times 25)^{3/2} = 2^{3/2}\,25^{3/2} = 2^{1+1/2}\,(25^{1/2})^3 = 2\times 2^{1/2}\times 5^3 = 2\times\sqrt{2}\times 125 = 250\,\sqrt{2}$, where we have used $5^2 = 25$.

(c) $8^{5/6} = 8^{1/2+1/3} = 8^{1/2}\,8^{1/3} = (4\times 2)^{1/2}\,2 = 4^{1/2}\,2^{1/2}\,2 = 2\times 2^{1/2}\times 2 = 4\,\sqrt{2}$, where we have used that $2^2 = 4$ and $2^3 = 8$. □

To **extend the definition of a^x to all real exponents** x, we use that any positive irrational number $x \in \mathbb{R}\setminus\mathbb{Q}$ (for example, $\sqrt{2}$) can be approximated by a sequence $\{x_n\}$ or rational numbers, that is, $x_n \in \mathbb{Q}$ for all n and $\lim_{n\to\infty} x_n = x$. Then we can define

$$a^x = \lim_{n\to\infty} a^{x_n}.$$

All the properties that we have derived for integer and fractional/rational exponents (see Lemma 2.25) hold now as well, and we can finally define exponential functions.

Theorem 2.27 (properties of real powers of positive real numbers)
Let a and b be positive real numbers, and let $x, y \in \mathbb{R}$. Then

$$a^{xy} = (a^x)^y = (a^y)^x,$$

and

$$a^{x+y} = a^x\,a^y, \qquad a^{x-y} = a^x\,a^{-y} = \frac{a^x}{a^y}.$$

Furthermore, we have

$$(a\,b)^x = a^x\,b^x, \qquad \left(\frac{a}{b}\right)^x = \frac{a^x}{b^x}.$$

Since $a > 0$ we have that,

$$a^x > 0 \qquad \textit{for all } x \in \mathbb{R}.$$

Definition 2.28 (exponential functions)
Let $a \in \mathbb{R}$ be a positive real number. Any function of the form

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = a^x,$$

*is called an **exponential function**. From Theorem 2.27, $a^x > 0$ for all $x \in \mathbb{R}$.*

We note that for $a = 1$, we have the special case

$$f(x) = 1^x = 1 \qquad \text{for all } x \in \mathbb{R},$$

that is, we have obtained a constant function.

The exponential function with the Euler's number e as base is of particular importance.

Definition 2.29 (exponential function with Euler number e as base)
Euler's number *is defined by*

$$e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = 2.71828\ldots,$$

and the exponential function with ***with base*** e,

$$\exp : \mathbb{R} \to \mathbb{R}, \qquad \exp(x) = e^x,$$

is usually referred to as ***the (natural) exponential function****.*

We observe that by definition

$$\exp(0) = e^0 = 1.$$

Determining a suitable set of function values of the (natural) exponential function $\exp(x) = e^x$, we get the following plot.

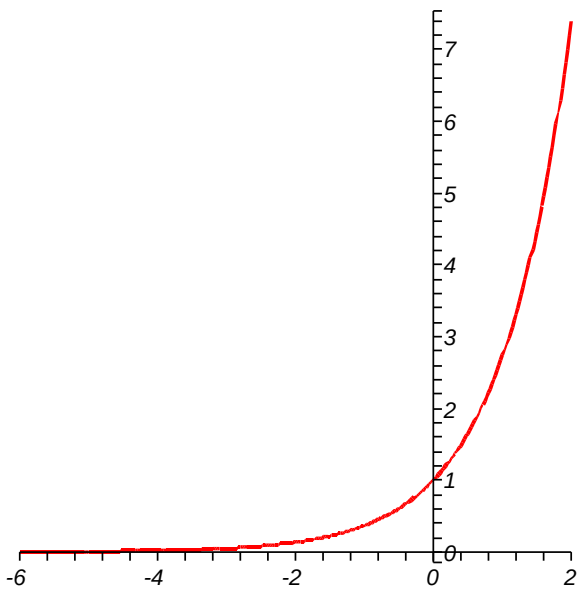


Figure 2.8: Graph of the (natural) exponential function $f(x) = e^x$.

We list some **properties of the (natural) exponential function**:

- The exponential function $\exp(x) = e^x$ **assumes only positive values** (and is thus **never zero**), that is,

$$\exp(x) = e^x > 0 \qquad \text{for all } x \in \mathbb{R}.$$

- $\exp(x) = e^x$ satisfies

$$\lim_{x\to-\infty} \exp(x) = \lim_{x\to-\infty} e^x = 0 \qquad \text{and} \qquad \lim_{x\to\infty} \exp(x) = \lim_{x\to\infty} e^x = \infty.$$

  Thus the x-axis is an **asymptote** for $\exp(x) = e^x$.

- The exponential function $\exp(x) = e^x$ is **strictly monotonically increasing on** $\mathbb{R}$, that is for any $x_1, x_2 \in \mathbb{R}$ with $x_1 < x_2$, we have $e^{x_1} < e^{x_2}$.
- Because $\exp(x) = e^x$ is strictly monotonically increasing on $\mathbb{R}$, the exponential function $\exp(x) = e^x$ is **one-to-one**. Indeed,

$$e^{x_1} = e^{x_2} \qquad \Rightarrow \qquad \frac{e^{x_1}}{e^{x_2}} = 1 \qquad \Rightarrow \qquad e^{x_1-x_2} = 1,$$

  and we know that $e^0 = 1$. Since $e^x < e^0$ for $x < 0$ and $e^0 < e^x$ for $0 < x$ (because $\exp(x) = e^x$ is strictly monotonically increasing on $\mathbb{R}$), we see that $e^{x_1-x_2} = 1 = e^0$ implies that $x_1 - x_2 = 0$, and thus $x_1 = x_2$.

- The **range** of $\exp(x) = e^x$ is given by $\exp(\mathbb{R}) = (0, \infty)$.
- **Euler's number** e is also the value of the infinite sum

$$1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \ldots = 1 + \sum_{k=1}^{\infty} \frac{1}{k!}.$$

2.4 Logarithmic Functions

The **(natural) exponential function** $\exp : \mathbb{R} \to \mathbb{R}$, $\exp(x) = e^x$, is **one-to-one**, and thus it has an **inverse function**.

Definition 2.30 (natural logarithm)
*The **inverse function** $\exp^{-1} : (0, \infty) \to \mathbb{R}$ **of the (natural) exponential function** $\exp : \mathbb{R} \to (0, \infty)$, $\exp(x) = e^x$, is called the **(natural) logarithm**, denoted by $\ln : (0, \infty) \to \mathbb{R}$. From the properties of the inverse function, the natural logarithm is defined by the properties*

$$\ln\big(\exp(x)\big) = \ln(e^x) = x \ \textit{ for all } x \in \mathbb{R}, \qquad \exp\big(\ln(y)\big) = e^{\ln(y)} = y \ \textit{ for all } y > 0.$$

By either evaluating $\ln(x)$ at a suitable set of values of $x > 0$ or by reflecting the graph of the function $\exp(x) = e^x$ on the diagonal $x = y$, we obtain a plot of the graph of the natural logarithm.

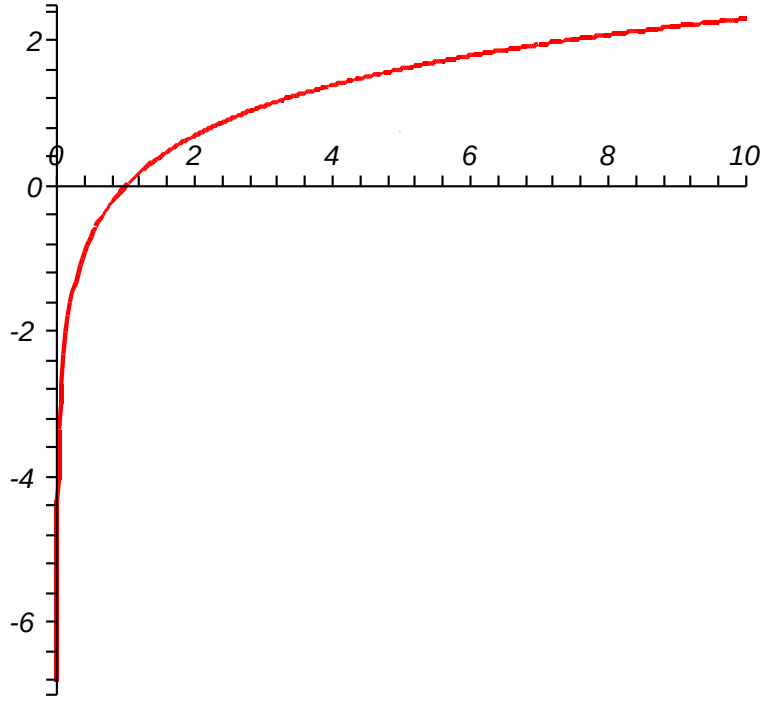


Figure 2.9: Graph of the (natural) logarithm $\ln(x)$.

We list some basic **properties of** $\ln(x)$:

- From $e^0 = 1$ we see that $\ln(1) = 0$, and we have that

$$\ln(x) > 0 \quad \text{for all } x > 1 \qquad \text{and} \qquad \ln(x) < 0 \quad \text{for all } 0 < x < 1.$$

- We have

$$\lim_{x\to\infty} \ln(x) = \infty \qquad \text{and} \qquad \lim_{x>0,\, x\to 0} \ln(x) = -\infty.$$

  Thus the y-axis is an **asymptote** for $\ln(x)$.

- The (natural) logarithm is **strictly monotonically increasing on** $(0, \infty)$, that is for any $x_1, x_2 \in (0, \infty)$ we have

$$0 < x_1 < x_2 \qquad \Rightarrow \qquad \ln(x_1) < \ln(x_2).$$

- As the inverse function of $\exp(x) = e^x$, the (natural) logarithm $\ln : (0, \infty) \to \mathbb{R}$ is **one-to-one**, and its **inverse function** is the (natural) exponential function $\exp : \mathbb{R} \to \mathbb{R}$, $\exp(x) = e^x$.

Furthermore, the (natural) logarithm has the following important properties.

Lemma 2.31 (properties of the (natural) logarithm)
The (natural) logarithm $\ln : (0, \infty) \to \mathbb{R}$ *satisfies*

$$\ln(x\,y) = \ln(x) + \ln(y) \qquad \textit{and} \qquad \ln\left(\frac{x}{y}\right) = \ln(x) - \ln(y), \tag{2.8}$$

and

$$\ln(x^y) = y\,\ln(x), \qquad \textit{and} \qquad \ln\left(x^{1/y}\right) = \frac{1}{y}\,\ln(x). \tag{2.9}$$

Proof: The identities follow from the definition of the (natural) logarithm and the properties of the (natural) exponential function (see Lemma 2.27)

$$e^{\ln(xy)} = x\,y = e^{\ln(x)}\,e^{\ln(y)} = e^{\ln(x)+\ln(y)},$$

and since the exponential function is one-to-one, we may conclude that the identity $\ln(x\,y) = \ln(x) + \ln(y)$ holds. The second property in (2.8) follows analogously. We have

$$e^{\ln(x^y)} = x^y = \left(e^{\ln(x)}\right)^y = e^{y\,\ln(x)},$$

and since the exponential function is one-to-one, we can conclude that $\ln(x^y) = y\,\ln(x)$. The second property in (2.9) follows analogously. □

Like the (natural) exponential function **any exponential function** $f : \mathbb{R} \to \mathbb{R}$, $f(x) = a^x$**, with** $a \neq 1$**, is one-to-one**.

Definition 2.32 (logarithm with base $a \neq 1$)
Let $a \neq 1$*. The* ***logarithm with base*** a*, denoted by* $\log_a : (0, \infty) \to \mathbb{R}$*, is defined as the inverse function of the function* $f : \mathbb{R} \to (0, \infty)$*,* $f(x) = a^x$*, that is,* $\log_a$ *satisfies*

$$\log_a(a^x) = x \quad \textit{for all} \;\; x \in \mathbb{R} \qquad \textit{and} \qquad a^{\log_a(y)} = y \quad \textit{for all} \;\; y \in (0, \infty).$$

We note that for the base $a = e$ (of the (natural) exponential function), we have the natural logarithm, that is, $\boxed{\log_e(y) = \ln(y).}$ Thus Definition 2.30 of the natural logarithm is just a special case of Definition 2.32. We have an analogous lemma to Lemma 2.31, which can be seen as the counterpart to Lemma 2.27.

Lemma 2.33 (properties of logarithmic functions)
Let $a \in \mathbb{R}$ with $a > 0$ and $a \neq 1$. The natural logarithm $\log_a : (0, \infty) \to \mathbb{R}$ satisfies

$$\log_a(x\,y) = \log_a(x) + \log_a(y) \qquad \textit{and} \qquad \log_a\left(\frac{x}{y}\right) = \log_a(x) - \log_a(y), \tag{2.10}$$

and

$$\log_a(x^y) = y\,\log_a(x), \qquad \textit{and} \qquad \log_a\left(x^{1/y}\right) = \frac{1}{y}\,\log_a(x). \tag{2.11}$$

Remark 2.34 (change of base for exponential functions with base $a \neq e$)
Our knowledge of the properties of $\exp(x) = e^x$ and $\ln(x)$ allows us to analyze exponential functions with bases $a \neq 1$ (that differ from e) with a simple ***'change of the base'****. We have*

$$\boxed{a^x = \left(e^{\ln(a)}\right)^x = e^{x\ln(a)} = \exp\big(x\,\ln(a)\big).}$$

Since $\ln(a) > 0$ for all $a > 1$, and $\ln(a) < 0$ for all $a < 1$, we see that

$$a^x = e^{\ln(a)\,x} = \begin{cases} e^{|\ln(a)|\,x} & \textit{if} \quad a > 1, \\ e^{-|\ln(a)|\,x} & \textit{if} \quad a < 1. \end{cases}$$

For illustration purposes we have plotted $f_1(x) = 2^x$ and $f_2(x) = (1/2)^x = 2^{-x}$ in the left picture in Figure 2.10.

Remark 2.35 (change of base for logarithms with base $a \neq e$)
Our knowledge of the properties of $\exp(x) = e^x$ and $\ln(x)$ allows us to analyze logarithms with bases $a \neq 1$ (that differ from e) with a simple ***'change of the base'****. We have*

$$\boxed{\log_a(x) = \log_a\left(b^{\log_b(x)}\right) = \log_b(x)\,\log_a(b) = \log_a(b)\,\log_b(x),}$$

and for the special case of $b = e$ we obtain

$$\boxed{\log_a(x) = \log_a(e)\,\log_e(x) = \log_a(e)\,\ln(x).}$$

Since

$$\log_a(x) \text{ is } \begin{cases} > 0 & \text{if } 0 < x < 1 \text{ and } a < 1, \\ < 0 & \text{if } x > 1 \text{ and } a < 1, \\ < 0 & \text{if } 0 < x < 1 \text{ and } a > 1, \\ > 0 & \text{if } x > 1 \text{ and } a > 1, \end{cases}$$

we have (note $1 < e$*)*

$$\log_a(x) = \log_a(e)\,\ln(x) = \begin{cases} |\log_a(e)|\,\ln(x) & \text{if } a > 1, \\ -|\log_a(e)|\,\ln(x) & \text{if } a < 1. \end{cases}$$

For illustration purposes we have plotted $g_1(x) = \log_2(x)$ *and* $g_2(x) = \log_{1/2}(x)$ *in the right picture in Figure 2.10.*

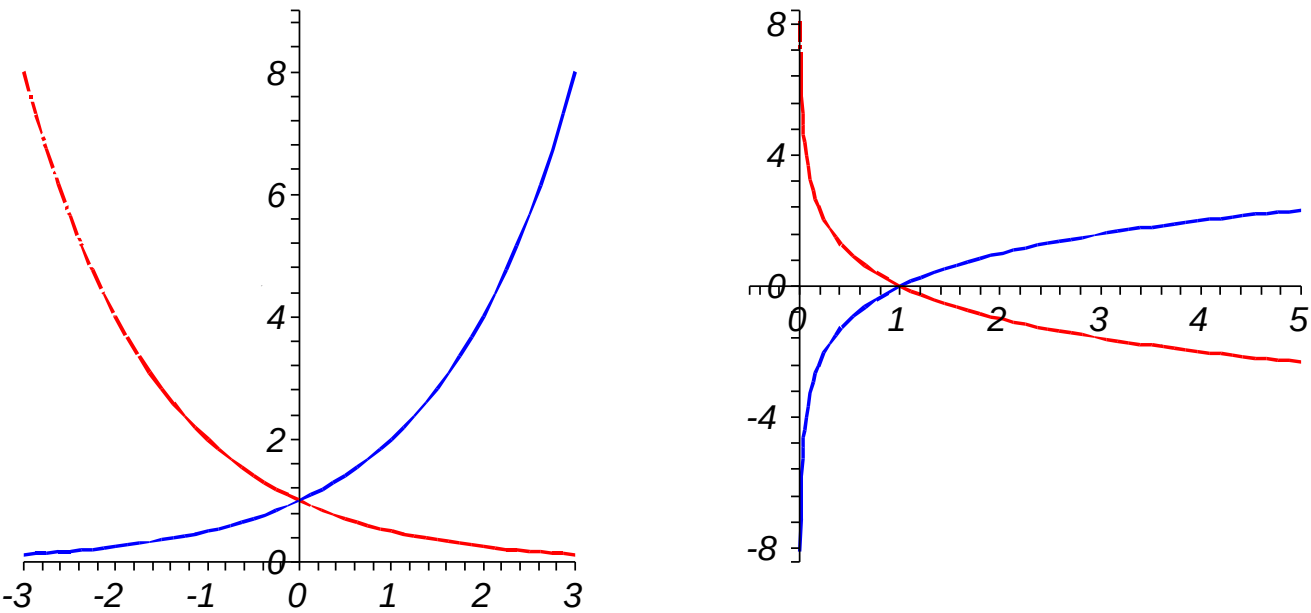


Figure 2.10: Graphs of $f_1(x) = 2^x$ (blue) and $f_2(x) = (1/2)^x$ (red) on the left, and the graphs of $g_1(x) = \log_2(x)$ (blue) and $g_2(x) = \log_{1/2}(x)$ (red) on the right.

Application 2.36 (Kepler's third law)
Kepler's Third Law *says that* T^2 ***is proportional to*** R^3*, where* T *is the time it takes a planet to orbit the sun, and* R *is its mean distance from the sun:*

$$T^2 \propto R^3,$$

that is, for some constant $K \in \mathbb{R}$*, with* $K > 0$*,*

$$T^2 = K\,R^3.$$

Taking the natural logarithm on both sides gives that

$$\ln(T^2) = \ln(K\,R^3) \qquad \Rightarrow \qquad 2\ln(T) = \ln(K) + \ln(R^3) = \ln(K) + 3\ln(R).$$

It follows that

$$\ln(T) = \frac{\ln(K)}{2} + \frac{3}{2}\ln(R),$$

that is, $\ln(T)$ *is an affine linear function*

$$f\big(\ln(R)\big) = m\,\ln(R) + c$$

of $\ln(R)$*, where* $m = 3/2$ *and* $c = \ln(K)/2$*. Thus a plot of* $\ln(T)$ *a gainst* $\ln(R)$*, which Kepler could draw, showed a* ***set of points on a straight line of slope*** $3/2$*. That's how Kepler's third law was discovered with the help of astronomical observation.* □

From this example we see that for relations between variables x and y of the form $c_1\,x^p = c_2\,y^q$ (where c_1 and c_2 are constants), taking the logarithm on both sides gives us a straight line, which is easier to draw and analyze than a curve.

2.5 Hyperbolic Functions

The last topic in this chapter are so-called **hyperbolic functions**.

Definition 2.37 (hyperbolic sine and hyperbolic cosine)
We define the ***hyperbolic sine*** $\sinh : \mathbb{R} \to \mathbb{R}$ *and the* ***hyperbolic cosine*** $\cosh : \mathbb{R} \to \mathbb{R}$ *by*

$$\sinh(x) = \frac{e^x - e^{-x}}{2} \qquad \textit{and} \qquad \cosh(x) = \frac{e^x + e^{-x}}{2}.$$

We observe from the definitions of $\sinh(x)$ and $\cosh(x)$ that (since $\lim_{x\to\infty} e^{-x} = 0$) for large x, $\sinh(x) \approx e^x/2$ and $\cosh(x) \approx e^x/2$. Likewise (since $\lim_{x\to-\infty} e^x = 0$) for small x, $\sinh(x) \approx -e^{-x}/2$ and $\cosh(x) \approx e^{-x}/2$. The hyperbolic sine and the hyperbolic cosine have the graphs below.

Lemma 2.38 (equation with $\sinh(x)$ and $\cosh(x)$)
We have that

$$\big(\cosh(x)\big)^2 - \big(\sinh(x)\big)^2 = 1 \qquad \textit{for all } x \in \mathbb{R}.$$

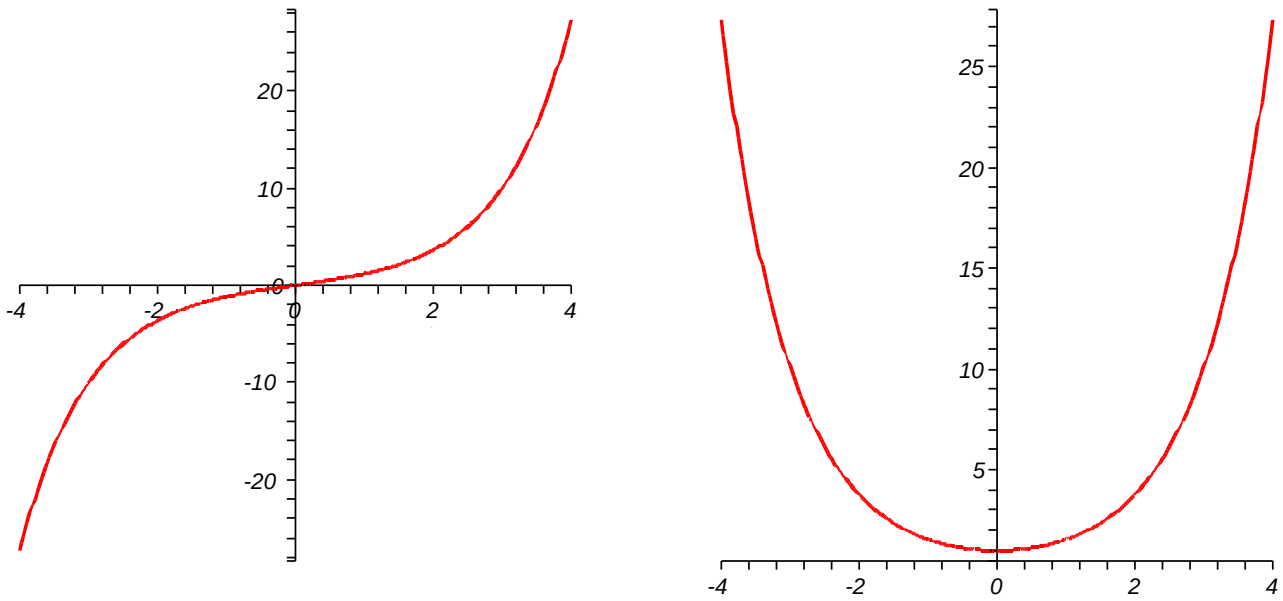


Figure 2.11: The graph of $\sinh(x)$ on the left, and the graph of $\cosh(x)$ on the right.

Proof: We simplify the the expression on the right-hand side.

$$\begin{aligned}
\cosh^2(x) - \sinh^2(x) &= \left(\frac{e^x + e^{-x}}{2}\right)^2 - \left(\frac{e^x - e^{-x}}{2}\right)^2 \\
&= \frac{(e^x + e^{-x})^2 - (e^x - e^{-x})^2}{4} \\
&= \frac{(e^{2x} + 2 + e^{-2x}) - (e^{2x} - 2 + e^{-2x})}{4} = 1,
\end{aligned}$$

where we have used the fact that $e^x\, e^{-x} = e^{x-x} = e^0 = 1$. □

Example 2.39 ($\sinh(x)$ is odd and $\cosh(x)$ is even)
Show formally that $\sinh(x)$ is an odd function and that $\cosh(x)$ is an even function.

Solution:

$$\sinh(-x) = \frac{e^{-x} - e^{-(-x)}}{2} = \frac{e^{-x} - e^x}{2} = -\frac{e^x - e^{-x}}{2} = -\sinh(x), \qquad x \in \mathbb{R},$$

and

$$\cosh(x) = \frac{e^x + e^{-x}}{2} = \frac{e^{-x} + e^{-(-x)}}{2} = \cosh(-x), \qquad x \in \mathbb{R},$$

as claimed. □

The next lemma lists some useful formulas for the functions $\sinh(x)$ and $\cosh(x)$

Lemma 2.40 (addition theorems)
We have for all $x \in \mathbb{R}$ that

$$\begin{aligned} \sinh(x+y) &= \sinh(x)\cosh(y) + \sinh(y)\cosh(x), \\ \cosh(x+y) &= \cosh(x)\cosh(y) + \sinh(x)\sinh(y). \end{aligned}$$

Next we introduce some more hyperbolic functions.

Definition 2.41 (hyperbolic tangent and hyperbolic cotangent function)
*The **hyperbolic tangent function** $\tanh : \mathbb{R} \to \mathbb{R}$ and the **hyperbolic cotangent function** $\coth : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ are defined by*

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \qquad \coth(x) = \frac{\cosh(x)}{\sinh(x)} = \frac{1}{\tanh(x)} = \frac{e^x + e^{-x}}{e^x - e^{-x}}.$$

The hyperbolic tangent function $\tanh(x)$ and the hyperbolic cotangent function $\coth(x)$ are plotted in Figure 2.12 below.

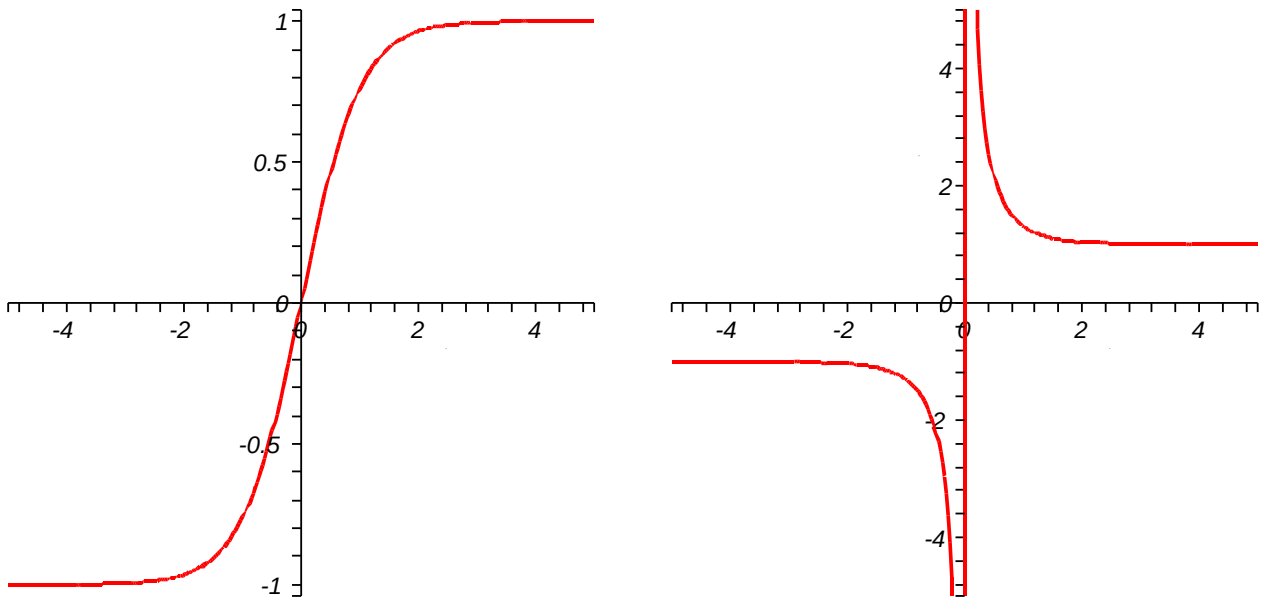


Figure 2.12: The graph of $\tanh(x)$ on the left and the graph of $\coth(x)$ on the right.

We observe that $g(x) = -1$ and $h(x) = 1$ are asymptotes for $\tanh(x)$ and $\coth(x)$, because for $|x|$ large we have

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \approx \frac{e^x}{e^x} = 1 \qquad \text{as } x \to \infty,$$
$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \approx \frac{-e^{-x}}{e^{-x}} = -1 \quad \text{as } x \to -\infty$$

and

$$\coth(x) = \frac{e^x + e^{-x}}{e^x - e^{-x}} \approx \frac{e^x}{e^x} = 1 \qquad \text{as } x \to \infty,$$
$$\cot(x) = \frac{e^x + e^{-x}}{e^x - e^{-x}} \approx \frac{e^{-x}}{-e^{-x}} = -1 \quad \text{as } x \to -\infty.$$

You may also encounter the hyperbolic functions $\operatorname{cosech} : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, and $\operatorname{sech} : \mathbb{R} \to \mathbb{R}$, defined by

$$\operatorname{cosech}(x) = \frac{1}{\sinh(x)}, \qquad \operatorname{sech}(x) = \frac{1}{\cosh(x)}.$$

Finally we want to discuss the **inverse hyperbolic functions**. From Figures 2.11 and 2.12, it is relatively clear that the functions $\sinh(x)$ and $\tanh(x)$ are one-to-one on $\mathbb{R}$, and $\cosh(x)$ is one-to-one for $x \geq 0$, and likewise for $x \leq 0$, but not on all of $\mathbb{R}$. The hyperbolic cotangent function $\coth(x)$ is one-to-one for $\mathbb{R} \setminus \{0\}$.

Definition 2.42 (inverse hyperbolic functions)

(i) The inverse function of $\sinh : \mathbb{R} \to \mathbb{R}$ *is called* $\sinh^{-1} = \operatorname{arcsinh} : \mathbb{R} \to \mathbb{R}$.

(ii) The inverse function of $\cosh : [0, \infty) \to [1, \infty)$ *is called* $\cosh^{-1} = \operatorname{arccosh} : [1, \infty) \to [0, \infty)$.

(iii) The inverse function of $\tanh : \mathbb{R} \to (-1, 1)$ *is called* $\tanh^{-1} = \operatorname{arctanh} : (-1, 1) \to \mathbb{R}$.

(iv) The inverse function of $\coth : \mathbb{R} \setminus \{0\} \to \mathbb{R} \setminus [-1, 1]$ *is called* $\coth^{-1} = \operatorname{arccoth} : \mathbb{R} \setminus [-1, 1] \to \mathbb{R} \setminus \{0\}$.

We plot the inverse hyperbolic functions in Figures 2.13 and 2.14. Their properties and graphs can easily by derived from the properties of the functions $\sinh(x)$, $\cosh(x)$, $\tanh(x)$, and $\coth(x)$. For example, we know that $g(x) = -1$ and $h(x) = 1$ are asymptotes for $\tanh(x)$ and $\coth(x)$. Therefore, we know that the vertical line

through $(-1, 0)$ and the vertical line through $(1, 0)$ are asymptotes of $\operatorname{arctanh}(x)$ and $\operatorname{arccoth}(x)$.

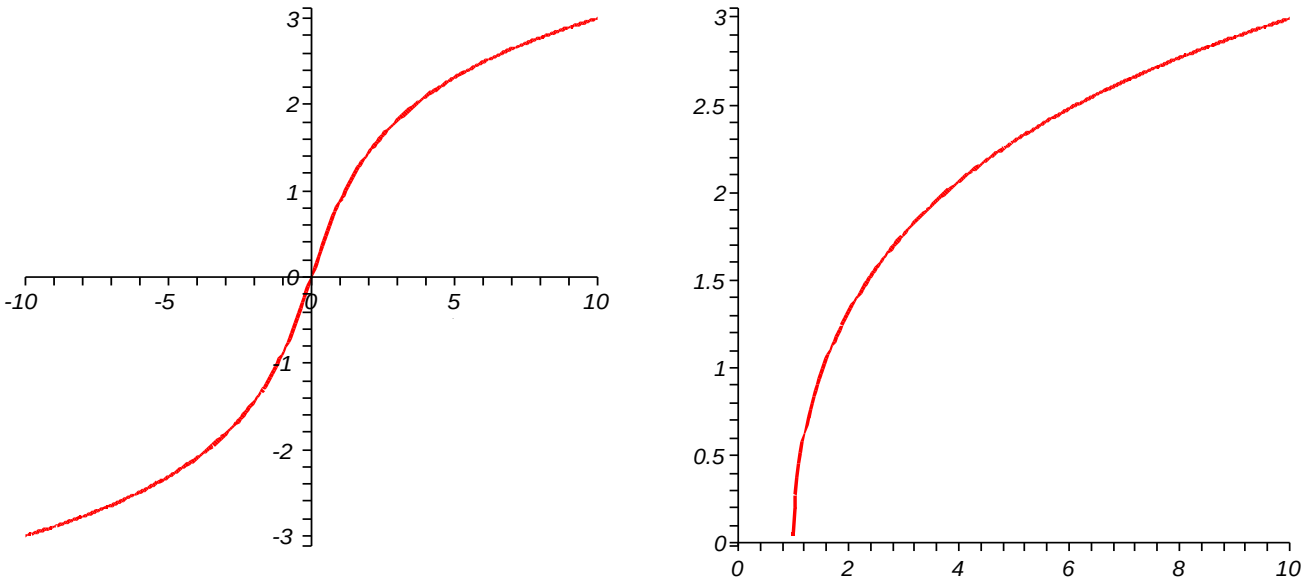


Figure 2.13: The graph of $\operatorname{arcsinh}(x)$ on the left, and the graph of $\operatorname{arccosh}(x)$ on the right.

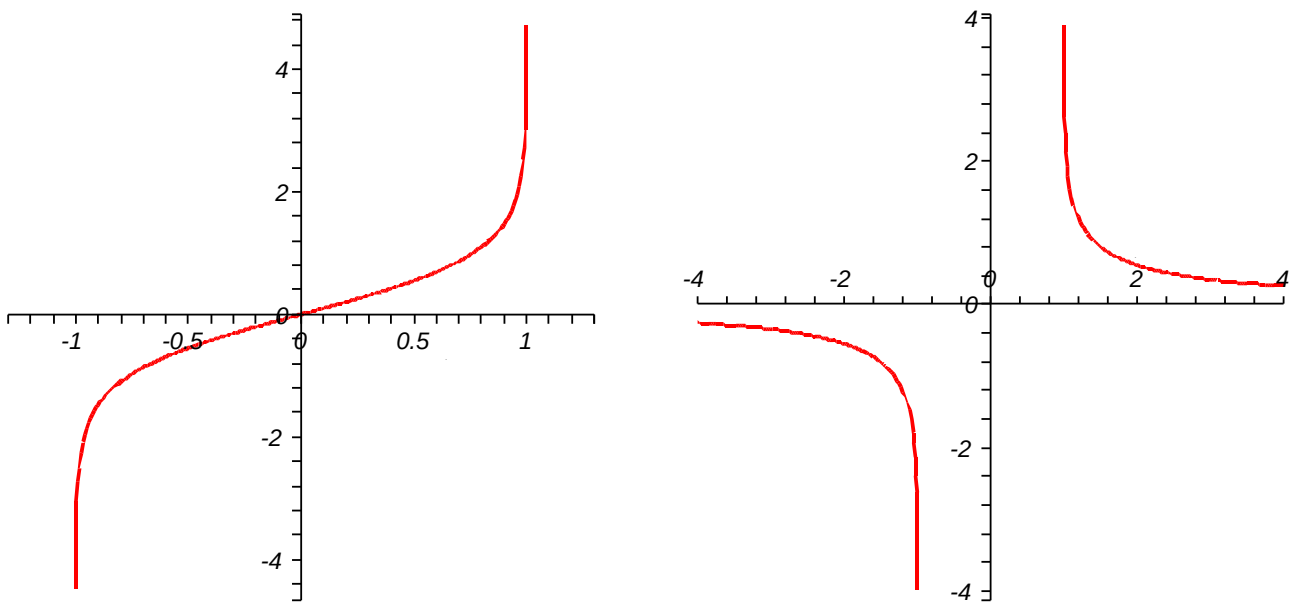


Figure 2.14: The graph of $\operatorname{arctanh}(x)$ on the left and the graph of $\operatorname{arccoth}(x)$ on the right.

Example 2.43 (representation of the inverse function of $\sinh(x)$)
Find an explicit representation of $\operatorname{arcsinh}(x)$.

Solution: We set $\sinh(x) = y$ and solve for x.

$$\sinh(x) = \frac{e^x - e^{-x}}{2} = y \quad \Rightarrow \quad e^x - e^{-x} = 2\,y$$

and after replacing $z = e^x$, we get

$$\begin{aligned} z - \frac{1}{z} = 2\,y \quad &\Rightarrow \quad z^2 - 1 = 2\,y\,z \quad \Rightarrow \quad z^2 - 2\,y\,z - 1 = 0 \\ \Rightarrow \quad (z-y)^2 - (1+y^2) = 0 \quad &\Rightarrow \quad \big(z - y + \sqrt{1+y^2}\big)\big(z - y - \sqrt{1+y^2}\big) = 0 \\ \Rightarrow \quad z = y \pm \sqrt{1+y^2}. \end{aligned}$$

Thus we see that

$$e^x = y \pm \sqrt{1+y^2},$$

and since $e^x > 0$ and $\sqrt{1+y^2} \geq \sqrt{y^2} = y$, we have $y - \sqrt{1+y^2} < y - y = 0$ and we can rule out the solution with the minus sign. Thus

$$e^x = y + \sqrt{1+y^2} \quad \Rightarrow \quad x = \ln\big(y + \sqrt{1+y^2}\big) \quad \Rightarrow \quad \operatorname{arcsinh}(y) = \ln\big(y + \sqrt{1+y^2}\big),$$

and we have obtained a representation of $\arcsin(x)$ in terms of other classical functions. □

Chapter 3

Differentiation

In this chapter we learn more properties of functions, namely continuity and differentiability. In Section 3.1, we briefly discuss continuity. If a function is **continuous**, then its graph is one 'continuous' curve. The functions that you will encounter in physics will mostly be continuous, but you may occasionally encounter functions with jumps or possibly with a singularity. Most of the functions that we discuss in this class will be continuous and even much smoother than just continuous. In Section 3.2, we introduce the **derivative** of a smooth enough function, and we discuss its **geometric interpretation**. In Section 3.3 we will learn the derivatives of most of the classical functions from Chapter 2. In Section 3.4, 3.5, and 3.6, we learn rules/techniques for computing derivatives, most notably, the **product rule**, the **quotient rule**, and the **chain rule**.

Functions and derivatives play a fundamental role in physical sciences. We will encounter two applications from physics in this chapter to illustrate this.

The first application is the **motion** of a car along a straight line as a function of the time. The **first derivative** describes the **velocity** of the car, and the **second derivative** describes the **acceleration** of the car.

The second application is the **radioactive decay** of radioactive materials such as uranium. This can be described by a very simple **(ordinary) differential equation** that we will solve by 'inspection'/guessing.

Intuitively **differential equations** are equations that express relations between an unknown function and some of its derivatives, and they arise in almost all areas of physics and technical processes. The aim is to **solve the equation**, that is, to **find the unknown function that describes the physical/technical process**.

For example, **electro-magnetic phenomena** are descried by the **Maxwell equations** (a system of partial differential equations) which describe the interrelationship between electric field, magnetic field, electric charge, and electric current. The **Navier-Stokes equations** describe the **motion of fluid substances such as liquids and gases**. For example, they are used to model **weather**, **ocean currents**, **water flow in a pipe**, and the **flow around an airfoil (wing)**. But also much simpler phenomena such as the **oscillation of the strings of a guitar** or the **vibration of the membrane of a drum** are described by a partial differential equation, the so-called **wave equation**. The **temperature distribution** in a metal bar is described by a partial differential equation, the so-called **heat equation**. **Population dynamics** can also be modeled with a differential equation.

The great importance of (partial and ordinary) differential equations for the description of technical processes in engineering, physics, and other sciences, motivates why we will devote a substantial part of this course to the discussion of functions and how to analyze, differentiate, and integrate them. This knowledge is a **prerequisite** for learning later-on **how to describe and model technical processes** with the help of functions, differential equations, and other mathematical tools.

3.1 Continuity

We briefly discuss the notion of **continuity of a function** which means that the graph of the function is a 'continuous curve'.

Definition 3.1 (continuous function)
*Let $A, B \subset \mathbb{R}$. A function in $f : A \to B$ is **continuous at the point** $x_0 \in A$ if we have that*

$$\lim_{x \in A,\, x \to x_0} f(x) = f(x_0).$$

*The function f is **continuous on** A, if it is continuous at all $x_0 \in A$. If a function f is **not continuous at a point** $x_0 \in A$ then we also say that f is **discontinuous at** x_0.*

We can easily see from the graph of a function if the function is continuous. If $f : A \to B$ it is continuous on its domain A, then the graph of f is one 'continuous' connected curve. On the other hand, if the graph of f breaks of at some point and continues at another, then the function is not continuous at the point where this 'jump' occurs. A function f is continuous at $x_0 \in A$ if the graph is continuous in

a neighborhood of x_0, that is, we can follow the graph starting from $(x_0, f(x_0))$ in both directions (for increasing x and for declining x) for a short distance. This is illustrated in Figure 3.1.

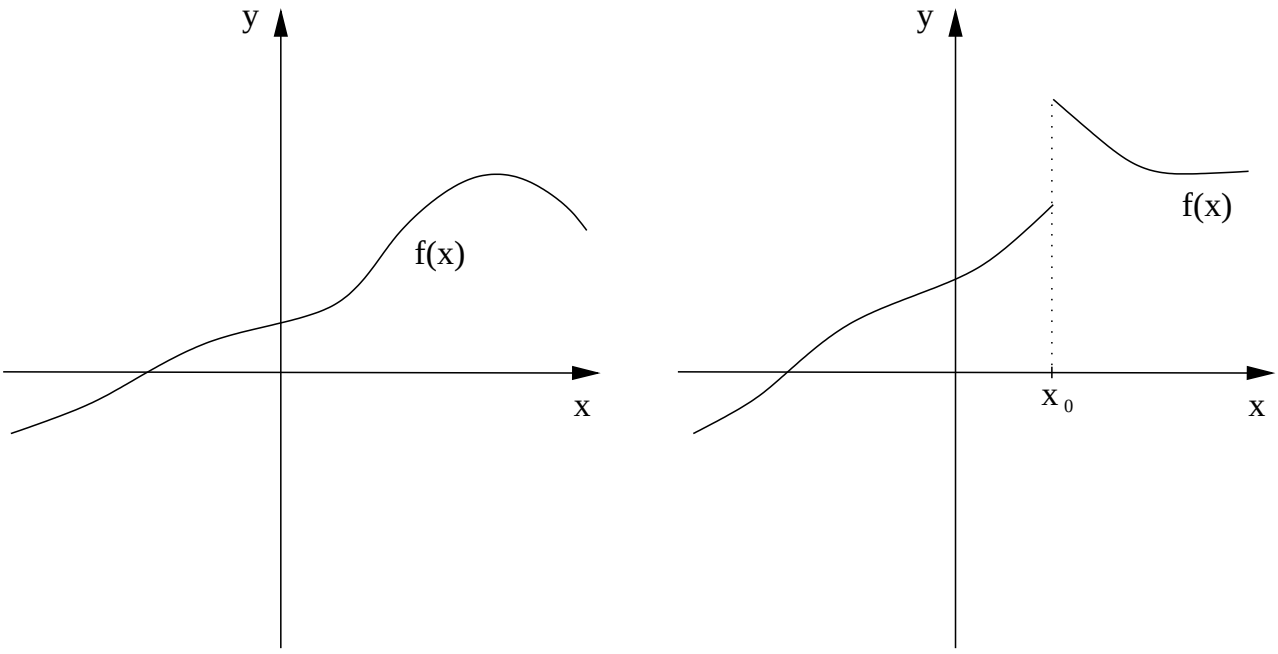


Figure 3.1: The function on the left is continuous everywhere. The function on the right is discontinuous at $x = x_0$, because the function has a jump.

We will not go into rigorously proving whether a function is continuous or not, but you should be able to determine from the graph whether the function is continuous at all points, and if not at which point the function is discontinuous.

All the classical functions that we encountered in the previous chapter are continuous and are, in fact, much smoother than just continuous.

The function that you encounter in physics will usually be either continuous, or they may have a number of jumps, or even a singularity (that is, a point x_0 at which $\lim_{x\to x_0} |f(x)| = \infty$).

Example 3.2 (continuous and discontinuous functions)

(a) The functions $\sin(x)$, $\cos(x)$, and e^x are continuous on $\mathbb{R}$.

(b) Affine linear functions

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = m\,x + c,$$

and quadratic functions

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = a\,x^2 + b\,x + c,$$

and, more general, all polynomials are continuous on $\mathbb{R}$.

(c) The natural logarithm $\ln : (0, \infty) \to \mathbb{R}$ is continuous on $(0, \infty)$.

(d) The function $f : \mathbb{R} \to \mathbb{R}$, defined by

$$f(x) = \begin{cases} -1 & \text{if} \quad x < 0, \\ 0 & \text{if} \quad x = 0, \\ 1 & \text{if} \quad x > 0, \end{cases} \tag{3.1}$$

is discontinuous at $x = 0$, but continuous at all $x \in \mathbb{R} \setminus \{0\}$. This can be seen from the graph in Figure 3.2. In Figure 3.2, the empty dots $\circ$ indicate that the value at $x = 0$ is not given by -1 or 1, whereas the filled dot $\bullet$ indicates that the function value at $x = 0$ is given by $y = 0$. $\square$

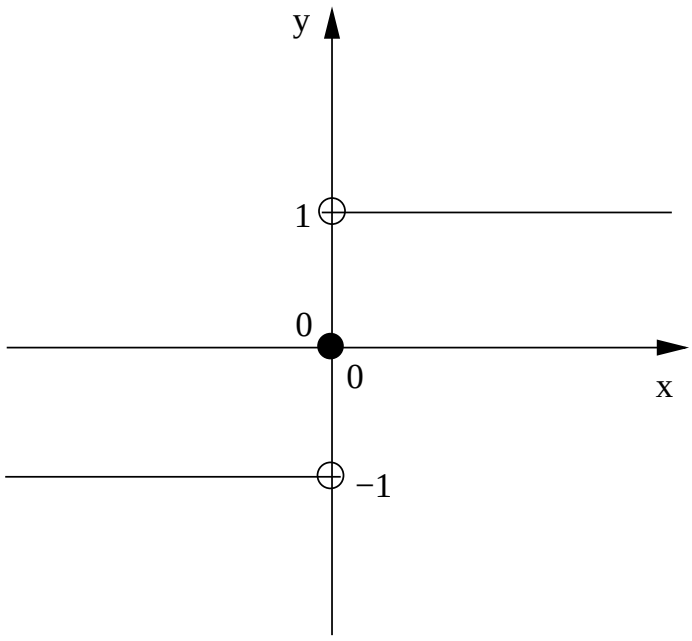


Figure 3.2: Graph of $f : \mathbb{R} \to \mathbb{R}$ defined by (3.1).

3.2 Differentiation

The **derivative** of a function f at a point x_0 is an accurate way of measuring the **rate of change** at this point. By the rate of change we mean here how much $f(x)$ grows or declines if x gets smaller than x_0 and larger than x_0, respectively.

A basic way to measure how a function $f : A \to B$ changes as x gets larger than x_0 is to consider the change in the y-direction $\Delta y = f(x_0 + \Delta x) - f(x_0)$ with a small positive Δx. Since we want the rate of change of f at x_0 relative to the

corresponding distance on the x-axis, an **approximation of the rate of change of f at x_0** is given by

$$\frac{f(x_0+\Delta x)-f(x_0)}{(x_0+\Delta x)-x_0}=\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}=\frac{\Delta y}{\Delta x}. \tag{3.2}$$

In fact, $\Delta y/\Delta x$ is the **average rate of change of f over the interval** $[x_0, x_0+\Delta x]$ $=\{x\in\mathbb{R} : x_0\le x\le x_0+\Delta x\}$. See Figure 3.3 for an illustration.

Likewise we can approximate the rate of change of $f(x)$ at x_0 as x declines by

$$\frac{f(x_0)-f(x_0-\Delta x)}{x_0-(x_0-\Delta x)}=\frac{f(x_0)-f(x_0-\Delta x)}{\Delta x}=\frac{\Delta y}{\Delta x}. \tag{3.3}$$

The expression (3.3) is the **average rate of change of f over the interval** $[x_0-\Delta x, x_0]=\{x\in\mathbb{R} : x_0-\Delta x\le x\le x_0\}$.

The expressions (3.2) and (3.3) are called a **Newton quotients** or **difference quotients** of the function f.

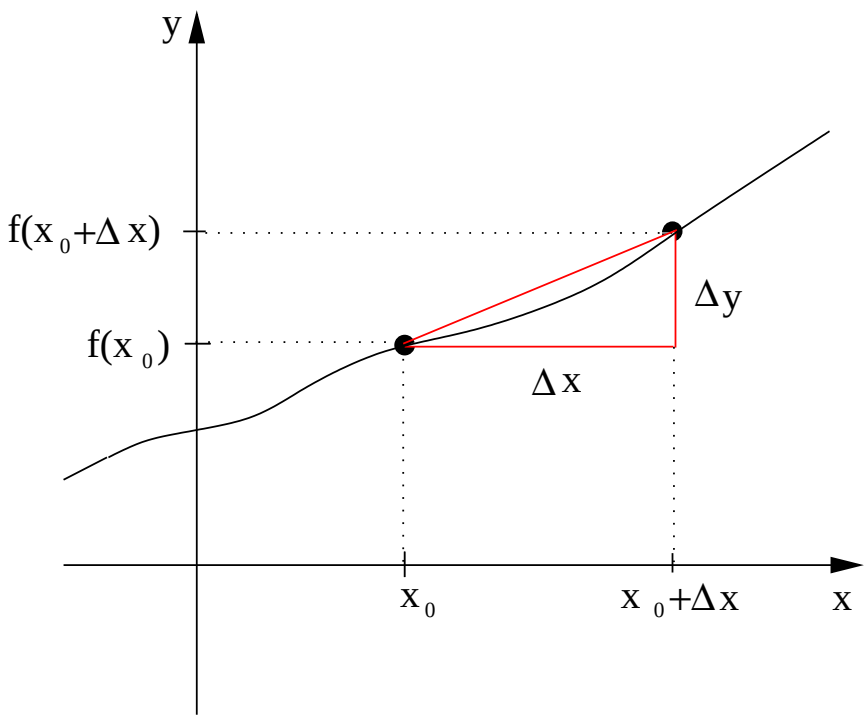


Figure 3.3: Approximation of the rate of change by the Newton quotient.

Off course we can do **better than to just get an approximation** of the rate of change of f at x_0! The idea is to let Δx in (3.2) and (3.3) shrink to zero, that is, we **take the limit for $\Delta x\to 0$**.

$$\text{rate of change of } f \text{ at } x_0 \;=\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}=\lim_{\Delta x\to 0}\frac{f(x_0)-f(x_0-\Delta x)}{\Delta x}.$$

That it does make **no difference whether we approach x_0 from the left or from the right** is only true if our function f **is 'smooth enough'**. Most functions that you encounter in physics are such smooth enough functions. 'Smooth enough' means more than continuous.

These heuristic observations lead to the definition of the derivative.

Definition 3.3 (derivative)
*Let $A, B \subset \mathbb{R}$. The **derivative** $df(x_0)/dx = f'(x_0)$ of a (smooth enough) function $f : A \to B$ **at the point** $x_0 \in A$ is defined by*

$$\frac{df(x_0)}{dx} = f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h},$$

*where h can be positive or negative, subject to $x_0 + h \in A$. The value $f'(x_0)$ measures the **rate of change** of the function f at the point $x = x_0$. The rate of change is the **slope of the tangent** of the graph at the point $(x_0, f(x_0))$.*

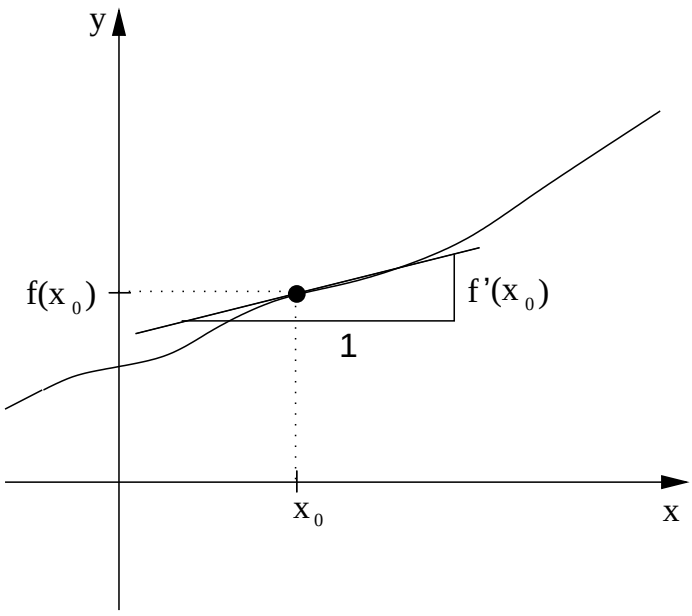


Figure 3.4: The derivative value $f'(x_0)$ is the rate of change of the function f at $x = x_0$, and the tangent of the graph at $(x_0, f(x_0))$ has the slope $f'(x_0)$.

*The function $f' : A \to \mathbb{R}$, defined by mapping $x \in A$ onto the derivative $f'(x)$ at the point x, is called the **derivative of** f.*

Note that this definition just formalizes our considerations above with $h = \pm\Delta x$ and $x = x_0$. The process of determining the derivative of f is called **differentiating the function** f.

It seems plausible that we may repeat the procedure if f is smooth enough to allow this. And we can do this again and again. This is formalized in the next definition. The **first two derivatives of physical quantities often have a physical meaning** as we will see in Examples 3.11 and 3.12 later in this section.

Definition 3.4 (higher order derivatives)
Let $f : A \to B$ be a smooth function. Once we have computed the derivative $f' : A \to \mathbb{R}$ of f, we can differentiate the derivative again. The derivative of the (first oder) derivative $f' : A \to \mathbb{R}$ is denoted by $f'' = d^2 f/dx^2$ and is given by

$$f''(x) = \frac{d^2 f(x)}{dx^2} = \frac{d}{dx}\left(\frac{df(x)}{dx}\right) = \big(f'(x)\big)',$$

*and we called it the **second (order) derivative of** f. We can repeat this process again and again, and so we can differentiate f m-times, and we denote the resulting function by*

$$f^{(m)}(x) = \frac{d^m f}{dx^m} = \frac{d}{dx}\left(f^{(m-1)}(x)\right),$$

*and call it the m**th (order) derivative of** f. With this notation $f^{(1)} = f'$ and $f^{(2)} = f''$.*

We give some examples.

Example 3.5 (derivative of a constant function)
The constant function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = c$, with an arbitrary constant $c \in \mathbb{R}$, has the derivative $f'(x) = 0$. Indeed, for any $x \in \mathbb{R}$

$$\frac{f(x+h) - f(x)}{h} = \frac{c - c}{h} = \frac{0}{h} = 0 \quad \Rightarrow \quad f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} 0 = 0,$$

as claimed. □

Example 3.6 (1st and 2nd order derivative of an affine linear function)
Consider the **affine linear function** $f : \mathbb{R} \to \mathbb{R}$, defined by

$$f(x) = m\,x + c,$$

whose graph is a straight line with slope m which intersects the y-axis at $y = c$. Because we know that the slope of the straight line is m, we expect that $f'(x) = m$. Indeed, we have that

$$\frac{f(x+h) - f(x)}{h} = \frac{[m\,(x+h) + c] - [m\,x + c]}{h} = \frac{m\,h}{h} = m.$$

Thus we find that

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} m = m,$$

as expected. The second order derivative $f''(x)$ is given by $f''(x) = 0$ because we know from the previous example that the derivative of the constant function $f'(x) = m$ is the zero function. □

An important and useful result for analyzing functions is the following lemma, which establishes a relation between the sign of the derivative and the growth of the function. Since $f'(x)$ is the **slope of the tangent** of the graph at $(x, f(x))$, the result below is not particularly surprising.

Lemma 3.7 (derivative and growth of a function)
Let $A, B \subset \mathbb{R}$, and let $[a, b]$ be a subinterval of A. Let $f : A \to B$ be a sufficiently smooth function. Then the following statements hold true:

*(i) The function f is **strictly monotonically increasing on** $[a, b]$ if*

$$f'(x) > 0 \qquad \text{for all } x \in (a, b).$$

*(ii) The function f is **monotonically increasing on** $[a, b]$ if*

$$f'(x) \geq 0 \qquad \text{for all } x \in (a, b).$$

*(iii) The function f is **strictly monotonically decreasing on** $[a, b]$ if*

$$f'(x) < 0 \qquad \text{for all } x \in (a, b).$$

*(iv) The function f is s**monotonically decreasing on** $[a, b]$ if*

$$f'(x) \leq 0 \qquad \text{for all } x \in (a, b).$$

We illustrate Lemma 3.7 for the functions from Examples 3.5 and 3.6.

Example 3.8 (growth of a constant function)
In Example 3.5, we saw that the **constant function** $f : \mathbb{R} \to \mathbb{R}$, $f(x) = c$, with an arbitrary constant $c \in \mathbb{R}$, has the derivative $f'(x) = 0$. Since $f'(x) = 0$, we know from Lemma 3.7 that the function $f(x) = c$ is both monotonically increasing on $\mathbb{R}$ and monotonically decreasing on $\mathbb{R}$. □

Example 3.9 (growth of an affine linear function)
In Example 3.6, we saw that an **affine linear function** $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = m\,x + c$, has the derivative $f'(x) = m$. From Lemma 3.7, we know that f is strictly monotonically increasing on $\mathbb{R}$ if $m > 0$, and that f is strictly monotonically decreasing on $\mathbb{R}$ if $m < 0$. □

Next we want to compute the first and second order derivative of a quadratic function.

Example 3.10 (1st and 2nd order derivative of a quadratic function)
Consider the **quadratic function** $f : \mathbb{R} \to \mathbb{R}$, defined by

$$f(x) = a\,x^2 + b\,x + c.$$

Then we have

$$\begin{aligned} f(x+h) - f(x) &= \big[a\,(x+h)^2 + b\,(x+h) + c\big] - \big[a\,x^2 + b\,x + c\big] \\ &= \big[a\,x^2 + 2\,a\,h\,x + a\,h^2 + b\,x + b\,h + c\big] - \big[a\,x^2 + b\,x + c\big] \\ &= 2\,a\,h\,x + a\,h^2 + b\,h. \end{aligned}$$

Thus we see that

$$\frac{f(x+h) - f(x)}{h} = \frac{2\,a\,h\,x + a\,h^2 + b\,h}{h} = 2\,a\,x + a\,h + b.$$

Now we let $h \to 0$, and obtain

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} (2\,a\,x + a\,h + b) = 2\,a\,x + b.$$

Thus $f'(x) = 2\,a\,x + b$. From Example 3.6, we know that

$$f''(x) = \frac{d}{dx}\big(2\,a\,x + b\big) = 2\,a.$$ □

To round of this section we will discuss the application of **motion in time along a straight line**.

Application 3.11 (velocity and acceleration)
Think of a car moving in a straight line. Then the ***velocity*** *of the car is defined to be the* ***rate of change of position****. Let $s = s(t)$ denote the position at time t, and in our coordinate plane the horizontal axis measures the time t and the vertical axis the position s. Then the* ***average velocity of the car*** *moving from position*

$s(t_1)$ to $s(t_2)$ during the time interval $[t_1, t_2] = \{t \in \mathbb{R} : t_1 \le t \le t_2\}$ is the quotient of the displacement $\Delta s = s(t_2) - s(t_1)$ divided by the time difference $\Delta t = t_2 - t_1$, that is,

$$v = \frac{\Delta s}{\Delta t} = \frac{s(t_2) - s(t_1)}{t_2 - t_1}.$$

By shrinking the length of the time interval to zero, we find that the ***velocity*** *$v(t)$ at the time t is given by the derivative $ds(t)/dt$. We summarize:*

$$s(t) = \textit{position at time } t, \qquad v(t) = \frac{ds(t)}{dt} = \textit{velocity at time } t. \tag{3.4}$$

When the driver starts his car, then he has to accelerate at the beginning to reach the desired velocity. The ***average acceleration*** *a over a time interval $[t_1, t_2]$ is defined as*

$$\textit{average acceleration} \;=\; \frac{\textit{velocity change}}{\textit{time taken for change}} = \frac{\Delta v}{\Delta t} = \frac{v(t_2) - v(t_1)}{t_2 - t_1}.$$

To get the ***acceleration*** *$a(t)$* ***at the time*** *t, we shrink the time interval in the quotient above, that is, $\Delta t \to 0$, and obtain $a(t) = dv(t)/dt$. Thus*

$$v(t) = \textit{velocity at time } t, \qquad a(t) = \frac{dv(t)}{dt} = \textit{acceleration at time } t. \tag{3.5}$$

We summarize what we have learned: If $s : [0, T] \to \mathbb{R}$ describes a motion during the time interval $[0, T] = \{t \in \mathbb{R} : 0 \le t \le T\}$, then we have

$$\begin{aligned} s(t) &= \textit{position at time } t, \\ v(t) &= \frac{ds(t)}{dt} = \textit{velocity at time } t, \\ a(t) &= \frac{dv(t)}{dt} = \frac{d^2 s(t)}{dt^2} = \textit{acceleration at time } t, \end{aligned}$$

where $d^2 s(t)/dt^2$ means that we take two times the derivative.

Application 3.12 (motion with uniform acceleration)
As in the previous example let $s : [0, T] \to \mathbb{R}$ denote the motion of a car along a straight line during the time interval $[0, T]$, and let $v(t)$ denote the velocity of the car during the time interval $[0, T]$. If we have that

$$\frac{v(t) - v(t_0)}{t - t_0} = a_0 \qquad \textit{for all } t, t_0 \in [0, T] \textit{ and with some constant } a_0,$$

then we say the the car has the ***uniform acceleration*** *a_0. From (3.6), we find*

$$v(t) = v(t_0) + a_0\,(t - t_0) \qquad \textit{for all } t \in [0, T], \tag{3.6}$$

that is, the ***velocity is an affine linear function****. The distance traveled over the time interval* $[t_0, t]$ *(where we from now assume that* $t > t_0$*) is given by*

$$s(t) - s(t_0) = \textit{average velocity } \times (t - t_0) = \frac{v(t) + v(t_0)}{2}\,(t - t_0).$$

That the formula for the average velocity over the time interval $[t_0, t]$ *is simply the mean value* $\big[v(t)+v(t_0)\big]/2$ *of* $v(t_0)$ *and* $v(t)$ *is due to the fact that the velocity (3.6) is an affine linear funciton. Substituting* $v(t)$ *by (3.6) yields*

$$s(t) - s(t_0) = \frac{1}{2}\left[\big(v(t_0) + a_0\,(t - t_0)\big) + v(t_0)\right](t - t_0) = v(t_0)\,(t - t_0) + \frac{1}{2}\,a_0\,(t - t_0)^2,$$

and thus we see that

$$s(t) = s(t_0) + v(t_0)\,(t - t_0) + \frac{1}{2}\,a_0\,(t - t_0)^2.$$

Thus the ***motion along a straight line with uniform acceleration is a quadratic function****.*

In the last section of this chapter we will encounter another application: radioactive decay of a radioactive material over time.

3.3 Standard Derivatives

You need to know the derivatives listed in Table 3.1 below of your head. They will **not be provided in the final exam!**

We will not derive all the derivatives in Table 3.1 below, but we later explicitly compute the derivatives of $\tan(x)$, $\cosh(x)$, and $\sinh(x)$, based on the knowledge of the derivatives of $\cos(x)$, $\sin(x)$, and e^x.

It is important to note that the natural exponential function $\exp(x) = e^x$ is the **only** function with the property that $f'(x) = f(x)$ for all x.

The functions whose derivatives are listed in Table 3.1, will often be the building blocks of more complicated functions. The elementary properties of the derivative, the product rule, the quotient rule, and the chain rule, will allow us to differentiate sums, differences, products, quotients, and compositions two or more functions from those listed in Table 3.1.

Function $f(x)$	Derivative $f'(x)$
$c = \text{constant}$	0
x^p	$p\,x^{p-1}$
e^x	e^x
$\ln(x)$	$\dfrac{1}{x}$
$\sin(x)$	$\cos(x)$
$\cos(x)$	$-\sin(x)$
$\tan(x)$	$\dfrac{1}{\big(\cos(x)\big)^2} = \big(\sec(x)\big)^2$
$\cot(x)$	$-\dfrac{1}{\big(\sin(x)\big)^2} = -\big(\mathrm{cosec}(x)\big)^2$
$\sinh(x)$	$\cosh(x)$
$\cosh(x)$	$\sinh(x)$
$\tanh(x)$	$\dfrac{1}{\big(\cosh(x)\big)^2} = \big(\mathrm{sech}(x)\big)^2$
$\coth(x)$	$-\dfrac{1}{\big(\sinh(x)\big)^2} = -\big(\mathrm{cosech}(x)\big)^2$

Table 3.1: Important derivatives.

3.4 Elementary Rules for Derivatives

The most simple form of a **composite function** is the sum (or difference) of two functions. If we know the derivatives of the two individual functions then we can easily compute the derivative of the sum/difference. This is stated in the next lemma.

Lemma 3.13 (derivative of sum or difference of two functions)
The ***derivative of the sum*** $(g+f)(x) = f(x) + g(x)$ *of two smooth enough functions* $f(x)$ *and* $g(x)$ *is given by the sum of the derivatives, that is,*

$$(f+g)'(x) = \big(f(x) + g(x)\big)' = f'(x) + g'(x). \tag{3.7}$$

For the ***derivative of the difference*** $(g-f)(x) = f(x) - g(x)$ *of two smooth enough functions* $f(x)$ *and* $g(x)$ *we have analogously*

$$(f-g)'(x) = \big(f(x) - g(x)\big)' = f'(x) - g'(x). \tag{3.8}$$

We observe here that the rules (3.7) and (3.8) extend obviously to the sum/difference of more than two functions. For example,

$$\big(f(x) + g(x) - h(x)\big)' = f'(x) + g'(x) - h'(x).$$

Example 3.14 (derivatives of sums and differences)

(a) $\big(\sin(x) + x^2 + x\big)' = \big(\sin(x)\big)' + (x^2)' + (x)' = \cos(x) + 2\,x + 1;$

(b) $\big(\tan(x) + e^x\big)' = \big(\tan(x)\big)' + (e^x)' = \big(\cos(x)\big)^{-2} + e^x = \big(\sec(x)\big)^2 + e^x;$

(c) $\big(\cos(x) + 3\big)' = \big(\cos(x)\big)' + (3)' = -\sin(x) + 0 = -\sin(x).$ □

3.5 Product Rule and Quotient Rule

The product of two functions can be easily differentiated if we know the derivatives of the two individual functions.

Lemma 3.15 (product rule)
The ***derivative of the product*** $f(x)\,g(x)$ *of two smooth enough functions* $f(x)$ *and* $g(x)$ *is given by*

$$\frac{d}{dx}\big(f(x)\,g(x)\big) = \big(f(x)\,g(x)\big)' = f'(x)\,g(x) + f(x)\,g'(x).$$

We give some examples.

Example 3.16 (derivative of the product of functions)

(a) $\big(\sin(x)\,\cos(x)\big)' = \big(\sin(x)\big)'\cos(x) + \sin(x)\,\big(\cos(x)\big)' = \big(\cos(x)\big)^2 - \big(\sin(x)\big)^2$, where we have used that $\big(\sin(x)\big)' = \cos(x)$ and $\big(\cos(x)\big)' = -\sin(x)$.

(b) $(x\,e^x)' = (x)'\,e^x + x\,(e^x)' = 1\,e^x + x\,e^x = (1+x)\,e^x$, where we have used that $(e^x)' = e^x$ and $(x)' = 1$.

(c) $(c\,f(x))' = (c)'\,f(x) + c\,f'(x) = 0\,f(x) + c\,f'(x) = c\,f'(x)$, where $c \in \mathbb{R}$ is a constant. □

From the Example 3.16 (c), we see the following:

Corollary 3.17 (product of function and constant)
The derivative of the product $c\,f(x)$ of a constant $c \in \mathbb{R}$ and a function $f(x)$ is given by

$$\big(c\,f(x)\big)' = c\,f'(x).$$

We also have a rule for differentiating the quotient of two functions.

Lemma 3.18 (quotient rule)
*The **derivative of the quotient** $f(x)/g(x)$ of two smooth enough functions $f(x)$ and $g(x)$, where $g(x) \neq 0$, is given by*

$$\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)\,g(x) - g'(x)\,f(x)}{\big(g(x)\big)^2}.$$

As an example we will compute the derivative of the tangent function $\tan(x)$ making only use of $\big(\sin(x)\big)' = \cos(x)$ and $\big(\cos(x)\big)' = -\sin(x)$.

Example 3.19 (derivative of the tangent function)
To differentiate $h(x) = \tan(x)$, we remember its definition as a quotient of the sine and cosine function

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

and then use the quotient rule with $f(x) = \sin(x)$, $f'(x) = \cos(x)$, and $g(x) = \cos(x)$, $g'(x) = -\sin(x)$. Then

$$\big(\tan(x)\big)' = \frac{d}{dx}\left(\frac{\sin(x)}{\cos(x)}\right) = \frac{f'(x)\,g(x) - g'(x)\,f(x)}{\big(g(x)\big)^2}$$

$$
\begin{aligned}
&= \frac{\cos(x)\,\cos(x) - \big(-\sin(x)\big)\,\sin(x)}{\big(\cos(x)\big)^2} \\
&= \frac{\big(\cos(x)\big)^2 + \big(\sin(x)\big)^2}{\big(\cos(x)\big)^2} \\
&= \frac{1}{\big(\cos(x)\big)^2} \;=\; \sec^2(x),
\end{aligned}
$$

where we have used $\big(\cos(x)\big)^2 + \big(\sin(x)\big)^2 = 1$ in the second last step. □

Example 3.20 (differentiation with the quotient rule)
The derivative of the function

$$h(x) = \frac{e^x}{e^x + 1}$$

can be computed with the quotient rule. We have $f(x) = e^x$, $f'(x) = e^x$, and $g(x) = e^x + 1$, $g'(x) = e^x$, and thus we find

$$h'(x) = \left(\frac{e^x}{e^x+1}\right)' = \frac{f'(x)\,g(x) - g'(x)\,f(x)}{\big(g(x)\big)^2} = \frac{e^x\,(e^x+1) - e^x\,e^x}{(e^x+1)^2} = \frac{e^x}{(e^x+1)^2}. \qquad \square$$

We discus another example.

Example 3.21 (derivative of $1/g(x)$)
Assume $g(x) \neq 0$. Then find the derivative of the function

$$h(x) = \frac{1}{g(x)}.$$

Solution: This can be done with the quotient rule with $f(x) = 1$. Since $f'(x) = 0$, we find from the quotient rule that

$$h'(x) = \left(\frac{1}{g(x)}\right)' = \frac{0 \times g(x) - g'(x) \times 1}{\big(g(x)\big)^2}. = \frac{-g'(x)}{\big(g(x)\big)^2}. \qquad \square$$

From the last example we obtain the following useful corollary.

Corollary 3.22 (special case of quotient rule)
The derivative of the quotient $1/g(x)$, where $g(x) \neq 0$, is given by

$$\frac{d}{dx}\left(\frac{1}{g(x)}\right) = \left(\frac{1}{g(x)}\right)' = \frac{-g'(x)}{\big(g(x)\big)^2}.$$

3.6 Chain Rule

With the rules learnt so far we can differentiate sums, products, and quotients of the functions in Table 3.1. But how do we differentiate functions such as $\sin(x^2)$ or $(\sin(x)+1)^{3/2}$ or $e^{\cos(x)}$ that cannot be written as sums, products, or quotients of functions whose derivatives we know?

To do this we need another rule, the so-called **chain rule**. We observe that all the mentioned functions are **compositions of two functions** or, in other words, **'functions of other functions'**. We say that a function $h : A \to C$ is a **composition** of two functions $f : B \to C$ and $g : A \to B$, if we can write h as

$$h(x) = f\big(g(x)\big),$$

that is, the variable y of $f(y)$ is replaced by $y = g(x)$. Indeed, we have that

$$\sin(x^2) = f\big(g(x)\big), \qquad \text{with} \qquad f(y) = \sin(y), \quad g(x) = x^2; \tag{3.9}$$

$$\big(\sin(x)+1\big)^{3/2} = f\big(g(x)\big), \qquad \text{with} \qquad f(y) = y^{3/2}, \quad g(x) = \sin(x)+1; \tag{3.10}$$

$$e^{\cos x} = f\big(g(x)\big), \qquad \text{with} \qquad f(y) = e^y, \quad g(x) = \cos(x). \tag{3.11}$$

Note that $f\big(g(x)\big)$ is not the same as $g\big(f(x)\big)$! For example, for $g(x) = x^2$ and $f(x) = \sin(x)$

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

The chain rule tells us how to differentiate the composition $f\big(g(x)\big)$, if we know the derivatives of $f(y)$ and $g(x)$.

Lemma 3.23 (chain rule)
Let $A, B, C \subset \mathbb{R}$ and let $f : B \to C$ and $g : A \to B$ be two sufficiently smooth functions. Then

$$\frac{d}{dx}\big[f\big(g(x)\big)\big] = \big[f\big(g(x)\big)\big]' = f'\big(g(x)\big)\, g'(x).$$

Example 3.24 (differentiation with chain rule)
We want to use the chain rule to differentiate the composite functions h given by (3.9), (3.10), and (3.11). The first thing to do is to work out the composition as $h(x) = f\big(g(x)\big)$. For the functions (3.9), (3.10), and (3.11) we have already done this.

(a) From (3.9), the function $h(x) = \sin(x^2)$, has the composition $h(x) = f\big(g(x)\big)$ with $f(y) = \sin(y)$ and $g(x) = x^2$. We have $f'(y) = \cos(y)$ and $g'(x) = 2\,x$, and the chain rule gives

$$\frac{d}{dx}\sin(x^2) = f'\big(g(x)\big)\,g'(x) = f'(x^2)\,2\,x = \cos(x^2)\,2\,x = 2\,x\,\cos(x^2).$$

(b) From (3.10), the function $h(x) = \big(\sin(x)+1\big)^{3/2}$, has the composition $h(x) = f\big(g(x)\big)$ with $f(y) = y^{3/2}$ and $g(x) = \sin(x)+1$. We have $f'(y) = (3/2)\,y^{1/2}$ and $g'(x) = \cos(x)$, and from the chain rule

$$\frac{d}{dx}\big(\sin(x)+1\big)^{3/2} = f'\big(g(x)\big)\,g'(x) = f'\big(\sin(x)+1\big)\,\cos(x) = \frac{3}{2}\,\cos(x)\,\big(\sin(x)+1\big)^{1/2}.$$

(c) From (3.10), the function $h(x) = e^{\cos(x)}$, has the composition $h(x) = f\big(g(x)\big)$ with $f(y) = e^y$ and $g(x) = \cos(x)$. We have $f'(y) = e^y$ and $g'(x) = -\sin(x)$, and from the chain rule

$$\begin{aligned}\frac{d}{dx}\,e^{\cos(x)} &= f'\big(g(x)\big)\,g'(x)\\ &= f'\big(\cos(x)\big)\,\big(-\sin(x)\big) = e^{\cos(x)}\,\big(-\sin(x)\big) = -\sin(x)\,e^{\cos(x)}.\end{aligned}$$ □

We will look at some more examples. Unlike before, when we have first substituted the expressions for $g(x)$ and $g'(x)$ into $f'\big(g(x)\big)\,g'(x)$, and subsequently substituted the expression for $f'(y)$ one step later, we will from now on do both substitutions at once.

Example 3.25 (differentiation with the chain rule)
Differentiate the functions

$$\text{(a)}\quad h(x) = \ln\big(\cos(x)\big), \qquad\qquad \text{(b)}\quad h(x) = e^{-x^2}.$$

(a) The function $h(x) = \ln\big(\cos(x)\big)$ is the composition $h(x) = f(g(x))$ with $f(y) = \ln(y)$ and $g(x) = \cos(x)$. We have $f'(y) = 1/y$ and $g'(x) = -\sin(x)$, and the chain rule yields that

$$\frac{d}{dx}\ln\big(\cos(x)\big) = f'\big(g(x)\big)\,g'(x) = \frac{1}{\cos(x)}\,\big(-\sin(x)\big) = -\,\frac{\sin(x)}{\cos(x)} = -\tan(x).$$

(b) The function $h(x) = e^{-x^2}$ is the composition $h(x) = f(g(x))$ with $f(y) = e^y$ and $g(x) = -x^2$. We have $f'(y) = e^y$ and $g'(x) = -2\,x$, giving that

$$\frac{d(e^{-x^2})}{dx} = f'\big(g(x)\big)\,g'(x) = e^{-x^2}\,(-2\,x) = -2\,x\,e^{-x^2}.$$ □

With the help of the chain rule we can find a **rule for finding the derivative of the inverse function** f^{-1} of a one-to-one function f if we know the derivative of f. We remark here that we will usually **denote the variable of the inverse function** f^{-1} **by** y and the **variable of the function** f **by** x. This is motivated by the fact that we set $y = f(x)$ and solve for x, that is, find $x = f^{-1}(y)$, if we want to find the inverse function. It also helps to avoid confusion and mistakes if we give the variable of the function f and the variable of its inverse function f^{-1} different names.

Corollary 3.26 (derivative of the inverse function)
Let $A, B \subset \mathbb{R}$. Let $f : A \to B$ be a smooth enough one-to-one function, and let $f^{-1} : f(A) \to A$ be the inverse function of f. Then we have that

$$\frac{df^{-1}(y)}{dy} = (f^{-1})'(y) = \frac{1}{f'\big(f^{-1}(y)\big)}.$$

Proof: From the properties of the inverse function, we have that

$$f\big(f^{-1}(y)\big) = y. \tag{3.12}$$

We can apply the chain rule to differentiate $h(y) = f\big(f^{-1}(y)\big)$, and we obtain from the chain rule

$$h'(y) = \big[f\big(f^{-1}(y)\big)\big]' = f'\big(f^{-1}(y)\big)\,(f^{-1})'(y). \tag{3.13}$$

However, we also have from (3.12) that $h(y) = y$ and thus $h'(y) = 1$. Substituting $h'(y) = 1$ into (3.13) yields

$$1 = f'\big(f^{-1}(y)\big)\,(f^{-1})'(y) \qquad \Rightarrow \qquad (f^{-1})'(y) = \frac{1}{f'\big(f^{-1}(y)\big)},$$

where we have divided by $f'\big(f^{-1}(y)\big)$ to get the second formula. □

Example 3.27 (derivative of the natural logarithm)
We want to derive that $\big(\ln(y)\big)' = 1/y$ with the help of Corollary 3.26. By definition, the natural logarithm $\ln : (0, \infty) \to \mathbb{R}$ is the inverse function of $\exp(x) = e^x$, and thus $e^{\ln(y)} = y$ and $\ln(e^x) = x$. Therefore we can apply Corollary 3.26 with $f(x) = e^x$ and and $f^{-1}(y) = \ln(y)$. Since $f'(x) = (e^x)' = e^x$, we have

$$\big(\ln(y)\big)' = (f^{-1})'(y) = \frac{1}{f'\big(f^{-1}(y)\big)} = \frac{1}{f'\big(\ln(y)\big)} = \frac{1}{e^{\ln(y)}} = \frac{1}{y}. \qquad \square$$

Remark 3.28 ($(f^{-1})^{-1} = f$)
We note here that the equations

$$f\big(f^{-1}(y)\big) = y \quad \textit{for all } y \in f(A) \qquad \textit{and} \qquad f^{-1}\big(f(x)\big) = x \quad \textit{for all } x \in A.$$

for a one-to-one function $f : A \to B$ and its inverse function $f^{-1} : f(A) \to A$, also imply that the ***inverse function*** $(f^{-1})^{-1}$ ***of*** f^{-1} ***is given by*** f*. Thus we could also have worked out that $(e^x)' = e^x$ based on the knowledge that $\big(\ln(y)\big)' = 1/y$.*

Example 3.29 (derivative of $\arcsin(x)$)
In the previous chapter we have introduced the function $\arcsin : [-1, 1] \to [\pi/2, \pi/2]$ as the inverse of the function $\sin : [-\pi/2, \pi/2] \to [-1, 1]$. Now we want to compute the derivative of the $\arcsin(y)$ with the help of Corollary 3.26.

Since $(\sin(x))' = \cos(x)$, we have

$$\big(\arcsin(y)\big)' = (f^{-1})'(y) = \frac{1}{f'\big(f^{-1}(y)\big)} = \frac{1}{\cos\big(\arcsin(y)\big)}. \tag{3.14}$$

To simplify the denominator $\cos\big(\arcsin(y)\big)$ further, we remember that

$$\big[\sin(x)\big]^2 + \big[\cos(x)\big]^2 = 1 \qquad \Rightarrow \qquad \cos(x) = \sqrt{1 - \big[\sin(x)\big]^2},$$

where we get only the positive root $+\sqrt{1 - \big[\sin(x)\big]^2}$ of $\big[\cos(x)\big]^2 = 1 - \big[\sin(x)\big]^2$ because $\cos(x) \geq 0$ for all $x \in [-\pi/2, \pi/2]$. Substituting $x = \arcsin(y)$ into the last formula yields

$$\cos\big(\arcsin(y)\big) = \sqrt{1 - \big[\sin\big(\arcsin(y)\big)\big]^2} = \sqrt{1 - y^2}, \tag{3.15}$$

where we used $\sin\big(\arcsin(y)\big) = y$ since $\arcsin(y)$ is the inverse function of $\sin(x)$. Substituting (3.15) into (3.14), we obtain

$$\frac{d}{dy}\arcsin(y) = \big(\arcsin(y)\big)' = \frac{1}{\cos\big(\arcsin(y)\big)} = \frac{1}{\sqrt{1 - y^2}}. \qquad \square$$

Application 3.30 (radioactive decay)
Radioactivity is caused by the disintegration of the nuclei of atoms of the radioactive material. The disintegration obeys the ***law of chance****, that is, we do not know which individual atom disintegrates next. However, we know that the* ***number of disintegrating atoms per second is proportional to the number of atoms****.*

Thus, if $N(t)$ denotes the number of atoms at the time t, then $dN(t)/dt$ is the ***disintegration rate*** *of the atoms, and we have*

$$\frac{dN(t)}{dt} = -\lambda\, N(t), \tag{3.16}$$

where $\lambda > 0$ is the constant of proportionality. We have a minus sign on the right-hand side, because the number of atoms declines over time. The equation (3.16) is a ***differential equation*** *and gives a relation between $N(t)$ and its first derivative $dN(t)/dt$. What we clearly would like to know is the* ***unknown function*** *$N(t)$. If we can find the unknown function $N(t)$ then we have* ***solved the differential equation****.*

For the differential equation (3.16), we can 'guess' the solution by inspection. We know that the only function that satisfies $f'(x) = f(x)$ is the natural exponential function $\exp(x) = e^x$. Compared to $f'(x) = f(x)$, (3.16) has an additional constant factor on the right-hand side. This leads us to consider $g(x) = \exp(\alpha x) = e^{\alpha x}$ for which we have, from the chain rule,

$$\frac{dg(x)}{dx} = \frac{d(e^{\alpha x})}{dx} = \alpha\, e^{\alpha x} = \alpha\, g(x).$$

Thus we predict that ***(3.16) has the solution***

$$N(t) = N_0\, e^{-\lambda t}, \tag{3.17}$$

where N_0 is a constant. We differentiate to test our guess and find, from the chain rule,

$$\frac{dN(t)}{dt} = \frac{d}{dt}\big[N_0\, e^{-\lambda t}\big] = -\lambda\, N_0\, e^{-\lambda t} = -\lambda\, N(t).$$

Thus (3.17) is indeed a solution to (3.16), and it is possible to show that all solutions of (3.16) are of the form (3.17).

Now we want to interpret and discuss the solution (3.17). We observe that for $t = 0$ we have $N(0) = N_0$. Thus the constant N_0 is the ***initial number of atoms of the radioactive material****. As you may know radioactive substances have a* ***half-life****, which is defined as the time $T_{1/2}$ after which half of the atoms have disintegrated. In formulas,*

$$N(T_{1/2}) = N_0\, e^{-\lambda T_{1/2}} = \frac{N_0}{2} \quad \Rightarrow \quad e^{-\lambda T_{1/2}} = 2^{-1} \quad \Rightarrow \quad T_{1/2} = \frac{\ln\left(2^{-1}\right)}{-\lambda} = \frac{\ln(2)}{\lambda}.$$

Thus the ***half-life is given by*** *$T_{1/2} = \ln(2)/\lambda$, and it can be easily computed if the constant of proportionality λ is known.*

Finally we will briefly discuss differentiating exponential functions $h(x) = a^x$ and logarithmic functions $h(x) = \log_a(x)$ with a base $a > 0$, with $a \neq 1$ and $a \neq e$.

Example 3.31 (derivatives of exponential functions)
Compute the first, second, and third (order) derivative of the **exponential function** $h : \mathbb{R} \to \mathbb{R}$, $h(x) = a^x$, where $a > 0$ and $a \neq 1$.

Solution: We know from Remark 2.34 in Section 2.3 that

$$h(x) = a^x = \big(e^{\ln(a)}\big)^x = e^{\ln(a)\,x}. \tag{3.18}$$

For $h(x) = a^x = f\big(g(x)\big)$ with $f(y) = e^y$, $f'(y) = e^y$, and $g(x) = \ln(a)\,x$, $g'(x) = \ln(a)$, the chain rule yields

$$h'(x) = f'\big(g(x)\big)\,g'(x) = e^{\ln(a)\,x}\,\ln(a) = \ln(a)\,a^x, \tag{3.19}$$

where we have used (3.18) in the last step. We observe from (3.19) that

$$h'(x) = \frac{d(a^x)}{dx} = \ln(a)\,a^x = \ln(a)\,h(x). \tag{3.20}$$

To find the second derivative of $h(x) = a^x$, we use (3.20) and find

$$\begin{aligned} h''(x) &= \frac{dh'(x)}{dx} = \frac{d}{dx}\big[\ln(a)\,h(x)\big] = \ln(a)\,h'(x) \\ &= \ln(a)\,\big[\ln(a)\,h(x)\big] = \big[\ln(a)\big]^2\,h(x) = \big[\ln(a)\big]^2\,a^x, \end{aligned} \tag{3.21}$$

where we have used (3.20) to get from the first line to the second line. To find the third derivative we proceed analogously and use (3.21) and (3.20).

$$\begin{aligned} h^{(3)}(x) &= \frac{dh''(x)}{dx} = \frac{d}{dx}\big(\big[\ln(a)\big]^2\,h(x)\big) = \big[\ln(a)\big]^2\,h'(x) \\ &= \big[\ln(a)\big]^2\,\big[\ln(a)\,h(x)\big] = \big[\ln(a)\big]^3\,h(x) = \big[\ln(a)\big]^3\,a^x. \end{aligned} \tag{3.22}$$

Based our knowledge of the first three derivatives (see (3.20), (3.21), and (3.22)), we suspect that the mth derivative of $h(x) = a^x$ is given by

$$h^{(m)}(x) = \frac{d^m(a^x)}{dx^m} = \big[\ln(a)\big]^m\,h(x) = \big[\ln(a)\big]^m\,a^x = \big[\ln(a)\big]^m\,e^{\ln(a)\,x},$$

which is indeed true. □

Example 3.32 (derivatives of logarithmic functions)
Compute the derivative of the logarithmic function $h : (0, \infty) \to \mathbb{R}$, $h(x) = \log_a(x)$, where $a > 0$ and $a \neq 1$.

Solution: We can compute the derivative in two different ways. Firstly we have from Remark 2.35 in Section 2.4 that

$$h(x) = \log_a(x) = \log_a(e)\,\ln(x) = \frac{1}{\ln(a)}\,\ln(x), \tag{3.23}$$

where $\log_a(e) = 1/\ln(a)$ follows from $1 = \log_a(a) = \log_a(e)\,\ln(a)$ with division by $\ln(a)$. Differentiating the last expression in (3.23) yields

$$h'(x) = \frac{d}{dx}\log_a(x) = \frac{d}{dx}\left(\frac{1}{\ln(a)}\,\ln(x)\right) = \frac{1}{\ln(a)}\,\frac{1}{x} = \frac{1}{\ln(a)\,x}.$$

Alternatively we can use that fact that $f^{-1}(y) = \log_a(y)$ is the inverse function of $f(x) = a^x$. Thus we may apply Corollary 3.26, making use of the fact that

$$f'(x) = \ln(a)\,a^x$$

from the previous example. Corollary 3.26 yields

$$\frac{d}{dy}\log_a(y) = (f^{-1})'(y) = \frac{1}{f'\big(f^{-1}(y)\big)} = \frac{1}{\ln(a)\,a^{\log_a(y)}} = \frac{1}{\ln(a)\,y},$$

where in the last step $a^{\log_a(y)} = y$ holds because $f^{-1}(y) = \log_a(y)$ is the inverse function of $f(x) = a^x$. $\square$

Chapter 4

Curves and Functions

In this chapter we will discuss in more detail how to **analyze and subsequently sketch functions**. In Chapter 1 we have learned to sketch functions by evaluating them at some points and using knowledge such as where the function is zero, has positive or negative values, whether it has asymptotes, and also where it is (strictly) monotonically increasing or decreasing. In this chapter, we will discuss **more properties of functions**, in particular, **local and global minima and maxima**, **points of inflection**, and the **curvature** of the function. The discussion of these topics will heavily rely on analyzing the **first and second (order) derivative** of the function. In some cases we need even **higher order derivatives**.

We discuss several examples in this chapter to illustrate the new topics. It is **essential that you solve analogous problems from the exercise sheets on your own**, because familiarity with applying the new concepts covered in this chapter can only be achieved through practice.

4.1 Roots of Functions

The first topic is **roots** of a function, that is, the points where the function has the value zero. We have already used roots of functions in Chapter 1 to determine where the function intersects the x-axis.

Definition 4.1 (roots of a function)
*Let $A, B \subset \mathbb{R}$. The **roots** of a function $f : A \to B$ are the values $x \in A$ at which $f(x) = 0$. We **find the roots of** f **by solving** $f(x) = 0$ **for** x.*

Example 4.2 (roots of a function)

(a) The affine linear function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = m\,x + c$, with $m \neq 0$, intersects the x-axis at the point(s) satisfying

$$f(x) = m\,x + c = 0 \qquad \Rightarrow \qquad m\,x = -c \qquad \Rightarrow \qquad x = -\frac{c}{m}.$$

Thus the affine linear function has the root $x = -c/m$.

(b) In Section 1.4, we have seen that the quadratic function $f(x) = a\,x^2 + b\,x + c$, with $a \neq 0$, has only root(s) if $b^2 - 4\,a\,c \geq 0$, and if this condition is satisfied then the root(s) are

$$x = \frac{-b \pm \sqrt{b^2 - 4\,a\,c}}{2a}.$$

(c) The function $f(x) = \sin(x)$ has the roots

$$\ldots, -3\pi, -2\pi, -\pi, 0, \pi, 2\pi, 3\pi, \ldots.$$

In shorter notation, the roots are $x = k\,\pi$, with $k \in \mathbb{Z}$.

(d) The natural exponential function $\exp(x) = e^x$ has no roots because we have

$$\exp(x) = e^x > 0 \qquad \text{for all} \quad x \in \mathbb{R}.$$

(e) The natural logarithm $\ln(x)$ has only one root at $x = 1$, since $\ln(1) = 0$ and since $\ln(x) < 0$ for $x \in (0, 1)$ and $\ln(x) > 0$ for $x > 1$. □

Remark 4.3 (roots of polynomials of degree n)
A polynomial $p : \mathbb{R} \to \mathbb{R}$ of degree n,

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

where $a_n \neq 0$, ***has at most n real roots****.*

4.2 Turning Points/Extrema: Local Maxima and Minima

In this section we define **local and global maxima and minima**, the so-called **turning points** or **extrema** of a function, and we will learn a **criterion for finding them with the help of the first and second derivative of the function**. Discussing local maxima and minima will involve some additional new terminology, namely, **stationary points** of a function.

In this chapter we need also the following notation for sets:

Notation 4.4 (intersection of two sets)
Let A and B be two sets. The set $A \cap B$ is called the ***intersection of*** *A* ***and*** *B, and it is defined to be the* ***set of all elements that are in both sets*** *A* ***and*** *B. In formulas,*

$$A \cap B = \{a \in A \,:\, a \text{ is also in } B\} = \{b \in B \,:\, b \text{ is also in } A\}\,,$$

and we see that $A \cap B = B \cap A$.

Now we define the terms local minimum and local maximum of a function.

Definition 4.5 (local maximum and minimum)
Let $A, B \subset \mathbb{R}$, and let $f : A \to B$ be a function.

(i) We say that f has a ***local maximum*** *at a point $x_0 \in A$ if*

$$f(x) \leq f(x_0) \qquad \text{for all } x \in (x_0 - \varepsilon, x_0 + \varepsilon) \cap A, \tag{4.1}$$

with some suitable $\varepsilon > 0$.

(ii) We say that f has a ***local minimum*** *at a point $x_0 \in A$ if*

$$f(x) \geq f(x_0) \qquad \text{for all } x \in (x_0 - \varepsilon, x_0 + \varepsilon) \cap A, \tag{4.2}$$

with some suitable $\varepsilon > 0$.

(iii) We say that f has a ***turning point*** *or* ***extremum*** *at $x_0 \in A$, if f has at x_0 either a local maximum or a local minimum.*

In words, if a function $f : A \to B$ has a **local maximum at a point** $x = x_0$, then there exists a neighborhood of x_0 such that $f(x_0)$ is the largest value among all $f(x)$ for x in this neighborhood, that is, $f(x) \leq f(x_0)$ for all x from the neighborhood.

Definition 4.6 (global maximum and minimum)
Let $A, B \subset \mathbb{R}$, and let $f : A \to B$ be a function.

(i) We say that f has a ***global maximum*** *at a point $x_0 \in A$ if*

$$f(x) \leq f(x_0) \qquad \text{for all } x \in A. \tag{4.3}$$

(ii) We say that f has a ***global minimum*** *at a point $x_0 \in A$ if*

$$f(x) \geq f(x_0) \qquad \text{for all } x \in A. \tag{4.4}$$

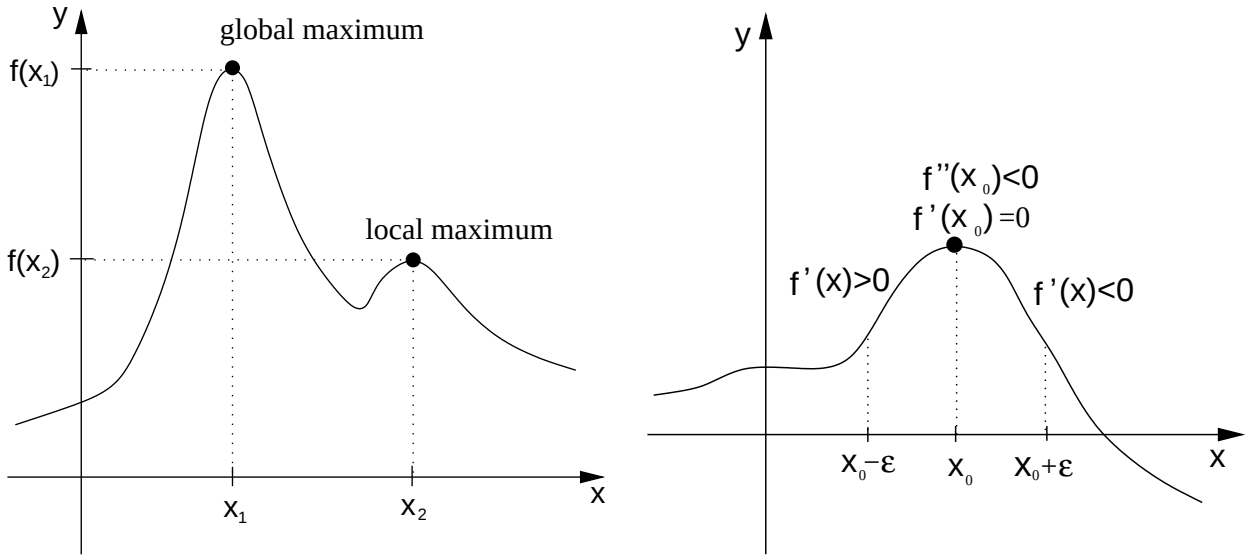


Figure 4.1: On the left an example of a function with a global maximum and a local maximum. On the right the behavior of f close to a (global/local) maximum.

A global maximum is also a local maximum, but the reverse is **not true**. Likewise, **a global minimum is also a local minimum**.

Definitions 4.5 and 4.6 are illustrated in Figure 4.1 for the case of the local and global maximum.

Example 4.7 (local and global minimum)
The function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, has a global (and local) minimum at $x = 0$, because

$$f(0) = 0 \leq x^2 = f(x) \qquad \text{for all} \quad x \in \mathbb{R}.$$

We observe that here

$$f(0) = 0 < x^2 = f(x) \qquad \text{for all} \quad x \in \mathbb{R} \quad \text{with} \quad x \neq 0,$$

that is, $x = 0$ is the **only point** at which the minimal value $f(0) = 0$ is attained. □

Example 4.8 (local and global maximum)
The function $\sin : \mathbb{R} \to [-1, 1]$ has a global and a local maximum all points

$$x = \frac{\pi}{2} + 2\pi\, k = \frac{(4k+1)\,\pi}{2}, \quad k \in \mathbb{Z}, \quad \text{that is,} \quad x \in \left\{ \ldots, -\frac{7\pi}{2}, -\frac{3\pi}{2}, \frac{\pi}{2}, \frac{5\pi}{2}, \frac{9\pi}{2}, \ldots \right\},$$

and the value of $\sin(x)$ at all these points is 1. For the maximum at $x = \pi/2$, we have

$$\sin(x) < \sin(\pi/2) \qquad \text{for all} \quad x \in [0, \pi] \quad \text{with} \quad x \neq \pi/2,$$

that is, in the interval $[0, \pi]$, the point $x = \pi/2$ is the **only point** at which we have $\sin(x) = 1$. □

We saw the following property in the last examples: **In a suitable neighborhood of a point** x_0 at which f has a maximum (minimum), this point is the **only point in that neighborhood** at which the value of the maximum (minimum) is attained. This property is **not always true** but will hold in almost all examples that we discuss. An example where this property does not hold are constant functions.

Example 4.9 (degenerate case: all points are maxima and minima)
Consider the constant function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = c$, with some constant $c \in \mathbb{R}$. Then for every (fixed) $x_0 \in \mathbb{R}$,

$$f(x_0) = c \leq c = f(x) \quad \text{for all} \quad x \in \mathbb{R} \quad \text{and} \quad f(x_0) = c \geq c = f(x) \quad \text{for all} \quad x \in \mathbb{R}.$$

Thus we see that x_0 is both a local and global minimum and a local and global maximum. Since $x_0 \in \mathbb{R}$ was arbitrary, all points $x \in \mathbb{R}$ are both a local and global maximum and a local and global minimum of the constant function $f(x) = c$. This is obviously a 'degenerate' case! □

With the terminology below we obtain an easy way of determining with the help of the first and second derivative at which points a function has a local maximum or a local minimum.

Definition 4.10 (stationary points)
Consider a smooth enough function $f : A \to B$. If for some $x_0 \in A$ we have $f'(x_0) = 0$, then x_0 is called a ***stationary point*** *of f.*

The next lemma provides the useful information that a turning point/extremum is also a stationary point. Thus we need only investigate for each stationary point whether the function f attains at this point a local maximum or minimum.

Lemma 4.11 (turning point/extremum ⇒ stationary point)
A turning point/extremum of a function f is also a stationary point of f.

Remark 4.12 (how to determine candidates for local maxima/minima)
From the last lemma we see that it is ***enough to check for all stationary points whether they are turning points/extrema****. A point which is* ***not*** *a stationary point* ***cannot*** *be a turning point/extremum.*

The statement of Lemma 4.11 is intuitively clear, because the **sign of the derivative $f'(x)$ determines whether the function is monotonically increasing or**

decreasing. For example, if we have a local maximum at $x = x_0$ then we have the following on a suitable interval $(x_0 - \varepsilon, x_0 + \varepsilon)$: On the left of x_0, the function f is monotonically increasing, that is, $f'(x) \geq 0$ for $x_0 - \varepsilon < x < x_0$. On the right of x_0 the function f is monotonically decreasing, that is, $f'(x) \leq 0$ for $x_0 < x < x_0 + \varepsilon$. Thus we can conclude that $f'(x_0) = 0$. This is illustrated in Figure 4.1.

We go now a step further in this analysis, and we will assume that we have no 'degenerate case', that is, we assume the following: If f has a local maximum (minimum) at x_0, then for some suitable $\varepsilon > 0$,

$$f(x) < f(x_0) \quad \big(f(x) > f(x_0), \text{ respectively} \big) \quad \text{for all } x \in (x_0 - \varepsilon, x_0 + \varepsilon) \text{ with } x \neq x_0.$$

We have seen that the derivative f' changes its sign at a local maximum or local minimum. Assume that f has a local maximum at $x = x_0$. Then f' changes at $x = x_0$ from being positive to being negative. If f' changes from being positive to being negative, then f'' must be **strictly monotonically decreasing in a neighborhood of** x_0 and therefore $f''(x_0) < 0$. If f has a local minimum at x_0, then f' changes at $x = x_0$ from being negative to being positive. If f' changes from being negative to being positive, then f'' is **strictly monotonically increasing in a neighborhood of** x_0, and therefore $f''(x_0) > 0$.

These observations motivate to the following theorem.

Theorem 4.13 (2nd derivative test for turning points/extrema)
Let $A, B \subset \mathbb{R}$, let $f : \mathbb{R} \to \mathbb{R}$ be a sufficiently smooth function, and let $x_0 \in A$ be a ***stationary point*** *of f, that is, $f'(x_0) = 0$. Then the following holds:*

(i) If $f''(x_0) < 0$, then the function f has a ***local maximum*** *at $x = x_0$.*

(ii) If $f''(x_0) > 0$, then the function f has a ***local minimum*** *at $x = x_0$.*

(iii) If $f''(x_0) = 0$, then we ***cannot conclude anything****: f could have a local maximum, or a local minimum, or neither at $x = x_0$.*

We observe that the third statement (iii) is not so satisfactory since we gain no information. Later-on in Section 4.5, we will discuss how to proceed in case (iii).

Now we will discuss some examples.

Example 4.14 (stationary points and turning points of $f(x) = x^2$)
The function $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = x^2$, has the derivative $f'(x) = 2\,x$. Stetting the $f'(x) = 0$ yields

$$f'(x) = 2\,x = 0 \qquad \Rightarrow \qquad x = 0.$$

We compute the second derivative $f''(x) = 2$ and evaluate at the stationary point $x = 0$. We find $f''(0) = 2 > 0$. Thus the function $f(x) = x^2$ has at $x = 0$ a local minimum. The coordinates of the global minimum are $(0, 0)$, and since $f(x) = x^2 \geq 0 = f(0)$, we see that the local minimum is also as global minimum. $\square$

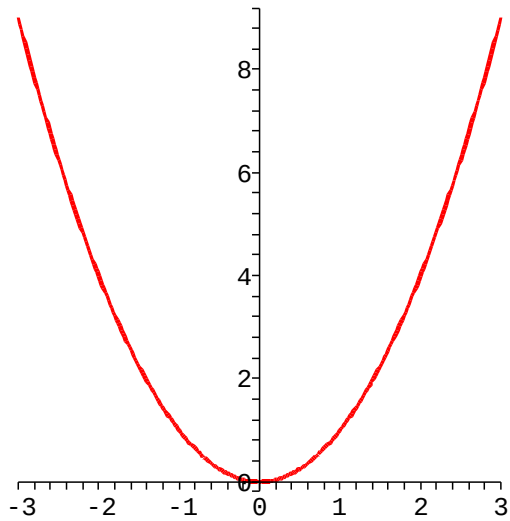


Figure 4.2: Graph of $f(x) = x^2$. The function has a global minimum at $x = 0$.

Example 4.15 (turning points/extrema of the sine function)
Find the turning points/extrema, that is, the local minima and maxima, of the sine function.

Solution: We have $f'(x) = \big(\sin(x)\big)' = \cos(x)$, and the stationary points are those for which $f'(x) = \cos(x) = 0$. Thus the stationary points of $\sin(x)$ are

$$x = \frac{(2k+1)\pi}{2}, \quad k \in \mathbb{Z}, \qquad \text{that is,} \qquad x \in \left\{ \ldots, -\frac{5\pi}{2}, -\frac{3\pi}{2}, -\frac{\pi}{2}, \frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}, \ldots \right\}.$$

The second derivative of f is given by $f''(x) = \big(\cos(x)\big)' = -\sin(x)$, and we find that

$$f''\left(\frac{(2k+1)\pi}{2}\right) = -\sin\left(\frac{(2k+1)\pi}{2}\right) = \begin{cases} -1 & \text{if} \quad k \text{ is even,} \\ 1 & \text{if} \quad k \text{ is odd.} \end{cases}$$

Since we can describe even k by $k = 2\ell$ and odd k by $k = 2\ell + 1$, $\ell \in \mathbb{Z}$, we have

$$f''\left(\frac{(4\ell+1)\pi}{2}\right) = -1, \qquad f''\left(\frac{(4\ell+3)\pi}{2}\right) = 1, \qquad \ell \in \mathbb{Z}.$$

Therefore we have local maxima at all points

$$x = \frac{(4\ell+1)\pi}{2}, \qquad \ell \in \mathbb{Z}, \qquad \text{that is,} \qquad x \in \left\{ \ldots, -\frac{7\pi}{2}, -\frac{3\pi}{2}, \frac{\pi}{2}, \frac{5\pi}{2}, \ldots \right\},$$

and the value of $\sin(x)$ at these points is 1. We have local minima at all points

$$x = \frac{(4\ell+3)\pi}{2}, \qquad \ell \in \mathbb{Z}, \qquad \text{that is,} \qquad x \in \left\{ \ldots, -\frac{5\pi}{2}, -\frac{\pi}{2}, \frac{3\pi}{2}, \frac{7\pi}{2}, \ldots \right\},$$

and the value of $\sin(x)$ are these points is -1. □

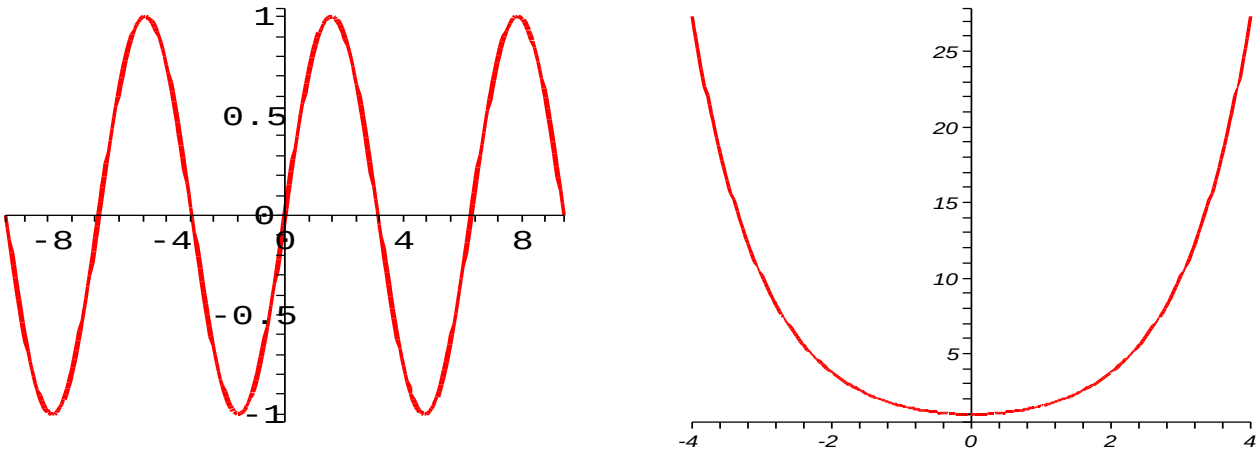


Figure 4.3: The graph of $\sin(x)$ on the left and the graph of $\cosh(x)$ on the right.

Example 4.16 (turning points/extrema of $\cosh(x)$)
Find the turning points/extrema of $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \cosh(x)$.

Solution: We compute the stationary points, that is, the points at which we have $f'(x) = \big(\cosh(x)\big)' = 0$. We find

$$f'(x) = \frac{d}{dx}\cosh(x) = \sinh(x) = \frac{e^x - e^{-x}}{2} = 0 \qquad \Rightarrow \qquad x = 0,$$

and $x = 0$ is the only stationary point. We differentiate again and get

$$f''(x) = \frac{d^2}{dx^2}\cosh(x) = \frac{d}{dx}\sinh(x) = \cosh(x) = \frac{e^x + e^{-x}}{2}$$

Evaluating at the stationary point $x = 0$ yields

$$f''(0) = \cosh(0) = \frac{e^0 + e^{-0}}{2} = \frac{2}{2} = 1 > 0.$$

Thus, from Theorem 4.13, we know that $f(x) = \cosh(x)$ has a local minimum at $x = 0$, and the coordinates of the local minimum are $\big(0, \cosh(0)\big) = (0, 1)$. □

4.3 Curvature: Convex Upward and Convex Downward

For plotting a function it is helpful to know 'which way the graph curves'. We will first explain and define what this means **geometrically**. Then we will give a formal definition of curvature with the help of the second derivative.

Definition 4.17 (geometric definition of convex upward/downward I)
Let $f : A \to B$ be a smooth enough function, and I be a subinterval of A.

*(i) The function f is **convex upward on** I if at every point $x_0 \in I$ the tangent of f at $x = x_0$ lies below the graph of f. If, in addition, close to $(x_0, f(x_0))$ this tangent touches the graph only at $(x_0, f(x_0))$, then f is **strictly convex upward on** I.*

*(ii) The function f is **convex downward on** I if at every point $x_0 \in I$ the tangent of f at $x = x_0$ lies above the graph of f. If, in addition, close to $(x_0, f(x_0))$ this tangent touches the graph only at $(x_0, f(x_0))$, then f is **strictly convex downward on** I.*

Definition 4.17 is illustrated in Figure 4.4 below. It is important to remember that the condition on the tangent needs to be fulfilled for **all** points x_0 in the given interval I.

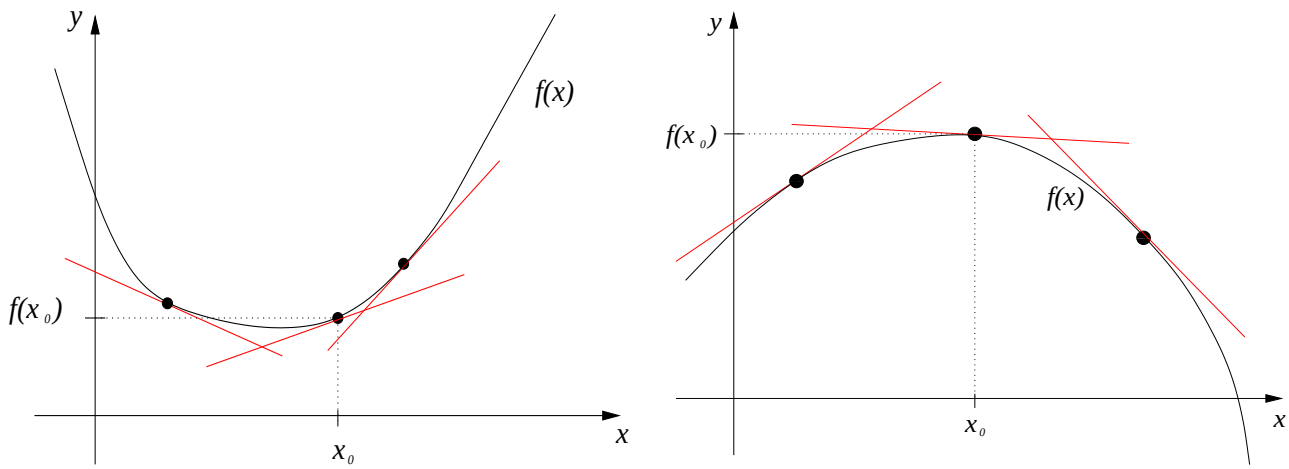


Figure 4.4: Illustration of the geometric definition of (strictly) convex upward and (strictly) convex downward from Definition 4.17.

Considering the function in the left picture of Figure 4.4, we see that if f is strictly

convex upward, then the derivative $f'(x)$ increases as x increases. Thus we expect that $f''(x) > 0$. Likewise, in the right picture, the derivative $f'(x)$ decreases as x increases, and thus we expect that $f''(x) < 0$. This is indeed the case and will give us an easy criterion for checking whether a function is (strictly) convex upward or (strictly) convex downward (see Definition 4.20 below).

Example 4.18 (strictly convex upward/downward function)
The function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, is strictly convex upward on $\mathbb{R}$, and the function $g : \mathbb{R} \to \mathbb{R}$, $g(x) = -x^2$, is a strictly convex downward on $\mathbb{R}$. This can be directly seen from the geometric definition with the tangent given above. □

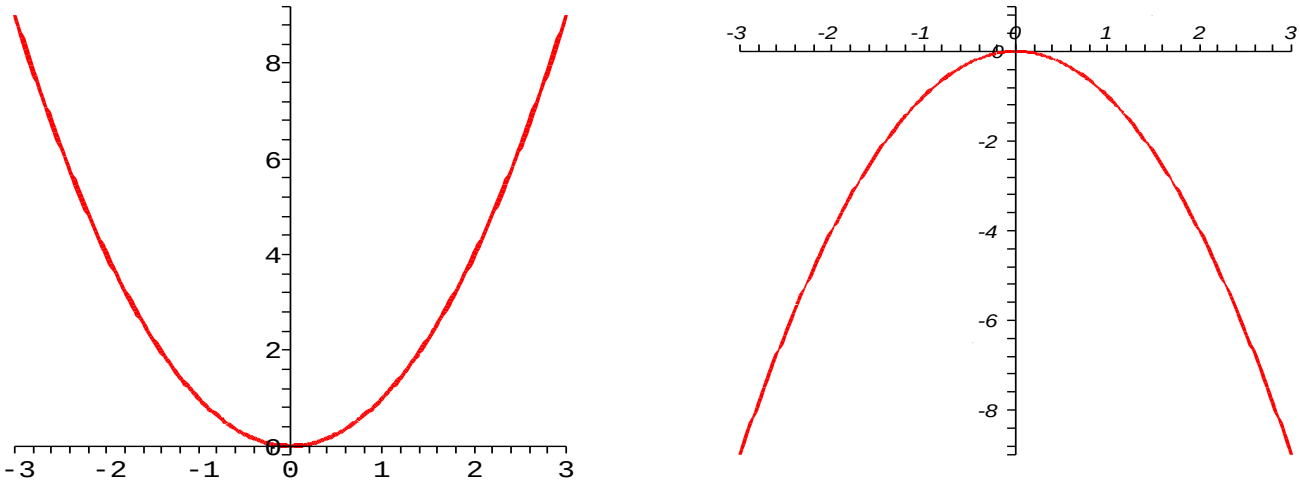


Figure 4.5: The function $f(x) = x^2$ (on the left) is strictly convex upward on $\mathbb{R}$, and the function $g(x) = -x^2$ (on the right) is strictly convex downward on $\mathbb{R}$.

Definition 4.19 (geometric definition of convex upward/downward II)
Let $f : A \to B$ be a smooth enough function, and I be a subinterval of A.

*(i) The function f is **convex upward on** I, if for every two points x_1 and x_2 in I the straight line that connects $(x_1, f(x_1))$ and $(x_2, f(x_2))$ lies above the graph of f. If, in addition, this straight line only touches the graph of f at $(x_1, f(x_1))$ and $(x_2, f(x_2))$, then f is **strictly convex upward on** I.*

*(ii) The function f is **convex downward on** I, if for every two points x_1 and x_2 in I the straight line that connects $(x_1, f(x_1))$ and $(x_2, f(x_2))$ lies below the graph of f. If, in addition, this straight line only touches the graph of f at $(x_1, f(x_1))$ and $(x_2, f(x_2))$, then f is **strictly convex downward on** I.*

Definition 4.19 is illustrated in Figure 4.6 below. It is important to note that the straight line that connects $(x_1, f(x_1))$ with $(x_2, f(x_2))$ needs to lie above the graph (and below the graph, respectively) for **all** points x_1 and x_2 in I.

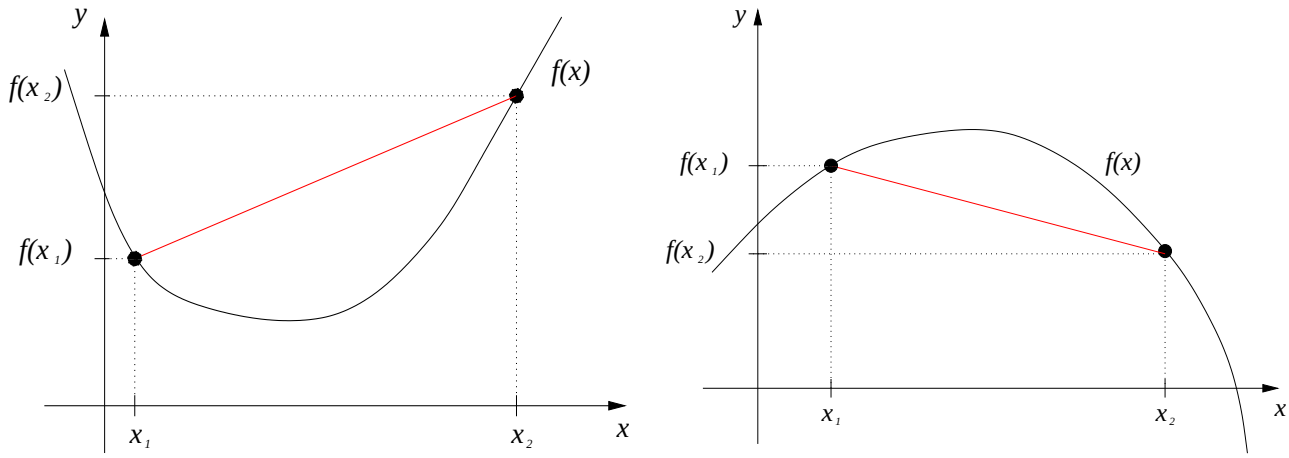


Figure 4.6: Illustration of the geometric definition of (strictly) convex upward and (strictly) convex downward from Definition 4.19.

Now we give a very **useful formal characterization of (strictly) convex upward and (strictly) convex downward** with the help of the **second derivative**.

Definition 4.20 ((strictly) convex upward/downward)
Let $A, B \subset \mathbb{R}$, and let $[a, b] \subset A$ be a closed subinterval of A. Consider a smooth enough function $f : A \to B$. Then the following holds true:

*(i) The function f is **strictly convex upward** on $[a, b]$ if*

$$f''(x) > 0 \qquad \text{for all } x \in (a, b).$$

*(ii) The function f is **convex upward** on $[a, b]$ if*

$$f''(x) \geq 0 \qquad \text{for all } x \in (a, b).$$

*(iii) The function f is **strictly convex downward** on $[a, b]$ if*

$$f''(x) < 0 \qquad \text{for all } x \in (a, b).$$

*(iv) The function f is **convex downward** on $[a, b]$ if*

$$f''(x) \leq 0 \qquad \text{for all } x \in (a, b).$$

With this lemma we can now easily investigate where a function is (strictly) convex upward or (strictly) convex downward. We demonstrate this for the functions $f(x) = x^2$ and $g(x) = -x^2$ from Example 4.18.

Example 4.21 (Example 4.18 continued)
Consider the functions $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, and $g : \mathbb{R} \to \mathbb{R}$, $g(x) = -x^2$, from Example 4.18. Verify with the help of Lemma 4.20 that f is strictly convex upward on $\mathbb{R}$, and that g is strictly convex downward on $\mathbb{R}$.

Solution: We compute the second derivative of each function, and we find

$$f'(x) = 2\,x, \qquad f''(x) = 2, \qquad \text{and} \qquad g'(x) = -2\,x, \qquad g''(x) = -2.$$

Since $f''(x) = 2 > 0$ for all $x \in \mathbb{R}$, the function $f(x) = x^2$ is strictly convex upward on $\mathbb{R}$. Since $g''(x) = -2 < 0$ for all $x \in \mathbb{R}$, the function $g(x) = -x^2$ is strictly convex downward on $\mathbb{R}$. □

Now we consider constant functions and affine linear functions. Since they are straight lines (and not curved) we expect, from Definition 4.19, that they are both convex upward and convex downward on $\mathbb{R}$. (Note that the straight line connecting two points $(x_1, f(x_1))$ and $(x_2, f(x_2))$ on the graph of a constant or affine linear function lies on the graph itself.)

Example 4.22 (constant functions and affine linear functions)
Show with the help of Lemma 4.20 that $f : \mathbb{R} \to \mathbb{R}$, defined $f(x) = m\,x + c$, with constants $m \in \mathbb{R}$ and $c \in \mathbb{R}$, is both convex upward and convex downward on $\mathbb{R}$.

Solution: We calculate the second derivative and find $f'(x) = m$ and $f''(x) = 0$. Thus

$$f''(x) = 0 \geq 0 \qquad \text{and} \qquad f''(x) = 0 \leq 0 \qquad \text{for all } x \in \mathbb{R}.$$

Thus, from Lemma 4.20, the function $f(x) = m\,x + c$ is both convex upward on $\mathbb{R}$ and convex downward on $\mathbb{R}$. □

Finally we will discuss three functions each of which is on some intervals strictly convex upward and on some intervals strictly convex downward.

Example 4.23 (curvature of the polynomial $f(x) = x^3$)
Investigate the curvature of the polynomial $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = x^3$.

Solution: We compute the second derivative and find

$$f'(x) = 3\,x^2 \qquad \text{and} \qquad f''(x) = 6\,x.$$

Thus we have

$$f''(x) = 6\,x < 0 \quad \text{for all } x < 0, \qquad \text{and} \qquad f''(x) = 6\,x > 0 \quad \text{for all } x > 0.$$

Thus we see that $f(x) = x^3$ is strictly convex downward on $(-\infty, 0]$, and that $f(x) = x^3$ is strictly convex upward on $[0, \infty)$. We have plotted $f(x) = x^3$ in the left picture of Figure 4.7. □

Example 4.24 (curvature of $\sinh(x)$)
Determine the curvature of $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \sinh(x)$.

Solution: We compute the second derivative of $\sinh(x)$.

$$f'(x) = \frac{d}{dx}\sinh(x) = \cosh(x), \qquad f''(x) = \frac{d^2}{dx^2}\sinh(x) = \frac{d}{dx}\cosh(x) = \sinh(x).$$

Since

$$f''(x) = \sinh(x) = \frac{e^x - e^{-x}}{2} < 0 \qquad \text{for} \quad x < 0,$$
$$f''(x) = \sinh(x) = \frac{e^x - e^{-x}}{2} > 0 \qquad \text{for} \quad x > 0,$$

we see that $f(x) = \sinh(x)$ is strictly convex downward on $(-\infty, 0]$ and strictly convex upward on $[0, \infty)$. We have plotted $f(x) = \sinh(x)$ in the middle picture in Figure 4.7. □

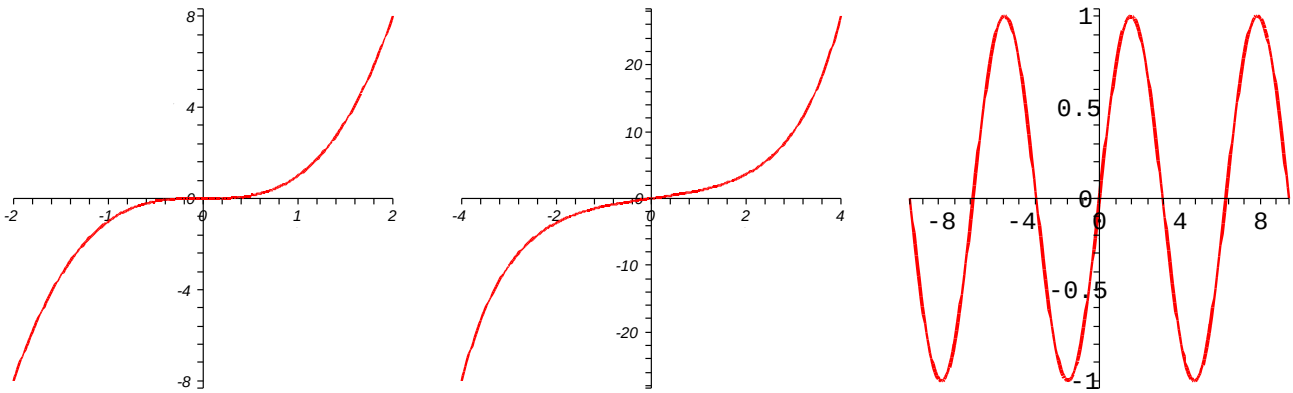


Figure 4.7: The graph of the function $f(x) = x^3$ on the left, the graph of the function $f(x) = \sinh(x)$ in the middle, and the graph of the function $f(x) = \sin(x)$ on the right.

Example 4.25 (curvature of $\sin(x)$)
Investigate the curvature of the sine function $\sin : \mathbb{R} \to \mathbb{R}$.

Solution: We want to use Lemma 4.20. Thus we calculate the second derivative of $\sin(x)$.

$$\frac{d}{dx}\sin(x) = \cos(x), \qquad \frac{d^2}{dx^2}\sin(x) = \frac{d}{dx}\cos(x) = -\sin(x).$$

From the geometric definition of the sine function we see that $-\sin(x) < 0$ for $x \in (0, \pi)$ and that $-\sin(x) > 0$ for $x \in (-\pi, 0)$. Because of the periodicity of the sine function with period 2π, we may add integer multiples of 2π, and thus we obtain

$$\frac{d^2}{dx^2}\sin(x) = -\sin(x) \quad \begin{cases} < 0 & \text{if} \quad x \in \big(2k\pi, (2k+1)\pi\big), \ k \in \mathbb{Z}, \\ > 0 & \text{if} \quad x \in \big((2k-1)\pi, 2k\pi\big), \ k \in \mathbb{Z}. \end{cases}$$

Thus we see that the function $\sin(x)$ is strictly convex upward on all closed intervals $\big[(2k-1)\pi, 2k\pi\big]$, $k \in \mathbb{Z}$, and that $\sin(x)$ is strictly convex downward on all closed intervals $\big[2k\pi, (2k+1)\pi\big]$, $k \in \mathbb{Z}$. The sine function is plotted in the right picture in Figure 4.7. □

The discussed examples were elementary in the sense that the curvature could be guessed from our knowledge of the graph of the function. This is intentional, since the knowledge of the graph helps with the understanding of the concept of curvature. In some examples that we get at the end of this chapter and in some problems on Exercise Sheet 4, you will not know the graph of the function beforehand and thus it will not be possible to determine the curvature without doing some calculations.

4.4 Changes of Curvature: Points of Inflection

We have already seen in the previous section that, if we want to analyze and plot a function, it is important to know about the **curvature** of f. Especially it is useful to know at which points (if any) f **changes its curvature**, that is, changes from being strictly convex upward to being strictly convex downward, or vice versa. In the previous section we have seen that the curvature can be characterized with the second derivative as follows:

A function f is **strictly convex upward** on $[a, b]$ if $f''(x) > 0$ for all $x \in (a, b)$, and f is **strictly convex downward** on $[a, b]$ if $f''(x) < 0$ for all $x \in (a, b)$.

Thus a **change of curvature is characterized by the fact that the second derivative changes its sign**. In particular, the **second derivative will be zero at the point where the change of curvature occurs**.

After these observations, the next definition is quite natural.

Definition 4.26 (point of inflection)
Let $A, B \subset \mathbb{R}$, and let $f : A \to B$ be a sufficiently smooth function. A ***point*** *$x_0 \in A$* ***where the second derivative*** *f''* ***changes its sign*** *is called a* ***point of inflection*** *of f. That is, if x_0 is a* ***point of inflection*** *of f, then we have with some suitable $\varepsilon > 0$*

$$f''(x) < 0 \quad \text{for all} \quad x \in (x_0 - \varepsilon, x_0) \qquad \text{and} \qquad f''(x_0) = 0$$
$$\text{and} \qquad f''(x) > 0 \quad \text{for all} \quad x \in (x_0, x_0 + \varepsilon), \tag{4.5}$$

or

$$f''(x) > 0 \quad \text{for all} \quad x \in (x_0 - \varepsilon, x_0) \qquad \text{and} \qquad f''(x_0) = 0$$
$$\text{and} \qquad f''(x) < 0 \quad \text{for all} \quad x \in (x_0, x_0 + \varepsilon). \tag{4.6}$$

Remark 4.27 (behavior of the graph at a point of inflection)
From Definition 4.20, we see that (4.5) means that at x_0 the graph of f ***changes from being strictly convex downward*** *(for $x < x_0$)* ***to being strictly convex upward*** *(for $x > x_0$). Likewise (4.6) means that at x_0 the graph of f* ***changes from being strictly convex upward*** *(for $x < x_0$)* ***to being strictly convex downward*** *(for $x > x_0$). This means that at x_0 the* ***curvature*** *of the graph* ***changes****.*

The definition of a point of inflection immediately implies a way of finding the points of inflection of a function.

Remark 4.28 (how to find the points of inflection)
Compute f'' and find its ***roots****, that is, the values of x for which $f''(x) = 0$. If for a root x_0 of f'' the* ***sign of*** *$f''(x)$* ***changes at*** *$x = x_0$ from being negative to being positive or vise versa, then the point x_0 is a* ***point of inflection****. Note that there are functions with points x_0 such that $f''(x_0) = 0$, but x_0 is* ***not a point of inflection****. So if $f''(x_0) = 0$, its is important to* ***check whether*** *f''* ***changes its sign at*** *$x = x_0$* ***or not****!*

Example 4.29 (analysis and sketch of $f(x) = x^3$)
Analyze the behavior of the function $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = x^3$, and sketch the function.

Solution: To be able to sketch the function $f(x) = x^3$, we determine its minima and maxima, points of inflection, and its behavior as $x \to \pm\infty$. We start with the behavior of $f(x)$ as $x \to \infty$,

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} x^3 = \infty, \qquad \lim_{x\to-\infty} f(x) = \lim_{x\to-\infty} x^3 = -\infty.$$

Furthermore, $f(x) = x^3 < 0$ for $x < 0$ and $f(x) = x^3 > 0$ for $x > 0$. Now we find the stationary points: if $f'(x) = 3\,x^2 = 0$, then $x = 0$. We compute the second derivative $f''(x) = 6\,x$. Since $f''(0) = 0$, from Theorem 4.13, the point $x = 0$ could be a local maximum or a local minimum or neither. The point $x = 0$ could also be a point of inflection, and to test this we investigate whether f'' changes its sign at $x = 0$. Since

$$f''(x) = 6\,x < 0 \quad \text{for all } x < 0 \qquad \text{and} \qquad f''(x) = 6\,x > 0 \quad \text{for all } x > 0,$$

we see that f'' changes its sign at $x = 0$. Thus $x = 0$ is a point of inflection, and $f(x) = x^3$ is strictly convex downward on $(-\infty, 0]$ and strictly convex upward on $[0, \infty)$. With this information and the values of $f(x)$ at some suitable points, we can sketch $f(x) = x^3$ and the graph is shown in Figure 4.8. □

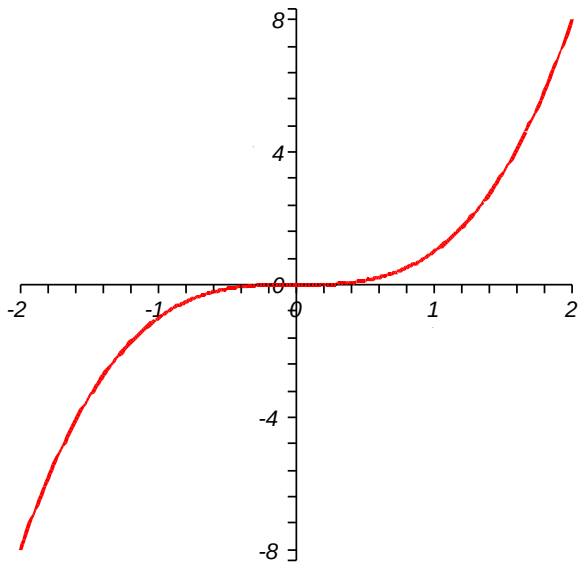


Figure 4.8: Graph of $f(x) = x^3$. The function has a point of inflection at $x = 0$.

Example 4.30 (analysis and sketch of polynomial of degree 3)
Analyze and sketch the function $f : \mathbb{R} \to \mathbb{R}$, given by

$$f(x) = x^3 - 4\,x = x\,(x^2 - 4) = x(x-2)(x+2).$$

Solution: From the last representation of f we see that its roots are at $x = 0$, $x = -2$ and $x = 2$. For very large $|x|$ the function $f(x) = x^3 - 4\,x$ behaves like x^3 because

$$f(x) = x^3 - 4\,x = x^3\left(1 - \frac{4}{x^2}\right) \approx x^3, \qquad \text{due to} \qquad \lim_{|x|\to\infty} \frac{4}{x^2} = 0.$$

In particular, $f(x) \approx x^3$ for large $|x|$ implies that

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} \left(x^3 - 4\,x\right) = \lim_{x\to\infty} x^3 = \infty,$$

and

$$\lim_{x\to-\infty} f(x) = \lim_{x\to-\infty} \left(x^3 - 4\,x\right) = \lim_{x\to-\infty} x^3 = -\infty.$$

From $f'(x) = 3\,x^2 - 4$, we find that the stationary points satisfy

$$f'(x) = 3\,x^2 - 4 = 0 \qquad \Rightarrow \qquad 3\,x^2 = 4 \qquad \Rightarrow \qquad x^2 = \frac{4}{3},$$

and thus the stationary points are

$$x = \sqrt{\frac{4}{3}} = \frac{2}{\sqrt{3}} \qquad \text{and} \qquad x = -\sqrt{\frac{4}{3}} = -\frac{2}{\sqrt{3}}.$$

The second derivative of $f(x) = x^3 - 4\,x$ is given by

$$f''(x) = \left(3\,x^2 - 4\right)' = 6\,x,$$

and we have

$$f''\left(\frac{2}{\sqrt{3}}\right) = 6\,\frac{2}{\sqrt{3}} = 4\,\sqrt{3} > 0, \qquad f''\left(-\frac{2}{\sqrt{3}}\right) = -6\,\frac{2}{\sqrt{3}} = -4\,\sqrt{3} < 0.$$

Thus the function $f(x) = x^3 - 4\,x$ has a local minimum at $x = 2/\sqrt{3}$ and it has a local maximum at $x = -2/\sqrt{3}$. From

$$f(2/\sqrt{3}) = \left(\frac{2}{\sqrt{3}}\right)^3 - 4\left(\frac{2}{\sqrt{3}}\right) = \frac{8}{3\sqrt{3}} - \frac{8}{\sqrt{3}} = \frac{8-24}{3\sqrt{3}} = -\frac{16}{3\sqrt{3}},$$

$$f(-2/\sqrt{3}) = \left(-\frac{2}{\sqrt{3}}\right)^3 - 4\left(-\frac{2}{\sqrt{3}}\right) = -\frac{8}{3\sqrt{3}} + \frac{8}{\sqrt{3}} = \frac{-8+24}{3\sqrt{3}} = \frac{16}{3\sqrt{3}},$$

we see that coordinates of the local maximum are $\left(-2/\sqrt{3}, 16/(3\sqrt{3})\right)$ and the coordinates of the local minimum are $\left(2/\sqrt{3}, -16/(3\sqrt{3})\right)$.

To find the points of infection (if any) we set the second derivative zero.

$$f''(x) = 6\,x = 0 \qquad \Rightarrow \qquad x = 0,$$

and since $f''(x) = 6\,x < 0$ for all $x < 0$ and $f''(x) = 6\,x > 0$ for all $x > 0$, we see that the point $x = 0$ is a point of inflection. The function $f(x) = x^3 - 4\,x$ is strictly convex downward on $(-\infty, 0]$ and strictly convex upward on $[0, \infty)$.

With this information and some suitable additional points $(x, f(x))$, we can now sketch $f(x) = x^3 - 4\,x$. The plot is in Figure 4.9. □

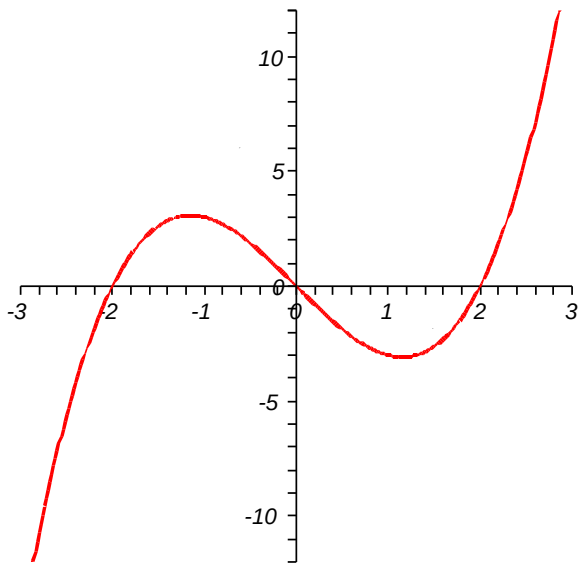


Figure 4.9: The graph of $f(x) = x^3 - 4\,x$. The function f has a point of inflection at $(0, 0)$ and a local maximum at $(-2/\sqrt{3}, 16/(3\sqrt{3})) \approx (-1.2, 3.1)$ and a local minimum at $(2/\sqrt{3}, -16/(3\sqrt{3})) \approx (1.2, -3.1)$

Example 4.31 (points of inflection of the sine function)
Find the points of inflection of the sine function.

Solution: We compute the second derivative of $\sin(x)$, and obtain

$$\frac{d}{dx}\sin(x) = \cos(x), \qquad \frac{d^2}{dx^2}\sin(x) = \frac{d}{dx}\cos(x) = -\sin(x),$$

and we see that the second derivative has the same zeros as the sine function itself. The roots of the second derivative of $\sin(x)$ are at

$$x = k\,\pi, \quad \text{where } k \in \mathbb{Z}, \qquad \text{that is,} \qquad x \in \{\ldots, -3\pi, -2\pi, -\pi, 0, \pi, 2\pi, 3\pi, \ldots\}\,.$$

From Example 4.25, we know that

$$\frac{d^2}{dx^2}\sin(x) = -\sin(x) \quad \begin{cases} < 0 & \text{if} \quad x \in \big(2k\,\pi, (2k+1)\,\pi\big), \ k \in \mathbb{Z}, \\ > 0 & \text{if} \quad x \in \big((2k-1)\,\pi, 2k\,\pi\big), \ k \in \mathbb{Z}, \end{cases}$$

and we see that the sine function changes sign at every point $x = k\,\pi$. Thus all the points $k\,\pi$, $k \in \mathbb{Z}$, are points of inflection. □

4.5 Turning Points/Extrema Revisited

With what we have learned so far, we do not know how to interpret the situation that $f'(x_0) = 0$, $f''(x_0) = 0$, and f'' does not change sign at $x = x_0$. We know only that the point x_0 is not a point of inflection. We will first discuss a simple example where such a situation occurs.

Example 4.32 (maxima and minima of $f(x) = x^4$ and $g(x) = -x^4$)
Consider the functions $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = x^4$, and $g : \mathbb{R} \to \mathbb{R}$, defined by $g(x) = -x^4$. Since $x^4 = (x^2)^2 > 0$ for all $x \neq 0$, we see that

$$\begin{aligned} f(0) = 0 < x^4 = f(x) \qquad & \text{for all} \quad x \in \mathbb{R} \setminus \{0\}, \\ g(0) = 0 > -x^4 = g(x) \qquad & \text{for all} \quad x \in \mathbb{R} \setminus \{0\}. \end{aligned}$$

Thus the function $f(x) = x^4$ has at $x = 0$ a global minimum, and the function $g(x) = -x^4$ has at $x = 0$ a global maximum.

However, we cannot determine this with the help of Theorem 4.13. Indeed, we have

$$f'(x) = 4\,x^3 = 0 \quad \Rightarrow \quad x = 0 \qquad \text{and} \qquad g'(x) = -4\,x^3 = 0 \quad \Rightarrow \quad x = 0,$$

and in both cases the only stationary point is $x = 0$. The second derivatives are given by

$$f''(x) = 12\,x^2 \qquad \text{and} \qquad g''(x) = -12\,x^2,$$

and we have that $f''(0) = 0$ and $g''(0) = 0$. Theorem 4.13 only tells us that f could have a local maximum or a local minimum or neither at $x = 0$. We also see that $x = 0$ is not a point of inflection of either of the two functions because the second derivatives do not change sign at $x = 0$.

How do we now determine that $f(x) = x^4$ has a local (and global) minimum at $x = 0$, and that $g(x) = -x^4$ has a local (and global) maximum at $x = 0$?

We continue to differentiate and find

$$f^{(3)}(x) = 24\,x \qquad \text{and} \qquad g^{(3)}(x) = -24\,x,$$

and

$$f^{(4)}(x) = 24 \qquad \text{and} \qquad g^{(4)}(x) = -24.$$

We see that the third derivatives are again zero at $x = 0$, that is, $f^{(3)}(0) = 0$ and $g^{(3)}(0) = 0$. However, the forth derivatives are different from zero, and we have

$$f^{(4)}(0) = 24 > 0 \qquad \text{and} \qquad g^{(4)}(0) = -24 < 0.$$

Based on this we would formulate the tentative rule that if at a stationary point x_0 of f, we have that $f''(x_0) = 0$ and $f^{(3)}(x_0) = 0$, but $f^{(4)}(x_0) \neq 0$, then x_0 is a turning point/extremum. If $f^{(4)}(x_0) > 0$, then f assumes at x_0 a local minimum; and if $f^{(4)}(x_0) < 0$, then f assumes at x_0 a local maximum. □

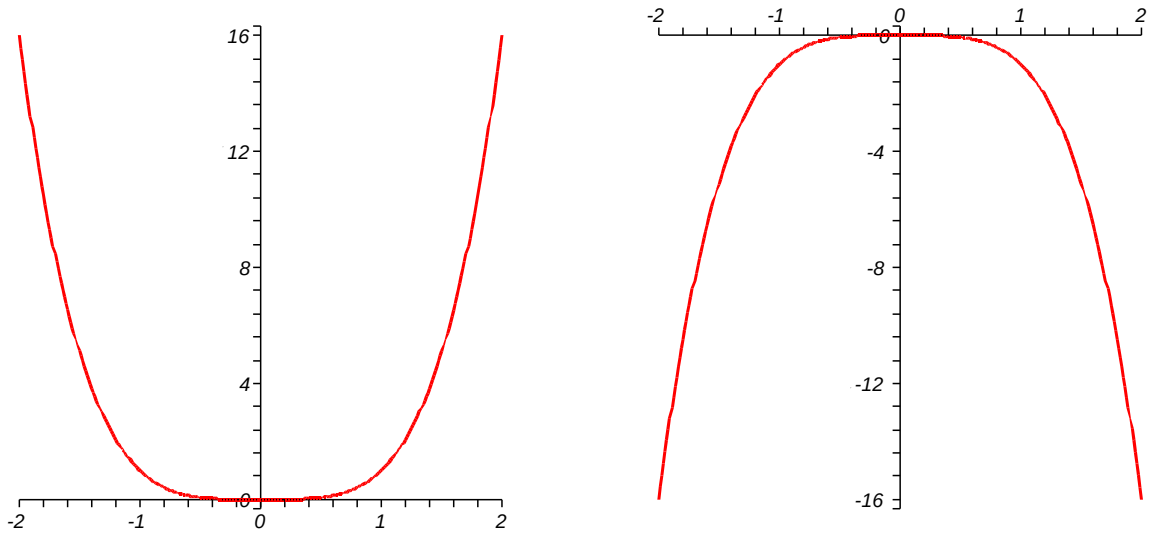


Figure 4.10: The graph of $f(x) = x^4$ on the left and the graph of $g(x) = -x^4$ on the right.

The tentative rule formulated in the previous example is indeed true and holds in an even more general form that covers all cases which we could not handle so far.

Theorem 4.33 (even order derivative test for extrema)
Let $A, B \subset \mathbb{R}$, and let $f : A \to B$ be a smooth enough function. If $x_0 \in A$ is a ***stationary point*** *of f (that is, $f'(x_0) = 0$) and if*

$$f''(x_0) = f^{(3)}(x_0) = \ldots = f^{(n-1)}(x_0) = 0 \qquad \text{and} \qquad f^{(n)}(x_0) \neq 0,$$

with an ***even*** *integer $n \geq 2$, then f has a* ***local maximum*** *at x_0 if*

$$f^{(n)}(x_0) < 0,$$

and f has a ***local minimum*** *at x_0 if*

$$f^{(n)}(x_0) > 0.$$

We observe that Theorem 4.33 includes Theorem 4.13 as the special case $n = 2$. Note that it is **essential that n is an even integer!** For odd n the statement is **not** true!

We will apply Theorem 4.33 in the next example.

Example 4.34 (polynomial of degree 5)
Analyze and sketch the graph of the polynomial $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = x\,(x-1)^4.$$

Solution: We see that the polynomial has the roots $x = 0$ and $x = 1$, that is, it intersects the x-axis at these points. We compute the first derivative of f with the product rule

$$f'(x) = (x-1)^4 + x\,4\,(x-1)^3 = (x-1)^3(x-1+4\,x) = (x-1)^3\,(5\,x-1),$$

and we see that the stationary points are $x = 1/5$ and $x = 1$. We compute the second derivative

$$\begin{aligned} f''(x) &= \left[(x-1)^3\,(5\,x-1)\right]' = 3\,(x-1)^2(5\,x-1) + (x-1)^3\,5 \\ &= (x-1)^2\,(15\,x-3+5\,x-5) = (x-1)^2\,(20\,x-8) = 4\,(x-1)^2\,(5\,x-2). \end{aligned}$$

We find that

$$f''(1/5) = 4\left(\frac{1}{5}-1\right)^2\left(5\times\frac{1}{5}-2\right) = 4\left(-\frac{4}{5}\right)^2(-1) = 4\times\frac{16}{25}\,(-1) = -\frac{64}{25} < 0,$$

and thus the polynomial has a local maximum at $x = 1/5$. The coordinates of this local maximum are $(1/5, 4^4/5^5) \approx (0.2, 0.08)$. At $x = 1$, we find on the other hand $f''(1) = 0$. The point $x = 1$ is not a point of inflection because $f''(x) \geq 0$ on $(0.4, \infty)$. (Indeed $(x-1)^2 \geq 0$ for all $x \in \mathbb{R}$ and $5\,x - 2 \geq 0$ is equivalent to $x \geq 2/5 = 0.4$.) Thus we suspect that the point $x = 1$ may also be a maximum or minimum and we keep on differentiating.

$$\begin{aligned} f^{(3)}(x) &= \left[4\,(x-1)^2\,(5\,x-2)\right]' = 8\,(x-1)\,(5\,x-2) + 4\,(x-1)^2\,5 \\ &= 4\,(x-1)\,(10\,x - 4 + 5\,x - 5) = 12\,(x-1)\,(5\,x-3), \end{aligned}$$

and we see that still $f^{(3)}(1) = 0$. We differentiate again and find that

$$\begin{aligned} f^{(4)}(x) &= \left[12\,(x-1)\,(5\,x-3)\right]' = 12\,(5\,x-3) + 12\,(x-1)\,5 \\ &= 12\,(5\,x - 3 + 5\,x - 5) = 12\,(10\,x - 8) = 24\,(5\,x - 4). \end{aligned}$$

Now we have

$$f^{(4)}(1) = 24\,(5-4) = 24 > 0,$$

and we know from Theorem 4.33 that the polynomial f attains a local minimum at $x = 1$. The coordinates of this minimum are $(1, 0)$.

We note that we could also have deduced the local minimum at $x = 1$ from the fact that

$$\begin{aligned} f'(x) &= (x-1)^3\,(5\,x-1) < 0 \qquad \text{for all } \; x \in (1/5, 1) \\ f'(x) &= (x-1)^3\,(5\,x-1) > 0 \qquad \text{for all } \; x > 1. \end{aligned}$$

Finally we discuss whether the function has any points of inflection. The equation $f''(x) = 0$ yields

$$f''(x) = 4\,(x-1)^2\,(5\,x-2) = 0 \qquad \Rightarrow \qquad x = 1 \quad \text{or} \quad x = \frac{2}{5} = 0.4.$$

Since f has a local minimum at $x = 1$, the point $x = 1$ is not a point of inflection. At the point $x = 2/5$ however, we have a point of inflection because $f''(x)$ changes its sign. Indeed, $(x-1)^2 > 0$ for $x \neq 1$, but $(5\,x - 2)$ changes its sign at $x = 2/5$. Since $5\,x - 2 < 0$ for $x < 2/5$ and $5\,x - 2 > 0$ for $x > 2/5$, we see that at $x = 2/5$ the function $f(x)$ changes from being strictly convex downward to being strictly convex upward. The coordinates of the point of inflection are $(2/5, (2 \times 3^4)/5^5) \approx (0.4, 0.05)$.

With these preparations, we can easily sketch the graph of $f(x) = x\,(x-1)^4$ which is shown in Figure 4.11 below.

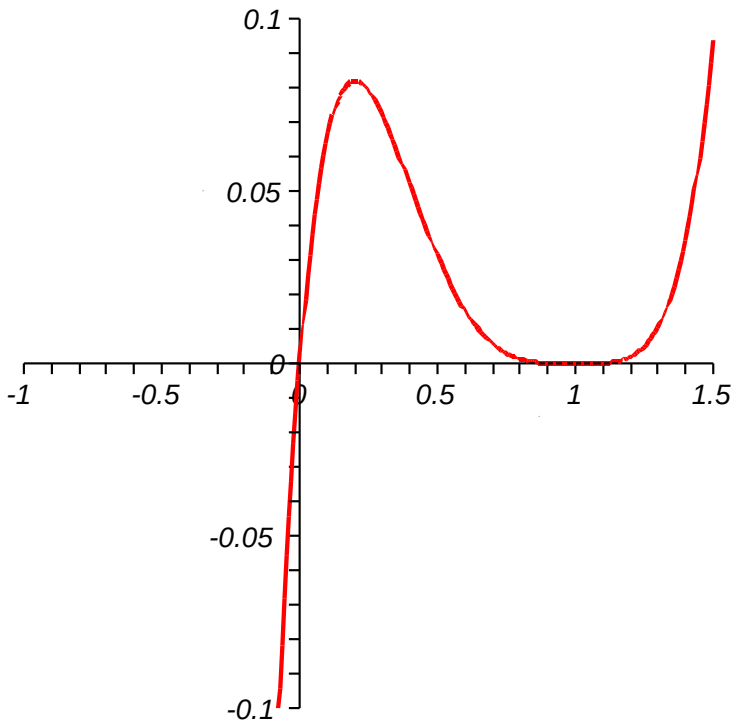


Figure 4.11: Graph of $f(x) = x\,(x-1)^4$. The function f has a local maximum at $(1/5, 4^4/5^5) \approx (0.2, 0.08)$, a local minimum at $(1, 0)$, and a point of inflection at $\big(2/5, (2 \times 3^4)/5^5\big) \approx (0.4, 0.05)$.

4.6 Analyzing and Sketching Functions

In this section we will discuss some more examples of analyzing and sketching functions. While discussing these examples, we will make use of the new knowledge gained in this chapter as well as of the old knowledge about sketching and analyzing functions from Chapters 1, 2, and 3.

At the end of this section we will discuss **damped oscillations** as an application. Damped oscillations describe, for example, the motion of a pendulum or the motion of a weight attached to an elastic spring.

Before we discuss some more examples, we summarize what we have learned that can help in analyzing and sketching functions.

Summary 4.35 (how to analyze and sketch a function)
Let $A, B \subset \mathbb{R}$, and let $f : A \to B$ be a smooth enough function. For ***analyzing*** *f* ***and sketching its graph****, the following may be useful:*

- *Check whether the function is* ***even*** *or* ***odd****.*
- *Compute the* ***roots*** *of f, that is, the points x at which $f(x) = 0$.*
- *Determine for which x the function f has* ***positive values*** *and for which x the function f has* ***negative values****.*
- *Determine where the function f is* ***(strictly) monotonically increasing*** *and where it is* ***(strictly) monotonically decreasing****.*
- *If the numbers in A get arbitrary large, then determine the* ***limit*** $\lim_{x\to\infty} f(x)$*, and if the numbers in A get arbitrarily small then determine the* ***limit*** $\lim_{x\to-\infty} f(x)$*. If the function is undefined at a single point x_0, then determine the* ***limits*** $\lim_{x<x_0,\, x\to x_0} f(x)$ *and* $\lim_{x>x_0,\, x\to x_0} f(x)$*, as appropriate. We would also like to know which* ***'pattern' the growth of*** *$f(x)$* ***follows*** *(that is, how fast f grows) as $x \to \pm\infty$ or $x \to x_0$. Find the* ***asymptotes*** *of f if it has any.*
- *Find the* ***stationary points*** *of f, that is, compute f' and determine the points x at which $f'(x) = 0$.*
- *For each stationary point check whether f has at this point a* ***local maximum*** *or a* ***local minimum*** *or* ***neither*** *by using Theorems 4.13 and 4.33.*
- *Determine the* ***curvature*** *of f, that is, where f is* ***(strictly) convex upward*** *and where f is* ***(strictly) convex downward****, and determine the* ***points of inflection****, that is, the points x_0 at which $f''(x_0) = 0$ and at which f'' changes its sign.*

Note that for plotting or even analyzing a function it is mostly not necessary to investigate all items listed above in Summary 4.35! Often it is sufficient for the task at hand to consider only a suitable selection of the items listed above.

Example 4.36 Let $a > 0$ be fixed. Analyze the function $f : \mathbb{R} \to \mathbb{R}$, defined by

$$f(x) = \sqrt{x^2 + a^2} = \left(x^2 + a^2\right)^{1/2},$$

and sketch its graph.

Solution: The function is even because

$$f(-x) = \sqrt{(-x)^2 + a^2} = \sqrt{x^2 + a^2} = f(x) \qquad \text{for all} \quad x \in \mathbb{R},$$

and therefore we know that its graph is mirror symmetric with respect to the y-axis. Since

$$x^2 + a^2 \geq a^2 > 0 \qquad \Rightarrow \qquad f(x) = \sqrt{x^2 + a^2} \geq a > 0,$$

we see that the function has only positive values, and thus it has no roots. We note that its lowest value and global minimum is attained at $x = 0$ and the value is $f(0) = \sqrt{a^2} = a$. Thus the global (and local) minimum has the coordinates $(0, a)$.

When $|x|$ is large, we have

$$f(x) = \sqrt{x^2 + a^2} \approx \sqrt{x^2} = \sqrt{|x|^2} = |x|$$

Thus $y(x) = x$ and $g(x) = -x$ are asymptotes of f. We have

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} |x| = \infty \qquad \text{and} \qquad \lim_{x\to-\infty} f(x) = \lim_{x\to-\infty} |x| = \infty.$$

Differentiating with the chain rule gives that

$$f'(x) = \frac{1}{2}\left(x^2 + a^2\right)^{-1/2} 2\,x = \frac{x}{\sqrt{x^2 + a^2}}.$$

Thus the only stationary point is at $x = 0$, where f assumes its global minimum. That f has a global minimum at $x = 0$ can also be seen from $f'(x) < 0$ for all $x < 0$ and $f'(x) > 0$ for all $x > 0$. We observe that for $|x|$ large, we have that

$$f'(x) = \frac{x}{\sqrt{x^2 + a^2}} \approx \frac{x}{|x|} = \begin{cases} 1 & \text{if} \quad x > 0 \text{ is very large,} \\ -1 & \text{if} \quad x < 0 \text{ is very small.} \end{cases}$$

We have previously derived that $f(x) \approx |x|$ for large $|x|$, and we see this reflected in the fact that $f'(x) \approx 1$ for very large $x > 0$ and that $f'(x) \approx -1$ for very small $x < 0$.

We compute the second derivative with the chain rule and the quotient rule

$$f''(x) = \frac{(x^2 + a^2)^{1/2} - x\,(1/2)\,(x^2 + a^2)^{-1/2}\,2\,x}{(x^2 + a^2)} = \frac{(x^2 + a^2) - x^2}{(x^2 + a^2)^{3/2}} = \frac{a^2}{(x^2 + a^2)^{3/2}}.$$

Since $f''(x) > 0$ for all $x \in \mathbb{R}$, f has no point of inflection. We see that $f''(0) = a^2/a^3 = 1/a > 0$, which gives another proof that $f(x)$ attains at the stationary point $x = 0$ a local minimum. With the information gained so far we can easily sketch $f(x) = \sqrt{x^2 + a^2}$, and the plot of the graph is shown in Figure 4.12 below. $\square$

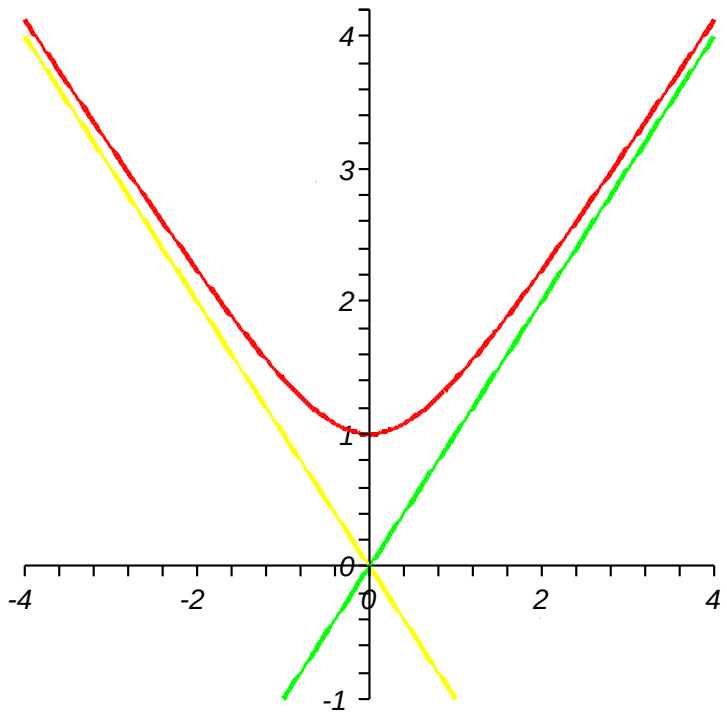


Figure 4.12: Graph of $f(x) = \sqrt{x^2 + a^2}$ and its asymptotes $y(x) = x$ and $g(x) = -x$.

Example 4.37 Analyze and plot the function $f : \mathbb{R} \setminus \{1\} \to \mathbb{R}$, defined by

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

Solution: We have that $f(0) = 0$, and that $f(x) > 0$ when $x > 1$, and $f(x) < 0$ when $x < 1$. As x approaches the point $x_0 = 1$, we have

$$\lim_{x>1,\, x\to 1} f(x) = \infty \qquad \text{and} \qquad \lim_{x<1,\, x\to 1} f(x) = -\infty.$$

Thus the vertical line through $(1, 0)$ is an asymptote of $f(x) = x^2/(x-1)$. When $|x|$ is large, we find

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

Let $y(x) = x + 1$. Since

$$\lim_{x\to\pm\infty} \big[f(x) - y(x)\big] = \lim_{x\to\pm\infty} \left[x + 1 + \frac{1}{x-1} - \big(x+1\big)\right] = \lim_{x\to\pm\infty} \frac{1}{x-1} = 0,$$

we see that the straight line $y(x) = x + 1$ is an asymptote of $f(x) = x^2/(x-1)$ (see Definition 1.11 for the definition of an asymptote). This implies

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} (x+1) = \infty \qquad \text{and} \qquad \lim_{x\to-\infty} f(x) = \lim_{x\to-\infty} (x+1) = -\infty.$$

Differentiating using the quotient rule gives that

$$f'(x) = \frac{2\,x\,(x-1) - 1\,x^2}{(x-1)^2} = \frac{2\,x^2 - 2\,x - x^2}{(x-1)^2} = \frac{x^2 - 2\,x}{(x-1)^2} = \frac{x\,(x-2)}{(x-1)^2}.$$

Thus the derivative $f'(x)$ is zero at $x = 0$ and at $x = 2$.

From the information obtained so far, we can already conclude that the function has a local maximum at $x = 0$ and a local minimum at $x = 2$. Indeed, since

$$\lim_{x>1,\,x\to 1} f(x) = \infty \qquad \text{and} \qquad \lim_{x\to\infty} f(x) = \infty$$

and since f has finite values on $(1, \infty)$, the function f must have at least one local minimum on $(1, \infty)$. Every local minimum is a stationary point, and we have only found the stationary point $x = 2$ in $(1, \infty)$. Thus this stationary point $x = 2$ is the point where f has the local minimum. Likewise we know, from

$$\lim_{x<1,\,x\to 1} f(x) = -\infty \qquad \text{and} \qquad \lim_{x\to-\infty} f(x) = -\infty$$

and from the fact that the function f has only finite values on $(-\infty, 1)$, that the function f has at least one local maximum on $(-\infty, 1)$. Since a point at which f has a local maximum is a stationary point and since f has only the stationary point $x = 0$ in $(-\infty, 1)$, we know that f assumes a local maximum at the point $x = 0$. We also know that there are no more turning points/extrema other than $x = 0$ and $x = 2$, because these are the only stationary points.

If you find these considerations too complicated, you can alternatively use the second derivative, once the stationary points $x = 0$ and $x = 2$ have been found. Before we compute the second derivative, we rewrite f' as

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

Now we can compute the second derivative easily

$$f''(x) = \big(1 - (x-1)^{-2}\big)' = (-1)(-2)(x-1)^{-3} = 2\,(x-1)^{-3}. \tag{4.7}$$

We find that $f''(0) = 2\,(-1)^3 = -2 < 0$, and thus f has a local maximum at $x = 0$. We find that $f''(2) = 2\,(1)^3 = 2 > 0$, and thus f has a local minimum at $x = 2$. Since $f''(x) \neq 0$ for all $x \in \mathbb{R} \setminus \{1\}$, the function f has no points of inflection.

The local maximum has the coordinates $(0, 0)$ and the local minimum has the coordinates $(2, 4)$, and we can plot f.

From (4.7) we see that $f''(x) < 0$ for $x < 1$ and that $f''(x) > 0$ for $x > 1$. Thus $f(x) = x^2/(x-1)$ is strictly convex downward on $(-\infty, 1)$ and strictly convex upward on $(1, \infty)$.

With all this information we can easily sketch $f(x) = x^2/(x-1)$ and the plot of f is shown in Figure 4.13. □

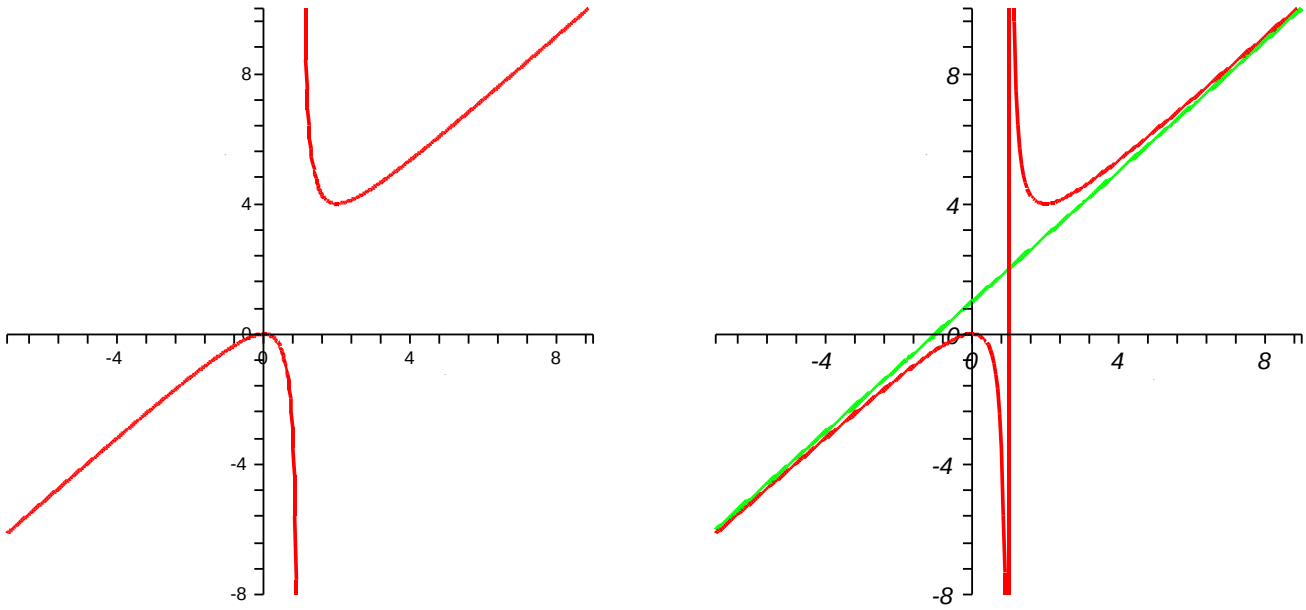


Figure 4.13: Graph of $f(x) = x^2/(x-1)$ in the left picture. In the right picture the asymptotes of f, which are $y(x) = x + 1$ and the vertical line through $(1, 0)$, have been included in the plot. The function f has a local maximum at $(0, 0)$ and a local minimum at $(2, 4)$.

Example 4.38 (analyze and sketch $f(x) = x\,e^{-x^2}$)
Analyze and sketch the function

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = x\,\exp(-x^2) = x\,e^{-x^2}.$$

Solution: Since $e^{-x^2} > 0$ for all $x \in \mathbb{R}$, we find that $f(x) = x\,e^{-x^2} < 0$ for all $x < 0$, $f(0) = 0$, and $f(x) = x\,e^{-x^2} > 0$ for all $x > 0$.

We have

$$\lim_{x\to\infty} x\,e^{-x^2} = 0 \qquad \text{and} \qquad \lim_{x\to-\infty} x\,e^{-x^2} = 0, \tag{4.8}$$

which can be shown with **de l'Hospital's rule** which you may know form school. In our situation we need the following version of de l'Hospital's rule: If $\lim_{x\to\pm\infty} f(x) = \infty$

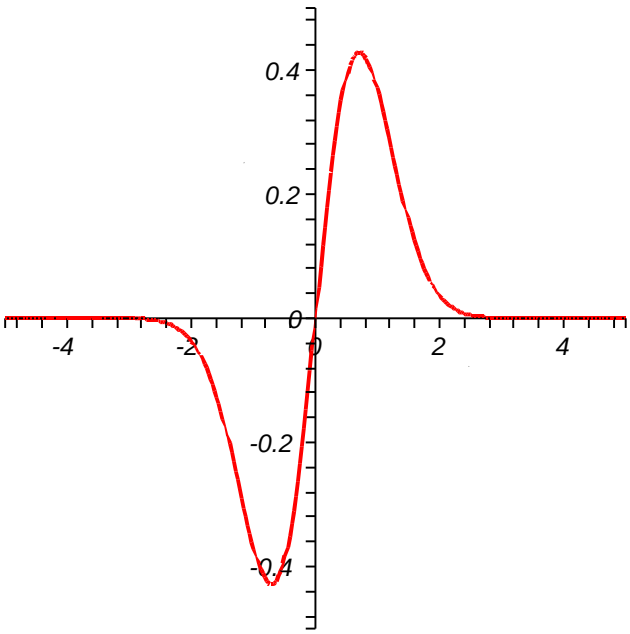


Figure 4.14: Graph of $f(x) = x\,e^{-x^2}$.

(or $\lim_{x\to\pm\infty} f(x) = -\infty$) and $\lim_{x\to\pm\infty} g(x) = \infty$, then

$$\lim_{x\to\pm\infty} \frac{f(x)}{g(x)} = \lim_{x\to\pm\infty} \frac{f'(x)}{g'(x)}.$$

Application in our case yields

$$\lim_{x\to\pm\infty} x\,e^{-x^2} = \lim_{x\to\pm\infty} \frac{x}{e^{x^2}} = \lim_{x\to\pm\infty} \frac{1}{2\,x\,e^{x^2}} = 0,$$

as claimed. From (4.8), we also see that $y(x) = 0$ is an asymptote of $f(x) = x\,e^{-x^2}$.

Next we determine the extrema/turning points and the points of inflection. We calculate f' with the product rule and the chain rule, and get

$$\begin{aligned} f'(x) &= \frac{d}{dx}\left[x\,e^{-x^2}\right] = e^{-x^2} + x\,e^{-x^2}\,(-2\,x) = (1-2\,x^2)\,e^{-x^2} \\ &= -2\left(x^2 - \frac{1}{2}\right)e^{-x^2} = -2\left(x - \frac{1}{\sqrt{2}}\right)\left(x + \frac{1}{\sqrt{2}}\right)e^{-x^2}. \end{aligned}$$

Thus $f'(x) = 0$ is satisfied for $x = -1/\sqrt{2}$ and $x = 1/\sqrt{2}$. Inspection of $f'(x)$ shows that $f'(x) < 0$ for all $x < -1/\sqrt{2}$, that $f'(x) > 0$ for all $x \in (-1/\sqrt{2}, 1/\sqrt{2})$, and that $f'(x) < 0$ for all $x > 1/\sqrt{2}$. Thus we may conclude that $f(x) = x\,e^{-x^2}$ has a local minimum at $x = -1/\sqrt{2}$ and a local maximum at $x = 1/\sqrt{2}$. The

coordinates of the local minimum are $(-1/\sqrt{2}, -e^{-1/2}/\sqrt{2}) \approx (-0.71, -0.43)$, and the coordinates of the local minimum are $(1/\sqrt{2}, e^{-1/2}/\sqrt{2}) \approx (0.71, 0.43)$.

We calculate the second derivative of f and find

$$\begin{aligned} f''(x) &= \frac{d}{dx}\left[(1-2\,x^2)\,e^{-x^2}\right] = -4\,x\,e^{-x^2} + (1-2\,x^2)\,e^{-x^2}\,(-2\,x) \\ &= \left[-4\,x - 2\,x + 4\,x^3\right] e^{-x^2} = \left[4\,x^3 - 6\,x\right] e^{-x^2} \\ &= 4\,x\left[x^2 - \frac{3}{2}\right] e^{-x^2} = 4\,x\left(x - \sqrt{\frac{3}{2}}\right)\left(x + \sqrt{\frac{3}{2}}\right) e^{-x^2}. \end{aligned} \tag{4.9}$$

By evaluating the second derivative at the stationary points $x = -1/\sqrt{2}$ and $x = 1/\sqrt{2}$ we could also have determined that they are extrema (instead of inspecting the sign of the first derivative). We do this now:

$$\begin{aligned} f''(-1/\sqrt{2}) &= -4\,\frac{1}{\sqrt{2}}\left(-\frac{1}{\sqrt{2}} - \sqrt{\frac{3}{2}}\right)\left(-\frac{1}{\sqrt{2}} + \sqrt{\frac{3}{2}}\right) e^{-1/2} \\ &= -\frac{4}{\sqrt{2}}\,\frac{(-1-\sqrt{3})}{\sqrt{2}}\,\frac{(-1+\sqrt{3})}{\sqrt{2}}\,e^{-1/2} \\ &= -\sqrt{2}\left((-1)^2 - (\sqrt{3})^2\right) e^{-1/2} = 2\sqrt{2}\,e^{-1/2} > 0, \end{aligned}$$

and we see that f has a local minimum at $x = -1/\sqrt{2}$.

$$\begin{aligned} f''(1/\sqrt{2}) &= 4\,\frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}} - \sqrt{\frac{3}{2}}\right)\left(\frac{1}{\sqrt{2}} + \sqrt{\frac{3}{2}}\right) e^{-1/2} \\ &= \frac{4}{\sqrt{2}}\,\frac{(1-\sqrt{3})}{\sqrt{2}}\,\frac{(1+\sqrt{3})}{\sqrt{2}}\,e^{-1/2} \\ &= \sqrt{2}\left(1^2 - (\sqrt{3})^2\right) e^{-1/2} = -2\sqrt{2}\,e^{-1/2} < 0, \end{aligned}$$

and we see that f has a local maximum at $x = 1/\sqrt{2}$.

Now we discuss the curvature of $f(x) = x\,e^{-x^2}$ and the points of inflection. From (4.9), we see that $f''(x) < 0$ for all $x < -\sqrt{3/2}$, and $f''(x) > 0$ for all $x \in (-\sqrt{3/2}, 0)$, and $f''(x) < 0$ for all $x \in (0, \sqrt{3/2})$, and $f''(x) > 0$ for all $x > \sqrt{3/2}$. Thus

- f is strictly convex downward on $(-\infty, -\sqrt{3/2}]$,
- f is strictly convex upward on $[-\sqrt{3/2}, 0]$,
- f is strictly convex downward on $[0, \sqrt{3/2}]$, and

- f is strictly convex upward on $[\sqrt{3/2}, \infty)$.

In particular, the roots of the second derivative, $x = 0$, $x = -\sqrt{3/2}$ and $x = \sqrt{3/2}$, are points of inflection of f, because f'' changes its sign at these points.

With this information we can easily sketch the function $f(x) = x\, e^{-x^2}$, and the plot is shown in Figure 4.14. □

Finally we close the chapter with the example of **damped oscillation**.

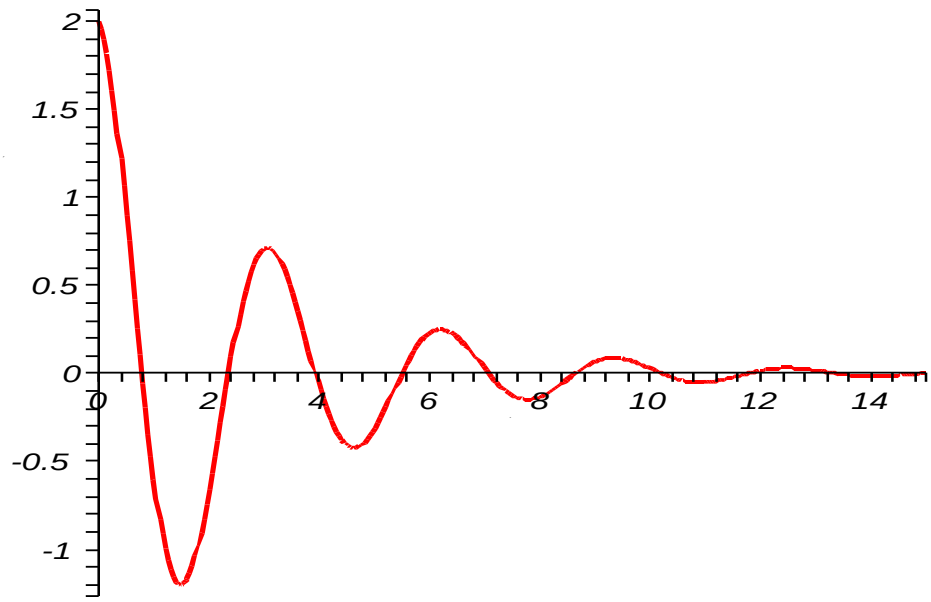


Figure 4.15: Graph of the damped oscillation $x(t) = 2\, e^{-t/3} \cos(2t)$.

Application 4.39 (damped oscillation)
*A **simple harmonic motion** occurs when the **force acting on the object is directly proportional to the displacement** x **from a fixed point and is always directed towards this point**. The displacement x of such a simple harmonic motion is of the form*

$$x(t) = a\, \sin(\omega\, t),$$

*where a is the **amplitude**, that is, the greatest displacement from equilibrium position, and the variable t is the time. The constant ω is given by*

$$\omega = 2\pi\, f, \qquad \text{where } f \text{ is the } \textbf{\textit{frequency of oscillation}},$$

*that is, the **number of cycles per second**. Examples of a simple harmonic motion are the motion of a pendulum or the motion of a mass attached to an elastic spring, if we neglect all external influences such as friction. The simple harmonic*

motion could also describe an electromagnetic wave. Since a pendulum on the earth is subjected to viscosity (friction) of the air, the amplitude of the pendulum becomes progressively smaller. We describe this by saying that the motion of the pendulum is ***damped****. Such a damped oscillation can be of the form*

$$x(t) = a\, e^{-b\,t}\, \cos(\omega\, t), \tag{4.10}$$

where the factor $e^{-b\,t}$ *with* $b > 0$ *controls the dampening effect. We have plotted (4.10) with* $a = 2$, $b = 1/3$, *and* $\omega = 2$ *in Figure 4.15 below.*

Chapter 5

Basic Integration

Integration is the **reverse operation to differentiation**. In this chapter we will introduce integration and discuss all standard techniques of integration. In Chapter 6, we will use integration to tackle various applications, such as, determining the volume and centre of mass of a body of revolution, as well as calculating moments of inertia and average values of functions.

In Section 5.1, we define the integral geometrically as the **area under the graph**, and then we derive a formal definition of the area under graph with the help of the **lower sum, upper sum**, and the **trapezium sum**. In Section 5.2, we introduce **primitives** (or **antiderivatives**) and **indefinite integrals** (which provide a means of getting a primitive). We discuss the **fundamental theorem of calculus** which links integration with differentiation and which is the 'backbone of calculus'. In Section 5.3, we get a **list of primitives/antiderivatives** (or indefinite integrals) of a set of standard functions. This list of primitives/antiderivatives follows directly from our knowledge of derivatives with the help of the fundamental theorem of calculus. In Section 5.4, we learn elementary properties of the integral, namely the **linear properties** and the **domain splitting property**. In Sections 5.5 and 5.6, we learn the two most important techniques of integration: **integration by substitution** and **integration by parts**, respectively. In Section 5.7, we discuss some examples where a combination of integration by substitution and integration by parts is needed to solve the integral, and we also discuss a **physical application of integration in gravitation**.

5.1 Geometric Definition and Interpretation of the Integral

The integral $\int_a^b f(x)\,dx$ of a function f over an interval $[a,b]$ is geometrically defined as follows:

Definition 5.1 (geometric definition of the integral)
*Let $A, B \subset \mathbb{R}$. The **integral of a function** $f : A \to B$ **over the interval** $[a,b] \subset A$, denoted by*

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

*is the **area under the graph** from $x = a$ to $x = b$, as indicated in Figure 5.1.*

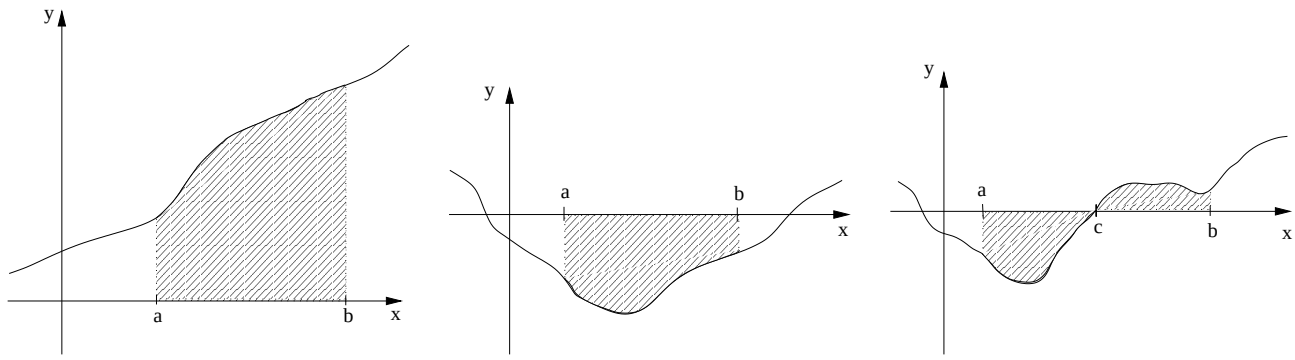


Figure 5.1: The definition of $\int_a^b f(x)\,dx$ as the area under the graph.

The area under the graph can be **positive**, if the graph of f for $x \in [a,b]$ lies above the x-axis, that is, if $f(x) \geq 0$ for all $x \in [a,b]$ as in the left picture in Figure 5.1. The area under the graph can be **negative**, if the the graph of f for $x \in [a,b]$ lies below the x-axis, that is, if $f(x) \leq 0$ for all $x \in [a,b]$ as in the middle picture in Figure 5.1. In the right picture in Figure 5.1, the graph of f for $x \in [a,b]$ lies partly above and partly below the x-axis. In this case for $x \in [a,c]$ the graph lies below the x-axis and the area under the graph for $x \in [a,c]$ is negative. And for $x \in [c,b]$ the graph lies above the x-axis and the area under the graph for $x \in [c,b]$ is positive. In the right picture of Figure 5.1, the area under the graph for $x \in [a,b]$ is the **sum of the negative area** under the graph for $x \in [a,c]$ **and the positive area** under the graph for $x \in [c,b]$.

For the most simple example of the integral of a constant function over an interval $[a,b]$, it is relatively clear how to measure the area under the graph.

Example 5.2 (integral of a constant function)
For the constant function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = c$, with some constant $c \in \mathbb{R}$ the integral

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

can be easily calculated: Due to the fact that graph of the constant function $f(x) = c$ is just a straight line parallel to the x-axis and through $(0, c)$, the area under the graph from $x = a$ to $x = b$ is just the rectangle with the corners (see Figure 5.2)

$$(a, 0), \quad (b, 0), \quad (b, c), \quad (a, c).$$

The area of the rectangle is

$$\text{length of bottom side} \;\times\; \text{height} \;=\; (b - a)\,c.$$

Thus

$$\int_a^b f(x)\,dx = \int_a^b c\,dx = (b-a)\,c. \qquad \square$$

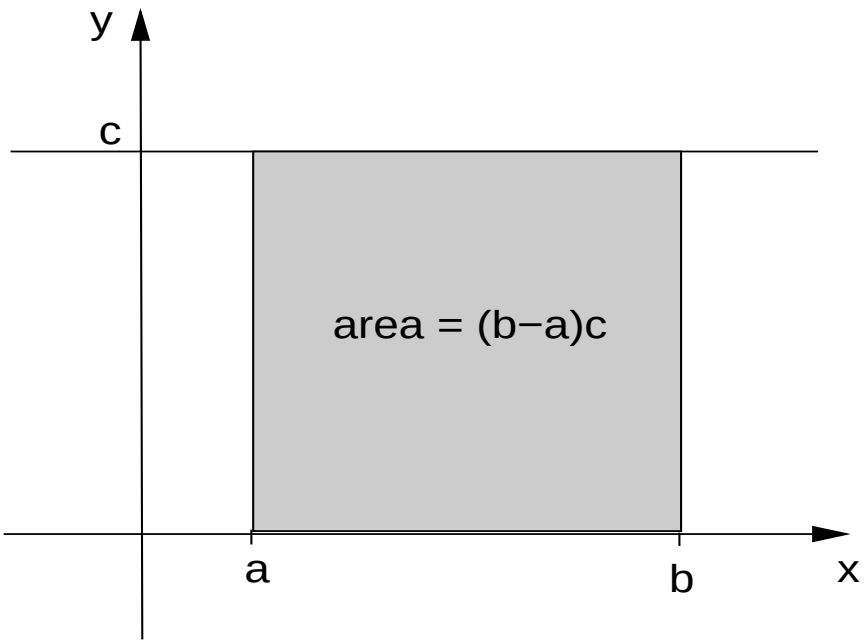


Figure 5.2: Area under the graph of the constant function $f(x) = c$ from $x = a$ to $x = b$.

Now we want to generalize this idea to the case of arbitrary continuous functions, where we initially not expect to get the area under the graph exactly, but to get an approximation of the exact value.

How would one elementary approximate the area under the graph?

Since we can easily compute the area of rectangles, an natural idea is to **fill the area under the graph with a suitable collection of rectangles** and then **take the sum of the areas of the rectangles as an approximation of the area of under the graph**. How this is done and what a suitable collection of rectangles is will be explained in the rest of this section.

For simplicity we will now assume that the function $f : A \to B$ which we want to integrate is **continuous**.

We want to **approximate the integral**

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

of a continuous function $f : A \to B$ over the interval $[a, b] \subset A$, that is, we want to **approximate the area under the graph from $x = a$ to $x = b$**. As a starting point, we **split the interval $[a, b]$ into n subintervals of equal length** (see Figure 5.3).

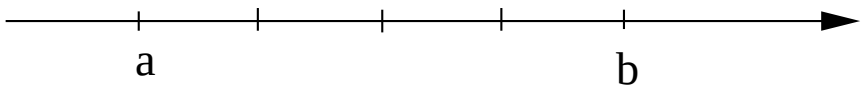


Figure 5.3: Division of the interval $[a, b]$ into $n = 4$ subintervals of equal length.

Since the length of the interval $[a, b]$ is $b-a$, each of the subintervals of equal length has then length $(b - a)/n$. The n subintervals are the given by

$$\left[a, a + \frac{b-a}{n}\right], \quad \left[a + \frac{b-a}{n}, a + 2\,\frac{b-a}{n}\right], \quad \left[a + 2\,\frac{b-a}{n}, a + 3\,\frac{b-a}{n}\right],$$
$$\ldots, \quad \left[a + (n-2)\,\frac{b-a}{n}, a + (n-1)\,\frac{b-a}{n}\right], \quad \left[a + (n-1)\,\frac{b-a}{n}, b\right],$$

or in shorter notation

$$\left[a + (k-1)\,\frac{b-a}{n}, a + k\,\frac{b-a}{n}\right], \qquad \text{for} \quad k = 1, 2, 3, \ldots, n-1, n.$$

For simplicity, we will use the abbreviated notation

$$x_k = a + k\,\frac{b-a}{n}, \qquad k = 0, 1, 2, 3, \ldots, n-1, n,$$

for the endpoints of the subintervals. The n subintervals are in this new notation

$$[x_0, x_1], \quad [x_1, x_2], \quad [x_2, x_3], \quad \ldots, \quad [x_{n-2}, x_{n-1}], \quad [x_{n-1}, x_n].$$

Now we take over each subinterval the rectangle whose height is either the minimal value of the function or the maximal value of the function assumed over this subinterval, respectively, as indicated in Figure 5.4. The sum of the area of all these rectangles in the left and right picture in Figure 5.4, respectively, gives an approximation of the area under the graph.

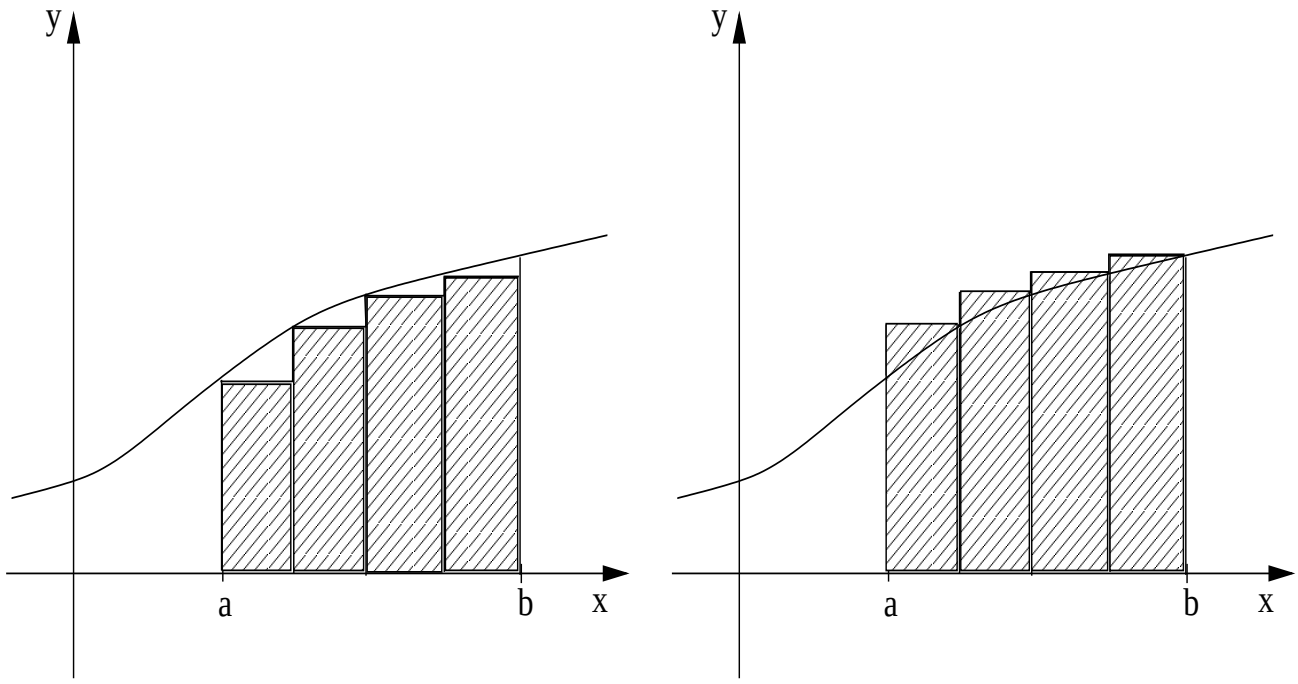


Figure 5.4: The lower sum of f over the interval $[a, b]$ with respect to a division of $[a, b]$ into $n = 4$ subintervals of equal length is the sum of the areas of the rectangles (see left picture). The upper sum of f over the interval $[a, b]$ with respect to a division of $[a, b]$ into $n = 4$ subintervals of equal length is the sum of the areas of the rectangles (see right picture).

The formal definition below uses the notation of the **maximum** and **minimum** of a continuous function on a closed interval $[c, d]$, defined by

$$\max_{x \in [c,d]} f(x) \;=\; \max\{f(x) \,:\, x \in [c, d]\} \;=\; \text{largest value of } f(x) \text{ for } x \in [c, d],$$

and

$$\min_{x \in [c,d]} f(x) \;=\; \min\{f(x) \,:\, x \in [c, d]\} \;=\; \text{smallest value of } f(x) \text{ for } x \in [c, d],$$

respectively.

Definition 5.3 (lower sum and upper sum)
Let $A, B \subset \mathbb{R}$, and let $f : A \to B$ be a continuous function. Let $[a, b] \subset A$, and let

$$x_k = a + k\,\frac{b-a}{n}, \qquad k = 0, 1, 2, 3, \ldots, n-1, n,$$

*be the endpoints of the intervals of the subdivision of $[a, b]$ into n subintervals of equal length. The **lower sum** $L_n(f, [a, b])$ of f with respect to the subdivision of $[a, b]$ into n subintervals of equal length is defined by*

$$L_n(f, [a, b]) = \frac{b-a}{n} \sum_{k=1}^{n} \min_{x \in [x_{k-1}, x_k]} f(x)$$

$$= \frac{b-a}{n} \left(\min_{x \in [x_0, x_1]} f(x) \right) + \frac{b-a}{n} \left(\min_{x \in [x_1, x_2]} f(x) \right) + \ldots + \frac{b-a}{n} \left(\min_{x \in [x_{n-1}, x_n]} f(x) \right).$$

*The **upper sum** $U_n(f, [a, b])$ of f with respect to the subdivision of $[a, b]$ into n subintervals of equal length is defined by*

$$U_n(f, [a, b]) = \frac{b-a}{n} \sum_{k=1}^{n} \max_{x \in [x_{k-1}, x_k]} f(x)$$

$$= \frac{b-a}{n} \left(\max_{x \in [x_0, x_1]} f(x) \right) + \frac{b-a}{n} \left(\max_{x \in [x_1, x_2]} f(x) \right) + \ldots + \frac{b-a}{n} \left(\max_{x \in [x_{n-1}, x_n]} f(x) \right).$$

We note here that if the function f is not continuous then the minimum needs to be replaced by the so-called **infimum** and the maximum needs to be replaced by the so-called **supremum**. The infimum and supremum are essentially generalizations of the notion of a minimum and a maximum. However, we will not discuss them here.

From Figure 5.4, it is immediately clear that the lower sum $L_n(f, [a, b])$ and the upper sum $U_n(f, [a, b])$ are an approximation of the integral $\int_a^b f(x)\,dx$. It is intuitively clear that this **approximation gets better if the subintervals get smaller**, that is, **if n gets larger**.

From Figure 5.4, it is also immediately clear the sum of the areas of the rectangle in the left picture is less then the area under the graph and that the sum of the areas of the rectangles in the right picture is more than the area under the graph. This implies that the lower sum is less than the area under the graph and that the upper sum is more than the area under the graph.

Lemma 5.4 (lower bound and upper bound for the intergral)
Let $A, B \subset \mathbb{R}$, and let $[a, b] \subset A$ be a closed interval. Let $f : A \to B$ be a continuous function. Then for all $n \in \mathbb{N}$,

$$L_n(f, [a, b]) \leq \int_a^b f(x)\, dx \leq U_n(f, [a, b]).$$

To get a better approximation than either the lower sum $L_n(f, [a, b])$ or the upper sum $U_n(f, [a, b])$, we can take their **mean value**, which means that the minimum and the maximum of the function values on each subinterval, respectively, is replaced by the mean value of these two. Thus

$$\begin{aligned}
&\frac{1}{2}\left[L_n(f, [a, b]) + U_n(f, [a, b])\right] && (5.1)\\
&= \frac{1}{2}\left[\frac{b-a}{n}\sum_{k=1}^{n}\min_{x\in[x_{k-1},x_k]} f(x) + \frac{b-a}{n}\sum_{k=1}^{n}\max_{x\in[x_{k-1},x_k]} f(x)\right]\\
&= \frac{b-a}{n}\sum_{k=1}^{n}\frac{1}{2}\left[\min_{x\in[x_{k-1},x_k]} f(x) + \max_{x\in[x_{k-1},x_k]} f(x)\right]. && (5.2)
\end{aligned}$$

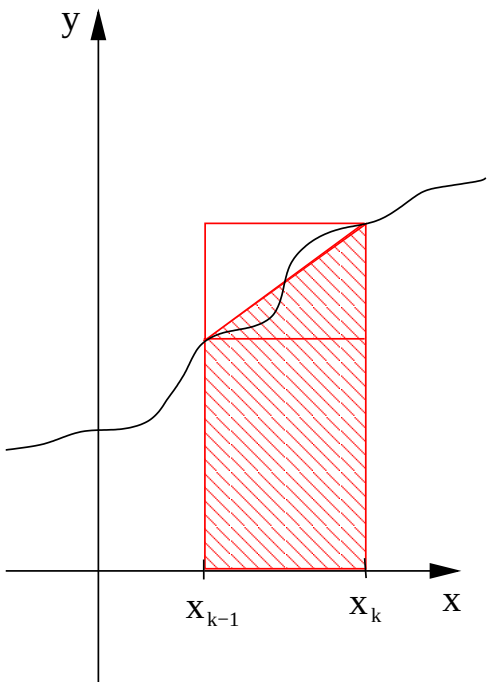


Figure 5.5: One trapezium from the trapezium sum.

One individual term in the sum in (5.2) is given by

$$\frac{b-a}{n}\,\frac{1}{2}\left[\min_{x\in[x_{k-1},x_k]} f(x) + \max_{x\in[x_{k-1},x_k]} f(x)\right],$$

and is the area of the **trapezium with the corners** (see Figure 5.5)

$$(x_{k-1},0),\quad (x_k,0),\quad \left(x_{k-1},\min_{x\in[x_{k-1},x_k]} f(x)\right),\quad \left(x_k,\max_{x\in[x_{k-1},x_k]} f(x)\right).$$

Thus the sum (5.2) is called the **trapezium sum**.

Definition 5.5 (trapezium sum)
Let $A, B \subset \mathbb{R}$, and let $[a,b] \subset A$. Let $f : A \to B$ be a continuous function. The ***mean value of the lower sum*** $L_n(f,[a,b])$ ***and the upper sum*** $U_n(f,[a,b])$ *is called the* ***trapezium sum*** *and is denoted by*

$$T_n(f,[a,b]) = \frac{b-a}{n}\sum_{k=1}^{n}\frac{1}{2}\left[\min_{x\in[x_{k-1},x_k]} f(x) + \max_{x\in[x_{k-1},x_k]} f(x)\right], \tag{5.3}$$

where as before

$$x_k = a + k\,\frac{b-a}{n}, \qquad k = 0,1,2,3,\ldots,n-1,n.$$

The **trapezium sum is an approximation of the area under the graph of** f **from** $x = a$ **to** $x = b$, and thus it is an approximation of the integral

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

If we increase n then the subintervals of $[a,b]$ become smaller and smaller, and the trapezium rule $T_n(f,[a,b])$ becomes a better and better approximation of the area under the graph, that is, of the integral $\int_a^b f(x)\,dx$. Thus we get the integral as the **limit** of the sequence $\{T_n(f,[a,b])\}$ for $n \to \infty$.

Definition 5.6 (formal definition of the integral)
Let $A, B \subset \mathbb{R}$, and let $[a,b] \subset A$. Let $f : A \to B$ be a continuous function. The ***integral of*** f ***over the interval*** $[a,b]$ *is defined as the limit*

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} T_n(f,[a,b]). \tag{5.4}$$

On the left-hand side of (5.4), we call $f(x)$ the ***integrand*** *and 'a' the* ***lower bound*** *and 'b' the* ***upper bound*** *of the integral.*

Of course working out integrals with Definition 5.6 is rather laborious, and in the following sections we will learn the integrals of some standard functions and rules to compute integrals of composite functions of such standard functions.

We observe that the trapezium rule $T_n(f,[a,b])$ can be easily programmed on a computer and gives for large n a fairly good basic approximation of the integral.

We will discuss two examples, where we use Definition 5.6 to compute the integral.

Example 5.7 (integral of a constant function)
Find the integral of the constant function $f:\mathbb{R}\to\mathbb{R}$, $f(x)=c$, where $c\in\mathbb{R}$ is a constant, over the interval $[0,1]$.

Solution: The n subintervals of $[0,1]$ of equal length have now length $(1-0)/n=1/n$ and are given by

$$\left[0,\frac{1}{n}\right],\quad \left[\frac{1}{n},\frac{2}{n}\right],\quad \ldots,\quad \left[\frac{(n-2)}{n},\frac{(n-1)}{n}\right],\quad \left[\frac{n-1}{n},1\right],$$

or equivalently

$$\left[\frac{k-1}{n},\frac{k}{n}\right],\qquad \text{for}\quad k=1,2,3,\ldots,n-1,n.$$

Since f is constant, we have

$$\min_{x\in\left[\frac{k-1}{n},\frac{k}{n}\right]} f(x)=c \qquad \text{and} \qquad \max_{x\in\left[\frac{k-1}{n},\frac{k}{n}\right]} f(x)=c \qquad \text{for}\quad k=1,2,3,\ldots,n-1,n.$$

Thus we have that

$$\frac{1}{2}\left[\min_{x\in\left[\frac{k-1}{n},\frac{k}{n}\right]} f(x)+\max_{x\in\left[\frac{k-1}{n},\frac{k}{n}\right]} f(x)\right]=\frac{1}{2}\left[c+c\right]=c \qquad \text{for}\quad k=1,2,3,\ldots,n-1,n,$$

and the trapezium sum is given by

$$T_n(f,[0,1])=\frac{1}{n}\sum_{k=1}^{n}\frac{1}{2}\left[\min_{x\in\left[\frac{k-1}{n},\frac{k}{n}\right]} f(x)+\max_{x\in\left[\frac{k-1}{n},\frac{k}{n}\right]} f(x)\right]=\frac{1}{n}\sum_{k=1}^{n}c=\frac{1}{n}\,n\,c=c.$$

This implies that the integral is given by

$$\int_0^1 f(x)\,dx=\int_0^1 c\,dx=\lim_{n\to\infty}T_n(f,[0,1])=\lim_{n\to\infty}c=c.$$

Of course, this is what we expected since the area under graph is the rectangle with corners $(0,0)$, $(1,0)$, $(1,c)$, and $(0,c)$, which has the area $(1-0)\times c=c$. $\square$

Example 5.8 (integral of an affine linear function)
Find the integral of the affine linear function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x$, over the interval $[0, 1]$.

Solution: The n subintervals of $[0, 1]$ of equal length have now length $(1-0)/n = 1/n$ and are given by

$$\left[0, \frac{1}{n}\right], \quad \left[\frac{1}{n}, \frac{2}{n}\right], \quad \ldots, \quad \left[\frac{(n-2)}{n}, \frac{(n-1)}{n}\right], \quad \left[\frac{n-1}{n}, 1\right],$$

or equivalently

$$\left[\frac{k-1}{n}, \frac{k}{n}\right], \qquad \text{for} \quad k = 1, 2, 3, \ldots, n-1, n.$$

Since $f(x) = x$ is strictly monotonically increasing on $\mathbb{R}$ (that is, $f(x_1) < f(x_2)$ if $x_1 < x_2$) we have for any interval $[c, d]$, that

$$\min_{x\in[c,d]} f(x) = \min_{x\in[c,d]} x = c \qquad \text{and} \qquad \max_{x\in[c,d]} f(x) = \max_{x\in[c,d]} x = d,$$

and thus for all $k = 1, 2, 3, \ldots, n-1, n$

$$\min_{x\in\left[\frac{k-1}{n},\frac{k}{n}\right]} f(x) = \min_{x\in\left[\frac{k-1}{n},\frac{k}{n}\right]} x = \frac{k-1}{n} \qquad \text{and} \qquad \max_{x\in\left[\frac{k-1}{n},\frac{k}{n}\right]} f(x) = \max_{x\in\left[\frac{k-1}{n},\frac{k}{n}\right]} x = \frac{k}{n}.$$

Thus we have that for all $k = 1, 2, 3, \ldots, n-1, n$,

$$\frac{1}{2}\left[\min_{x\in\left[\frac{k-1}{n},\frac{k}{n}\right]} f(x) + \max_{x\in\left[\frac{k-1}{n},\frac{k}{n}\right]} f(x)\right] = \frac{1}{2}\left[\frac{k-1}{n} + \frac{k}{n}\right] = \frac{2k-1}{2n}.$$

Thus the trapezium sum is given by

$$\begin{aligned} T_n(f, [0, 1]) &= \frac{1}{n}\sum_{k=1}^{n} \frac{1}{2}\left[\min_{x\in\left[\frac{k-1}{n},\frac{k}{n}\right]} f(x) + \max_{x\in\left[\frac{k-1}{n},\frac{k}{n}\right]} f(x)\right] \\ &= \frac{1}{n}\sum_{k=1}^{n} \frac{2k-1}{2n} = \frac{1}{2\,n^2}\sum_{k=1}^{n}(2k-1) = \frac{1}{2\,n^2}\left(2\sum_{k=1}^{n} k - \sum_{k=1}^{n} 1\right). \end{aligned}$$

Since

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

we find that

$$T_n(f, [0, 1]) = \frac{1}{2\,n^2}\left(2\sum_{k=1}^{n} k - \sum_{k=1}^{n} 1\right) = \frac{1}{2\,n^2}\big[n(n+1) - n\big] = \frac{1}{2\,n^2}\,n^2 = \frac{1}{2}.$$

Thus the integral is given by

$$\int_0^1 f(x)\,dx = \int_0^1 x\,dx = \lim_{n\to\infty} T_n(f,[0,1]) = \lim_{n\to\infty} \frac{1}{2} = \frac{1}{2}.$$

Of course, this is what we expected since the area under graph is the area of the triangle with corners $(0,0)$, $(1,0)$, and $(1,1)$, and this area is $(1\times 1)/2 = 1/2$. $\square$

5.2 Primitives, the Fundamental Theorem of Calculus, and Indefinite Integrals

In this section we will establish a **connection between integration and differentiation**. We will see that taking a so-called indefinite integral and differentiating are **'reverse operations' to each other**.

Definition 5.9 (primitive/antiderivative)
Let $A, B \subset \mathbb{R}$, and let $f : A \to B$ be a continuous function. A continuous function $F : A \to B$ with

$$\frac{dF(x)}{dx} = F'(x) = f(x) \qquad \text{for all } x \in A$$

*is called a **primitive of** f or **antiderivative of** f.*

We consider some examples.

Example 5.10 (primitives/antiderivatives of functions)

(a) The function $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = 2\,x$, has the primitive/antiderivative $F(x) = x^2$. Indeed,

$$F'(x) = (x^2)' = 2\,x = f(x) \qquad \text{for all } x \in \mathbb{R}.$$

We note that, since the derivative of a constant function is zero, the function $\widetilde{F}(x) = x^2 + 3$ is also a primitive/antiderivative for $f(x) = x^2$.

(b) A primitive/antiderivative of the function $f : (0,\infty) \to \mathbb{R}$, $f(x) = 1/x$, is given by the function $F : (0,\infty) \to \mathbb{R}$, $F(x) = \ln(x)$. Indeed,

$$F'(x) = \big(\ln(x)\big)' = \frac{1}{x} = f(x) \qquad \text{for all } x \in (0,\infty).$$

Since for any constant function $g(x) = c$ we have $g'(x) = 0$, the function $\widetilde{F} : (0, \infty) \to \mathbb{R}$, $\widetilde{F}(x) = \ln(x) + c$, where $c \in \mathbb{R}$ is an arbitrary constant, is also a primitive/antiderivative of $f(x) = 1/x$. □

Lemma 5.11 (primitives of the same function differ by a constant)
Let $A, B \subset \mathbb{R}$. Let $f : A \to B$ be a continuous function, and let $F : A \to B$ be a primitive/antiderivative of f. Then, for every constant $c \in \mathbb{R}$, the function

$$\widetilde{F} : A \to B, \qquad \widetilde{F}(x) = F(x) + c, \tag{5.5}$$

is also a primitive/antiderivative of f. ***The difference of any two primitives/antiderivatives of f is a constant function.***

Proof: Differentiating (5.5) yields

$$\widetilde{F}'(x) = \big(F(x) + c\big)' = F'(x) + 0 = F'(x) = f(x),$$

and thus $\widetilde{F}$ is indeed a primitive/antiderivative of f.

It remains to show that the difference of any two primitives/antiderivatives of f is a constant function. Let F and $\widetilde{F}$ be two primitives/antiderivative of f. Then

$$F'(x) = f(x) \qquad \text{and} \qquad \widetilde{F}'(x) = f(x)$$

imply, by subtracting the second equation from the first, that

$$\big(F(x) - \widetilde{F}(x)\big)' = F'(x) - \widetilde{F}'(x) = f(x) - f(x) = 0.$$

Since the only functions whose derivatives are zero are constant functions we can conclude that there exists some constant $c \in \mathbb{R}$ such that

$$F(x) - \widetilde{F}(x) = c \qquad \text{for all} \quad x \in A.$$

Thus we see that any two primitives/antiderivative of f differ only by a constant function, as claimed. □

We observe that Lemma 5.11 implies that **given one primitive/antiderivative of f, we can obtain any other primitive of f by adding a suitable constant**.

We can obtain some elementary rules for primitives/antiderivatives of functions from the properties of the derivative.

Lemma 5.12 (elementary properties of the primitive/antiderivative)
Let $A, B \subset \mathbb{R}$, and let $\alpha \in \mathbb{R}$. Let $f : A \to B$ and $g : A \to B$ be continuous functions and let $F : A \to B$ and $G : A \to B$ be primitives/antiderivatives of f and g, respectively. Then the function $F(x) + G(x)$ is a primitive of $f(x) + g(x)$, and the function $\alpha\, F(x)$ is a primitive of $\alpha\, f(x)$.

Proof: From the properties of the derivative

$$\big(F(x)+G(x)\big)' = F'(x)+G'(x) = f(x)+g(x) \quad \text{and} \quad \big(\alpha\, F(x)\big)' = \alpha\, F'(x) = \alpha\, f(x),$$

which implies that $F + G$ is a primitive/antiderivative of $f + g$ and $\alpha\, F$ is a primitive/antiderivative of $\alpha\, f$. □

We note that (in the notation of Lemma 5.12) the combination of both properties yields

$$\boxed{\big(\alpha\, F(x) + \beta\, G(x)\big)' = \alpha\, F'(x) + \beta\, G'(x) = \alpha\, f(x) + \beta\, g(x),}$$

that is, $\alpha\, F + \beta\, G$ is a primitive/antiderivative of $\alpha\, f + \beta\, g$. We could also have the sum of more than two functions.

We apply this new knowledge to find some primitives/antiderivatives.

Example 5.13 (primitives of sums and multiples of functions)
Determine a primitive/antiderivative of the function $f : \mathbb{R} \to \mathbb{R}$, defined by

$$f(x) = 2\, e^x + 5\, \cos(x).$$

Solution: We know that $(e^x)' = e^x$ and $\big(\sin(x)\big)' = \cos(x)$. Thus we have that a primitive/antiderivative of f is given by

$$F : \mathbb{R} \to \mathbb{R}, \qquad F(x) = 2\, e^x + 5\, \sin(x),$$

where we have used Lemma 5.12. □

Example 5.14 (primitives of sums and multiples of functions)
Find a primitive/antiderivative of the function $f : (-\pi/2, \pi/2) \to \mathbb{R}$, defined by

$$f(x) = \frac{7}{\big(\cos(x)\big)^2} - \frac{\cosh(x)}{4}.$$

Solution: We know that $\big(\tan(x)\big)' = \big(\cos(x)\big)^{-2}$ and $\big(\sinh(x)\big)' = \cosh(x)$, and thus a primitive/antiderivative of f is given by

$$F : (-\pi/2, \pi/2) \to \mathbb{R}, \qquad F(x) = 7\,\tan(x) - \frac{\sinh(x)}{4},$$

where we have used Lemma 5.12. □

Now we can formulate the fundamental theorem of calculus.

Theorem 5.15 (fundamental theorem of calculus)
Let $A, B \subset \mathbb{R}$. Let $f : A \to B$ be a continuous function, and let $F : A \to B$ be a primitive/antiderivative of f. Then for any interval $[a, b] \subset A$

$$\int_a^b f(x)\,dx = F(b) - F(a). \tag{5.6}$$

Often we find also the abbreviated notation $F(x)|_a^b$, defined by

$$F(x)|_a^b = F(b) - F(a),$$

instead of $F(b) - F(a)$ on the right-hand side of (5.6).

The fundamental theorem of calculus is of paramount importance, because it links integration and differentiation!

We show in two examples how to make use of the fundamental theorem of calculus.

Example 5.16 (application of the fundamental theorem of calculus)
Evaluate the integral

$$\int_1^2 \frac{1}{x}\,dx.$$

Solution: Since $\big(\ln(x)\big)' = 1/x$, the function $F : (0, \infty) \to \mathbb{R}$, $F(x) = \ln(x)$, is a primitive/antiderivative of $f : (0, \infty) \to \mathbb{R}$, $f(x) = 1/x$. Thus from (5.6) in the fundamental theorem of calculus

$$\int_1^2 \frac{1}{x}\,dx = \ln(x)|_1^2 = \ln(2) - \ln(1) = \ln(2).$$ □

Example 5.17 (application of the fundamental theorem of calculus)
Evaluate the integral

$$\int_0^\pi \cos(x)\,dx.$$

Solution: Since $\big(\sin(x)\big)' = \cos(x)$, the function $F : \mathbb{R} \to \mathbb{R}$, $F(x) = \sin(x)$, is a primitive/antiderivative of $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \cos(x)$. Thus from (5.6) in the fundamental theorem of calculus

$$\int_0^\pi \cos(x)\,dx = \sin(x)|_0^\pi = \sin(\pi) - \sin(0) = 0 - 0 = 0.$$

Should we be surprised that the integral is zero? The answer is no! In the graph the same area lies for $x \in [\pi/2, \pi]$ below the x-axis and above the graph of $\cos(x)$ that lies for $x \in [0, \pi/2]$ above the x-axis an below the graph of $\cos(x)$. Thus negative and positive area cancel each other out, and the integral is zero. □

We point out here that it **does not matter which primitive/antiderivative F of f is used on the right-hand side of (5.6)**. If we replace F in (5.6) by another primitive/antiderivative $\widetilde{F}$ of f, then we know by Lemma 5.11 that $\widetilde{F}(x) = F(x) + c$ with some constant $c \in \mathbb{R}$. Thus, from (5.6),

$$\int_a^b f(x)\,dx = \widetilde{F}(b) - \widetilde{F}(a) = \big(F(b) + c\big) - \big(F(b) - c\big) = F(b) - F(a).$$

Before we continue, we introduce the following **convention of notation**: for a continuous function $f : A \to B$ and $[a, b] \subset A$, we define

$$\boxed{\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx.} \tag{5.7}$$

If we replace in (5.6) the variable over which we integrate by t and subsequently replace the upper bound b by x and the lower bound a by x_0, then (5.6) reads

$$\int_{x_0}^x f(t)\,dt = F(x) - F(x_0) \qquad \Rightarrow \qquad F(x_0) + \int_{x_0}^x f(t)\,dt = F(x),$$

and we see that the primitive/antiderivative F of f can be represented as

$$F(x) = F(x_0) + \int_{x_0}^x f(t)\,dt. \tag{5.8}$$

Since adding of any constant to F yields another primitive/antiderivative of f we may add $-F(x_0)$ to (5.8) and get the **primitive/antiderivative**

$$F_0(x) = \int_{x_0}^x f(t)\,dt$$

of f. This leads to the definition of the so-called **indefinite integral** of a function which gives us a primitive/antiderivative of f.

Definition 5.18 (indefinite integral)
Let $f : (a, b) \to \mathbb{R}$ be a continuous function. For every fixed $x_0 \in (a, b)$, the function $F_0 : (a, b) \to \mathbb{R}$

$$F_0(x) = \int_{x_0}^{x} f(t)\, dt, \tag{5.9}$$

*is called an **indefinite integral** of f.*

We note that in Definition 5.18 we need the notation (5.7), since we can have in (5.9) that $x_0 > x$.

Lemma 5.19 (indefinite integral defines primitive/antiderivative)
Let $f : (a, b) \to \mathbb{R}$ be a continuous function. For every fixed $x_0 \in (a, b)$, the indefinite integral $F_0 : (a, b) \to \mathbb{R}$

$$F_0(x) = \int_{x_0}^{x} f(t)\, dt,$$

*is a **primitive/antiderivative** of f.*

Later-on when we will have learnt some techniques for computing integrals, we will make use of Lemma 5.19 to compute primitives of composite functions.

We give an elementary example of the use of Lemma 5.19.

Example 5.20 (indefinite integral)
All primitives/antiderivatives of the function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 3\, e^x$, are given by

$$F : \mathbb{R} \to \mathbb{R}, \qquad F(x) = \int_0^x 3\, e^t\, dt + c,$$

where $c \in \mathbb{R}$ is an arbitrary constant. Since $(3\, e^t)' = 3\, e^t$, we have from the (5.6) that

$$F(x) = \int_0^x 3\, e^t\, dt = 3\, e^t|_0^x + c = 3\, e^x - 3\, e^0 + c = 3\, e^x + (c - 3) = 3\, e^x + \widetilde{c},$$

with an arbitrary constant $\widetilde{c} = c - 3$. Of course, the fact $(3\, e^t)' = 3\, e^t$, tells us already that $3\, e^x$, and thus also $3\, e^x + \widetilde{c}$ is a primitive of f. □

5.3 Standard Integrals

In the last section we have seen that a primitive/antiderivative $F : A \to B$ of a continuous function $f : A \to B$ allows us to evaluate an integral with the help of the **fundamental theorem of calculus** (see Theorem 5.15), which says that, for any $[a, b] \subset A$, we have

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

We have further learnt (see Lemma 5.19 above) that for any $x_0 \in A$ the **indefinite integral**

$$F_0(x) = \int_{x_0}^x f(t)\,dt$$

defines a primitive/antiderivative of f, and thus, from Lemma 5.11, **any primitive/antiderivative of f is of the form**

$$F(x) = F_0(x) + c = \int_{x_0}^x f(t)\,dt + c.$$

Thus the knowledge of the primitives/antiderivatives of standard functions allows us to evaluate integrals. In Chapter 3, we have discussed derivatives and learned some standard functions f and their derivatives f'. **The function f is a primitive/antiderivative of its derivative** f', and thus we can use Table 3.1 from Chapter 3 to establish a table of functions and their primitives/antiderivatives (see Table 5.1).

Note that you are expected to **know the primitives/antiderivatives in Table 5.1 from memory!** They will **not** be provided in the final exam.

For example we know that $(x^{p+1})' = (p+1)\,x^p$, and dividing by $p+1$, we find that

$$\left(\frac{x^{p+1}}{p+1}\right)' = x^p \qquad \Rightarrow \qquad \frac{x^{p+1}}{p+1} \quad \text{is a primitive/antiderivative for} \quad x^p.$$

From $(\cos(x))' = -\sin(x)$ and thus $(-\cos(x))' = \sin(x)$, we see that $-\cos(x)$ is a primitive/antiderivative of $\sin(x)$.

Function $f(x)$	Primitive/Antiderivative $F(x)$
0	$c = \text{constant}$
x^n (with $n \neq -1$)	$\dfrac{x^{n+1}}{n+1} + c$
e^x	$e^x + c$
$\dfrac{1}{x}$	$\ln(\lvert x\rvert) + c$
$\sin(x)$	$-\cos(x) + c$
$\cos(x)$	$\sin(x) + c$
$\dfrac{1}{\big(\cos(x)\big)^2} = \big(\sec(x)\big)^2$	$\tan(x) + c$
$\dfrac{1}{\big(\sin(x)\big)^2} = \big(\operatorname{cosec}(x)\big)^2$	$-\cot(x) + c$
$\sinh(x)$	$\cosh(x) + c$
$\cosh(x)$	$\sinh(x) + c$
$\dfrac{1}{\big(\cosh(x)\big)^2} = \big(\operatorname{sech}(x)\big)^2$	$\tanh(x) + c$
$\dfrac{1}{\big(\sinh(x)\big)^2} = \big(\operatorname{cosech}(x)\big)^2$	$-\coth(x) + c$

Table 5.1: Important primitives/antiderivatives.

We give two examples in which we compute integrals with the help of Table 5.1 and the fundamental theorem of calculus (see Theorem 5.15).

Example 5.21 (application of fundamental theorem of calculus)
Since $\big(-\cos(x)\big)' = \sin(x)$, we have

$$\int_0^{2\pi} \sin(x)\,dx = -\cos(x)|_0^{2\pi} = -\cos(2\pi) - \big(-\cos(0)\big) = -1 - (-1) = 0,$$

where we have used $\sin(0) = 1$ and $\sin(2\pi) = 1$. That the integral is zero is not surprising, since the positive and negative area under the graph of $\sin(x)$ for $x \in [0, 2\pi]$ just cancel each other out. □

Example 5.22 (application of fundamental theorem of calculus)
Since $(x^4/4)' = x^3$, we have from the fundamental theorem of calculus (see Theorem 5.15) that

$$\int_1^4 x^3\,dx = \left.\frac{x^4}{4}\right|_1^4 = \frac{4^4}{4} - \frac{1^4}{4} = 64 - \frac{1}{4} = \frac{255}{4}.$$ □

5.4 Elementary Properties of the Integral

The most basic but nevertheless quite important properties of the integral are the so-called **linear properties**.

Lemma 5.23 (linear properties of the integral)
Let $A, B \subset \mathbb{R}$, and let $[a, b] \subset A$. Let $f : A \to B$ and $g : A \to B$ be continuous functions, and let $\alpha \in \mathbb{R}$. Then

$$\int_a^b \big(f(x) + g(x)\big)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx. \tag{5.10}$$

and

$$\int_a^b \alpha\, f(x)\,dx = \alpha \int_a^b f(x)\,dx. \tag{5.11}$$

Proof: This can be easily verified with the properties of the derivative and the fundamental theorem of calculus (see Theorem 5.15) as follows. Let F be a primitive/antiderivative of f and let G be a primitive/antiderivative of g. Then $(F+G)' = F' + G' = f + g$ and $(\alpha\, F)' = \alpha\, F' = \alpha\, f$ from the properties of the derivative, and thus $F+G$ is a primitive/antiderivative of $f+g$ and $\alpha\, F$ is a primitive/antiderivative of $\alpha\, f$. Now the left-hand side and the right-hand side of (5.10) can be evaluated with the fundamental theorem of calculus (see Theorem 5.15).

$$\int_a^b \big(f(x)+g(x)\big)\,dx = \big(F(b)+G(b)\big)-\big(F(a)+G(a)\big) = \big(F(b)-F(a)\big)+\big(G(b)-G(a)\big),$$

and

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

We see that the left-hand side and the right-hand side of (5.10) do indeed coincide. Likewise, we have for (5.11)

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

and we see that (5.11) holds true. □

We note that (5.10) and (5.11) are intuitively clear if we think of the geometric interpretation of the integral $\int_a^b f(x)\, dx$ as the area under the graph of f for $x \in [a, b]$. We discuss some examples.

Example 5.24 (linear properties of the integral)
Compute the integral

$$\int_1^e \left(3\, x^2 - 4\, x + \frac{3}{x}\right) dx.$$

Solution: We use the linear properties of the integral and obtain

$$\int_1^e \left(3\, x^2 - 4\, x + \frac{3}{x}\right) dx = 3 \int_1^e x^2\, dx - 4 \int_1^e x\, dx + 3 \int_1^e \frac{1}{x}\, dx.$$

Since $(x^3/3) = x^2$, $(x^2/2)' = x$ and $\big(\ln(|x|)\big)' = 1/x$, it follows from the fundamental theorem of calculus (see Theorem 5.15) that

$$\begin{aligned} 3 \int_1^e x^2\, dx - 4 \int_1^e x\, dx + 3 \int_1^e \frac{1}{x}\, dx \;&=\; 3\, \frac{x^3}{3}\bigg|_1^e - 4\, \frac{x^2}{2}\bigg|_1^e + 3\, \ln(|x|)\big|_1^e \\ =\; \big(x^3 - 2\, x^2 + 3\, \ln(|x|)\big)\big|_1^e \;&=\; \big(e^3 - 2\, e^2 + 3\, \ln(e)\big) - \big(1 - 2 + 3\, \ln(1)\big) \\ =\; e^3 - 2\, e^2 + 3 + 1 \;&=\; e^3 - 2\, e^2 + 4, \end{aligned}$$

where we have used that $\ln(1) = 0$ and $\ln(e) = 1$. Thus we find that

$$\int_1^e \left(3\, x^2 - 4\, x + \frac{3}{x}\right) dx = e^3 - 2\, e^2 + 4.$$ □

We note here that we can alternatively use Lemma 5.12 to work out a primitive/antiderivative of the integrand first and subsequently use the fundamental theorem of calculus (see Theorem 5.15). A primitive/antiderivative of the integrand $f(x) = 3\, x^2 - 4\, x + 3/x$ is given by $F(x) = x^3 - 2\, x^2 + 3\, \ln(|x|)$, where we have used Lemma 5.12 and the fact that $(x^3/3) = x^2$, $(x^2/2)' = x$ and $\big(\ln(|x|)\big)' = 1/x$. From the fundamental theorem of calculus (see Theorem 5.15) we obtain now that

$$\int_1^e \left(3\, x^2 - 4\, x + \frac{3}{x}\right) dx = \big[x^3 - 2\, x^2 + 3\, \ln(|x|)\big]\big|_1^e,$$

and simplifying of the right-hand side yields the same result that we obtained in the previous example.

Example 5.25 (linear properties of the integral)
For $x > 0$, evaluate the indefinite integral

$$\int_1^x \frac{t^2 - t}{t^2}\, dt.$$

Solution: In this case is it not directly possible to determine a primitive/antiderivative of the integrand with the help of Table 5.1. Therefore we try to simplify the integrand.

$$\frac{t^2 - t}{t^2} = \frac{t(t-1)}{t^2} = \frac{t-1}{t} = 1 - \frac{1}{t},$$

and we know a primitive/antiderivative of each of the two individual terms. From $(t)' = 1$ and $\big(\ln(|t|)\big)' = 1/t$ we get

$$\int_1^x \frac{t^2 - t}{t^2}\, dt = \int_1^x \left(1 - \frac{1}{t}\right) dt = \int_1^x 1\, dt - \int_1^x \frac{1}{t}\, dt = t|_1^x - \ln(|t|)|_1^x = x - 1 - \ln(x),$$

where we have used $\ln(1) = 0$ in the last step. Since the indefinite integral defines a primitive/antiderivative of the integrand, we know that a primitive/antiderivative of $f(x) = (x^2 - x)/x^2$ is given by $F(x) = x - 1 - \ln(x)$, where $x > 0$. □

Another very useful elementary property of the integral is the **domain splitting property**.

Lemma 5.26 (domain splitting property)
Let $f : [a, c] \to \mathbb{R}$ be a continuous function, and let $a < b < c$. Then

$$\int_a^c f(x)\, dx = \int_a^b f(x)\, dx + \int_b^c f(x)\, dx. \tag{5.12}$$

The domain splitting property (5.12) will be verified in one problem on Exercise Sheet 5, and it follows relatively straight-forward from the geometric definition of the integral as the area under the graph. We illustrate the use of the domain splitting property with an example.

Example 5.27 (application of domain splitting property)
Compute the integral

$$\int_{-1}^1 |x|\, dx$$

of the absolute value function $f(x) = |x|$ over the interval $[-1, 1]$.

Solution: From the definition of the absolute value $|x|$ (by $|x| = -x$ for $x < 0$ and $|x| = x$ for $x \geq 0$) and the domain splitting property (5.12), we have

$$\int_{-1}^{1} |x|\,dx = \int_{-1}^{0} |x|\,dx + \int_{0}^{1} |x|\,dx = \int_{-1}^{0} (-x)\,dx + \int_{0}^{1} x\,dx = -\int_{-1}^{0} x\,dx + \int_{0}^{1} x\,dx.$$

Since $(x^2/2)' = x$, we find

$$\begin{aligned}\int_{-1}^{1} |x|\,dx &= -\int_{-1}^{0} x\,dx + \int_{0}^{1} x\,dx. \\ &= -\left.\frac{x^2}{2}\right|_{-1}^{0} + \left.\frac{x^2}{2}\right|_{0}^{1} = -\left(0 - \frac{(-1)^2}{2}\right) + \left(\frac{1^2}{2} - 0\right) = \frac{1}{2} + \frac{1}{2} = 1.\end{aligned}$$

From sketching the graph of $f(x) = |x|$, we can see that 1 is indeed the correct area, since the sum of the two triangles given by the area under the graph is just a square with sides of length 1. □

Remark 5.28 (integral is non-negative for non-negative function)
Another useful property of the integral is that $f(x) \geq 0$ for all $x \in [a, b]$ implies

$$\int_{a}^{b} f(x)\,dx \geq 0.$$

This is intuitively clear from the interpretation of the integral as the area under the graph.

In the next two sections we will learn two important methods for evaluating integrals: integration by substitution and integration by parts.

5.5 Integration by Substitution

Before we introduce integration by parts we start with two motivating examples.

Example 5.29 (primitive of $4\,e^{4x}$)
To find a primitive of $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 4\,e^{4x}$, we observe that from the chain the chain rule $(e^{4x})' = 4\,e^{4x}$ and thus the function

$$F : \mathbb{R} \to \mathbb{R}, \qquad F(x) = e^{4x}$$

is a primitive/antiderivative of $f(x) = 4\,e^{4x}$. Thus the integral of $f(x) = 4\,e^{4x}$ over an interval $[a, b]$ can be evaluated as follows

$$\int_a^b 4\,e^{4x}\,dx = e^{4x}\big|_a^b = e^{4b} - e^{4a}, \tag{5.13}$$

where we have used the fundamental theorem of calculus (see Theorem 5.15). □

Example 5.30 (primitive of $2\,x\,e^{x^2}$)
We want to compute the integral

$$\int_0^b 2\,x\,e^{x^2}\,dx.$$

We cannot determine a primitive/antiderivative directly from Table 5.1, but inspection of the the integrand with the chain rule in mind leads us to guess that $F(x) = e^{x^2}$ is a primitive/antiderivative of $f(x) = 2\,x\,e^{x^2}$, which is indeed true: from the chain rule

$$F'(x) = \big(e^{x^2}\big)' = e^{x^2}\,2\,x = 2\,x\,e^{x^2} = f(x).$$

Thus we obtain that

$$\int_0^b 2\,x\,e^{x^2}\,dx = e^{x^2}\big|_0^b = e^{b^2} - 1, \tag{5.14}$$

where we have used the fundamental theorem of calculus (see Theorem 5.15). □

For finding a primitive/antiderivative of the integrand in each of the integrals (5.13) and (5.14), we have used the **chain rule** (see Lemma 3.23) for differentiating functions of the form $H : A \to B$, $H(x) = (F \circ g)(x) = F\big(g(x)\big)$. The chain rule says that the derivative of the composite function H is given by

$$H'(x) = \big[F\big(g(x)\big)\big]' = F'\big(g(x)\big)\,g'(x) = f\big(g(x)\big)\,g'(x) \qquad \text{with} \quad f(y) = F'(y). \tag{5.15}$$

In Example 5.30, we have in this notation $F(y) = e^y$, $g(x) = x^2$ and $g'(x) = 2\,x$, and $F'(y) = f(y) = e^y$. From (5.15), we can derive the following general formula by integrating (5.15) over an interval $[a, b] \subset A$ and using the fundamental theorem of calculus (see Theorem 5.15).

$$F\big(g(x)\big)\big|_a^b = \int_a^b \big[F\big(g(x)\big)\big]'\,dx = \int_a^b f\big(g(x)\big)\,g'(x)\,dx, \tag{5.16}$$

and we can rewrite the left-hand side as

$$F\big(g(x)\big)\big|_a^b = F\big(g(b)\big) - F\big(g(a)\big) = F(y)\big|_{g(a)}^{g(b)} = \int_{g(a)}^{g(b)} f(y)\,dy. \tag{5.17}$$

In the last step we have used the fundamental theorem of calculus (see Theorem 5.15) and the fact that F is a primitive/antiderivative of f. Combining (5.16) and (5.17), yields

$$\int_{g(a)}^{g(b)} f(y)\,dy = \int_a^b f\big(g(x)\big)\,g'(x)\,dx, \tag{5.18}$$

which is the formula for **integration by substitution**.

Theorem 5.31 (integration by substitution)
Let $f : B \to C$ be a continuous function and let $g : A \to B$ be a sufficiently smooth function. Then for any interval $[a, b] \subset A$,

$$\int_{g(a)}^{g(b)} f(y)\,dy = \int_a^b f\big(g(x)\big)\,g'(x)\,dx. \tag{5.19}$$

If $F : B \to C$ is a primitive of f, then the left-hand side can be expressed as

$$\int_{g(a)}^{g(b)} f(y)\,dy = F(y)\Big|_{g(a)}^{g(b)} = F\big(g(b)\big) - F\big(g(a)\big).$$

In Theorem 5.31 above 'sufficiently smooth' means that g is continuous and that it has a derivative g' which is continuous.

How is integration by substitution performed in practice?

As we already saw in the examples discussed as motivation, we will very often start with an integral that can be interpreted as the right-hand side of (5.19). If this is the case, we set $y = y(x) = g(x)$, but we also need to substitute $g'(x)\,dx$ and the boundaries of the integral. To do this we differentiate $y = y(x) = g(x)$ with respect to x, that is,

$$\frac{dy}{dx} = \frac{dg(x)}{dx} = g'(x) \qquad \Rightarrow \qquad dy = g'(x)\,dx,$$

and observe that the new boundaries are $y(a) = g(a)$ and $y(b) = g(b)$. Thus

$$\int_a^b f\big(\underbrace{g(x)}_{=y}\big)\,\underbrace{g'(x)\,dx}_{=dy} = \int_{g(a)}^{g(b)} f(y)\,dy. \tag{5.20}$$

If we know a primitive of f then we can now easily evaluate the integral on the right-hand side of (5.20).

If we want to start with the left-hand side of (5.19), then the procedure is similar. We have

$$\int_c^d f(y)\,dy,$$

but we cannot find a primitive of f. Thus we try to get the integrand in a more convenient form with a suitable substitution $y = y(x) = g(x)$, where we now assume g to be one-to-one so that it has an inverse g^{-1}. Then the substitution

$$y = y(x) = g(x) \qquad \Leftrightarrow \qquad x = x(y) = g^{-1}(y),$$

and, from differentiating,

$$\frac{dy}{dx} = \frac{dg(x)}{dx} = g'(x) \qquad \Rightarrow \qquad dy = g'(x)\,dx,$$

yield

$$\int_c^d f(y)\,dy = \int_{g^{-1}(c)}^{g^{-1}(d)} f\big(g(x)\big)\,g'(x)\,dx,$$

where the new boundaries $x(c) = g^{-1}(c)$ an $x(d) = g^{-1}(d)$ have been determined with the inverse function $x = x(y) = g^{-1}(y)$.

We observe here that, of course, we can also **use integration by substitution to find a primitive of a function f by evaluating an indefinite integral of f.**

Now we will discuss many examples. In doing so we will encounter different types of integrals that can be evaluated with integration by substitution. The trick is to **find a suitable substitution** and this is largely **practice and experience**. For this reason it is essential that you attempt as many problems from the exercise sheets as possible. If you are already confident about the new methods it is also a good idea to not just read the examples in the lecture notes but instead to attempt to solve them yourself before you look at the provided solution.

Example 5.32 (integration by substitution)
Evaluate the integral

$$\int_1^3 \frac{1}{2\,x-1}\,dx.$$

Solution: Since $1/y$ is the derivative of the (natural) logarithm $\ln(|y|)$, we choose the substitution

$$y = y(x) = 2\,x - 1 \qquad \Rightarrow \qquad \frac{dy}{dx} = 2 \qquad \Rightarrow \qquad dx = \frac{1}{2}\,dy,$$

which yields the new boundaries $y(1) = 2 \times 1 - 1 = 1$ and $y(3) = 2 \times 3 - 1 = 5$. Performing the substitution in the integral yields

$$\int_1^3 \frac{1}{2\,x-1}\,dx = \int_1^5 \frac{1}{y}\,\frac{1}{2}\,dy = \frac{1}{2}\,\ln(|y|)\big|_1^5 = \frac{\ln(5)-\ln(1)}{2} = \frac{\ln(5)}{2}. \qquad \square$$

Example 5.33 (real powers of affine linear functions)
Find all primitives/antiderivatives to the functions

(a) $f(x) = (3\,x+5)^7$, where $x > -\dfrac{5}{3}$; (b) $h(x) = (1+2x)^{-1/4}$, where $x > -\dfrac{1}{2}$.

Solution: Since an indefinite integral defines a primitive/antiderivative, we determine

$$\text{(a)} \quad \int_0^x (3\,t+5)^7\,dt \qquad\qquad \text{(b)} \quad \int_0^x (1+2\,t)^{-1/4}\,dt,$$

where it should be noted that the choice of the lower boundary value of the integrals is arbitrary. To evaluate the integral in (a), we choose the substitution

$$y = y(t) = 3\,t+5 \qquad \Rightarrow \qquad \frac{dy}{dt} = 3 \qquad \Rightarrow \qquad dt = \frac{1}{3}\,dy,$$

and obtain new boundaries $y(0) = 3\times 0+5 = 5$ and $y(x) = 3\,x+5$. Thus we obtain from integration by substitution and $(y^8)' = 8\,y^7$

$$\begin{aligned}\int_0^x (\underbrace{3\,t+5}_{=y})^7 \underbrace{dt}_{=(1/3)\,dy} &= \int_5^{3x+5} y^7\,\frac{1}{3}\,dy = \frac{1}{24}\int_5^{3x+5} 8\,y^7\,dy \\ &= \frac{y^8}{24}\bigg|_5^{3x+5} = \frac{(3\,x+5)^8}{24} - \frac{5^8}{24}.\end{aligned}$$

Since all primitives/antiderivatives of f differ only by constants we see that all primitives of f are of the form

$$F(x) = \frac{(3\,x+5)^8}{24} + c, \qquad \text{with a constant } \; c \in \mathbb{R}.$$

To evaluate the second integral, we set

$$y = y(t) = 1+2\,t \qquad \Rightarrow \qquad \frac{dy}{dt} = 2 \qquad \Rightarrow \qquad dt = \frac{1}{2}\,dy,$$

and the new boundaries of the integral are $y(0) = 1+2\times 0 = 1$ and $y(x) = 1+2\,x$. Thus we obtain from $(y^{3/4})' = (3/4)\,y^{-1/4}$ and integration by substitution

$$\begin{aligned}\int_0^x (\underbrace{1+2\,t}_{=y})^{-1/4} \underbrace{dt}_{=(1/2)\,dy} &= \int_1^{1+2x} y^{-1/4}\,\frac{1}{2}\,dy = \frac{2}{3}\int_1^{1+2x} \frac{3}{4}\,y^{-1/4}\,dy \\ &= \frac{2}{3}\,y^{3/4}\Big|_1^{1+2x} = \frac{2}{3}\,(1+2\,x)^{3/4} - \frac{2}{3}.\end{aligned}$$

Since all primitives/antiderivatives of h differ only by constants we see that all primitives of h are of the form

$$H(x) = \frac{2}{3}\,(1+2\,x)^{3/4} + c, \qquad \text{with a constant } \; c \in \mathbb{R}. \qquad \square$$

Remark 5.34 (check your primitive/antiderivative by differentiating!)
After working out a primitive, it is a good idea to ***test you primitive by differentiating it****. For example, differentiating the answer in Example 5.33 (a) with the chain rule gives*

$$\frac{d}{dx}\left[\frac{(3\,x+5)^8}{24}+c\right]=\frac{8\,(3x+5)^7}{24}\,3=(3x+5)^7,$$

which is the integrand in Example 5.33 (a) as required.

Example 5.35 (integral of $\big(\sin(x)\big)^5\,\cos(x)$)
Show that

$$\int_0^{\pi/2}\big(\sin(x)\big)^5\,\cos(x)\,dx=\frac{1}{6}.$$

Solution: We substitute $y=y(x)=\sin(x)$. Then

$$\frac{dy}{dx}=\cos(x)\qquad\Rightarrow\qquad dy=\cos(x)\,dx$$

and $y(0)=\sin(0)=0$, $y(\pi/2)=\sin(\pi/2)=1$. Performing the substitution, we obtain

$$\int_0^{\pi/2}\big(\sin(x)\big)^5\,\cos(x)\,dx=\int_0^{\pi/2}\underbrace{\big(\sin(x)\big)^5}_{=y^5}\,\underbrace{\cos(x)\,dx}_{=dy}=\int_0^1 y^5\,dy=\frac{y^6}{6}\bigg|_0^1=\frac{1}{6},$$

as claimed. □

Remark 5.36 (integral boundaries can be dropped if we want a primitive)
If we want a primitive/antiderivative, then we may drop the lower and upper boundary of the integral, but we have to ***keep track of our substitutions and have to substitute 'backwards' after evaluation of the integral****!*

We demonstrate this for the function $f(x)=(3\,x+5)^7$ for which we want to find a primitive in Example 5.33 (a). If we write the integral without the lower and upper boundary and use as before the substitution

$$y=y(x)=3\,x+5,\qquad\frac{dy}{dx}=3\quad\Rightarrow\quad\frac{1}{3}\,dy=dx,$$

then we find that

$$\int(3\,x+5)^7\,dx=\int y^7\,\frac{1}{3}\,dy=\frac{1}{24}\int 8\,y^7\,dy=\frac{y^8}{24}+c=\frac{(3\,x+5)^8}{24}+c,$$

where it is ***crucial that we have substituted back*** $y=3\,x+5$ ***in the last step!*** *If we use the laxer notation for indefinite integrals without boundaries then we will* ***automatically add an (integration) constant*** $c\in\mathbb{R}$.

From now on we will also use the laxer notation of the indefinite integral without a lower and upper boundary, as explained in the last remark. We use this notation in the next examples.

Example 5.37 (primitives/antiderivatives of $\big(\sin(x)\big)^2 \cos(x)$)
Find all primitives/antiderivatives of the function $f : \mathbb{R} \to \mathbb{R}$, given by

$$f(x) = \big(\sin(x)\big)^2 \cos(x).$$

Solution: A primitive/antiderivative of the function $f(x) = \big(\sin(x)\big)^2 \cos(x)$ is given by the indefinite integral

$$\int \big(\sin(x)\big)^2 \cos(x)\, dx.$$

We use the substitution

$$y = y(x) = \sin(x), \qquad \frac{dy}{dx} = \cos(x) \quad \Rightarrow \quad dy = \cos(x)\, dx,$$

and get

$$\int \big(\sin(x)\big)^2 \cos(x)\, dx = \int \underbrace{\big(\sin(x)\big)^2}_{=y^2} \underbrace{\cos(x)\, dx}_{=dy} = \int y^2\, dy = \frac{y^3}{3} + c = \frac{1}{3}\big(\sin(x)\big)^3 + c.$$

Thus all primitives/antiderivatives of $f(x) = \big(\sin(x)\big)^2 \cos(x)$ are of the form

$$F : \mathbb{R} \to \mathbb{R}, \qquad F(x) = \frac{1}{3}\big(\sin(x)\big)^3 + c,$$

with some constant $c \in \mathbb{R}$. □

Example 5.38 (primitives/antiderivatives and integral of $1/\sqrt{1-x^2}$)
Find all primitives/antiderivatives of the function $f : (-1, 1) \to \mathbb{R}$, given by

$$f(x) = \frac{1}{\sqrt{1-x^2}}, \qquad x \in (-1, 1),$$

and evaluate the integral

$$\int_{-1/2}^{1/2} \frac{1}{\sqrt{1-x^2}}\, dx.$$

Solution: We compute the indefinite integral

$$\int \frac{1}{\sqrt{1-x^2}}\, dx$$

with the substitution

$$x = x(y) = \sin(y) \quad \Leftrightarrow \quad y = y(x) = \arcsin(x); \qquad \frac{dx}{dy} = \cos(y) \quad \Rightarrow \quad dx = \cos(y)\,dy. \tag{5.21}$$

Note that $\sin : [-\pi/2, \pi/2] \to [-1, 1]$ is one-to-one and has range $[-1, 1]$, and thus the substitution (5.21) is legitimate. We find that

$$\int \frac{1}{\sqrt{1-x^2}}\,dx = \int \frac{1}{\sqrt{1-\big(\sin(y)\big)^2}}\,\cos(y)\,dy. \tag{5.22}$$

We have $1 - \big(\sin(y)\big)^2 = \big(\cos(y)\big)^2$, and thus $\sqrt{1-\big(\sin(y)\big)^2} = \cos(y)$, where we have used that $\cos(y) \geq 0$ for all $y \in [-\pi/2, \pi/2]$ when taking the root. Thus we obtain

$$\int \frac{1}{\sqrt{1-\big(\sin(y)\big)^2}}\,\cos(y)\,dy = \int \frac{\cos(y)}{\cos(y)}\,dy = \int 1\,dy = y = \arcsin(x) + c. \tag{5.23}$$

From (5.22) and (5.23), we see that all primitives/antiderivatives of the function $f(x) = 1/\sqrt{1-x^2}$ are given by

$$F(x) = \arcsin(x) + c,$$

with some constant $c \in \mathbb{R}$.

Now we can use the fundamental theorem of calculus (see Theorem 5.15) to evaluate the definite integral

$$\int_{-1/2}^{1/2} \frac{1}{\sqrt{1-x^2}}\,dx = \arcsin(x)\big|_{-1/2}^{1/2} = \arcsin(1/2) - \arcsin(-1/2) = \frac{\pi}{6} - \frac{-\pi}{6} = \frac{\pi}{3},$$

and find that the value of the definite integral is $\pi/3$. □

Remark 5.39 (integrand $1/\sqrt{a^2-x^2}$)
In analogy to the last example, the indefinite integral

$$\int \frac{1}{\sqrt{a^2-x^2}}\,dx, \qquad x \in (-a, a), \qquad \textit{with a constant } \ a \in \mathbb{R},\ a > 0.$$

can be computed with the substitution $x = x(y) = a\,\sin(x)$.

Example 5.40 (indefinite integral of $1/\sqrt{x^2-a^2}$)
Evaluate the indefinite integral

$$\int \frac{1}{\sqrt{x^2-a^2}}\,dx, \qquad \text{where } \ x > a \ \text{ with } \ a \in \mathbb{R},\ a > 0,$$

with the substitution $x = a\,\cosh(y)$, where $y > 0$.

Solution: We first discuss why this is useful. Letting $x = a\,\cosh(y)$, where $y > 0$, the denominator becomes

$$\sqrt{x^2 - a^2} = \sqrt{a^2\,\big(\cosh(y)\big)^2 - a^2} = \sqrt{a^2\,\big[\big(\cosh(y)\big)^2 - 1\big]}$$

and using $\big(\cosh(y)\big)^2 - \big(\sinh(y)\big)^2 = 1$, thus $\big(\cosh(y)\big)^2 - 1 = \big(\sinh(y)\big)^2$, we find that (use $\sinh(y) > 0$ for $y > 0$)

$$\sqrt{x^2 - a^2} = \sqrt{a^2\,\big[\big(\cosh(y)\big)^2 - 1\big]} = \sqrt{a^2\,\big(\sinh(y)\big)^2} = a\,\sinh(y).$$

We see that the denominator simplifies considerably. For $y > 0$, the function $x = x(y) = a\,\cosh(y)$ is one-to-one with range $(0, \infty)$, and so the substitution $x = x(y) = a\,\cosh(y)$, or equivalently $y = \operatorname{arccosh}(x/a)$, makes sense. We find

$$\frac{dx}{dy} = a\,\sinh(y) \qquad \Rightarrow \qquad dx = a\,\sinh(y)\,dy,$$

and thus for $y > 0$

$$\int \frac{1}{\sqrt{x^2 - a^2}}\,dx = \int \frac{1}{a\,\sinh(y)}\,a\,\sinh(y)\,dy = \int 1\,dy = y + c = \operatorname{arccosh}(x/a) + c,$$

where $c \in \mathbb{R}$ is an arbitrary constant. □

Remark 5.41 (integral of $1/\sqrt{x^2 + a^2}$)
To evaluate the indefinite integral

$$\int \frac{1}{\sqrt{x^2 + a^2}}\,dx, \qquad \textit{where} \;\; a > 0,$$

use the substitution $x = x(y) = a\,\sinh(y)$ *or equivalently* $y = y(x) = \operatorname{arcsinh}(x/a)$.

Example 5.42 (indefinite integral of $1/(x^2 + a^2)$)
To evaluate the indefinite integral

$$\int \frac{1}{x^2 + a^2}\,dx \qquad \text{where} \quad x \in \mathbb{R} \quad \text{and} \quad a > 0,$$

use the substitution $x = a\,\tan(y)$, or equivalently $y = \arctan(x/a)$.

Solution: We first work out the denominator with the substitution $x = a\,\tan(y)$.

$$x^2 + a^2 = a^2\,\big(\tan(y)\big)^2 + a^2 = a^2\,\big[\big(\tan(y)\big)^2 + 1\big] = a^2\,\frac{\big(\sin(y)\big)^2 + \big(\cos(y)\big)^2}{\big(\cos(y)\big)^2},$$

and using $\big(\sin(y)\big)^2 + \big(\cos(y)\big)^2 = 1$, we find that

$$x^2 + a^2 = \frac{a^2}{\big(\cos(y)\big)^2} \qquad \Rightarrow \qquad \frac{1}{x^2+a^2} = \frac{\big(\cos(y)\big)^2}{a^2}.$$

The substitution $x = x(y) = a\,\tan(y)$, or equivalently $y = y(x) = \arctan(x/a)$, is legitimate because $\tan : (-\pi/2, \pi/2) \to \mathbb{R}$ is one-to-one with range $\mathbb{R}$. We have

$$\frac{dx}{dy} = \frac{a}{\big(\cos(y)\big)^2} \qquad \Rightarrow \qquad dx = \frac{a}{\big(\cos(y)\big)^2}\,dy.$$

Thus the substitution yields

$$\int \frac{1}{x^2+a^2}\,dx = \int \frac{\big(\cos(y)\big)^2}{a^2}\,\frac{a}{\big(\cos(y)\big)^2}\,dy = \int \frac{1}{a}\,dy = \frac{y}{a} + c = \frac{\arctan(x/a)}{a} + c,$$

where $c \in \mathbb{R}$ is an arbitrary constant. □

5.6 Integration by Parts

The **method of integration by parts** follows from the **product rule for differentiation** (see Lemma 3.15):

$$\frac{d}{dx}\big[f(x)\,g(x)\big] = \big[f(x)\,g(x)\big]' = f'(x)\,g(x) + f(x)\,g'(x).$$

Integrating on both sides over the interval $[a,b]$ yields

$$f(x)\,g(x)|_a^b = \int_a^b \big[f(x)\,g(x)\big]'\,dx = \int_a^b f'(x)\,g(x)\,dx + \int_a^b f(x)\,g'(x)\,dx,$$

and thus after rearranging

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

Often we find the abbreviated notation

$$f(x)\,g(x)|_a^b = f(b)\,g(b) - f(a)\,g(a).$$

Theorem 5.43 (integration by parts)
Let $f : A \to B$ and $g : A \to B$ be smooth enough functions, and let $[a,b] \subset A$. Then we have

$$\int_a^b f'(x)\,g(x)\,dx = f(b)\,g(b) - f(a)\,g(a) - \int_a^b f(x)\,g'(x)\,dx. \tag{5.24}$$

This result is **useful for integrating functions of the form** $h(x) = f'(x) \cdot g(x)$ **for which** $f(x) \cdot g'(x)$ **is easier to integrate than** $h(x)$.

Remark 5.44 (integration by parts for indefinite integrals)
Before we discuss examples, we mention that we can also use integration by parts for indefinite integrals, that is,

$$\int_{x_0}^{x} f'(t)\, g(t)\, dt = f(t)\, g(t)|_{x_0}^{x} - \int_{x_0}^{x} f(t)\, g'(t)\, dt,$$

or in the laxer notation without boundary values

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

Now we discuss several examples. To perform integration by parts for an integral of a function h, you need to decide **how you decompose** h **as a product** $h(x) = f'(x)\, g(x)$ **of suitable functions** f' **and** g. The choice is normally determined by the fact that we know the primitive/antiderivative f of f' and that it offers benefits for the remaining integral to differentiate g. However, we have only performed a 'useful' integration by parts if the integral on the right-hand side of (5.24) is simpler to integrate than the original integral. Often we need to integrate repeatedly by parts before the original integral is evaluated. As with integration by substitution, finding a suitable way of integrating by parts is largely experience and practice.

Example 5.45 (definite integral and primitive/antiderivative of $x\, e^{-x}$)
Evaluate the integral

$$\int_0^1 x\, e^{-x}\, dx,$$

and find all primitives/antiderivatives of $f(x) = x\, e^{-x}$.

Solution: We use integration by parts (5.24) with $f(x) = -e^{-x}$ and $g(x) = x$ and thus $f'(x) = e^{-x}$ and $g'(x) = 1$. We find

$$\int_0^1 x\, e^{-x}\, dx = -x\, e^{-x}|_0^1 + \int_0^1 e^{-x}\, dx = -x\, e^{-x}|_0^1 - e^{-x}|_0^1 = -e^{-1} + 0 - e^{-1} + 1 = 1 - 2\, e^{-1}.$$

With the same integration by parts, we find that evaluation of an indefinite integral of $x\, e^{-x}$ yields

$$\int x\, e^{-x}\, dx = -x\, e^{-x} + \int e^{-x}\, dx = -x\, e^{-x} - e^{-x} + c = -(x+1)\, e^{-x} + c,$$

and thus any primitive/antiderivative of $h(x) = x\,e^{-x}$ is of the form

$$H(x) = -(x+1)\,e^{-x} + c,$$

with a constant $c \in \mathbb{R}$. We differentiate to test our result: from the product rule $\big(-(x+1)\,e^{-x} + c\big)' = (-1)\,e^{-x} - (x+1)\,(-e^{-x}) = x\,e^{-x}$, as expected. □

Remark 5.46 (differentiate to test your primitive/antiderivative!)
If you determined a primitive/antiderivative H of a function h, it is always useful to differentiate the primitive/antiderivative and check that indeed $H' = h$.

Example 5.47 (integral of $x\,\sin x$)
Evaluate the integral

$$\int_0^\pi x\,\sin(x)\,dx.$$

Solution: We use integration by parts (5.24) with $f(x) = -\cos(x)$, $g(x) = x$ and thus $f'(x) = \sin(x)$, $g'(x) = 1$. Then, using $\cos(\pi) = -1$ and $\sin(0) = \sin(\pi) = 0$,

$$\begin{aligned}\int_0^\pi x\,\sin(x)\,dx &= -x\,\cos(x)|_0^\pi + \int_0^\pi \cos(x)\,dx \\ &= -\pi\,\cos(\pi) + \sin(x)|_0^\pi \;=\; -\pi\,\cos(\pi) + \sin(\pi) - \sin(0) \;=\; \pi.\end{aligned}$$

Thus we have

$$\int_0^\pi x\,\sin(x)\,dx = \pi.$$ □

Example 5.48 (integral of $e^x\,\sin x$)
Evaluate the integral

$$I = \int_0^\pi e^x\,\sin(x)\,dx.$$

Solution: This is an example of an integral that **cannot be directly evaluated**. Instead, it can be **'solved' via a simple equation**, as follows. We apply integration by parts (5.24) with $f(x) = e^x$, $g(x) = \sin(x)$, and thus $f'(x) = e^x$ and $g'(x) = \cos(x)$. Then

$$\begin{aligned}\int_0^\pi e^x\,\sin(x)\,dx &= e^x\,\sin(x)|_0^\pi - \int_0^\pi e^x\,\cos(x)\,dx \\ &= e^\pi\,\sin(\pi) - e^0\,\sin(0) - \int_0^\pi e^x\,\cos(x)\,dx \\ &= -\int_0^\pi e^x\,\cos(x)\,dx, \qquad (5.25)\end{aligned}$$

where we have used $\sin(0) = \sin(\pi) = 0$. Now we apply integration by parts (5.24) a second time with $f(x) = e^x$, $g(x) = \cos(x)$, and thus $f'(x) = e^x$ and $g'(x) = -\sin(x)$. Then, using $\cos(\pi) = -1$ and $\cos(0) = 1$,

$$\begin{aligned} -\int_0^\pi e^x \cos(x)\,dx &= -e^x \cos(x)|_0^\pi - \int_0^\pi e^x \sin(x)\,dx \\ &= -e^\pi \cos(\pi) + e^0 \cos(0) - \int_0^\pi e^x \sin(x)\,dx \\ &= e^\pi + 1 - \int_0^\pi e^x \sin(x)\,dx. \end{aligned} \tag{5.26}$$

Combining (5.25) and (5.26), we obtain

$$\int_0^\pi e^x \sin(x)\,dx = e^\pi + 1 - \int_0^\pi e^x \sin(x)\,dx. \tag{5.27}$$

The original integral I occurs in (5.27) again on the right-hand side with a negative sign. We add the original integral I on both sides and divide afterwards by 2 to obtain

$$2\,I = 2\int_0^\pi e^x \sin(x)\,dx = e^\pi + 1 \qquad \Rightarrow \qquad I = \int_0^\pi e^x \sin(x)\,dx = \frac{e^\pi + 1}{2}. \quad \square$$

The trick used to evaluate the integral in the last example is common for a certain class of (definite or indefinite) integrals. **One performs integration by parts twice and obtains**

$$I = C - \alpha\,I,$$

where I denotes the original integral, α is a positive real number, and C is a real value in the case of a definite integral or a function in the case of an indefinite integral, respectively. Now we can rearrange to determine I and find, by adding $\alpha\,I$ on both sides and then dividing by $(1+\alpha)$,

$$I = C - \alpha\,I \qquad \Rightarrow \qquad I + \alpha\,I = (1+\alpha)\,I = C \qquad \Rightarrow \qquad I = \frac{C}{1+\alpha}.$$

Example 5.49 (primitives/antiderivatives and integral of $\ln x$)
Find all primitives/antiderivatives of $h(x) = \ln(x)$ and evaluate the integral

$$\int_1^2 \ln(x)\,dx.$$

Solution: With $f(x) = x$ and $g(x) = \ln(x)$, and thus $f'(x) = 1$ and $g'(x) = 1/x$, we have from the integration by parts formula (5.24)

$$\int \ln(x)\,dx = x\,\ln(x) - \int x\,\frac{1}{x}\,dx = x\,\ln(x) - \int 1\,dx = x\,\ln(x) - x + c,$$

with an arbitrary constant c. Thus every primitive/antiderivative of $h(x) = \ln(x)$ is of the form

$$H(x) = x\,\ln(x) - x + c = x\,\big(\ln(x) - 1\big) + c,$$

with a constant $c \in \mathbb{R}$. From the fundamental theorem of calculus (see Theorem 5.15)

$$\begin{aligned}\int_1^2 \ln(x)\,dx &= H(x)|_1^2 = \big(x\,\ln(x) - x + c\big)\big|_1^2 \\ &= 2\,\ln(2) - 2 + c - \big(1\,\ln(1) - 1 + c\big) = 2\,\ln(2) - 1.\end{aligned}$$

□

Example 5.50 (primitives of $\big(\sin(x)\big)^2 + 2\,x + 3$)
Find all primitives/antiderivatives of the continuous function $h : \mathbb{R} \to \mathbb{R}$, given by

$$h(x) = \big(\sin(x)\big)^2 + 2\,x + 3.$$

Solution: We want to calculate the indefinite integral of $h(x) = \big(\sin(x)\big)^2 + 2\,x + 3$ given by

$$\int \big[\big(\sin(x)\big)^2 + 2\,x + 3\big]\,dx.$$

From the linear properties of the integral (see Lemma 5.23), from $(x)' = 1$ and $(x^2)' = 2x$, and from the fundamental theorem of calculus (see Theorem 5.15), we find that

$$\begin{aligned}\int \big[\big(\sin(x)\big)^2 + 2\,x + 3\big]\,dx &= \int \big(\sin(x)\big)^2\,dx + \int 2\,x\,dx + \int 3\,dx \quad (5.28)\\ &= \int \big(\sin(x)\big)^2\,dx + x^2 + 3\,x + c_1, \quad (5.29)\end{aligned}$$

with a constant $c_1 \in \mathbb{R}$. From the integration by parts formula (5.24) we have, with $f(x) = -\cos(x)$, $g(x) = \sin(x)$, and thus $f'(x) = \sin(x)$, $g'(x) = \cos(x)$,

$$\int \big(\sin(x)\big)^2\,dx = -\cos(x)\,\sin(x) + \int \big(\cos(x)\big)^2\,dx. \quad (5.30)$$

Now we use that $\big(\sin(x)\big)^2 + \big(\cos(x)\big)^2 = 1$ and replace $\big(\cos(x)\big)^2 = 1 - \big(\sin(x)\big)^2$ in the remaining integral in (5.30). Thus

$$\begin{aligned}\int \big(\sin(x)\big)^2\,dx &= -\cos(x)\,\sin(x) + \int \big[1 - \big(\sin(x)\big)^2\big]\,dx \\ &= -\cos(x)\,\sin(x) + \int 1\,dx - \int \big(\sin(x)\big)^2\,dx\end{aligned}$$

$$= \quad -\cos(x)\,\sin(x) + x + c_2 - \int \big(\sin(x)\big)^2\,dx,$$

with a constant $c_2 \in \mathbb{R}$. Adding the original integral on both sides yields

$$2\int_0^x \big(\sin(x)\big)^2\,dt = x - \cos(x)\,\sin(x) + c_2$$

$$\Rightarrow \qquad \int_0^x \big(\sin(x)\big)^2\,dt = \frac{1}{2}\big[x - \cos(x)\,\sin(x)\big] + \frac{c_2}{2}. \tag{5.31}$$

Thus we have from (5.29) and (5.31) that

$$\begin{aligned}\int \big[\big(\sin(x)\big)^2 + 2\,x + 3\big]\,dx &= \frac{1}{2}\big[x - \cos(x)\,\sin(x)\big] + \frac{c_2}{2} + x^2 + 3x + c_1 \\ &= x^2 + \frac{7}{2}\,x - \frac{1}{2}\,\cos x\,\sin x + c,\end{aligned}$$

with an arbitrary real constant $c = c_1 + c_2/2$. Therefore any primitive/antiderivative of $h(x) = \big(\sin(x)\big)^2 + 2\,x + 3$ is of the form

$$H(x) = x^2 + \frac{7}{2}\,x - \frac{1}{2}\,\cos x\,\sin x + c,$$

with some constant $c \in \mathbb{R}$. □

Example 5.51 (integral of $\big(\cos(x)\big)^{-2}$)
Find all primitives/antiderivatives of the function $h : (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$, given by

$$h(x) = \big(\sec(x)\big)^2 = \frac{1}{\big(\cos(x)\big)^2},$$

and evaluate the definite integral

$$\int_{-\pi/4}^{\pi/4} \frac{1}{\big(\cos(x)\big)^2}\,dx.$$

Solution: To calculate the indefinite integral of $h(x) = (\cos x)^{-2}$ given by

$$\int \frac{1}{\big(\cos(x)\big)^2}\,dx,$$

we use that $1 = \big(\cos(x)\big)^2 + \big(\sin(x)\big)^2$ to rewrite the indefinite integral as follows

$$\int \frac{1}{\big(\cos(x)\big)^2}\,x = \int \frac{\big(\cos(x)\big)^2 + \big(\sin(x)\big)^2}{\big(\cos(x)\big)^2}\,dx = \int \left(1 + \frac{\big(\sin(x)\big)^2}{\big(\cos(x)\big)^2}\right)\,dx.$$

Thus from the linear properties of the integral

$$\int \frac{1}{\big(\cos(x)\big)^2}\,dx = \int 1\,dx + \int \frac{\big(\sin(x)\big)^2}{\big(\cos(x)\big)^2}\,dx = x + c_1 + \int \frac{\big(\sin(x)\big)^2}{\big(\cos(x)\big)^2}\,dx, \tag{5.32}$$

with a constant $c_1 \in \mathbb{R}$. We apply the integration by parts formula (5.24) with $f(x) = 1/\cos(x)$, $g(x) = \sin(x)$, and thus $f'(x) = \sin(x)/\big(\cos(x)\big)^2$, $g'(x) = \cos(x)$ to the remaining integral, and obtain

$$\int \frac{\big(\sin(x)\big)^2}{\big(\cos(x)\big)^2}\,dx = \frac{\sin(x)}{\cos(x)} - \int \frac{1}{\cos(x)}\,\cos(x)\,dx = \frac{\sin(x)}{\cos(x)} - \int 1\,dx = \tan(x) - x + c_2, \tag{5.33}$$

with a constant $c_2 \in \mathbb{R}$. Substituting (5.33) into (5.32) yields

$$\int \frac{1}{\big(\cos(x)\big)^2}\,dx = x + c_1 + \tan(x) - x + c_2 = \tan(x) + c,$$

with an arbitrary constant $c = c_1 + c_2$. Thus all primitives/antiderivatives of the function $h(x) = 1/\big(\cos(x)\big)^2$ are of the form

$$H(x) = \tan(x) + c, \qquad x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right),$$

with some constant $c \in \mathbb{R}$. From the fundamental theorem of calculus (see Theorem 5.15), we have

$$\int_{-\pi/4}^{\pi/4} \frac{1}{\big(\cos(x)\big)^2}\,dx = H(x)\big|_{-\pi/4}^{\pi/4} = \big(\tan(x) + c\big)\big|_{-\pi/4}^{\pi/4} = \tan(\pi/4) - \tan(-\pi/4),$$

and, since $\sin(\pi/4) = \cos(\pi/4) = 1/\sqrt{2}$ and $\sin(-\pi/4) = -1/\sqrt{2}$, $\cos(-\pi/4) = 1/\sqrt{2}$, we have that $\tan(\pi/4) = 1$ and $\tan(-\pi/4) = -1$. Thus

$$\int_{-\pi/4}^{\pi/4} \frac{1}{\big(\cos(x)\big)^2}\,dx = \tan(\pi/4) - \tan(-\pi/4) = 1 - (-1) = 2. \qquad \square$$

5.7 Examples That Need More Than One Method of Integration

Often there is not one unique way to evaluate a definite integral or compute a primitive/antiderivative as an indefinite integral. A few integrals can be evaluated

with either integration by substitution or integration by parts, and to evaluate some integrals you will need both integration by substitution and integration by parts.

We now give two examples where **both integration by substitution and integration by parts** are needed.

Example 5.52 (integral of $x^5 \sin(x^3)$)
Evaluate the integral

$$\int_0^{\sqrt[3]{2\pi}} x^5 \sin(x^3)\, dx.$$

Solution: With the substitution

$$y = y(x) = x^3, \qquad \frac{dy}{dx} = 3\,x^2 \quad \Leftrightarrow \quad dy = 3\,x^2\, dx,$$

we obtain, with the new boundaries $y(0) = 0$ and $y(\sqrt[3]{2\pi}) = 2\pi$,

$$\int_0^{\sqrt[3]{2\pi}} x^5 \sin(x^3)\, dx = \frac{1}{3}\int_0^{\sqrt[3]{2\pi}} \underbrace{x^3}_{=y}\, \underbrace{\sin(x^3)}_{=\sin(y)}\, \underbrace{3x^2\, dx}_{=dy} = \frac{1}{3}\int_0^{2\pi} y\, \sin(y)\, dy.$$

Now we use the integration by parts formula (5.24) with $f(y) = -\cos(y)$, $g(y) = y$ and thus $f'(y) = \sin(y)$ and $g'(y) = 1$ and obtain

$$\begin{aligned}
\frac{1}{3}\int_0^{2\pi} y\, \sin(y)\, dy &= \frac{1}{3}\left(-y\, \cos(y)\big|_0^{2\pi} + \int_0^{2\pi} \cos(y)\, dy\right) \\
&= \frac{1}{3}\big[-2\pi\, \cos(2\pi) + 0\cos(0)\big] + \frac{1}{3}\, \sin(y)\big|_0^{2\pi} \\
&= -\frac{2\pi}{3} + \frac{1}{3}\big[\sin(2\pi) - \sin(0)\big] \;=\; -\frac{2\pi}{3},
\end{aligned}$$

where we have used $\cos(0) = \cos(2\pi) = 1$ and $\sin(0) = \sin(2\pi) = 0$. Thus

$$\int_0^{\sqrt[3]{2\pi}} x^5 \sin(x^3)\, dx = -\frac{2\pi}{3}. \qquad \square$$

Example 5.53 (integral of $e^{2x} \cos(e^x - 1)$)
Evaluate the integral

$$\int_0^1 e^{2x} \cos(e^x - 1)\, dx.$$

Solution: The substitution

$$y = y(x) = e^x - 1 \quad \Leftrightarrow \quad e^x = y + 1, \qquad \frac{dy}{dx} = e^x \quad \Leftrightarrow \quad dy = e^x\, dx,$$

with the new boundaries $y(0) = e^0 - 1 = 0$ and $y(1) = e - 1$, yields

$$\int_0^1 e^{2x} \cos(e^x - 1)\, dx = \int_0^1 \underbrace{e^x}_{=y+1} \underbrace{\cos(e^x - 1)}_{=\cos(y)} \underbrace{e^x\, dx}_{=dy} = \int_0^{e-1} (y+1) \cos(y)\, dy. \quad (5.34)$$

Now we use the integration by parts formula (5.24), with $f(y) = \sin(y)$, $g(y) = y+1$ and thus $f'(y) = \cos(y)$ and $g'(y) = 1$, and obtain

$$\begin{aligned} \int_0^{e-1} (y+1) \cos(y)\, dy &= (y+1) \sin(y)|_0^{e-1} - \int_0^{e-1} \sin(y)\, dy \\ &= e \sin(e-1) - \sin(0) + \cos(x)|_0^{e-1} \\ &= e \sin(e-1) + \cos(e-1) - \cos(0) \\ &= e \sin(e-1) + \cos(e-1) - 1, \end{aligned} \quad (5.35)$$

where we have used $\sin(0) = 0$ and $\cos(0) = 1$. From (5.34) and (5.35) we obtain

$$\int_0^1 e^{2x} \cos(e^x - 1)\, dx = e \sin(e-1) + \cos(e-1) - 1. \qquad \square$$

Application 5.54 (gravitational potential of the earth)
*The **gravitational potential of the earth** (which is caused by the gravitational force) at a point is defined numerically as equal to the work done in moving a unit mass from infinity to this point. The **gravitational force** (or **force of attraction**) on a unit mass outside the earth is given by*

$$F(r) = \frac{G\,M}{r^2},$$

*where r is the distance from the earth's centre, G is the gravitational constant, and M is the mass of the earth (assumed to be located at the earth's centre). The **work for moving the mass** is given by*

$$\text{work} \;=\; \text{force} \;\times\; \text{distance in the direction of the force}$$

*Since the gravitational force depends on the distance r and changes continuously, the **work done in moving a unit mass from infinity to the point with distance** r is given by the integral*

$$\begin{aligned} \int_\infty^r F(x)\, dx = \int_\infty^r \frac{G\,M}{x^2}\, dx &= \lim_{b\to\infty} \int_b^r \frac{G\,M}{x^2}\, dx = \lim_{b\to\infty} -\int_r^b \frac{G\,M}{x^2}\, dx \\ &= \lim_{b\to\infty} \frac{G\,M}{x}\bigg|_r^b = \lim_{b\to\infty} \left[\frac{G\,M}{b} - \frac{G\,M}{r}\right] = -\frac{G\,M}{r}. \end{aligned}$$

Thus the ***gravitational potential of the earth*** *at any point with distance r from the earth's centre is given by*

$$V(r) = -\frac{G\,M}{r}.$$

Chapter 6

Further Integration

In this chapter we learn several **applications of integration** that are important for engineering and physics. The first such application is finding the **average of a function** which is discussed in Section 6.1. For example we may want to know the average speed of a car driving from London to Brighton or the average temperature during the month January. The next application is the **calculation of the area between the graphs of two functions** with the help of the integral which is discussed in Section 6.2. This allows us to calculate areas bounded by rather general shapes. In Section 6.3, we discuss a physical application of integration, namely the **work done by moving an object in the presence of a force**. This application is directly derived from the geometric interpretation of the integral as the area under the curve. In Section 6.4, we learn the **calculation of volumes of revolution**. A volume of revolution is any three-dimensional body that is rotationally symmetric about an axis. Examples of a volumes of revolution are, for example, a cylinder, a sphere, and a cone. In Section 6.5, we discuss the **mass of a rod with constant cross-sectional area**, and in Sections 6.6 and 6.7, we learn how to find (the x-coordinate) of the **centre of mass** and the **moment of inertia** of three dimensional bodies with certain symmetries, respectively.

6.1 Average Value of a Function on an Interval

The simplest application of integration is to find the **average of a (continuous) function on an interval** $[a, b]$.

As a motivation, we want to determine the **average temperature over one calender week**. To get an approximation of the average temperature over one calender week, we might take the sum of the temperatures at the start of each hour and take the average, that is the sum of all 7×24 temperatures divided by 7×24. More precisely, if $f(t)$ describes the temperature at the time t measured in hours, then we would take the average

$$\frac{1}{7 \times 24} \sum_{k=1}^{7\times 24} f(k-1) = \frac{1}{(7 \times 24 - 0)} \sum_{k=1}^{7\times 24} f(k-1) \times 1. \tag{6.1}$$

(Note that we have $f(k-1)$ because we measure the temperature at the beginning of each hour $[k-1, k]$.) Keeping in mind that one week (measured in hours) is the time interval $[0, 7 \times 24]$, we can interpret (6.1) as the **sum of the areas of the rectangles with height** $f(k-1)$ **and width** 1, where $k = 1, 2, \ldots 7 \times 24$, divided by the length time interval $[0, 7 \times 24]$. The sum of the areas of these rectangles is an approximation of the area under the curve.

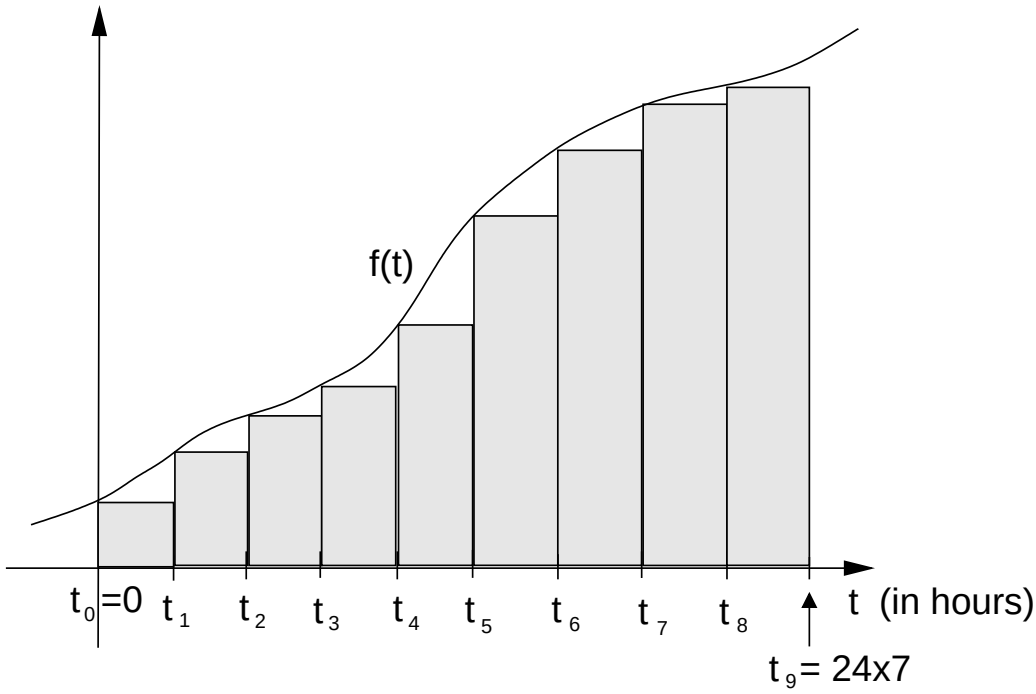


Figure 6.1: The approximation (6.2) (with $N = 9$) of the temperature over one calender week.

The average of the temperatures taken every minute would be an even better approximation of the temperature. In general, if we have N time measurements taken at the times $t_k = k\,\Delta t$, $k = 0, 1, 2, \ldots, N-1$, with $\Delta t = (7 \times 24)/N$, then an approximation of the average temperature is given by

$$\frac{1}{N} \sum_{k=1}^{N} f(t_{k-1}) \quad = \quad \frac{1}{(7 \times 24 - 0)} \sum_{k=1}^{N} f\left((k-1)\,\frac{7 \times 24}{N}\right) \times \frac{7 \times 24}{N}$$

$$= \frac{1}{(7 \times 24 - 0)} \sum_{k=1}^{N} f\big((k-1)\,\Delta t\big) \times \Delta t, \tag{6.2}$$

where we have used $t_{k-1} = (k-1)\,\Delta t = (k-1)\,(7 \times 24)/N$. From the last representation in (6.2), we see that, apart from the factor $1/(7 \times 24)$ in front of the sum, the sum can be interpreted as the sum of the areas of the rectangles over $\big[(k-1)\,\Delta t, k\,\Delta t\big]$ with bottom width Δt and height $f(t_{k-1}) = f\big((k-1)\Delta t\big)$. This is illustrated in Figure 6.1, and we see that the sum of the areas of these rectangles is an approximation for the integral of $f(t)$ over the time interval $[0, 7 \times 24]$.

Thus if we shrink the width Δt of the intervals (between the measured temperatures) to zero, the sum in (6.2) becomes the integral. Thus the **average temperature during the calender week** is given by the integral

$$\frac{1}{(T-0)} \int_0^T f(t)\,dt, \qquad \text{with} \quad T = 7 \times 24.$$

The considerations for finding the average temperature over a calender week motivate the following lemma.

Lemma 6.1 (average value of a function)
Let $A, B \subset \mathbb{R}$, and let $[a,b] \subset A$. Let $f : A \to B$ be a continuous function. The ***average value of*** f ***on the interval*** $[a,b]$ *is given by*

$$\text{average value of } f \text{ on } [a,b] \;=\; \frac{\text{area under graph over } [a,b]}{b-a} \;=\; \frac{1}{b-a} \int_a^b f(x)\,dx.$$

Remark 6.2 (another visualization of the average)
From rearranging we see that

$$\big(\text{average value of } f \text{ on } [a,b]\big) \;\times (b-a) \;=\; \int_a^b f(x)\,dx. \tag{6.3}$$

that is, the ***rectangular box with a horizontal side of length*** $(b-a)$ ***and height given by the average function value of*** f ***over*** $[a,b]$ ***has the same area as the area under the graph for*** $x \in [a,b]$.

Example 6.3 (average value of an affine linear function)
Find the average value of the affine linear function $f(x) = m\,x + c$ on an arbitrary interval $[a,b]$.

Solution: The average value of $f(x) = m\,x + c$ on the interval $[a, b]$ is given by

$$\begin{aligned}
\frac{1}{b-a}\int_a^b f(x)\,dx &= \frac{1}{b-a}\int_a^b \big(m\,x + c\big)\,dx = \frac{1}{b-a}\left(\frac{m}{2}\,x^2 + c\,x\right)\Big|_a^b \\
&= \frac{1}{b-a}\left[\left(\frac{m}{2}\,b^2 + c\,b\right) - \left(\frac{m}{2}\,a^2 + c\,a\right)\right] \\
&= \frac{1}{b-a}\left[\frac{m}{2}\,\big(b^2 - a^2\big) + c\,(b-a)\right] \\
&= \frac{1}{b-a}\left[\frac{m}{2}\,(b-a)(b+a) + c\,(b-a)\right] \\
&= \frac{m}{2}\,(b+a) + c.
\end{aligned}$$

Inspection of the average value of $f(x) = m\,x + c$ on $[a, b]$ shows that this value is also given by the average of $f(b)$ and $f(a)$, that is,

$$\frac{1}{2}\big[f(b) + f(a)\big] = \frac{1}{2}\left[\big(m\,b + c\big) + \big(m\,a + c\big)\right] = \frac{1}{2}\left[m\,(b+a) + 2\,c\right] = \frac{m}{2}\,(b+a) + c.$$

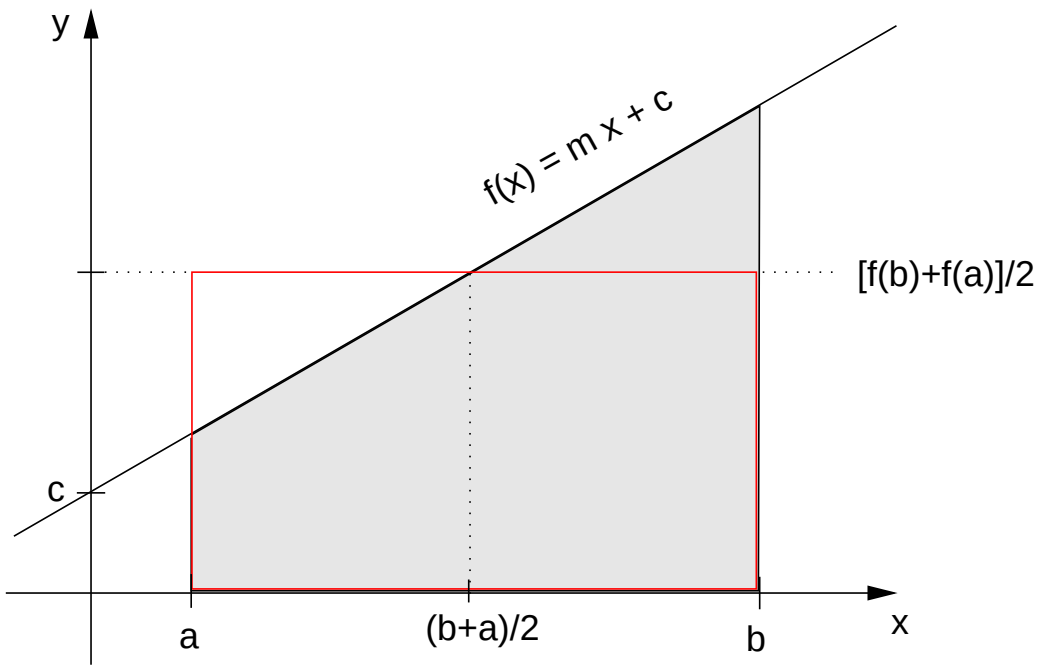


Figure 6.2: Average value $\frac{1}{b-a}\int_a^b f(x)\,dx = \big[f(b) + f(a)\big]/2$ of the affine linear function $f(x) = m\,x + c$ on the interval $[a, b]$. The picture shows the area under the graph for $x \in [a, b]$ and the rectangle with the same area, and it illustrates why in the special case of an affine linear function we have the average value $\big[f(b) + f(a)\big]/2$.

That we can obtain the average of an affine linear function $f(x) = m\,x + c$ on $[a, b]$ by taking the average $\big[f(b) + f(a)\big]/2$ of the the values $f(b)$ and $f(a)$ at the right

and left endpoint of the interval is due to the fact that the graph is a straight line (see Figure 6.2 for illustration). **For other functions this not true, that is, the average values on an interval cannot be obtained by taking the average of the values at the endpoints of the interval!** □

We give three examples.

Example 6.4 (average value of $f(x) = x^2$ on $[1, 2]$)
Calculating the average value of the function $f(x) = x^2$ on $[1, 2]$ yields

$$\frac{1}{2-1}\int_1^2 x^2\,dx = \frac{x^3}{3}\bigg|_1^2 = \frac{8}{3} - \frac{1}{3} = \frac{7}{3},$$

and we see that the average value of $f(x) = x^2$ on $[1, 2]$ is $7/3$. □

In Figure 6.3 we have plotted the area under the graph and the rectangle with the same area for Example 6.4.

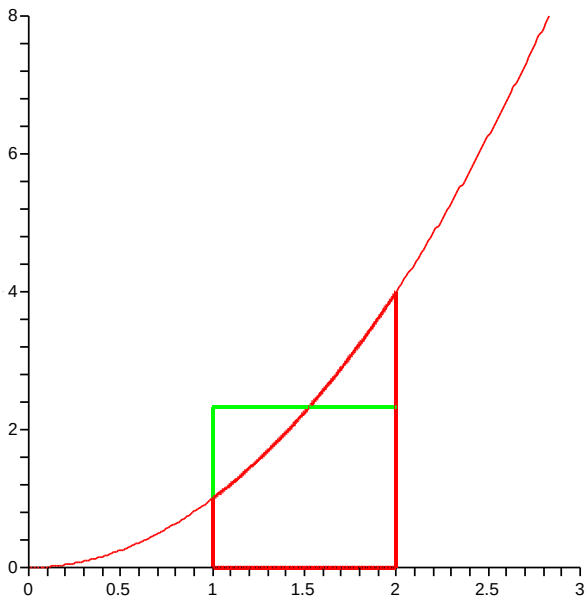


Figure 6.3: The area $\int_1^2 x^2\,dx$ under the graph of $f(x) = x^2$ for $x \in [1, 2]$, and the rectangle over $[1, 2]$ with height given by the average value of $f(x) = x^2$ on $[1, 2]$. This rectangle has also the area $\int_1^2 x^2\,dx$ (see formula (6.3)).

Example 6.5 (average value of $f(x) = \sin(x)$ on $[0, 2\pi]$)
Calculating the average value of the function $f(x) = \sin(x)$ on $[0, 2\pi]$ yields

$$\frac{1}{2\pi - 0}\int_0^{2\pi} \sin(x)\,dx = \frac{1}{2\pi}\big(-\cos(x)\big)\big|_0^{2\pi} = \frac{1}{2\pi}\big[-\cos(2\pi)+\cos(0)\big] = \frac{1}{2\pi}\big[-1+1\big] = 0.$$

and we see that the average value of $f(x) = \sin(x)$ on $[0, 2\pi]$ is 0. $\square$

In Figure 6.4 we have plotted the area under the graph and the rectangle with the same area for Example 6.5.

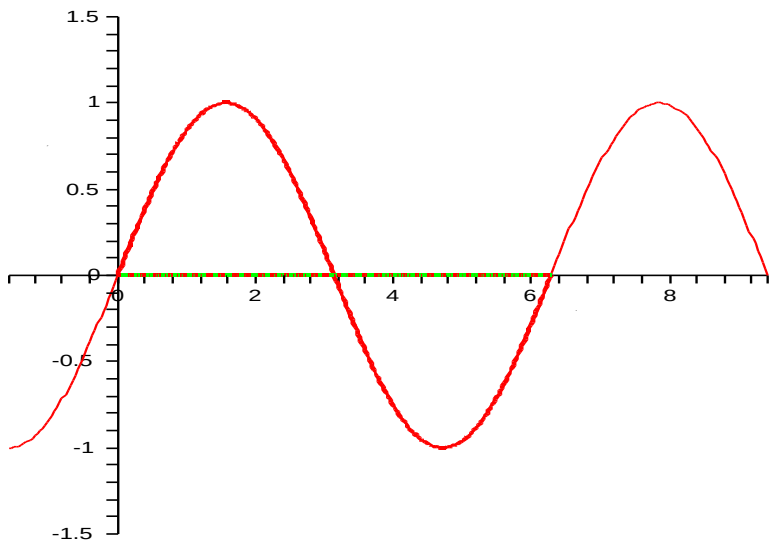


Figure 6.4: The area $\int_0^{2\pi} \sin(x)\,dx$ under the graph of $\sin(x)$ for $x \in [0, 2\pi]$, and the rectangle over $[0, 2\pi]$ with the height given by the average value of $f(x) = \sin(x)$ on $[0, 2\pi]$. This rectangle has also the area $\int_0^{2\pi} \sin(x)\,dx$ (see formula (6.3)). Note that since the average value of $f(x) = \sin(x)$ on $[0, 2\pi]$ is zero, the rectangle has height zero and is therefore simply the straight line $g(x) = 0$, $x \in [0, 2\pi]$.

Example 6.6 (average value of $f(x) = e^x$ on $[-1, 1]$)
Calculating the average value of the function $f(x) = e^x$ on $[-1, 1]$ yields

$$\frac{1}{1-(-1)} \int_{-1}^{1} e^x\,dx = \frac{1}{2}\, e^x\Big|_{-1}^{1} = \frac{1}{2}\left(e^1 - e^{-1}\right) = \frac{1}{2}\left(e - \frac{1}{e}\right),$$

and we see that the average value of $f(x) = e^x$ on $[-1, 1]$ is $(e - e^{-1})/2 \approx 1.18$. $\square$

In Figure 6.5 we have plotted the area under the graph and the rectangle with the same area for Example 6.6.

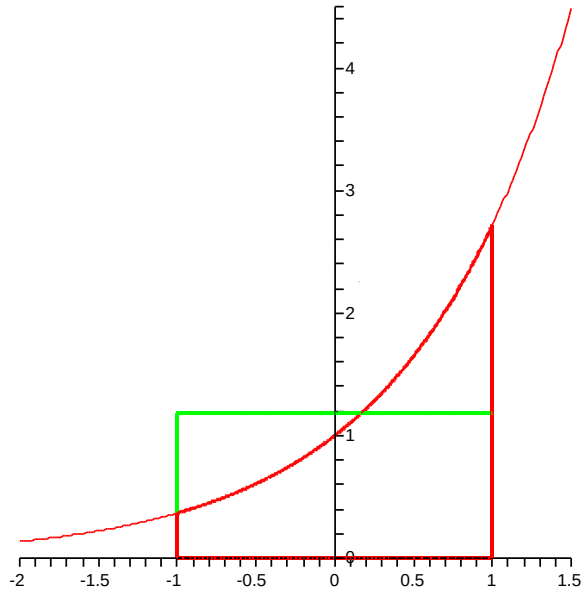


Figure 6.5: The area $\int_{-1}^{1} e^x \, dx$ under the graph of $f(x) = e^x$ for $x \in [-1, 1]$, and the rectangle over $[-1, 1]$ with height given by the average value of $f(x) = e^x$ on $[-1, 1]$. This rectangle has also the area $\int_{-1}^{1} e^x \, dx$ (see (6.3)).

6.2 Areas Bounded by the Graphs of Two Functions

In this section we discuss how to determine the **area between the graphs of two functions**. The method for doing this follows directly for the geometric interpretation of the integral as the area under the graph. Since the method is easiest to understand for a concrete example, we start with an example to illustrate the idea.

Example 6.7 (area between the graphs of two functions)
Find the area between the graphs of the two functions

$$g(x) = \frac{1}{x} \qquad \text{and} \qquad h(x) = 5 - 4\,x.$$

By the area between the two graphs we mean the **finite area bounded by the the two graphs between its points of intersection**.

Solution: A sketch of the graphs of the two functions is given in Figure 6.6.

We work out the points of intersection of the two functions, by determining all points for which $g(x) = h(x)$, that is,

$$\frac{1}{x} = 5 - 4\,x.$$

Multiplying by x and afterwards dividing by 4 gives

$$1 = 5\,x - 4\,x^2 \quad \Leftrightarrow \quad 0 = 4\,x^2 - 5\,x + 1 \quad \Leftrightarrow \quad 0 = x^2 - \frac{5}{4}\,x + \frac{1}{4} = \left(x - \frac{1}{4}\right)(x-1).$$

So the graphs of the two functions intersect at $x = 1/4$ and $x = 1$. We have that

$$g(x) = \frac{1}{x} < 5 - 4\,x = h(x) \qquad \text{for all } \frac{1}{4} < x < 1.$$

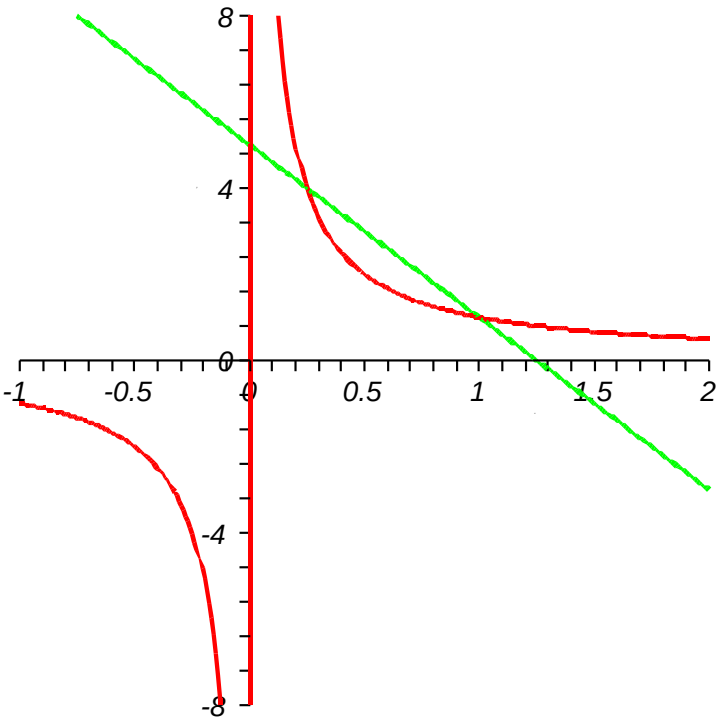


Figure 6.6: The area between the graphs of $g(x) = 1/x$ and $h(x) = 5 - 4\,x$.

This can be seen as follows: We have

$$g(1/2) = \frac{1}{1/2} = 2 < h(1/2) = 5 - 4\,\frac{1}{2} = 5 - 2 = 3,$$

and, since $x = 1/4$ and $x = 1$ are the only points of intersection, we know, from $g(1/2) < h(1/2)$, that $g(x) < h(x)$ holds for all points $x \in (1/4, 1)$.

The area A between the graphs is thus the area under the graph of $h(x) = 5 - 4\,x$ for $x \in [1/4, 1]$ minus the area under the graph of $g(x) = 1/x$ for $x \in [1/4, 1]$. Thus

$$\begin{aligned} A &= \int_{1/4}^{1} h(x)\,dx - \int_{1/4}^{1} g(x)\,dx = \int_{1/4}^{1} \big[h(x) - g(x)\big]\,dx = \int_{1/4}^{1} \left[(5 - 4\,x) - \frac{1}{x}\right] dx \\ &= \big[5\,x - 2\,x^2 - \ln(|x|)\big]\big|_{1/4}^{1} = \big[5 - 2 - \ln(1)\big] - \left[\frac{5}{4} - \frac{2}{16} - \ln\left(\frac{1}{4}\right)\right] \\ &= 3 - \left[\frac{9}{8} - \big(\ln(1) - \ln(4)\big)\right] = \frac{24}{8} - \frac{9}{8} - \ln(4) = \frac{15}{8} - \ln(4) \approx 0.49, \end{aligned}$$

where we have used $\ln(a/b) = \ln(a) - \ln(b)$. Thus the area between the graphs from $x = 1/4$ to $x = 1$ is $A = 15/8 - \ln(4) \approx 0.49$. □

We formulate the method for finding the finite area between the graphs of two functions that intersect each other at least twice.

Method 6.8 (for finding the area between the graphs of two functions)
Let g and h be two functions that intersect at two, or more, points.

- *Solve $g(x) = h(x)$ for x to find the* ***points of intersection*** *$x_1 < x_2 < \ldots < x_n < x_{n+1}$ of the two graphs (where by assumption $n \geq 1$).*
- *For each two adjacent points of intersection x_k and x_{k+1}, determine whether $g(x) < h(x)$ or $g(x) > h(x)$ for all $x \in (x_k, x_{k+1})$.*
- ***If*** *$g(x) < h(x)$* ***for all*** *$x \in (x_k, x_{k+1})$ calculate the* ***area*** *A_k* ***between the graphs for*** *$x \in [x_k, x_{k+1}]$ by evaluating the integral*

$$A_k = \int_{x_k}^{x_{k+1}} \big[h(x) - g(x)\big]\,dx.$$

  If *$g(x) > h(x)$* ***for all*** *$x \in (x_k, x_{k+1})$ calculate the* ***area*** *A_k* ***between the graphs for*** *$x \in [x_k, x_{k+1}]$ by evaluating the integral*

$$A_k = \int_{x_k}^{x_{k+1}} \big[g(x) - h(x)\big]\,dx.$$

- *The* ***area*** *A* ***between the graphs*** *is the sum of all the areas A_k, that is,*

$$A = \sum_{k=1}^{n} A_k.$$

Example 6.9 (area between the graphs of two functions)
Find the finite area between the graphs of the functions

$$f(x) = (x-1)^2 \qquad \text{and} \qquad g(x) = 1.$$

Solution: We find the the points of intersection by setting $f(x) = g(x)$ and solving for x.

$$(x-1)^2 = 1 \qquad \Rightarrow \qquad 0 = (x-1)^2 - 1 = (x-1-1)(x-1+1) = (x-2)\,x,$$

and we see that the points of intersection are $x = 0$ and $x = 2$. Since we have $(x-1)^2 < 1$ for $x \in (0,2)$, we see that $f(x) < g(x)$ for all $x \in (0,2)$. Thus the area between the graphs is given by

$$\begin{aligned} A &= \int_0^2 g(x)\,dx - \int_0^2 f(x)\,dx = \int_0^2 \big[g(x) - f(x)\big]\,dx = \int_0^2 \big[1-(x-1)^2\big]\,dx \\ &= \left[x - \frac{(x-1)^3}{3}\right]\Bigg|_0^2 = \left[2 - \frac{1}{3}\right] - \left[0 - \frac{-1}{3}\right] = 2 - \frac{2}{3} = \frac{4}{3}. \end{aligned}$$

Thus the area between the graphs of the two functions $f(x) = (x-1)^2$ and $g(x) = 1$ between $x = 0$ and $x = 2$ is given by $A = 4/3$. □

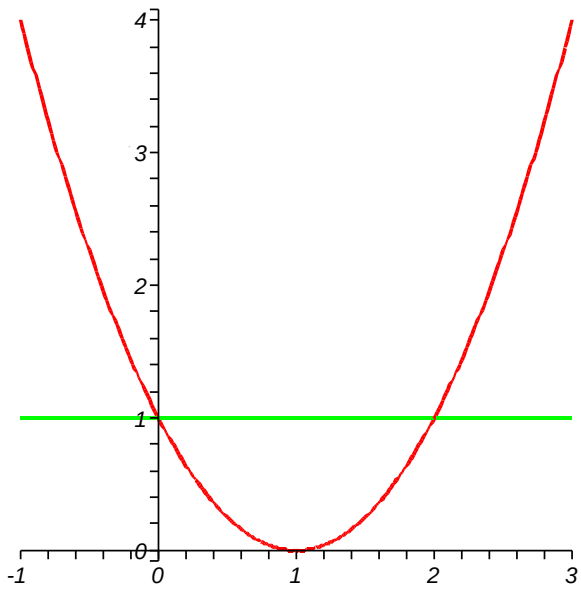


Figure 6.7: The area between the graphs of $f(x) = (x-1)^2$ and $g(x) = 1$.

Example 6.10 (area between the graphs of two functions)
Find the finite area between the graphs of the functions

$$g(x) = \sin(x) \qquad \text{and} \qquad h(x) = \frac{2}{\pi}\,x.$$

Solution: We have plotted the graphs of $g(x) = \sin(x)$ and $h(x) = 2\,x/\pi$ in Figure 6.8. From the picture we can already guess that $x = -\pi/2$, $x = 0$, and $x = \pi/2$ are the points of intersection of the two graphs. Indeed, if $g(x) = h(x)$, then

$$\sin(x) = \frac{2}{\pi}\,x, \tag{6.4}$$

and (6.4) is satisfied for $x = -\pi/2$, $x = 0$, and $x = \pi/2$, since $\sin(-\pi/2) = -1$, $\sin(0) = 0$, and $\sin(\pi/2) = 1$.

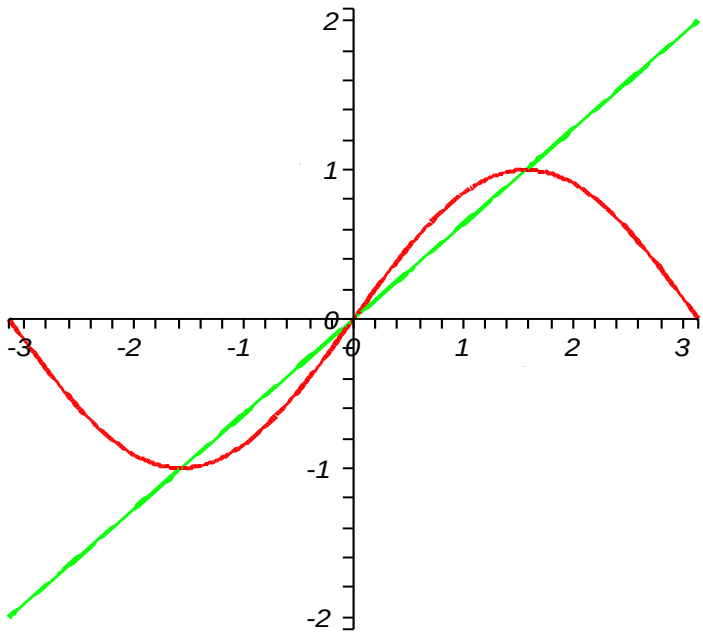


Figure 6.8: The area between the graphs of $g(x) = \sin(x)$ and $h(x) = 2\,x/\pi$.

Why are there no other points of intersection?

For $x > \pi/2$ we have $\sin(x) \le 1 < 2\,x/\pi$, and for $x < -\pi/2$ we have $\sin(x) \ge -1 > 2\,x/\pi$. Thus there are no points of intersection outside the interval $[-\pi/2, \pi/2]$. The points $x = -\pi/2$, $x = 0$, and $x = \pi/2$ are the only points of intersection in the interval $[-\pi/2, \pi/2]$ for the following reasons: Since $g'(x) = \big(\sin(x)\big)' = \cos(x)$ and $g''(x) = \big(\cos(x)\big)' = -\sin(x)$, we see that the function $g(x) = \sin(x)$ is strictly convex downward on $[0, \pi/2]$. Thus the straight line connecting $\big(0, \sin(0)\big) = (0, 0)$ with $\big(\pi/2, \sin(\pi/2)\big) = (\pi/2, 1)$ lies on the interval $(0, \pi/2)$ strictly below the graph of $g(x) = \sin(x)$. Since this straight line is just $h(x) = 2\,x/\pi$ for $x \in [0, \pi/2]$, we see that the functions $g(x) = \sin(x)$ and $h(x) = 2\,x/\pi$ intersect on $[0, \pi/2]$ only in the endpoints $x = 0$ and $x = \pi/2$ and that $g(x) > h(x)$ for all $x \in (0, \pi/2)$.

For $x \in [-\pi/2, 0]$ we can show analogously $g(x) < h(x)$ for all $x \in (-\pi/2, 0)$, and thus $x = -\pi/2$ and $x = 0$ are the only points of intersection in $[-\pi/2, 0]$.

We have already seen that $g(x) > h(x)$ for $x \in (0, \pi/2)$ and that $g(x) < h(x)$ for $x \in (-\pi/2, 0)$. Thus the area between the two graphs is given by

$$\begin{aligned}
A &= \int_{-\pi/2}^{0} \big[h(x) - g(x)\big]\, dx + \int_{0}^{\pi/2} \big[g(x) - h(x)\big]\, dx \\
&= \int_{-\pi/2}^{0} \left[\frac{2}{\pi}\, x - \sin(x)\right] dx + \int_{0}^{\pi/2} \left[\sin(x) - \frac{2}{\pi}\, x\right] dx \\
&= \left[\frac{1}{\pi}\, x^2 + \cos(x)\right]\Bigg|_{-\pi/2}^{0} + \left[-\cos(x) - \frac{1}{\pi}\, x^2\right]\Bigg|_{0}^{\pi/2} \\
&= \big[0 + \cos(0)\big] - \left[\frac{1}{\pi}\,\frac{\pi^2}{4} + \cos(-\pi/2)\right] + \left[-\cos(\pi/2) - \frac{1}{\pi}\,\frac{\pi^2}{4}\right] - \big[-\cos(0) - 0\big] \\
&= 1 - \frac{\pi}{4} - \frac{\pi}{4} + 1 \;=\; 2 - \frac{\pi}{2} \;\approx\; 0.43,
\end{aligned}$$

where we have used $\cos(0) = 1$ and $\cos(-\pi/2) = \cos(\pi/2) = 0$. Thus the area between the two graphs is given by $A = 2 - \pi/2 \approx 0.43$. □

6.3 Work Done While Moving an Object Acted on by a Force

In this section we discuss the work done while moving an object in the presence of a force that acts on this object.

Suppose that a **particle is acted on by a force** $f(x)$, where x is the position of the particle. If the distance Δx is small and the magnitude of the force on the particle hardly changes as it moves from x to $x + \Delta x$, then the **work done moving the particle from** x **to** $x + \Delta x$ is approximately

$$\text{work done} \;=\; \text{force} \;\times\; \text{distance moved} \;\approx\; f(x) \times [(x + \Delta x) - x] = f(x) \times \Delta x.$$

To find the total work done as the particle moves from $x = a$ to $x = b$ (which may be quite far apart), we split the interval $[a, b]$ up into a large number n of small pieces of equal length $\Delta x = (b - a)/n$, and sum up the work done while moving the particle along each subinterval. With the notation

$$x_k = a + k\,\frac{b - a}{n} = a + k\,\Delta x, \qquad \text{that is,}$$

$$x_0 = a, \quad x_1 = a + \frac{b-a}{n}, \quad \ldots \quad x_{n-1} = a + (n-1)\,\frac{b-a}{n}, \quad x_n = b,$$

an **approximation of the work done moving the particle from** $x = a$ **to** $x = b$ is given by

$$\begin{aligned} W \;\approx\; W_n \;&=\; \sum_{k=1}^{n} f(x_{k-1})\,\big(x_k - x_{k-1}\big) \;=\; \sum_{k=1}^{n} f(x_{k-1})\,\Delta x \\ &=\; \sum_{k=1}^{n} f\left(a + (k-1)\,\frac{b-a}{n}\right)\,\frac{b-a}{n}. \end{aligned}$$

We see that the individual terms in the sum are just the areas of the rectangles in Figure 6.9. Increasing n, or, in other words, shrinking the length $\Delta x = (b-a)/n$ of the subintervals to zero, yields the exact value of the work. From the introduction of the integral in the previous chapter, we see thus that the **exact value of the work done moving the particle from** $x = a$ **to** $x = b$ is given by the area under the graph of $f(x)$ for $x \in [a, b]$, that is, by the integral

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

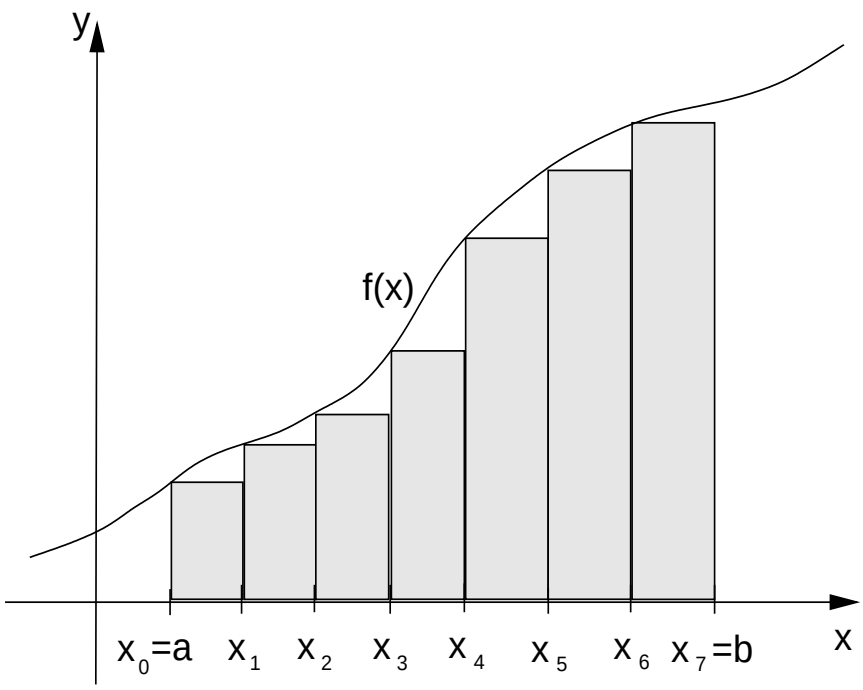


Figure 6.9: The area of the rectangle over $[x_{k-1}, x_k]$ is approximately the work done while moving the particle in the presence of the force $f(x)$ from $x = x_{k-1}$ to $x = x_k$.

We note that the movement of a **particle** in the presence of a force served just as an example. Whenever we move an object in the presence of a force $f(x)$ from $x = a$

to $x = b$, then $\int_a^b f(x)\,dx$ gives the **work done**. We note that the force f has to be a force acting along the straight line of movement.

Thus we have derived that:

$$\boxed{\begin{pmatrix}\text{work done moving an object acted on}\\ \text{by a force } f(x) \text{ from } x = a \text{ to } x = b\end{pmatrix} = W = \int_a^b f(x)\,dx.} \tag{6.5}$$

Example 6.11 (work done by repelling force of two charged particles)
Suppose that a positively charged particle is fixed at the point $x = 0$, while another positively charged particle moves (under the mutual repulsive force of the two particles) along the x-axis from 10 cm away to 1 m away. Assume that the force on the second particle at the position $x > 0$ in metres is

$$f(x) = \frac{C}{x^2}, \qquad \text{where} \quad C = 7.3 \times 10^{-26}\,\mathrm{N\,m^2}.$$

(Note that $\mathrm{N} = \mathrm{kg\,m/s^2}$ stands for Newton, the standard unit for measuring forces.) Find out the work done while the second particle moves from 10 cm away to 1 m away from the first particle.

Solution: From (6.5), the total work done is (use $10\,\mathrm{cm} = 0.1\,\mathrm{m}$)

$$\begin{aligned} W &= \int_{0.1}^{1} \frac{C}{x^2}\,dx = C\left[-\frac{1}{x}\right]\Bigg|_{0.1}^{1} = C\left(-1 + \frac{1}{0.1}\right) \\ &= C\,(-1 + 10) = 9\,C = 6.57 \times 10^{-25}. \end{aligned}$$

Thus the total work done is $W = 6.57 \times 10^{-25}\,\mathrm{N\,m}$. □

Example 6.12 (work done by stretching an elastic band)
An elastic band is pulled out steadily through the small distance of 1 cm, and the force exerted to achieve this increases linearly from 0 N to 10 N (where the physical unit $\mathrm{N} = \mathrm{kg\,m/s^2}$ is called Newton). That the force f exerted increases linearly from 0 N to 10 N means that $f(x) = \alpha\,x$ with a suitable constant $\alpha \in \mathbb{R}$. Determine the work that has been exerted to stretch the elastic band 1 cm.

Solution: After expressing $1\,\mathrm{cm} = 0.01\,\mathrm{m}$ in the standard unit meters, we find from $f(0.01) = 10$ and $f(x) = \alpha\,x$ that

$$10 = f(0.01) = \alpha \times 0.01 \qquad \Rightarrow \qquad \alpha = \frac{10}{0.01} = 1000.$$

Thus the constant α is given by

$$\alpha = 1000\,\frac{\mathrm{N}}{\mathrm{m}},$$

and the force is

$$f(x) = 1000\,x.$$

The work done is now

$$W = \int_0^{0.01} f(x)\,dx = \int_0^{0.01} 1000\,x\,dx = 1000\,\frac{x^2}{2}\bigg|_0^{0.01} = 1000\,\frac{10^{-4}}{2} = 0.05,$$

and we see that the work done is $W = 0.05\,\mathrm{N\,m}$. □

6.4 Volumes of Revolution

In this section we discuss **volumes of revolution** which are obtained by **rotating the graph of a function $f(x)$ about the x-axis** and considering the volume/solid body encased the by the rotated graph over an interval $[a, b]$. We will explain this now in detail and then discuss some examples.

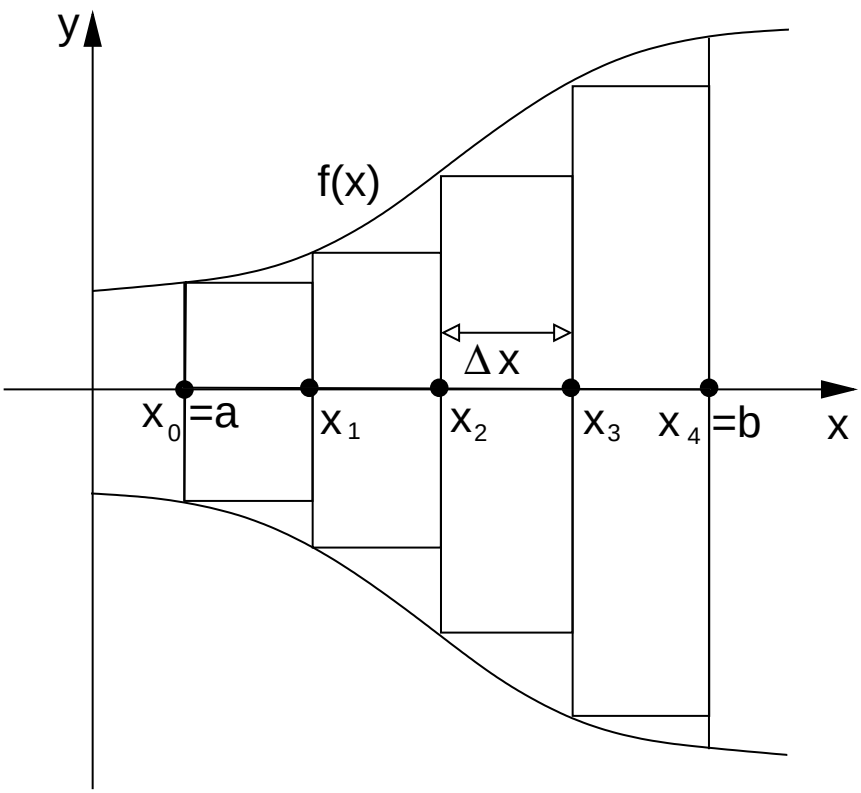


Figure 6.10: Cross-section (along the rotational axis) of the volume of revolution of f between $x = a$ and $x = b$. The 'rectangles' indicate the cross-sections of the discs from approximating the volume of revolution with four discs of equal thickness $\Delta x = (b - a)/4$.

Suppose we have a continuous function $f(x)$ whose graph lies above the x-axis in the range $x \in [a, b]$, that is, $f(x) \geq 0$ for $x \in [a, b]$. Then **rotating this part of the graph by** 360 **degrees**, or equivalently 2π radians, **about the x-axis** will produce a **volume/solid body** as illustrated in Figure 6.10, where we see a cross-section through the volume/solid body along the axis of revolution. Such a volume/solid body is called the **volume of revolution of** f **between** $x = a$ **and** $x = b$.

In order to calculate the **volume of revolution**, we subdivide $[a, b]$ into n subintervals of equal length, that is, into the n subintervals $[x_{k-1}, x_k]$, $k = 1, 2, \ldots, n$, of equal length $\Delta x = (b - a)/n$, where

$$x_k = a + k\,\frac{b-a}{n} = a + k\,\Delta x, \qquad k = 0, 1, 2, \ldots, n-1, n.$$

The volume V_k over the interval $[x_{k-1}, x_k]$ is for small Δx approximately the **disk with radius** $f(x_{k-1})$ **and thickness** Δx (see Figure 6.10). To help imagining this, we may think of slicing the volume of revolution vertically/perpendicular to the x-axis into n slices of equal thickness Δx. Each slice is then approximately the disk with radius $f(x_{k-1})$ and thickness Δx. **Each such disk has the volume**

$$V_k \;=\; \text{area of circular face of disk} \;\times\; \text{thickness of disk} \;=\; \pi\,[f(x_{k-1})]^2 \times \Delta x.$$

Thus an **approximation of the total volume** V **of revolution** is obtained by summing up the volumes $\pi\,[f(x_{k-1})]^2 \times \Delta x$ of all these disks, that is,

$$V \approx \sum_{k=1}^{n} V_k = \sum_{k=1}^{n} \pi\,[f(x_{k-1})]^2 \times \Delta x = \sum_{k=1}^{n} \pi\,\big[f(x_{k-1})\big]^2\,\Delta x. \tag{6.6}$$

If we increase the number n of subintervals, then the length of the subintervals $\Delta x = (b - a)/n$ shrinks, and the approximation of the volume of revolution gets better. In the **limit** $n \to \infty$, and thus $\Delta x \to 0$, we **recover the actual volume of revolution** and obtain mathematically the **integral**:

$$\boxed{\begin{pmatrix}\text{volume of revolution of } f\\ \text{between } x = a \text{ and } x = b\end{pmatrix} = V = \int_a^b \pi\,\big[f(x)\big]^2\,dx = \pi\int_a^b \big[f(x)\big]^2\,dx.} \tag{6.7}$$

We could also have obtained (6.7) with the following argumentation: After slicing the volume of revolution into slices of equal thickness Δx, we consider an arbitrary slice from x to $x + \Delta x$. This slice can be approximated by the disk with radius $f(x)$ and thickness Δx. Thus the volume of the slice is approximately

$$\text{volume of disk from } x \text{ to } x + \Delta x \;\approx\; \pi\,\big[f(x)\big]^2 \times \Delta x \;=\; \pi\,\big[f(x)\big]^2\,\Delta x$$

Now we sum up the volumes of all these disks which we will indicate with the symbolic notation

$$\sum \pi \left[f(x)\right]^2 \Delta x. \tag{6.8}$$

The sum (6.8) of the volumes of all these disks gives us an approximation of the volume of revolution. For the limit $\Delta x \to 0$, the summation sign $\sum$ is replaced by the integral from $x = a$ to $x = b$, and Δx is replaced by dx, and we obtain

$$\begin{pmatrix} \text{volume of revolution of } f \\ \text{between } x = a \text{ and } x = b \end{pmatrix} = V = \int_a^b \pi \left[f(x)\right]^2 dx = \pi \int_a^b \left[f(x)\right]^2 dx.$$

We discuss some examples.

Example 6.13 (cone as volume of revolution)
Find the volume of a (circular) cone of height h whose base has radius r.

Solution:

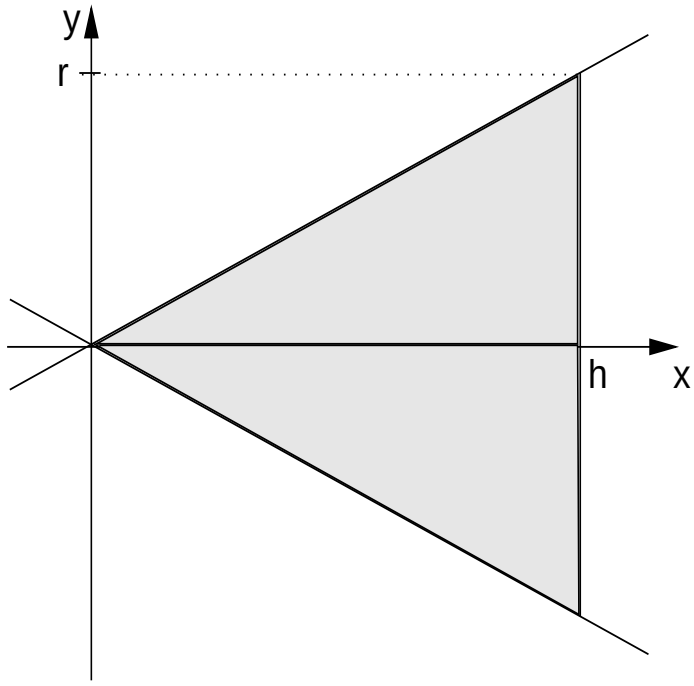


Figure 6.11: The cone generated by rotation of $f(x) = (r/h)\,x$ for $x \in [0, h]$ about the x-axis.

To find the volume of a cone of height h, we notice that such a cone can be generated by rotating the segment of the graph of $f(x) = m\,x$ between $x = 0$ and $x = h$ about the x-axis where m has to be chosen such that the cone has the desired base (see

Figure 6.11). Since the base of the cone should have radius r, we need to choose m such that

$$r = f(h) = m\,h \qquad \Rightarrow \qquad m = \frac{r}{h}.$$

Thus the function $f(x)$ is given by

$$f(x) = \frac{r}{h}\,x.$$

From (6.7) the volume of the cone is given by

$$\begin{aligned} V &= \pi \int_0^h \big[f(x)\big]^2\,dx = \pi \int_0^h \left(\frac{r}{h}\,x\right)^2 dx = \frac{\pi\,r^2}{h^2}\int_0^h x^2\,dx \\ &= \frac{\pi\,r^2}{h^2}\,\frac{x^3}{3}\bigg|_0^h = \frac{\pi\,r^2}{h^2}\,\frac{h^3}{3} = \frac{\pi\,r^2\,h}{3}. \end{aligned}$$

Thus the cone has the volume $V = (\pi\,r^2\,h)/3$. □

Example 6.14 (sphere as body of revolution)
Find the volume of the sphere with radius $r > 0$.

Solution:

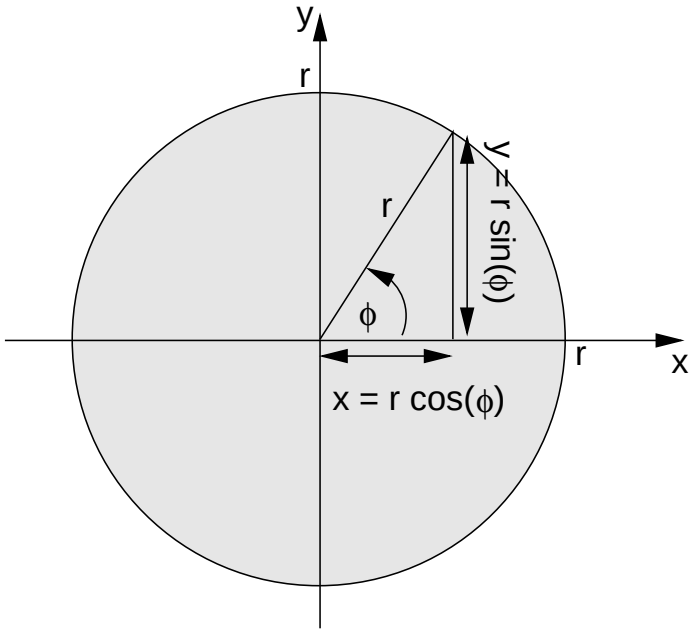


Figure 6.12: The sphere as a volume of revolution by rotating $y(x) = \sqrt{r^2 - x^2}$ for $x \in [-r, r]$ about the x-axis.

Consider Figure 6.12, where we show a cross section of the sphere through its center. Any point (x, y) on the boundary of this cross section has coordinates

$$x = r\,\cos(\phi), \qquad y = r\,\sin(\phi), \qquad \text{where} \quad \phi \in [0, 2\pi].$$

From $\big[\sin(\phi)\big]^2 + \big[\cos(\phi)\big]^2 = 1$, we find that for any point on the boundary

$$x^2 + y^2 = \big[r\,\sin(\phi)\big]^2 + \big[r\,\cos(\phi)\big]^2 = r^2\,\Big(\big[\sin(\phi)\big]^2 + \big[\cos(\phi)\big]^2\Big) = r^2,$$

and thus

$$y^2 = r^2 - x^2. \tag{6.9}$$

If we want to represent the sphere as a volume of revolution, the it is enough to rotate the upper part of the boundary, for which $\phi \in [0, \pi]$, that is, $x \in [-r, r]$ and $y \geq 0$, about the x-axis. Thus we obtain from (6.9) the y-coordinates of the upper hemi-circle in Figure 6.12 as a function $y(x)$ defined by

$$y(x) = \sqrt{r^2 - x^2}, \qquad x \in [-r, r].$$

By rotating $y(x)$ for $x \in [-r, r]$ about the x-axis we obtain the sphere of radius r as a volume of revolution. From (6.7), the volume of the sphere of radius r is given by

$$\begin{aligned} V &= \pi \int_{-r}^{r} \big[y(x)\big]^2\,dx \;=\; \pi \int_{-r}^{r} \Big[\sqrt{r^2 - x^2}\Big]^2\,dx \;=\; \pi \int_{-r}^{r} (r^2 - x^2)\,dx \\ &= \pi \left(r^2\,x - \frac{x^3}{3}\right)\Bigg|_{-r}^{r} \;=\; \pi \left[\left(r^3 - \frac{r^3}{3}\right) - \left(-r^3 + \frac{r^3}{3}\right)\right] \\ &= \pi \left[2\,r^3 - 2\,\frac{r^3}{3}\right] \;=\; 2\pi\,r^3 \left[1 - \frac{1}{3}\right] \;=\; 2\pi\,r^3\,\frac{3-1}{3} \;=\; \frac{4\pi}{3}\,r^3. \end{aligned}$$

Thus the volume of the sphere of radius of r is given by $V = (4\pi/3)\,r^3$. □

6.5 Mass of a Rod

In this section we learn how to determine the mass of a **rod with constant cross section** and a **density that varies only along the length of the rod**.

Consider a rod with **length** L and **constant cross-sectional area** A as shown in Figure 6.13 below. Let the x-axis run along the length of the rod with the point $x = 0$ at the left end of the rod and the point $x = L$ at the right end of the rod. We assume that the **density of the rod may vary along its length**, that is, the density is a function $\rho(x)$, $x \in [0, L]$, that depends only on the position x along the rod and that is constant for each cross-section.

We want to calculate the **mass of the rod**.

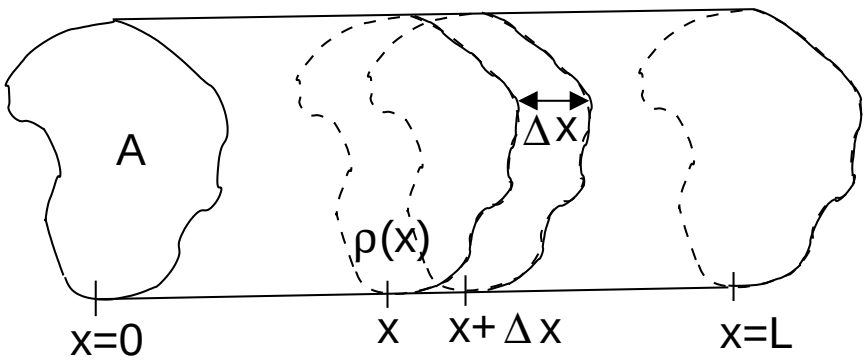


Figure 6.13: Rod with constant cross-section A whose density $\rho(x)$ varies only along the rod (but is constant for each cross section). To determine the mass of the rod we consider a thin slice of the rod with cross-section A and thickness Δx.

For a volume V of constant density, the **mass** is given by

$$\text{mass} \;=\; \text{constant density} \;\times\; \text{volume}$$

The fact that the density is constant on each cross section, gives us the idea to cut the rod vertically/perpendicular to its length into thin slices of thickness Δx. For each such a slice the density is almost constant, and the sum of the masses of all these thin slices gives an approximation of the mass of the rod.

The **slice of the rod between** x **and** $x + \Delta x$ has the **volume**

$$\text{cross-sectional area} \;\times\; \text{thickness of slice} \;=\; A \times \big[(x + \Delta x) - x\big] = A\,\Delta x.$$

The density of the slice is approximately $\rho(x)$. So the **mass of the slice** between x and $x + \Delta x$ is approximately

$$\text{mass of slice} \;\approx\; \text{density} \;\times\; \text{volume} \;\approx\; \rho(x) \times A\,\Delta x = A\,\rho(x)\,\Delta x.$$

The **total mass of the rod is approximately** the sum of the masses $A\,\rho(x)\,\Delta x$ of all the slices of thickness Δx into which we may have cut the rod, which we may write symbolically as

$$\text{mass of rod} \;\approx\; \sum A\,\rho(x)\,\Delta x.$$

If we shrink Δx further and further, that is, for the **limit** $\Delta x \to 0$, the sum becomes the integral and Δx becomes dx. Thus the mass of the rod is given by the **integral of** $A\,\rho(x)$ **from** $x = a$ **and** $x = b$.

$$\boxed{\left(\begin{array}{l}\text{mass of rod with length } L,\\ \text{constant cross-section } A,\\ \text{and density } \rho(x)\end{array}\right) = M = \int_0^L A\,\rho(x)\,dx = A\int_0^L \rho(x)\,dx.} \tag{6.10}$$

We discuss three examples.

Example 6.15 (rod with constant cross-section and constant density)
Find the mass of the rod with length L, constant cross-section A, and constant density $\rho(x) = c$, $x \in [0, L]$.

Solution: From (6.10), the mass is given by

$$M = \int_0^L A\,\rho(x)\,dx = A\int_0^L c\,dx = A\,c\int_0^L 1\,dx = A\,c\,x|_0^L = A\,c\,L.$$

This is what we expect for the constant density $\rho(x) = c$, since in this case

$$M \;=\; \text{volume} \times \text{density} \;=\big(A \times L\big) \times c = A\,L\,c. \qquad \square$$

Example 6.16 (rod with constant cross-section and varying density)
Find the mass of the rod with length L, constant cross-section A, and density $\rho(x) = e^x$, $x \in [0, L]$.

Solution: From (6.10), the mass is given by

$$M = \int_0^L A\,\rho(x)\,dx = A\int_0^L e^x\,dx = A\,e^x\big|_0^L = A\left[e^L - e^0\right] = A\left[e^L - 1\right]. \qquad \square$$

Example 6.17 (rod with constant cross-section and varying density)
Find the mass of the rod with length L, constant cross-section A, and density

$$\rho(x) = \frac{c}{L^2 + x^2}, \qquad 0 \le x \le L.$$

Solution: From (6.10), the mass of the rod is given by

$$M = \int_0^L A\,\rho(x)\,dx = A\int_0^L \frac{c}{L^2 + x^2}\,dx.$$

To find the mass M, we make the substitution

$$x = x(y) = L\,\tan(y) \qquad \Leftrightarrow \qquad y = y(x) = \arctan(x/L).$$

Since $x \in [0, L]$ and we find that $y \in [0, \pi/4]$, and, in particular, the new boundaries of the integral are

$$y(0) = \arctan(0) = 0 \qquad \text{and} \qquad y(L) = \arctan(L/L) = \arctan(1) = \pi/4.$$

Furthermore, we have

$$\frac{dx}{dy} = \frac{d}{dy}\big[L\,\tan(y)\big] = \frac{L}{\big[\cos(y)\big]^2} \qquad \Rightarrow \qquad dx = \frac{L}{\big[\cos(y)\big]^2}\,dy,$$

and the denominator simplifies through the substitution $x = L\,\tan(y)$ to

$$\begin{aligned} L^2 + x^2 &= L^2 + L^2\left[\tan(y)\right]^2 = L^2\left(1 + \frac{\left[\sin(y)\right]^2}{\left[\cos(y)\right]^2}\right) \\ &= L^2\,\frac{\left[\cos(y)\right]^2 + \left[\sin(y)\right]^2}{\left[\cos(y)\right]^2} = \frac{L^2}{\left[\cos(y)\right]^2}, \end{aligned}$$

where we have used $\left[\cos(y)\right]^2 + \left[\sin(y)\right]^2 = 1$. Thus we find with the substitution $x = L\,\tan(y)$

$$\begin{aligned} M &= A\int_0^L \frac{c}{L^2 + x^2}\,dx = A\int_0^{\pi/4} \frac{c\left[\cos(y)\right]^2}{L^2}\,\frac{L}{\left[\cos(y)\right]^2}\,dy \\ &= \frac{A\,c}{L}\int_0^{\pi/4} 1\,dy = \frac{A\,c}{L}\,y\big|_0^{\pi/4} = \frac{A\,c}{L}\,\frac{\pi}{4} = \frac{\pi\,A\,c}{4\,L}. \end{aligned}$$

Thus the mass of the rod is given by $M = (\pi\,A\,c)/(4\,L)$. □

6.6 Centre of Mass

In this section we discuss how to determine the **x-coordinate of the centre of mass** of a body (along the x-axis) for which the length L, the cross-sectional area $A(x)$, and the density $\rho(x)$ are given.

To introduce the concept of the centre of mass we start with the most simple case, namely, a collection of particles that are located on the x-axis: Let n particles with masses $m_1, m_2, \ldots, m_n$ be placed at positions $x_1, x_2, \ldots, x_n$ along the x-axis, then their **centre of mass** is defined to be at the position x_c given by

$$\boxed{\text{centre of mass} = x_c = \frac{\displaystyle\sum_{k=1}^{n} m_k\,x_k}{\displaystyle\sum_{j=1}^{n} m_j} = \sum_{k=1}^{n} \frac{m_k}{\left(\sum_{j=1}^{n} m_j\right)}\,x_k.} \tag{6.11}$$

We observe that the denominator just contains the sum of the masses of all particles, whereas in the numerator the masses are multiplied by their x-coordinates before the sum is taken. In the right-most representation of the centre of mass in (6.11), we can interpret $m_k/\left(\sum_{j=1}^{n} m_j\right)$ as a 'weighting factor' for the coordinate x_k. Thus

the x-coordinate x_k of a particle with a large mass m_k carries 'more weight' for determining the centre of mass than the x-coordinate of a particle with a small mass. This is just as we expect.

Now we want to extend this idea to a **continuous body** arranged along the x-axis, with length L, cross-sectional area $A(x)$, $x \in [0, L]$, and density $\rho(x)$, $x \in [0, L]$. See Figure 6.14 for illustration, and note that we do **not** assume that the body has constant cross-sectional area, but the **cross-sectional area** $A(x)$ **is allowed to vary with** x. However, we still assume that the density $\rho(x)$ depends only on the x-coordinate.

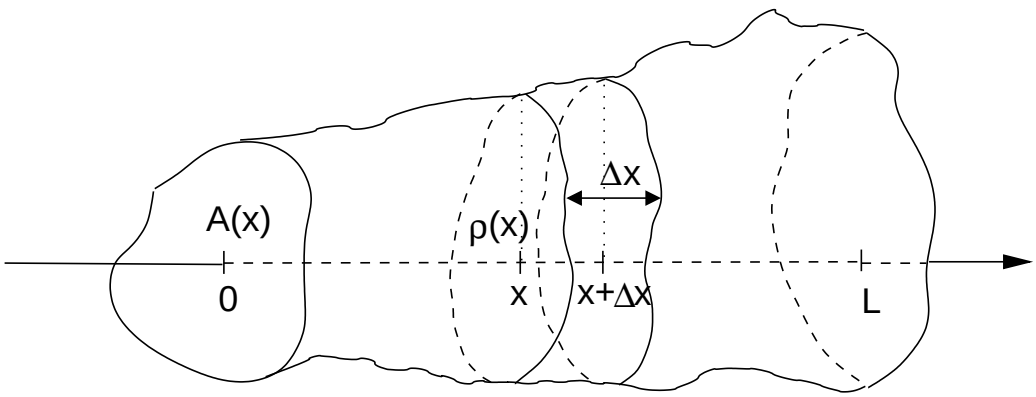


Figure 6.14: A body of length L aligned along the x-axis with cross-sectional area $A(x)$ and density $\rho(x)$.

Since the density depends only on x, we can get an **approximation of the mass of the body** by slicing the body perpendicular to the x-axis into n thin slices of thickness $\Delta x = L/n$ and **summing up the masses of all these slices**. For the slice between the x-coordinates x and $x + \Delta x$ (see Figure 6.14), the **volume of the slice** is approximately

$$\text{cross-sectional area at } x \;\times\; \text{thickness of slice} \;=\; A(x) \times \big[(x + \Delta x) - x\big] = A(x)\,\Delta x.$$

Since the density of the slice is approximately $\rho(x)$, the **mass of the slice** is approximately

$$\text{mass of slice} \;\approx\; \text{density} \;\times\; \text{volume} \;\approx\; \rho(x) \times A(x)\,\Delta x = A(x)\,\rho(x)\,\Delta x. \qquad (6.12)$$

Now we sum up the masses (6.12) of all the n slices into which we have cut the solid body, and we obtain, with $\Delta x = L/n$

$$\text{mass of solid body} \;=\; M \;\approx\; \sum_{k=1}^{n} A\big((k-1)\,\Delta x\big)\,\rho\big((k-1)\,\Delta x\big)\,\Delta x$$

$$= \sum_{k=1}^{n} A\left((k-1)\,\frac{L}{n}\right)\,\rho\left((k-1)\,\frac{L}{n}\right)\,\frac{L}{n}.$$

Less formal and symbolically we could just write

$$M \approx \sum A(x)\,\rho(x)\,\Delta x. \tag{6.13}$$

If we shrink the width of the slices, that is, take the limit $\Delta x \to 0$, then we obtain the exact mass of the solid body, and in (6.13) the sum is replaced by the integral over $[0, L]$ and Δx is replaced by dx. Thus the **mass of the solid body with length L, cross-sectional area $A(x)$, and density** $\rho(x)$ is given by

$$\boxed{\begin{pmatrix}\text{mass of solid body with length } L,\\ \text{cross-section } A(x), \text{ and density } \rho(x)\end{pmatrix} = M = \int_0^L A(x)\,\rho(x)\,dx.} \tag{6.14}$$

Coming back to (6.12), we see that the **contribution of one slice to the numerator in (6.11)** is approximately

$$\text{mass of the slice} \;\times\; x\text{-coordinate} \;=\; \big(A(x)\,\rho(x)\,\Delta x\big) \times x = A(x)\,\rho(x)\,x\,\Delta x. \tag{6.15}$$

Thus the **numerator in (6.11)** is approximately the sum of the quantities (6.15) for all slices, that is,

$$\begin{aligned}\text{numerator in (6.11)} \;&\approx\; \sum_{k=1}^{n} A\big((k-1)\,\Delta x\big)\,\rho\big((k-1)\,\Delta x\big)\,\big((k-1)\,\Delta x\big)\,\Delta x\\ &=\; \sum_{k=1}^{n} A\left((k-1)\,\frac{L}{n}\right)\,\rho\left((k-1)\,\frac{L}{n}\right)\,\left((k-1)\,\frac{L}{n}\right)\,\frac{L}{n},\end{aligned}$$

or in the laxer symbolic notation

$$\text{numerator in (6.11)} \;\approx\; \sum A(x)\,\rho(x)\,x\,\Delta x. \tag{6.16}$$

Taking as before the limit $\Delta x \to 0$ yields that in (6.16) the sum is replaced by the integral and Δx by dx, and thus we obtain

$$\int_0^L A(x)\,\rho(x)\,x\,dx. \tag{6.17}$$

From (6.14) and (6.17), we obtain that the **x-coordinate of the centre of mass of the solid body** is given by

$$\boxed{\left(\begin{array}{l} x\text{-coordinate of centre of mass} \\ \text{of a body with length } L\text{, cross-} \\ \text{section } A(x)\text{, and density } \rho(x) \end{array}\right) = x_c = \frac{\displaystyle\int_0^L A(x)\,\rho(x)\,x\,dx}{\displaystyle\int_0^L A(x)\,\rho(x)\,dx}.} \tag{6.18}$$

Example 6.18 (x-coordinate of the centre of mass of a cylinder)
Determine the x-coordinate of the centre of mass of a cylinder of radius r and length L (see Figure 6.15), whose density $\rho(x)$ varies only along the length $x \in [0, L]$ of the cylinder and is given by

$$\rho(x) = 1 + x^2, \qquad x \in [0, L].$$

Solution:

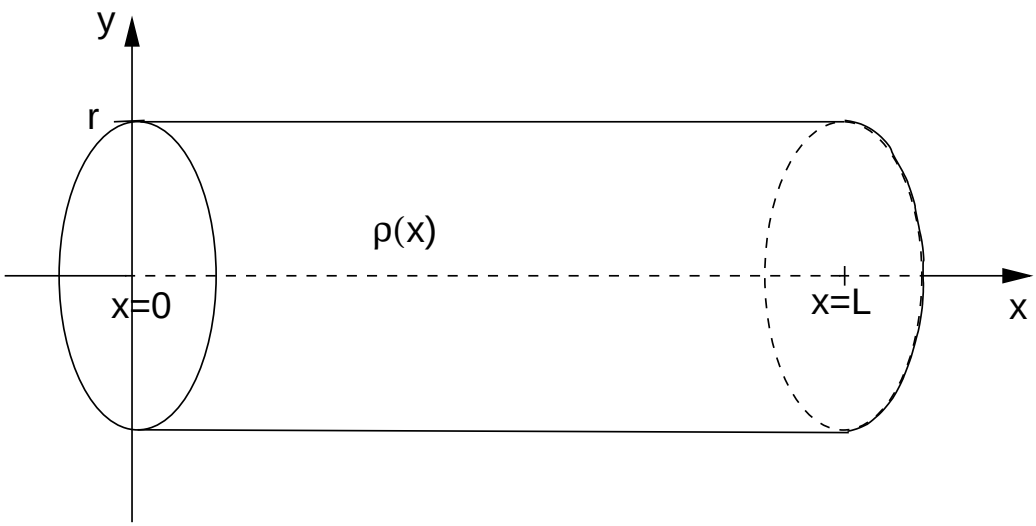


Figure 6.15: The cylinder generated by rotation of the constant function $f(x) = r$ for $x \in [0, L]$ about the x-axis.

The cross-sectional area of the cylinder is given by

$$A(x) = \pi\, r^2$$

and thus the mass of the cylinder is given by

$$\begin{aligned} M &= \int_0^L A(x)\,\rho(x)\,dx = \int_0^L \pi\, r^2 \left(1 + x^2\right) dx = \pi\, r^2 \int_0^L \left(1 + x^2\right) dx \\ &= \pi\, r^2 \left[x + \frac{x^3}{3}\right]\Bigg|_0^L = \pi\, r^2 \left(\left[L + \frac{L^3}{3}\right] - \left[0 + \frac{0^3}{3}\right]\right) \end{aligned}$$

$$= \pi r^2 L \left[1 + \frac{L^2}{3}\right] = \frac{\pi r^2 L}{3} \left[3 + L^2\right]. \tag{6.19}$$

Furthermore, the numerator of (6.18) becomes

$$\begin{aligned}
\int_0^L A(x)\,\rho(x)\,x\,dx &= \int_0^L \pi r^2 \left(1 + x^2\right) x\,dx = \pi r^2 \int_0^L \left(x + x^3\right) dx \\
&= \pi r^2 \left[\frac{x^2}{2} + \frac{x^4}{4}\right]\Bigg|_0^L = \pi r^2 \left(\left[\frac{L^2}{2} + \frac{L^4}{4}\right] - \left[\frac{0^2}{2} + \frac{0^4}{4}\right]\right) \\
&= \pi r^2 \frac{L^2}{2} \left[1 + \frac{L^2}{2}\right] = \frac{\pi r^2 L^2}{4} \left[2 + L^2\right]. \qquad (6.20)
\end{aligned}$$

Thus from (6.18) and (6.19) and (6.20), the x-coordinate of the centre of mass is given by

$$x_c = \frac{\displaystyle\int_0^L A(x)\,\rho(x)\,x\,dx}{\displaystyle\int_0^L A(x)\,\rho(x)\,dx} = \frac{\pi r^2 L^2 \left[2 + L^2\right]/4}{\pi r^2 L \left[3 + L^2\right]/3} = \frac{3\,L}{4} \frac{\left[2 + L^2\right]}{\left[3 + L^2\right]}. \qquad \square$$

Example 6.19 (x-coordinate of centre of mass of cone with const. density) Find the position of the centre of mass of a solid cone with base radius r, height h and constant density $\rho(x) = c$, $x \in [0, h]$.

Solution:

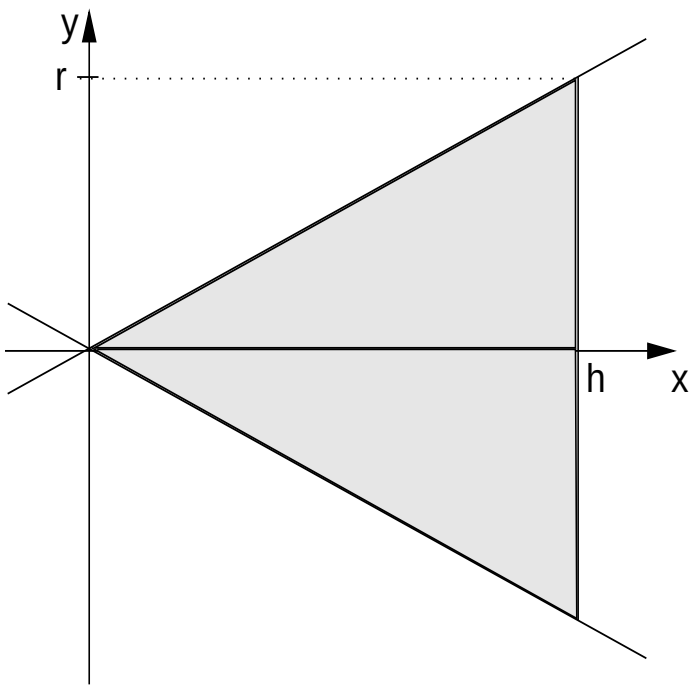


Figure 6.16: The cone generated by rotation of $f(x) = (r/h)\,x$ for $x \in [0, h]$ about the x-axis.

The solid cone with base radius r, height h can be generated as a volume of revolution by rotating the segment of the graph of

$$f(x) = \frac{r}{h}\,x$$

between $x = 0$ and $x = h$ about the x-axis. See Figure 6.16 and Example 6.13 for details. Thus the cross-sectional area is given by

$$A(x) = \pi\,\big[f(x)\big]^2 = \pi\,\left(\frac{r}{h}\,x\right)^2 = \pi\,\frac{r^2}{h^2}\,x^2.$$

Thus we find that the mass of the cone is

$$\begin{aligned} M \;=\; \int_0^h A(x)\,\rho(x)\,dx \;&=\; \int_0^h \pi\,\frac{r^2}{h^2}\,x^2\,c\,dx \;=\; \frac{c\,\pi\,r^2}{h^2}\int_0^h x^2\,dx \\ &=\; \frac{c\,\pi\,r^2}{h^2}\,\frac{x^3}{3}\bigg|_0^h \;=\; \frac{c\,\pi\,r^2}{h^2}\left(\frac{h^3}{3} - 0\right) \;=\; \frac{c\,\pi\,r^2\,h}{3}, \end{aligned}$$

and

$$\begin{aligned} \int_0^h A(x)\,\rho(x)\,x\,dx \;&=\; \int_0^h \pi\,\frac{r^2}{h^2}\,x^2\,c\,x\,dx \;=\; \frac{c\,\pi\,r^2}{h^2}\int_0^h x^3\,dx \\ &=\; \frac{c\,\pi\,r^2}{h^2}\,\frac{x^4}{4}\bigg|_0^h \;=\; \frac{c\,\pi\,r^2}{h^2}\left(\frac{h^4}{4} - 0\right) \;=\; \frac{c\,\pi\,r^2\,h^2}{4}. \end{aligned}$$

So the centre of mass is at

$$x_c = \frac{\displaystyle\int_0^L A(x)\,\rho(x)\,x\,dx}{\displaystyle\int_0^L A(x)\,\rho(x)\,dx} = \frac{c\,\pi\,r^2\,h^2/4}{c\,\pi\,r^2\,h/3} = \frac{h/4}{1/3} = \frac{3\,h}{4}.$$

So the centre of mass of the cone is 3/4 of the way down from the apex to the base of the cone, that is, 1/4 of the way up from the base to the apex of the cone. □

6.7 Moments of Inertia

In this section we discuss **moments of inertia**. Intuitively the moment of inertia of a body is the **'resistance' that the body offers** (and which we have to overcome) **to being set to rotate about a given axis**. The moment of inertia of a body

depends on the location of the axis of rotation and on the distribution of the mass over the body and its perpendicular distance to the axis of rotation.

For example, suppose we have two flywheels A and B of identical size and mass, but wheel A has most of its mass near its rim while wheel B has most of its mass near its centre. It is observed experimentally that a higher force is required to rotate wheel A at constant angular speed than to rotate wheel B at the same angular speed. This is a consequence of wheel A having a higher **moment of inertia** than wheel B.

If n particles with masses $m_1, m_2, \ldots, m_n$ are at perpendicular distances $r_1, r_2, \ldots, r_n$ from an axis of rotation, then their **moment of inertia** about that axis is defined to be

$$\boxed{\text{moment of inertia} \;=\; I \;=\; \sum_{k=1}^{n} m_k\, r_k^2.} \tag{6.21}$$

We may think of the moment of inertia as the **'resistance' of the arrangement of the n masses against being made to rotate about the axis of rotation**.

In reality, **mass is usually distributed over an extended body** and we want to **determine the moment of inertia of such a body about any axis of rotation inside or outside the body**. To derive a formula for the moment of inertia of an extended body about a given axis of rotation, we start with a standard example, namely a thin rectangular bar rotating about an axis located at one of its endpoints.

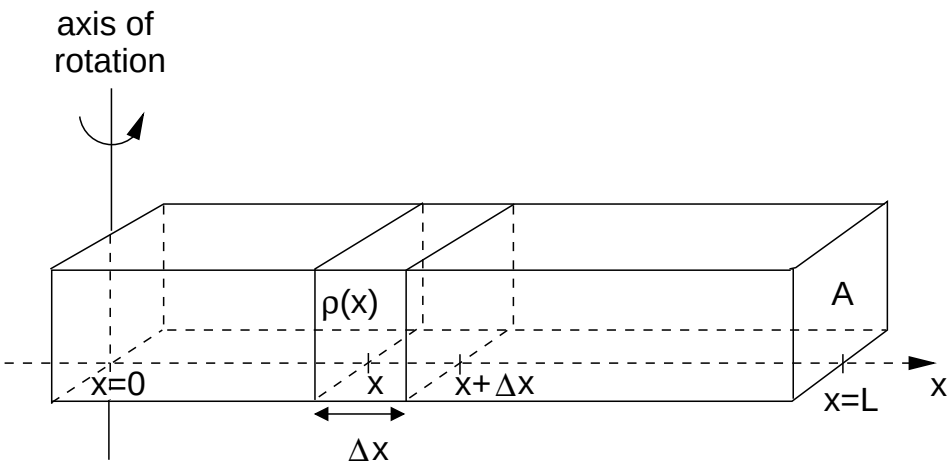


Figure 6.17: A thin rectangular bar of length L and constant cross-section A, whose density $\rho(x)$ varies only along the bar. To determine the moment of inertia of the bar about the axis through the left end of the bar, we slice the bar into thin slices of thickness Δx and sum up the moments of inertia of the thin slices.

Consider **a thin rectangular bar** of length L with **constant cross-section** A, and assume that the cross-section A is small compared to the length L of the bar. We lay our x-axis along the length of the bar such that the left end the bar is at $x = 0$ and the right end of the bar is at $x = L$. We assume that the **density of the bar varies only along the bar**, that is, the density is a function $\rho(x)$. See Figure 6.17 for illustration.

Now we want to determine the **moment of inertia of the bar about the vertical axis located at its left end**, that is, about the vertical axis through $x = 0$ (see Figure 6.17). Since the cross-section A is constant and small compared to the length of the bar and since the density $\rho(x)$ depends only on x, we may approximate the moment of inertia as follows: We slice the rectangular bar perpendicular to its length into n thin slices of thickness $\Delta x = L/n$ (with n large), and **sum up the moments of inertia for all the individual slices** (see Figure 6.17).

Such a slice, going from x-coordinates x to $x + \Delta x$, has cross-sectional area A, and approximately the density $\rho(x)$. Thus its mass is approximately

$$\text{mass of slice} \;\approx\; \text{volume} \;\times\; \text{density} \;\approx\; A\,\big[(x+\Delta x) - x\big] \times \rho(x) = A\,\rho(x)\,\Delta x.$$

Therefore the **moment of inertia of such a slice** is approximately

$$\text{mass of slice} \;\times\; \begin{pmatrix}\text{perpendicular distance}\\ \text{from axis of rotation}\end{pmatrix}^2 \approx \big[A\,\rho(x)\,\Delta x\big] \times x^2 = A\,\rho(x)\,x^2\,\Delta x.$$

To get an approximation of the moment of inertia of the rectangular bar about the vertical rotational axis through $x = 0$, we can **sum up the moments of inertia of the slices**. In the symbolic notation we have

$$\begin{pmatrix}\text{moment of inertia of the bar of length } L\\ \text{about the vertical axis through } x = 0\end{pmatrix} \;=\; I \;\approx\; \sum A\,\rho(x)\,x^2\,\Delta x.$$

If we shrink the thickness Δx of the slices then the approximation improves, and, for the **limit** $\Delta x \to 0$, the sum becomes the integral and Δx is replaced by dx. Thus the **moment of inertia** of the rectangular bar **about the vertical axis through** $x = 0$ is given by

$$\boxed{\begin{pmatrix}\text{moment of inertia of bar}\\ \text{of length } L \text{ about vertical}\\ \text{axis through } x = 0\end{pmatrix} = I = \int_0^L A\,\rho(x)\,x^2\,dx = A\int_0^L \rho(x)\,x^2\,dx.} \tag{6.22}$$

If we change the location of the axis of rotation from $x = 0$ to the point $x = a$ (that is, we shift the axis of rotation horizontally by a), then the factor x^2 in (6.22) needs

to be replaced by $(x-a)^2$, since $x-a$ is now the perpendicular distance from the axis of rotation. Thus the **moment or inertia** of the rectangular bar **about the vertical axis through** $x=a$ is given by

$$\boxed{\left(\begin{array}{l}\text{moment of inertia of bar}\\ \text{of length } L \text{ about vertical}\\ \text{axis through } x=a\end{array}\right) = I = \int_0^L A\,\rho(x)\,(x-a)^2\,dx.} \tag{6.23}$$

Example 6.20 (moment of inertia of rectangular bar)
Find the moment of inertia of a thin rectangular bar of length L, constant cross-sectional area A and constant density $\rho(x)=c$, for rotation about the vertical axis through (a) an end point of the bar and (b) the middle of the bar.

Solution: The moment of inertia is given by (6.23) with $\rho(x)=c$ and (a) $a=0$ and (b) $a=L/2$. Thus we find that the moment of inertia about the vertical axis through an endpoint is

$$I = \int_0^L A\,c\,x^2\,dx = A\,c\int_0^L x^2\,dx = A\,c\,\frac{x^3}{3}\bigg|_0^L = \frac{1}{3}\,A\,c\,L^3,$$

and the moment of inertia about the vertical axis through the middle of the bar is

$$\begin{aligned} I &= \int_0^L A\,c\left(x-\frac{L}{2}\right)^2 dx = A\,c\int_0^L \left(x-\frac{L}{2}\right)^2 dx = A\,c\,\frac{1}{3}\left(x-\frac{L}{2}\right)^3\bigg|_0^L \\ &= A\,c\,\frac{1}{3}\left[\left(\frac{L}{2}\right)^3 - \left(-\frac{L}{2}\right)^3\right] = A\,c\,\frac{1}{3}\left[\frac{L^3}{8}+\frac{L^3}{8}\right] = A\,c\,\frac{2}{3}\,\frac{L^3}{8} = \frac{1}{12}\,A\,c\,L^3. \end{aligned}$$

We note that the moment of inertia about a vertical rotational axis through the middle of the bar is only one quarter of the moment of inertia about the vertical rotational axis through one endpoint of the bar. □

The second application that we discuss is how to compute the the **moment of inertia of hollow and solid cylinders for rotation about their axis of symmetry**.

Consider a very **thin hollow cylinder (a tube)** with radius R and mass M which is uniformly distributed over the hollow cylinder. We want to compute the moment of inertia about the axis of circular symmetry as rotational axis. Since the thin hollow cylinder has all its mass at distance R from its axis of circular symmetry, we may calculate the moment of inertia by slicing the hollow cylinder vertically into n thin stripes each of equal volume and equal mass $\Delta m = M/n$ (see Figure 6.18).

Each of these stripes has perpendicular distance R to the axis of circular symmetry. Thus the **moment of inertia of the hollow cylinder about its axis of circular symmetry** is given by

$$\boxed{\begin{pmatrix} \text{moment of interia of the} \\ \text{hollow cylinder about its} \\ \text{axis of circular symmetry} \end{pmatrix} = I = \sum_{k=1}^{n} \Delta m\, R^2 = R^2 \sum_{k=1}^{n} \Delta m = R^2\, M.} \tag{6.24}$$

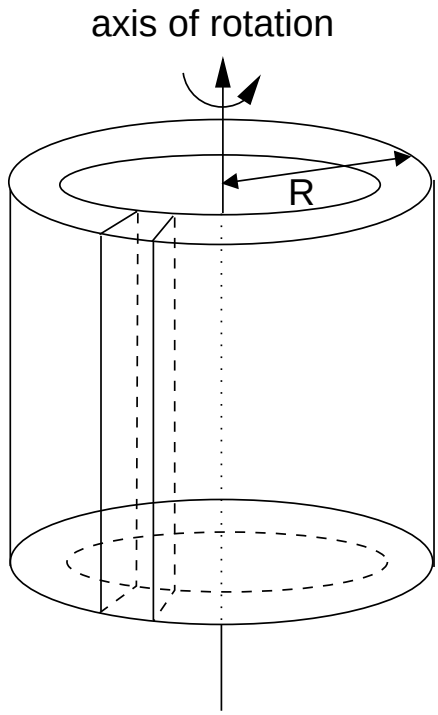


Figure 6.18: Thin hollow cylinder with radius R and mass M. To determine the moment of inertia about the axis of circular symmetry, we cut the cylinder vertically into n stripes of equal volume and mass and sum up their moments of inertia.

Now we consider a **solid cylinder** with length L, radius R, and constant density $\rho(x) = c$. The **cylindrical shell** with thickness Δx at a distance x from its axis of symmetry (see Figure 6.19) has the **volume**

$$\begin{aligned} \text{volume of cylindrical shell} &= \pi\,(x + \Delta x)^2\, L - \pi\, x^2\, L \\ &= \pi\left[x^2 + 2\,x\,\Delta x + (\Delta x)^2\right] L - \pi x^2 L \\ &= \pi\left[2\,x\,\Delta x + (\Delta x)^2\right] L \\ &\approx 2\,\pi\, L\, x\, \Delta x. \end{aligned}$$

In the last step we used the fact that for very small Δx, the term $(\Delta x)^2$ is much

smaller than $2\,x\,\Delta x$ and can thus be neglected. Hence the **mass of the shell** is

$$\text{mass of shell} \;\approx\; \text{volume} \;\times\; \text{density} \;\approx\; 2\,\pi\,L\,x\,\Delta x \times \rho(x) = 2\,\pi\,c\,L\,x\,\Delta x,$$

where we have used $\rho(x) = c$.

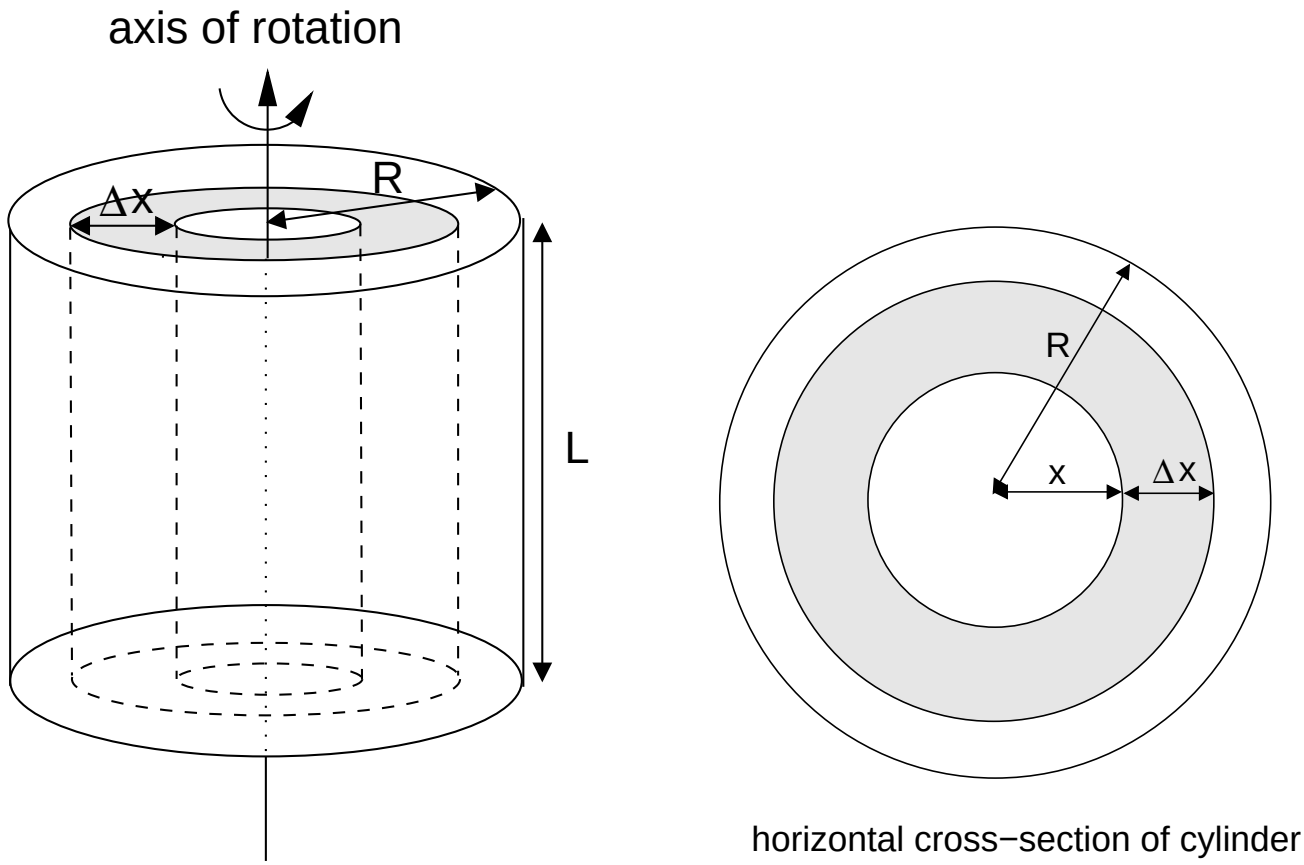


Figure 6.19: A cylindrical shell of the solid cylinder with radius R on the left. In the right picture, we see a horizontal cross-section through the cylinder.

Then the **moment of inertia of the cylinder** about its axis of circular symmetry is approximately the sum of the moments of inertia of all the individual thin shells, that is, the sum of

$$\text{moment of inertia of the shell} \;=\; \big(2\,\pi\,c\,L\,x\,\Delta x\big) \times x^2 = 2\,\pi\,c\,L\,x^3\,\Delta x$$

taken as x changes from $x = 0$ to $x = r - \Delta x$ in jumps of length $\Delta x = R/n$ with n large. In the laxer symbolic notation, the approximation of the moment of inertia of the cylinder about its axis of circular symmetry is given by

$$\sum 2\,\pi\,c\,L\,x^3\,\Delta x.$$

In the **limit** as the thickness Δx goes to zero, we recover the actual moment of inertia of the cylinder. The sum goes over into the integral and Δx is replaced by dx. Thus the **moment of inertia of the solid cylinder** with radius R, height L, and constant density $\rho(x) = c$ **about its axis of circular symmetry** is given by

$$\boxed{\left(\begin{array}{l}\text{moment of inertia of solid}\\ \text{cylinder about its axis of}\\ \text{circular symmetry}\end{array}\right) = \int_0^R 2\,\pi\,c\,L\,x^3\,dx = 2\,\pi\,c\,L\int_0^R x^3\,dx.} \tag{6.25}$$

We work out the integral in (6.25), and find

$$2\,\pi\,c\,L\int_0^R x^3\,dx = 2\,\pi\,c\,L\,\frac{x^4}{4}\bigg|_0^R = 2\,\pi\,c\,L\,\frac{R^4}{4} = \frac{\pi\,c\,L\,R^4}{2}. \tag{6.26}$$

The cross-sectional area of the cylinder is $A = \pi\,R^2$, and so the total mass of the cylinder is given by

$$M = \int_0^L A\,\rho(y)\,dy = \int_0^L \left(\pi\,R^2\right)c\,dy = \pi\,R^2\,c\int_0^L dy = \pi\,R^2\,c\,y|_0^L = \pi\,c\,L\,R^2.$$

Thus we find from (6.25) and (6.26) that

$$\boxed{\left(\begin{array}{l}\text{moment of inertia of solid}\\ \text{cylinder about its axis of}\\ \text{circular symmetry}\end{array}\right) = \frac{c\,\pi\,L\,R^4}{2} = \frac{1}{2}\,M\,R^2.} \tag{6.27}$$

Comparing the moment of inertia of a solid cylinder with radius R, mass M, and constant density $\rho(x) = c$, about its axis of symmetry with the moment of inertia of a hollow cylinder of the same mass and same radius about its axis of symmetry, we find that the **moment of inertia of the solid cylinder is half that of the hollow cylinder with the same radius and mass**.

Chapter 7

Series Expansions and Approximations

In this chapter we discuss **(finite) sums** and **(infinite) series**. An infinite series is essentially an infinite sum, and we will have to inspect the series to see whether it converges, that is, has a finite value. Among the infinite series, we will discuss both **series of real numbers** but also so-called **power series**, and we will learn how to **expand a function into a power series** about a given point. The first few terms of such an expansion provide often a **good approximation** of the function close to the point of expansion.

In Section 7.1, we will introduce **finite sums**, and we will discuss **arithmetic progression** and **geometric progression**. In Section 7.2, we will discuss the **binomial formula** and two ways to determine the coefficients obtained when multiplying the factors in $(a+b)^n$ and writing $(a+b)^n$ as a sum: **Pascal's triangle** which you may know from school, and the representation of $(a+b)^n$ as a sum with the help of the so-called **binomial coefficients** which can be computed directly (rather than recursively as done with Pascal's triangle).

In Section 7.3, we introduce **infinite series of real numbers**

$$\sum_{k=m}^{\infty} a_k = a_m + a_{m+1} + \ldots + a_k + \ldots = \lim_{n\to\infty} \underbrace{\sum_{k=m}^{n} a_k}_{= S_n} . \tag{7.1}$$

As the representation on the right in (7.1) implies you should really think of an

infinite series as the **limit** of the sequence $\{S_n\}$ of **partial sums**

$$S_n = \sum_{k=m}^{n} a_k.$$

The infinite series **converges if the sequence** $\{S_n\}$ converges, and the **infinite series does not converge** – we say it **diverges** – if the sequence $\{S_n\}$ does not converge. What does it intuitively mean if the series (7.1) converges? If the series (7.1) converges and $S = \lim_{n\to\infty} S_n$, then, roughly, the larger n the closer S_n gets to the value S of the series. Thus for large n, the finite partial sum S_n will be a good approximation of the value S of the series.

In Section 7.4, we introduce so-called **power series**

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k = a_0 + a_1 (x - x_0) + a_2 (x - x_0)^2 + \ldots + a_k (x - x_0)^k + \ldots, \quad (7.2)$$

where here x_0 and $a_0, a_1, a_2, \ldots$ are constants. The expression (7.2) is called a power series because we sum up **powers** of $(x - x_0)$ (multiplied with some constants). For each fixed x we have just an infinite sum of the real numbers $a_k (x - x_0)^k$ for $k = 0, 1, 2, 3, \ldots$. However, you should think of the power series (7.2) as a **function of the variable** x. The **partial sums** of the power series are

$$S_n(x) = \sum_{k=0}^{n} a_k (x - x_0)^k,$$

and we see that the n-th partial $S_n(x)$ sum is just a **polynomial of degree** n. The value of the power series at a fixed x is given by taking the limit of the sequence of partial sums $\{S_n(x)\}$ at x, that is,

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k = \lim_{n\to\infty} S_n(x) = \lim_{n\to\infty} \sum_{k=0}^{n} a_k (x - x_0)^k. \quad (7.3)$$

If we think in (7.3) of x as a variable we see that, for those x at which the power series converges, the function given by the power series is **approximated by the sequence** $\{S_n\}$ **of polynomials** S_n.

Very important examples of power series are the so-called **Taylor series** of a function about x_0 and, as a special case, the **Maclaurin series** which is the Taylor series about $x_0 = 0$. As the notion 'Taylor series of a function about x_0' implies a Taylor series is computed for a given function and its partial sums, the **Taylor**

polynomials, will provide approximations of the function. The simplest case of a Taylor series is the **Maclaurin series** of a smooth function f, which is given by

$$T(f,0)(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!}\, x^k = f(0) + f'(0)\, x + \frac{f''(0)}{2}\, x^2 + \frac{f'''(0)}{6}\, x^3 + \dots . \tag{7.4}$$

We see that the coefficients of the powers of x involve the values of f and its derivatives at $x_0 = 0$, and for $x = x_0 = 0$ the series has the value $f(0)$. In Section 7.6, we will learn that the **Taylor polynomials** of a function f about x_0 (that is, the partial sums of the Taylor/Maclaurin series) provide an approximation of $f(x)$ for x close to x_0. For the Maclaurin series (7.4), the Taylor polynomial of degree n is given by

$$S_n(f,0)(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}\, x^k = f(0) + f'(0)\, x + \frac{f''(0)}{2}\, x^2 + \dots + \frac{f^{(n)}(0)}{n!}\, x^n. \tag{7.5}$$

In engineering and physics, the **approximation of a function f by a Taylor polynomial of low degree** is a very useful and common tool. Low degree means here often degree zero, one, two, or three, that is, we approximate $f(x)$ close to x_0 by a constant, a linear function, a quadratic function, or a cubic function, respectively. We will learn a **'rule of thumb'** for estimating the error of the approximation of $f(x)$ by its Taylor polynomial of degree n.

In Section 7.7 we learn that we can **differentiate a power series term by term** and **integrate a power series term by term**, that is, we may move the differentiation and the integral, respectively, behind the summation sign.

7.1 Finite Sums, and Arithmetic and Geometric Progressions

Before we discuss two particular types of **finite sums** (or **finite series**), namely **arithmetic progression** and **geometric progression**, we will briefly introduce sums and the notation with the summation sign (even though we have used the summation sign already occasionally in the last chapters).

Definition 7.1 (finite sum (or finite series))
*Let m and n be integers, and let $m \leq n$. The **finite sum** (or **finite series**)*

$$\sum_{k=m}^{n} a_k$$

is defined by

$$\sum_{k=m}^{n} a_k = a_m + a_{m+1} + \ldots + a_k + \ldots + a_{n-1} + a_n, \tag{7.6}$$

*where the a_k, for $k = m, m+1, \ldots, n-1, n$, are real numbers. Usually but not always we will have that $m \geq 0$. The symbol $\sum$ is called the **summation sign**. The **index** k in (7.6) is a 'placeholder' for the values $m, m+1, \ldots, n-1, n$. The summation index could be given any other name, for example j, and we have*

$$\sum_{k=m}^{n} a_k = \sum_{j=m}^{n} a_j.$$

We give some examples of finite sums/finite series.

Example 7.2 (finite sums/finite series)
Compute the following finite sums/finite series:

(a) $\displaystyle\sum_{k=1}^{6} k$, (b) $\displaystyle\sum_{j=3}^{5} j^2$, (c) $\displaystyle\sum_{r=2}^{4} c$, where $c \in \mathbb{R}$ is a constant.

Solution: We work out the sums in (a) and (b) and find

$$\text{(a)} \quad \sum_{k=1}^{6} k = 1 + 2 + 3 + 4 + 5 + 6 = 21,$$

and

$$\text{(b)} \quad \sum_{j=3}^{5} j^2 = 3^2 + 4^2 + 5^2 = 9 + 16 + 25 = 50.$$

The third sum is a 'bit peculiar' in that we sum up the constant c which does **not** actually depend on r. This means that for each value of the index r we add the constant c, that is,

$$\sum_{r=2}^{4} c = \underset{\substack{\uparrow \\ r=2}}{c} + \underset{\substack{\uparrow \\ r=3}}{c} + \underset{\substack{\uparrow \\ r=4}}{c} = 3\,c. \qquad \square$$

From generalizing Example 7.2 (c), we find that for $m \leq n$ and any constant $c \in \mathbb{R}$

$$\boxed{\sum_{k=m}^{n} c = (n - m + 1)\, c.} \tag{7.7}$$

Lemma 7.3 (properties of sums)

$$\sum_{k=m}^{n} \big(a_k + b_k\big) = \sum_{k=m}^{n} a_k + \sum_{k=m}^{n} b_k, \tag{7.8}$$

and for any constant $c \in \mathbb{R}$

$$\sum_{k=m}^{n} \big(c\, a_k\big) = c \sum_{k=m}^{n} a_k. \tag{7.9}$$

In the next example, we will compute a well-known sum in which the summation index k goes up to an arbitrary integer n.

Example 7.4 (value of $\sum_{k=1}^{n} k$)
Let $n > 0$ be an arbitrary integer. Compute the sum

$$S_n = \sum_{k=1}^{n} k,$$

Solution: To find $S_n = \sum_{k=1}^{n} k$, we write the sum S_n explicitly as

$$S_n = 1 + 2 + 3 + \ldots + (n-1) + n, \tag{7.10}$$

and we rewrite the sum S_n also backwards

$$S_n = n + (n-1) + (n-2) + \ldots + 2 + 1. \tag{7.11}$$

Then we add (7.10) and (7.11) such that we add the first terms in both sums together, add the second terms in both sums together, and so on

$$\begin{aligned} 2\, S_n &= \big(1+n\big) + \big(2 + (n-1)\big) + \big(3 + (n-2)\big) + \ldots + \big((n-1) + 2\big) + \big(n+1\big) \\ &= \underbrace{(n+1) + (n+1) + (n+1) + \ldots + (n+1) + (n+1)}_{n \text{ of these}} \\ &= n\,(n+1). \end{aligned}$$

Dividing by 2, yields finally

$$\boxed{S_n = \sum_{k=1}^{n} k = \frac{1}{2}\, n\,(n+1).} \tag{7.12}$$

We will need (7.12) when we work out a formula for the arithmetic progression below. □

We call the expression on the right-hand side of (7.12) a **closed representation** or **closed formula** for the sum on the left-hand side. The term 'closed' means here that the new representation does not contain any summation but is just one expression.

Now we will discuss **arithmetic progression** and **geometric progression**.

Definition 7.5 (arithmetic progression)
A series of the form

$$S_n = \sum_{k=0}^{n} \big(a + k\,h\big) = a + \big(a+h\big) + \big(a+2\,h\big) + \ldots + \big(a+k\,h\big) + \ldots + \big(a+n\,h\big),$$

where $a \in \mathbb{R}$ and $h \in \mathbb{R}$ with $h > 0$, is called an ***arithmetic progression*** *(or* ***arithmetic sum****). We observe that the* ***difference*** *between any two consecutive terms (where we subtract the term with the larger index from its predecessor) is*

$$\big(a + k\,h\big) - \big(a + (k-1)\,h\big) = h.$$

For this reason h is called the ***common difference****.*

Now we want to work out a closed representation for the arithmetic progression. Using (7.8), (7.9), and (7.7), we find that

$$S_n = \sum_{k=0}^{n} \big(a + k\,h\big) = \sum_{k=0}^{n} a + \sum_{k=0}^{n} k\,h = \sum_{k=0}^{n} a + h \sum_{k=0}^{n} k = (n+1)\,a + h \sum_{k=1}^{n} k,$$

where we have used in the last step that we may start the second sum with $k = 1$, since for $k = 0$ the term k is zero. Now we use (7.12) to get a closed representation of the remaining sum and find

$$S_n = \sum_{k=0}^{n} \big(a+k\,h\big) = (n+1)\,a + h \sum_{k=1}^{n} k = (n+1)\,a + h\,\frac{n\,(n+1)}{2} = (n+1)\left[a + \frac{n\,h}{2}\right].$$

Thus the closed representation of the arithmetic series is given by

$$\boxed{S_n = \sum_{k=0}^{n} \big(a + k\,h\big) = (n+1)\left[a + \frac{n\,h}{2}\right] = (n+1)\left[\frac{a + (a + n\,h)}{2}\right]} \tag{7.13}$$

From the right-most representation in (7.13), we see that the term in the angular brackets is the mean value of the the first term a and the last term $a + n\,h$ of the sum. The formula (7.13) contains (7.12) as the special case $a = 0$ and $h = 1$.

Example 7.6 (application of arithmetic progression)
Calculate the value of the following finite sum

$$\sum_{k=11}^{50} \big(10 + 2\,k\big).$$

Solution: We write the the finite sum as

$$\sum_{k=11}^{50} \big(10 + 2\,k\big) = \sum_{k=0}^{50} \big(10 + k\,2\big) - \sum_{k=0}^{10} \big(10 + k\,2\big),$$

and then use the formula (7.13) to evaluate each individual sum (which is an arithmetic progression with $a = 10$ and $h = 2$) on the right-hand side.

$$\sum_{k=0}^{50} \big(10 + k\,2\big) = 51\left[10 + \frac{50 \times 2}{2}\right] = 51 \times 60 = 3060,$$

and

$$\sum_{k=0}^{10} \big(10 + k\,2\big) = 11\left[10 + \frac{10 \times 2}{2}\right] = 11 \times 20 = 220.$$

Thus we find

$$\sum_{k=11}^{50} \big(10{+}2\,k\big) = \sum_{k=0}^{50} \big(10{+}k\,2\big) - \sum_{k=0}^{10} \big(10{+}k\,2\big) = 3060{-}220 = 2840. \qquad \square$$

Definition 7.7 (geometric progression)
A series of the form

$$S_n = \sum_{k=0}^{n} a\,r^k = a + a\,r + a\,r^2 + \ldots + a\,r^k + \ldots + a\,r^n,$$

where $a \in \mathbb{R}$ and $r \in \mathbb{R}$, is called a ***geometric progression*** *(or* ***geometric sum****). We observe that the* ***ratio*** *of any two consecutive terms (where we divide the term with the larger index by its predecessor) is*

$$\frac{a\,r^k}{a\,r^{k-1}} = r.$$

For this reason r is called the ***common ratio****.*

We attempt to find a closed representation of the geometric progression. Multiplying the geometric progression

$$S_n = \sum_{k=0}^{n} a\,r^k. \tag{7.14}$$

by r gives that

$$r\,S_n = r\sum_{k=0}^{n} a\,r^k = \sum_{k=0}^{n} a\,r^{k+1}.$$

Now we substitute the summation index by $j = k+1$, and thus $k = j-1$. When doing this we have also to decide which values j can assume. Since we had $k = 0, 1, 2, \ldots, n$, we have that $j = k+1$ assumes the values $j = 1, 2, 3, \ldots, n+1$. Thus

$$r\,S_n = \sum_{k=0}^{n} a\,r^{k+1} = \sum_{j=1}^{n+1} a\,r^j.$$

Renaming j to k gives that

$$r\,S_n = \sum_{k=1}^{n+1} a\,r^k. \tag{7.15}$$

Subtracting $r\,S_n$ from S_n yields therefore, using (7.14) and (7.15),

$$S_n - r\,S_n = \sum_{k=0}^{n} a\,r^k - \sum_{k=1}^{n+1} a\,r^k = a\,r^0 + \sum_{k=1}^{n} a\,r^k - \left[\sum_{k=1}^{n} a\,r^k + a\,r^{n+1}\right] = a\,r^0 - a\,r^{n+1},$$

where we have used the fact that all terms, except the term for $k = 0$ in the first sum and the term for $k = n+1$ in the second sum, occur in both sums and thus are cancelled out. Simplifying the left-hand side and the right-hand side further yields

$$(1-r)\,S_n = S_n - r\,S_n = a\,r^0 - a\,r^{n+1} = a\left(1 - r^{n+1}\right).$$

If $r \neq 1$, then dividing by $(1-r)$ yields

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

If $r = 1$, then

$$S_n = \sum_{k=0}^{n} a\,1^k = \sum_{k=0}^{n} a = (n+1)\,a,$$

and this case is not of particular interest.

We summarize what we have learned as a lemma.

Lemma 7.8 (closed form of geometric progression)
Let $a \in \mathbb{R}$ and $r \in \mathbb{R}$ with $r \neq 1$. Then the ***geometric progression*** *(or* ***geometric sum****) has the closed form*

$$S_n = \sum_{k=0}^{n} a\,r^k = a\,\frac{1-r^{n+1}}{1-r}. \tag{7.16}$$

You should know the closed formula (7.16) of the geometric progression from memory! The closed formula (7.16) of the geometric progression is useful in many contexts and we will apply it later in this chapter.

Example 7.9 (application of geometric progression)
Compute the value of the following finite sums:

(a) $\displaystyle\sum_{k=0}^{3} 2^k$, (b) $\displaystyle\sum_{k=0}^{5} \left(-\frac{1}{2}\right)^k$, (c) $\displaystyle\sum_{k=0}^{4} 9 \times 10^{-k}$.

Solution: The finite sum in (a) is a geometric progression with $a = 1$, $r = 2$, and $n = 3$. Thus, from (7.16)

$$\sum_{k=0}^{3} 2^k = \frac{1-2^4}{1-2} = \frac{1-16}{-1} = 15.$$

The finite sum in (b) is a geometric progression with $a = 1$ and $r = -1/2$ and $n = 5$. From (7.16) we have

$$\sum_{k=0}^{5} \left(-\frac{1}{2}\right)^k = \frac{1-(-1/2)^6}{1-(-1/2)} = \frac{1-(1/2)^6}{3/2} = \frac{1-1/64}{3/2} = \frac{63}{64}\,\frac{2}{3} = \frac{21}{32}.$$

Writing the finite sum in (c) as

$$\sum_{k=0}^{4} 9 \times 10^{-k} = \sum_{k=0}^{4} 9 \times \left(\frac{1}{10}\right)^k,$$

we see that it is also a geometric progression with $a = 9$, $r = 1/10 = 10^{-1}$ and $n = 4$. Thus from (7.16)

$$\sum_{k=0}^{4} 9 \times 10^{-k} = 9\,\frac{1-(10^{-1})^5}{1-10^{-1}} = 9\,\frac{1-10^{-5}}{9/10} = 10\left(1-10^{-5}\right) = 10 \times 0.99999 = 9.9999.$$

We see how useful the closed formula (7.16) of the geometric progression is. □

7.2 Binomial Expansions

In this section we discuss **binomial expansions**.

We are familiar with the **first binomial formula**

$$\left(a+b\right)^2 = a^2 + 2\,a\,b + b^2, \tag{7.17}$$

and the **second binomial formula**

$$\left(a-b\right)^2 = a^2 - 2\,a\,b + b^2. \tag{7.18}$$

We want to derive an analogous formula for the case $(a+b)^n$ and $(a-b)^n$, where n can be an arbitrary non-negative integer. Before we do this, we observe that (7.18) follows immediately from (7.17) by

$$\left(a-b\right)^2 = \left(a+(-b)\right)^2 = a^2 + 2\,a\,(-b) + (-b)^2 = a^2 - 2\,a\,b + b^2.$$

Therefore, we will restrict our discussion to deriving expansions for $(a+b)^n$.

We start with working out $(a+b)^3$ by hand

$$\left(a+b\right)^3 = (a+b)\left(a+b\right)^2 = (a+b)\left(a^2+2\,a\,b+b^2\right) = a^3 + 3\,a^2\,b + 3\,a\,b^2 + b^3.$$

If we work out $(a+b)^4$, then we will find

$$(a+b)^4 = a^4 + 4\,a^3\,b + 6\,a^2\,b^2 + 4\,a\,b^3 + b^4.$$

We observe that $(a+b)^0 = 1$ and $(a+b)^1 = a+b$, and write the expressions for $(a+b)^n$, with $n = 0, 1, 2, 3, 4$ in the following form:

$$
\begin{array}{lcc}
(a+b)^0 & = & \boxed{1} \\
(a+b)^1 & = & \boxed{1}\,a + \boxed{1}\,b \\
(a+b)^2 & = & \boxed{1}\,a^2 + \boxed{2}\,a\,b + \boxed{1}\,b^2 \\
(a+b)^3 & = & \boxed{1}\,a^3 + \boxed{3}\,a^2 b + \boxed{3}\,a\,b^2 + \boxed{1}\,b^3 \\
(a+b)^4 & = & \boxed{1}\,a^4 + \boxed{4}\,a^3 b + \boxed{6}\,a^2 b^2 + \boxed{3}\,a\,b^3 + \boxed{1}\,b^4
\end{array}
\tag{7.19}
$$

We observe that the **coefficients**, which have each been put into a box, obey the following pattern: The first and the last coefficient have always the value one. The other coefficients can be be obtained from the coefficients in the previous line as follows: If we number the coefficients according to the order in which they appear in (7.19), then the kth coefficient is the sum of the $(k-1)$th and kth coefficient in the previous line. This pattern continues for $(a+b)^n$ with $n > 4$. The structure (7.19) is referred to as **Pascal's triangle**.

For general n, the expansion of $(a+b)^n$ consists of terms of the form

$$\text{coefficient} \;\times\; a^{n-r} b^r \qquad \text{with} \quad r = 0,\, 1,\, 2,\, \ldots,\, n,$$

which arise from taking b from r of the parentheses $(a+b)$ and taking a from each of the remaining $n-r$ of the parentheses $(a+b)$ and then multiplying these factors. Indeed, we have the **binomial formula**

$$\boxed{(a+b)^n = \sum_{r=0}^{n} \binom{n}{r}\, a^{n-r}\, b^r,} \tag{7.20}$$

where the **binomial coefficients** $\binom{n}{r}$ of $a^{n-r}\, b^r$ are defined by

$$\boxed{\binom{n}{r} = \frac{n!}{r!\,(n-r)!},} \tag{7.21}$$

with the **factorials** $n!$ and $r!$ defined by

$$0! = 1, \qquad 1! = 1, \qquad m! = (m-1)! \times m = 1 \times 2 \times 3 \times \ldots \times (m-1) \times m, \quad m \in \mathbb{N}.$$

Note that the **binomial coefficients are integers**, which is not obvious from their definition (7.21). The binomial coefficients are exactly the coefficients that occur

in Pascal's triangle. More precisely, the binomial coefficient $\binom{n}{r}$ is the $(r+1)$th coefficient in the row for $(a+b)^n$ in Pascal's triangle (7.19).

Example 7.10 (binomial expansion)
Determine an expansion of $(1+x)^4$ using the formulas (7.20) and (7.21).

Solution: We have to work out the binomial coefficients $\binom{4}{r}$ for $r = 0, 1, 2, 3, 4$.

$$\begin{aligned}
\binom{4}{0} &= \frac{4!}{0!\,(4-0)!} = \frac{4!}{1 \times 4!} = 1, \\
\binom{4}{1} &= \frac{4!}{1!\,(4-1)!} = \frac{4!}{3!} = 4, \\
\binom{4}{2} &= \frac{4!}{2!\,(4-2)!} = \frac{4!}{2!\,2!} = \frac{4!}{4} = 3! = 6, \\
\binom{4}{3} &= \frac{4!}{3!\,(4-3)!} = \frac{4!}{3!\,1!} = \frac{4!}{1!\,(4-1)!} = \binom{4}{1} = 4, \\
\binom{4}{4} &= \frac{4!}{4!\,(4-4)!} = \frac{4!}{4!\,0!} = \frac{4!}{0!\,(4-0)!} = \binom{4}{0} = 1,
\end{aligned}$$

and we obtain indeed the numbers from Pascal's triangle. Thus

$$(1+x)^4 = \sum_{r=0}^{4} \binom{4}{r} 1^{4-r}\, x^r = \sum_{r=0}^{4} \binom{4}{r} x^r = 1 + 4\,x + 6\,x^2 + 4\,x^3 + x^4. \qquad \square$$

The observation from the last example, that $\binom{4}{3} = \binom{4}{4-3} = \binom{4}{1}$ and $\binom{4}{4} = \binom{4}{0}$, holds in the following general form:

$$\binom{n}{r} = \frac{n!}{r!\,(n-r)!} = \frac{n!}{(n-r)!\,r!} = \frac{n!}{(n-r)!\,(n-(n-r))!} = \binom{n}{n-r}.$$

The way by which we generate the (binomial) coefficients in Pascal's triangle implies the following formula for the binomial coefficients:

$$\binom{n}{r} + \binom{n}{r+1} = \binom{n+1}{r+1}.$$

7.3 Infinite Series

In this section we discuss **(infinite) series**, by which we mean an expression of the form

$$\sum_{k=m}^{\infty} a_k = a_m + a_{m+1} + \ldots + a_k + a_{k+1} + \ldots. \tag{7.22}$$

In (7.22) we do not stop summing up at a certain index $k = n$ but go on summing up 'until infinity'. The **value of an infinite series** (7.22) is in fact the **limit**

$$\sum_{k=m}^{\infty} a_k = \lim_{n\to\infty} \sum_{k=m}^{n} a_k = \lim_{n\to\infty} S_n \qquad \text{of the sequence } \{S_n\}, \text{ where} \qquad S_n = \sum_{k=m}^{n} a_k.$$

The finite sums

$$S_n = \sum_{k=m}^{n} a_k, \qquad \text{where} \quad n \geq m$$

are called the **partial sums** of the series, and as we saw above the value of the series is the **limit of the sequence** $\{S_n\}$ **of partial sums**.

To clarify matters, let us go back to what it means if a **sequence** $\{S_n\}$ **of real numbers** S_n **converges to a real number** S, in formulas,

$$\lim_{n\to\infty} S_n = S.$$

It means roughly that **if** n **is large enough then** S_n **is very close to the limit** S, and **if we increase** n **then** S_n **gets closer and closer to the limit**. Let us look at a concrete example.

Example 7.11 (series of real numbers approaches its limit)
The sequence

$$\{S_n\} \quad \text{defined by} \quad S_n = 1 + \frac{1}{n} = \frac{n+1}{n},$$

converges to the number $S = 1$, since

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right) = 1 + \lim_{n\to\infty} \frac{1}{n} = 1.$$

Let us work out S_n for some samples of n, say $n = 1, 2, 10, 50, 100, 1000$, to see how the sequence $\{S_n\}$ approaches the limit $S = 1$ as n increases. We have

$$S_1 = 1 + \frac{1}{1} = 2, \qquad S_2 = 1 + \frac{1}{2} = \frac{3}{2} = 1.5, \qquad S_{10} = 1 + \frac{1}{10} = 1 + 0.1 = 1.1,$$
$$S_{50} = 1 + \frac{1}{50} = 1 + 0.02 = 1.02, \qquad S_{100} = 1 + \frac{1}{100} = 1 + 0.01 = 1.01,$$
$$S_{1000} = 1 + \frac{1}{1000} = 1 + 0.001 = 1.001,$$

and we see indeed that S_n approaches the limit $S = 1$ as n increases. □

Sequences of real numbers $\{S_n\}$ may **converge**, or **not**, and if a sequence of $\{S_n\}$ of real numbers **does not converge, then we say it diverges**.

In the Example 7.11 above we have seen a sequence of real numbers that converges. Now we also give an example of two **sequences of real numbers that diverge**.

Example 7.12 (sequences of real numbers that diverge)
The sequences

$$\{S_n\} \quad \text{with} \quad S_n = n \qquad \text{and} \qquad \{T_n\} \quad \text{with} \quad T_n = (-1)^n$$

diverge. Indeed, the members $S_n = n$ of the sequence $\{S_n\}$ get **arbitrarily large** as $n \to \infty$, and thus $\lim_{n\to\infty} S_n = \lim_{n\to\infty} n = \infty$. **If we find that the 'limit' is plus (or minus) infinity, then the sequence diverges!** In the other example the values of $T_n = (-1)^n$ **alternate** or **'oscillate'** between -1 and 1. Neither of these values can be the limit: for example 1 cannot be the limit, for if it were the limit we would have that for n large enough all T_n are close to 1, but some such T_n are still -1 which is not close to 1. Thus $\{T_n\}$ does not converge. □

Now that we have clarified the meaning of convergence for sequences of real numbers, we come back to infinite series of real numbers in the definition below.

Definition 7.13 (convergence of an infinite series)
Let m be an integer. The infinite series

$$\sum_{k=m}^{\infty} a_k = a_m + a_{m+1} + \ldots + a_k + a_{k+1} + \ldots \tag{7.23}$$

converges *if the sequence $\{S_n\}$ of* ***partial sums***

$$S_n = \sum_{k=m}^{n} a_k = a_m + a_{m+1} + \ldots + a_n, \qquad \textit{where} \quad n \geq m$$

converges*. The infinite series (7.23)* ***diverges*** *(that is,* ***does not converge****) if the sequence $\{S_n\}$ of partial sums* ***diverges*** *(that is,* ***does not converge****). If the sequence $\{S_n\}$ of partial sums converges to the value $S = \lim_{n\to\infty} S_n$, then S is the value of the series, that is, $S = \sum_{k=m}^{\infty} a_k$.*

We discuss an example.

Example 7.14 (geometric series)
Let $a \neq 0$. Investigate for which r the **geometric series**

$$\sum_{k=0}^{\infty} a\, r^k = a\, r^0 + a\, r^1 + a\, r^2 + \ldots + a\, r^k + \ldots \tag{7.24}$$

converges and for which r the series diverges (that is, does not converge).

Solution: First let $|r| \neq 1$. We consider the sequence S_n of partial sums

$$S_n = \sum_{k=0}^{n} a\, r^k = a\, \frac{1 - r^{n+1}}{1 - r}. \tag{7.25}$$

where the closed representation on the right follows from Lemma 7.8. (Note that the partial sums (7.25) are **geometric progressions**.)

If $|r| > 1$, then

$$|r^{n+1}| = |r|^{n+1} \to \infty \quad \text{as} \quad n \to \infty \qquad \Rightarrow \qquad \lim_{n\to\infty} S_n = \lim_{n\to\infty} a\, \frac{1 - r^{n+1}}{1 - r} = \infty.$$

Thus, for $|r| > 1$, the sequence $\{S_n\}$ does not converge, and consequently the infinite series (7.24) does not converge.

If $|r| < 1$, then

$$r^{n+1} \to 0 \quad \text{as} \quad n \to \infty \qquad \Rightarrow \qquad \lim_{n\to\infty} S_n = \lim_{n\to\infty} a\, \frac{1 - r^{n+1}}{1 - r} = a\, \frac{1}{1 - r} = \frac{a}{1 - r}.$$

Consequently, if $|r| < 1$, then the infinite series (7.24) converges and we have

$$\sum_{k=0}^{\infty} a\, r^k = \lim_{n\to\infty} S_n = \frac{a}{1 - r}.$$

If $r = 1$, then $r^k = 1^k = 1$ and we have

$$S_n = \sum_{k=0}^{n} a\, r^k = \sum_{k=0}^{n} a\, 1^k = \sum_{k=0}^{n} a = (n + 1)\, a,$$

and therefore

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} (n + 1)\, a = \begin{cases} \infty & \text{if} \quad a > 0, \\ -\infty & \text{if} \quad a < 0. \end{cases}$$

We see that for $r = 1$ the sequence $\{S_n\}$ of partial sums diverges, and thus, for $r = 1$, the geometric series (7.24) diverges.

If $r = -1$, then $r^k = (-1)^k$, and $(-1)^k = 1$ if k is even and $(-1)^k = -1$ if k is odd. Thus we find that

$$S_n = \sum_{k=0}^{n} a\,r^k = \sum_{k=0}^{n} a\,(-1)^k = a + (-a) + a + (-a) + \ldots + \begin{cases} a & \text{if } n \text{ is even,} \\ a + (-a) & \text{if } n \text{ is odd.} \end{cases}$$

Thus we have for odd n as many terms a as we have terms $-a$, and they cancel each other out. For even n we have one more term a, and the remaining terms cancel each other out. Thus

$$S_n = \sum_{k=0}^{n} a\,(-1)^k = \begin{cases} a & \text{if } n \text{ is even,} \\ 0 & \text{if } n \text{ is odd.} \end{cases}$$

Since the values of S_n alternate ('oscillate') between 0 and $a \neq 0$, the sequence $\{S_n\}$ does not converge, that is, it diverges. Thus the geometric series (7.24) diverges for $r = -1$. □

We summarize what we have learnt about the geometric series in the following lemma.

Lemma 7.15 (convergence of the geometric series)
Let $a \neq 0$ and let $r \in \mathbb{R}$. The ***geometric series***

$$\sum_{k=0}^{\infty} a\,r^k = a\,r^0 + a\,r^1 + a\,r^2 + \ldots + a\,r^k + \ldots$$

converges for $|r| < 1$ and diverges for $|r| \geq 1$. For $|r| < 1$, we have

$$\sum_{k=0}^{\infty} a\,r^n = \frac{a}{1-r}. \tag{7.26}$$

Example 7.16 (geometric series)
From the formula (7.26) for the geometric series with $a = 1$ and $r = 1/2$ (note $|r| = 1/2 < 1$), we have

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots = \sum_{k=0}^{\infty} \frac{1}{2^k} = \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k = \frac{1}{1-1/2} = \frac{1}{1/2} = 2.$$ □

Example 7.17 (geometric series)
Find the value of the series

$$1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \ldots = \sum_{k=0}^{\infty} \frac{(-1)^k}{2^k}.$$

Solution: Rewriting the series as

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{2^k} = \sum_{k=0}^{\infty} \left(-\frac{1}{2}\right)^k,$$

reveals that it is the geometric series with $a = 1$ and $r = -1/2$. Since $|r| = |-1/2| = 1/2 < 1$, we know that the geometric series converges, and from (7.26) its limit is given by

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{2^k} = \sum_{k=0}^{\infty} \left(-\frac{1}{2}\right)^k = \frac{1}{1-(-1/2)} = \frac{1}{3/2} = \frac{2}{3}.$$

To investigate 'how' the geometric series with $a = 1$ and $r = -1/2$ converges to its limit $2/3$, we evaluate the first four partial sums.

$$\begin{aligned}
S_0 &= 1 > \frac{2}{3}, \\
S_1 &= 1 - \frac{1}{2} = \frac{1}{2} < \frac{2}{3}, \\
S_2 &= 1 - \frac{1}{2} + \frac{1}{4} = \frac{3}{4} > \frac{2}{3}, \\
S_3 &= 1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} = \frac{5}{8} < \frac{2}{3},
\end{aligned}$$

and so the sequence $\{S_n\}$ of the partial sums converges to the limit $2/3$ in an 'oscillatory' fashion. □

Example 7.18 (geometric series)
Investigate the convergence of the series

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{10}.$$

Solution: We observe that this is a geometric series with $a = 1/10$ and $r = -1$. Since $|r| = |-1| = 1$, we know from Lemma 7.15 that the series diverges. □

Example 7.19 (series that diverge)
Show that the series

$$\text{(a)} \quad \sum_{k=1}^{\infty} k \qquad \text{and} \qquad \text{(b)} \quad \sum_{k=1}^{\infty} \frac{1}{k}$$

diverge.

Solution: (a) We have that

$$\sum_{k=1}^{\infty} k = \lim_{n\to\infty} \sum_{k=1}^{n} k = \lim_{n\to\infty} \frac{n(n+1)}{2} = \infty,$$

where we have used that the finite sum $\sum_{k=1}^{n} k$ is an arithmetic progression.

(b) We group the terms in the series together as indicated below

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \left(\frac{1}{9} + \ldots + \frac{1}{16}\right) + \ldots.$$

Each sum of the terms in one of the brackets is larger than 1/2, and thus

$$\sum_{k=1}^{\infty} \frac{1}{k} = \lim_{n\to\infty} \sum_{k=1}^{2^n} \frac{1}{k} \geq \lim_{n\to\infty} \left(1 + n \times \frac{1}{2}\right) = \infty.$$

Thus we have shown that both series diverge. □

7.4 Power Series

Power series are infinite series for which the **partial sums are polynomials**. Therefore a power series defines a **function**. In this section, we will define power series and look at some examples.

We start the section with a motivating example. From Lemma 7.15, we know that the **geometric series**

$$\sum_{k=0}^{\infty} x^k = 1 + x + x^2 + x^3 + x^4 + \ldots + x^k + \ldots \tag{7.27}$$

converges for $|x| < 1$ and diverges for $|x| \geq 1$. For $|x| < 1$ we have seen that

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}. \tag{7.28}$$

If we think in (7.27) and (7.28) of x as a variable, then we have our first example of a **power series**. From (7.28) we see that the power series

$$\sum_{k=0}^{\infty} x^k = 1 + x + x^2 + x^3 + x^4 + \ldots + x^k + \ldots$$

which we now consider as a **function of the variable** x, converges for $|x| < 1$ to the function $1/(1-x)$.

Definition 7.20 (power series about x_0)
An infinite series of the form

$$\sum_{k=0}^{\infty} a_k (x-x_0)^k = a_0 + a_1 (x-x_0) + a_2 (x-x_0)^2 + \ldots + a_k (x-x_0)^k + \ldots \quad (7.29)$$

is called a ***power series about*** x_0. *Here* $x \in \mathbb{R}$ *is the* ***variable*** *of the power series. The real numbers* $a_0, a_1, a_2, a_3 \ldots, a_k, \ldots$ *are called the* ***coefficients*** *and* x_0 *is the* ***centre of expansion*** *of the power series (7.29).*

From the definition of the power series it is clear that the power series is a **function of the variable** x.

The case that you will mostly encounter is $x_0 = 0$. Then power series (7.29) reads

$$\boxed{\sum_{k=0}^{\infty} a_k x^k = a_0 + a_1 x + a_2 x^2 + \ldots + a_k x^k + \ldots.}$$

The **partial sums** of the power series (7.29) are given by the **polynomials**

$$\boxed{S_n(x) = \sum_{k=0}^{n} a_k (x-x_0)^k = a_0 + a_1 (x-x_0) + a_2 (x-x_0)^2 + \ldots + a_n (x-x_0)^n.}$$

Note that **it is allowed that some or even all of the coefficients a_k are zero**.

We give some examples of power series.

Example 7.21 (power series)
The following expressions are power series:

(a) In the power series

$$\sum_{k=0}^{\infty} x^{2k} = 1 + x^2 + x^4 + x^6 + x^8 + \ldots, \quad (7.30)$$

about $x_0 = 0$, we have $a_k = 1$ if k is even and $a_k = 0$ if k is odd. All the odd powers of x are missing.

(b) In the power series

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \ldots, \tag{7.31}$$

about $x_0 = 0$ the coefficients are given by $a_k = 1/k!$ for all $k \in \mathbb{N}_0$.

(c) An example of a power series about $x_0 = 1$ is given by

$$\sum_{k=0}^{\infty} (x-1)^k.$$

Here we have the coefficients $a_k = 1$ for all $k \in \mathbb{N}_0$.

(d) Let $a > 0$ be a constant. The power series

$$\sum_{k=1}^{\infty} \frac{(-1)^k\, a^k}{k}\,(x+2)^k = \sum_{k=1}^{\infty} \frac{(-1)^k\, a^k}{k}\,\big(x-(-2)\big)^k = -a\,(x+2) + \frac{a^2}{2}\,(x+2)^2 + \ldots$$

is a power series about $x_0 = -2$ with the coefficients $a_k = (-1)^k\, a^k/k$, $k \in \mathbb{N}$. □

A crucial question about power series is the following:

For which values of x does the power series converge?

When we talk about the **convergence of a power series at a point** x, then we will think of x **as fixed** and the power series with this value of x,

$$\sum_{k=0}^{\infty} a_k\,(x-x_0)^k = a_0 + a_1\,(x-x_0) + a_2\,(x-x_0)^2 + \ldots + a_k\,(x-x_0)^k + \ldots, \tag{7.32}$$

is just a **series of real numbers**. For each fixed x, we can discuss its convergence. Thus a power series (7.32) **converges at a (fixed) point x if at this fixed point x the sequence of the partial sums** $\{S_n(x)\}$,

$$S_n(x) = \sum_{k=0}^{n} a_k\,(x-x_0)^k = a_0 + a_1\,(x-x_0) + a_2\,(x-x_0)^2 + \ldots + a_k\,(x-x_0)^n$$

converges. We observe that the power series **converges always at** $x = x_0$, since for $x = x_0$ all but the first term are zero.

Example 7.22 (convergence of geometric series)
From Lemma 7.15, we know that

$$\sum_{k=0}^{\infty} a\,x^k$$

converges to the function $a/(1-x)$ if $|x| < 1$ and diverges (does not converge) if $|x| \geq 1$. □

Example 7.23 (power series related to the geometric series)
Show that

$$\sum_{k=0}^{\infty}(x-1)^k$$

converges for $x \in (0,2)$ and diverges for x with $x \leq 0$ and $x \geq 2$.

Solution: We let $y = x - 1$, then from the geometric series

$$\sum_{k=0}^{\infty}(x-1)^k = \sum_{k=0}^{\infty} y^k = \frac{1}{1-y} = \frac{1}{1-(x-1)} = \frac{1}{2-x} \qquad \text{if } |y| < 1,$$

and for $|y| \geq 1$ the series diverges. Thus $\sum_{k=0}^{\infty}(x-1)^k$ converges for $|y| = |x-1| < 1$, that is, for $x \in (0,2)$, and it diverges for $|y| = |x-1| \geq 1$, that is, for $x \leq 0$ and $x \geq 2$. □

Example 7.24 (power series related to the geometric series)
Show that

$$\sum_{k=0}^{\infty} x^{2k}$$

converges for x with $|x| < 1$ and diverges for $|x| \geq 1$.

Solution: Letting $y = x^2$, we find, from the geometric series,

$$\sum_{k=0}^{\infty} x^{2k} = \sum_{k=0}^{\infty} y^k = \frac{1}{1-y} = \frac{1}{1-x^2} \qquad \text{if } |y| = |x^2| < 1,$$

and $\sum_{k=0}^{\infty} x^{2k} = \sum_{k=0}^{\infty} y^k$ diverges if $|y| = |x^2| \geq 1$. Thus the series converges for $|x| < 1$ and diverges for $|x| \geq 1$ □

In the last two examples we have exploited our knowledge of the geometric series to obtain information about the convergence a given series. This is not always possible, but a discussion how to investigate the convergence of an arbitrary power series goes beyond the scope of this course.

7.5 Maclaurin Series and Taylor series

In Lemma 7.15, we have seen that the geometric series satisfies

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \qquad \text{for all } x \text{ with } |x| < 1. \tag{7.33}$$

We can interpret (7.33) in the following way: The function $f(x) = 1/(1-x)$ has the **power series expansion** about x_0 given by the left-hand side of (7.33)

Can we expand every smooth function into a power series expansion about x_0, and if yes, how do we obtain the coefficients a_k?

This question leads us to the Taylor series of a function about a point x_0: If the function is so smooth that it has derivatives of arbitrarily high order, then we can indeed expand the function into a power series about x_0, the so-called **Taylor series about** x_0! The **Maclaurin series** mentioned in the section heading is a special case and the most common type of a Taylor series. We will find an explicit formula for the coefficients a_k that involves the values of f and its derivatives at x_0.

If we have for a **smooth function** (given by a closed non-series representation)

$$f(x) = \sum_{k=0}^{\infty} a_k\, x^k = a_0 + a_1\, x + a_2\, x^2 + \ldots + a_k\, x^k + \ldots \quad \text{for all x close to } x_0 = 0, \tag{7.34}$$

then **how are the values of the coefficients** a_k **related to the function** f**?**

Letting $x = x_0 = 0$ gives that

$$f(0) = a_0 \qquad \Leftrightarrow \qquad a_0 = f(0) = \frac{f(0)}{0!}. \tag{7.35}$$

To get rid of the constant term a_0 of the power series we differentiate (7.34), where we may differentiate the series term by term, that is, we may just move the differentiation behind the summation and differentiate the individual terms of the series. Differentiating with respect to x gives that

$$\begin{aligned} f'(x) &= \frac{d}{dx}\left[\sum_{k=0}^{\infty} a_k\, x^k\right] = \sum_{k=0}^{\infty} \frac{d}{dx}\left[a_k\, x^k\right] = \sum_{k=1}^{\infty} a_k\, k\, x^{k-1} \\ &= a_1 + a_2\, 2\, x + a_3\, 3\, x^2 + \ldots + a_k\, k\, x^{k-1} + \ldots, \end{aligned} \tag{7.36}$$

where we have used $(a_0)' = 0$. Letting $x = 0$ in (7.36) gives that that

$$f'(0) = a_1 \qquad \Leftrightarrow \qquad a_1 = f'(0) = \frac{f'(0)}{1!}. \tag{7.37}$$

We differentiate again and obtain

$$f''(x) = \frac{d}{dx}\left[\sum_{k=1}^{\infty} a_k\, k\, x^{k-1}\right] = \sum_{k=1}^{\infty} \frac{d}{dx}\left[a_k\, k\, x^{k-1}\right] = \sum_{k=2}^{\infty} a_k\, k\,(k-1)\, x^{k-2}$$

$$= a_2\,2 + a_3\,6\,x + \ldots + a_k\,\big(k(k-1)\big)\,x^{k-2} + \ldots, \tag{7.38}$$

where we have used $(a_1)' = 0$. Letting $x = 0$ in (7.38) gives that

$$f''(0) = a_2\,2 \qquad \Leftrightarrow \qquad a_2 = \frac{f''(0)}{2} = \frac{f''(0)}{2!}. \tag{7.39}$$

Substituting (7.35), (7.37), and (7.39) into the power series (7.34) yields

$$f(x) = f(0) + f'(0)\,x + \frac{f''(0)}{2!}\,x^2 + a_3x^3 + \ldots + a_k\,x^k + \ldots. \tag{7.40}$$

Continuing in this way, that is, differentiating 3-times, 4-times, ..., k-times, ..., gives that

$$a_k = \frac{f^{(k)}(0)}{k!} \qquad \text{for all} \quad k = 0, 1, 2, 3, 4, \ldots. \tag{7.41}$$

We see that (7.35), (7.37), and (7.39) are indeed special cases of (7.41). Substituting the coefficients (7.41) into (7.34) yields

$$f(x) = \sum_{k=0}^{\infty} a_k\,x^k = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!}\,x^k$$

which is the so-called **Maclaurin series of the function** f.

Definition 7.25 (Maclaurin series of a function)
*The **Maclaurin series** of a smooth function $f(x)$ is defined by*

$$T(f,0)(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!}\,x^k = f(0) + f'(0)\,x + \frac{f''(0)}{2!}\,x^2 + \frac{f'''(0)}{3!}\,x^3 + \ldots.$$

From setting $x = 0$, we see that always

$$T(f,0)(0) = f(0),$$

but for x different from zero, **we cannot know whether $T(f,0)(x) = f(x)$ holds or not without further investigation**.

We give some examples of Maclaurin series.

Example 7.26 (Maclaurin series of $\exp(x)$)
Find the Maclaurin series for the exponential function $f(x) = \exp(x) = e^x$.

Solution: We have that $f'(x) = (e^x)' = e^x = f(x)$, and from repeated application of $(e^x)' = e^x$ we find

$$f^{(k)}(x) = \frac{d^k}{dx^k}\, e^x = e^x, \qquad k = 0, 1, 2, 3, \ldots$$

Evaluation of the derivatives at $x = 0$ yields

$$f^{(k)}(0) = e^0 = 1, \qquad k = 0, 1, 2, 3, \ldots.$$

Hence the Maclaurin series of $f(x) = e^x$ is given by

$$T(f, 0)(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!}\, x^k = \sum_{k=0}^{\infty} \frac{1}{k!}\, x^k = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots.$$

It is possible to show that the Maclaurin series of $f(x) = e^x$ converges for all $x \in \mathbb{R}$ to the function $f(x) = e^x$. □

Example 7.27 (Maclaurin series of $\sin(x)$)
Find the Maclaurin series for the function $f(x) = \sin(x)$.

Solution: We have that

$$\begin{aligned} &f(x) = \sin(x), \qquad f'(x) = \cos(x), \qquad f''(x) = -\sin(x), \\ &f'''(x) = -\cos(x), \qquad f^{(4)}(x) = \sin(x) = f(x) \end{aligned}$$

If we continue differentiating we find that the even order derivatives (note $k = 2\ell$, $\ell = 0, 1, 2, 3, \ldots$, describes all even non-negative integers) are given by

$$f^{(2\ell)}(x) = (-1)^\ell \sin(x). \tag{7.42}$$

Analogously, the odd order derivatives (note that $k = 2\ell + 1$, $\ell = 0, 1, 2, 3 \ldots$, describes all odd positive integers) are given by

$$f^{(2\ell+1)}(x) = (-1)^\ell \cos(x). \tag{7.43}$$

Evaluation of the derivatives at $x = 0$ yields

$$\begin{aligned} &f(0) = \sin(0) = 0, \qquad f'(0) = \cos(0) = 1, \qquad f''(0) = -\sin(0) = 0, \\ &f'''(0) = -\cos(0) = -1, \qquad f^{(4)}(0) = \sin(0) = 0, \qquad \ldots. \end{aligned}$$

Using the general formulas (7.42) and (7.43) we find

$$f^{(2\ell)}(0) = (-1)^\ell \sin(0) = 0 \qquad \text{for all} \quad \ell = 0, 1, 2, 3 \ldots,$$

$$f^{(2\ell+1)}(0) = (-1)^\ell \cos(0) = (-1)^\ell \times 1 = (-1)^\ell \qquad \text{for all} \quad \ell = 0, 1, 2, 3 \ldots.$$

We see that all even order derivatives have at $x = 0$ the value zero and thus the Maclaurin series of $f(x) = \sin(x)$ is given by

$$\begin{aligned} T(f,0)(x) &= \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!}\, x^k = \sum_{\ell=0}^{\infty} \frac{f^{(2\ell+1)}(0)}{(2\ell+1)!}\, x^{2\ell+1} = \sum_{\ell=0}^{\infty} \frac{(-1)^\ell}{(2\ell+1)!}\, x^{2\ell+1} \\ &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \frac{x^{11}}{11!} + \ldots. \end{aligned}$$

It is possible to show that the Maclaurin series of $f(x) = \sin(x)$ converges for all $x \in \mathbb{R}$ to $\sin(x)$. □

Example 7.28 (Maclaurin series of $\cos(x)$)
The Maclaurin series for the function $f(x) = \cos(x)$ is

$$T(f,0)(x) = \sum_{\ell=0}^{\infty} \frac{(-1)^\ell}{(2\ell)!}\, x^{2\ell} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots.$$

We will show this on the exercise sheet for this chapter. The Maclaurin series of $f(x) = \cos(x)$ converges for all $x \in \mathbb{R}$ to the function $f(x) = \cos(x)$. □

Example 7.29 (Maclaurin series of $f(x) = 1/(1-x)$)
From Lemma 7.15, we have

$$\text{geometic series } = \sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \qquad \text{for all } x \text{ with } |x| < 1.$$

Thus we would expect the Maclaurin series of the function $f(x) = 1/(1-x)$ to be the geometric series. We will show that this is indeed the case! From the chain rule

$$f(x) = (1-x)^{-1}, \quad f'(x) = (1-x)^{-2}, \quad f''(x) = 2\,(1-x)^{-3}, \quad f'''(x) = 3!\,(1-x)^{-4},$$

and for general $k \in \mathbb{N}_0$

$$f^{(k)}(x) = k!\,(1-x)^{-(k+1)} \qquad \text{and thus} \qquad f^{(k)}(0) = k!\,(1-0)^{-(k+1)} = k!.$$

Thus the Maclaurin series of $f(x) = 1/(1-x)$ is given by

$$T(f,0)(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!}\, x^k = \sum_{k=0}^{\infty} \frac{k!}{k!}\, x^k = \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + x^3 + x^4 + x^5 + \ldots,$$

which is indeed the geometric series. □

Example 7.30 (Maclaurin series of $\ln(x+1)$)
Find the Maclaurin series for the function $f(x) = \ln(1+x)$.

Solution: We observe that $f(x) = \ln(x+1)$ is defined for $x > -1$ and so it makes sense to consider the Maclaurin series. We have that

$$f(x) = \ln(1+x), \quad f'(x) = \frac{1}{1+x}, \quad f''(x) = -\frac{1}{(1+x)^2},$$
$$f'''(x) = \frac{2}{(1+x)^3}, \quad f^{(4)}(x) = -\frac{2\times 3}{(1+x)^4}, \quad f^{(5)}(x) = \frac{2\times 3\times 4}{(1+x)^5}, \quad \ldots$$

We find the general formula

$$f(x) = \ln(1+x) \qquad \text{and} \qquad f^{(k)}(x) = \frac{(-k)^{k-1}(k-1)!}{(1+x)^k}, \quad k = 1,\, 2,\, 3,\, \ldots,$$

and evaluating at $x = 0$ yields

$$f(0) = \ln(1) = 0, \qquad \text{and} \qquad f^{(k)}(0) = (-1)^{k-1}(k-1)!, \quad k = 1,\, 2,\, 3,\, \ldots.$$

Thus the Maclaurin series of the function $f(x) = \ln(x+1)$ is given by

$$\begin{aligned} T(f,0)(x) &= f(0) + \sum_{k=1}^{\infty} \frac{f^{(k)}(0)}{k!}\, x^k = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}(k-1)!}{k!}\, x^k \\ &= \sum_{i=1}^{\infty} \frac{(-1)^{k-1}}{k}\, x^k = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots. \end{aligned}$$

It can be shown that the Maclaurin series of $f(x) = \ln(x+1)$ converges for all x with $|x| < 1$, that is, for $x \in (-1, 1)$ to the function $f(x) = \ln(x+1)$. □

In analogy to the Maclaurin series we can calculate the **power series expansion of a smooth function about** x_0, where now the values of the function and its derivatives at x_0 are used in the definition of the coefficients.

Definition 7.31 (Taylor series of a function)
*The **Taylor series** for a smooth function $f(x)$ **about** x_0 is defined by*

$$\begin{aligned} T(f,x_0)(x) &= \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}\,(x-x_0)^k \\ &= f(x_0) + f'(x_0)\,(x-x_0) + \frac{f''(x_0)}{2!}\,(x-x_0)^2 + \frac{f'''(x_0)}{3!}\,(x-x_0)^3 + \ldots. \end{aligned}$$

We observe that **for $x_0 = 0$ the Taylor series $T(f, x_0)$ becomes just the Maclaurin series $T(f, 0)$ of f.** So the Taylor series is the generalization of the Maclaurin series where we expand the function f about an arbitrary point x_0. We will mostly encounter Mclaurin series in this course and only very occasionally Taylor series about a point x_0 with $x_0 \neq 0$.

We give two examples of Taylor series about a point x_0 with $x_0 \neq 0$.

Example 7.32 (Taylor series of $\ln(x)$ about $x_0 = 1$)
Determine the Taylor series of $f(x) = \ln(x)$ about $x_0 = 1$.

Solution: We compute the derivatives of the function $f(x) = \ln(x)$.

$$f'(x) = \frac{1}{x}, \quad f''(x) = \frac{(-1)}{x^2}, \quad f'''(x) = \frac{2!}{x^3}, \quad \ldots, \quad f^{(k)}(x) = \frac{(-1)^{k-1}(k-1)!}{x^k}.$$

Evaluating at $x = 1$ yields

$$f(1) = \ln(1) = 0, \quad f'(1) = 1 = 0!, \quad f''(1) = -1 = -1!, \quad f'''(1) = 2!$$

and in the general case

$$f^{(k)}(1) = (-1)^{k-1}\,(k-1)! \qquad \text{for all } k = 1, 2, 3, \ldots.$$

Thus the Taylor series of $f(x) = \ln(x)$ about $x_0 = 1$ is given by

$$\begin{aligned} T(f,1)(x) &= \sum_{k=0}^{\infty} \frac{f^{(k)}(1)}{k!}\,(x-1)^k = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}(k-1)!}{k!}\,(x-1)^k = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}\,(x-1)^k \\ &= (x-1) - \frac{1}{2}\,(x-1)^2 + \frac{1}{3}\,(x-1)^3 - \frac{1}{4}\,(x-1)^4 + \ldots. \end{aligned}$$

It can be shown that the Taylor series of $f(x) = \ln(x)$ about $x_0 = 1$ converges for all x with $|x-1| < 1$ to $\ln(x)$, and that it diverges for all x with $|x-1| > 1$. □

Example 7.33 (Taylor series of $\sin(x)$ about $x_0 = \pi/4$)
Find the first five terms in the Taylor series for the function $f(x) = \sin(x)$ about $x_0 = \pi/4$.

Solution: We have that

$$\begin{aligned} &f(x) = \sin(x), \qquad f'(x) = \cos(x), \qquad f''(x) = -\sin(x), \\ &f'''(x) = -\cos(x), \qquad f^{(4)}(x) = \sin(x) = f(x), \qquad \ldots \end{aligned}$$

So taking $x = x_0 = \pi/4$ gives that

$$f(\pi/4) = \sin(\pi/4) = \frac{1}{\sqrt{2}}, \qquad f'(\pi/4) = \cos(\pi/4) = \frac{1}{\sqrt{2}},$$
$$f''(\pi/4) = -\sin(\pi/4) = -\frac{1}{\sqrt{2}}, \qquad f'''(\pi/4) = -\cos(\pi/4) = -\frac{1}{\sqrt{2}},$$
$$f^{(4)}(\pi/4) = \sin(\pi/4) = f(\pi/4) = \frac{1}{\sqrt{2}}, \qquad \ldots$$

That the Taylor series of $f(x) = \sin(x)$ about $x_0 = \pi/4$ is given by

$$\begin{aligned}
T(f,\pi/4)(x) &= \sum_{k=0}^{\infty} \frac{f^{(k)}(\pi/4)}{k!}\left(x-\frac{\pi}{4}\right)^k = f(\pi/4) + f'(\pi/4)\left(x-\frac{\pi}{4}\right) \\
&\quad + \frac{f''(\pi/4)}{2!}\left(x-\frac{\pi}{4}\right)^2 + \frac{f'''(\pi/4)}{3!}\left(x-\frac{\pi}{4}\right)^3 + \frac{f^{(4)(\pi/4)}}{4!}\left(x-\frac{\pi}{4}\right)^4 + \ldots \\
&= \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\left(x-\frac{\pi}{4}\right) - \frac{1}{2!}\frac{1}{\sqrt{2}}\left(x-\frac{\pi}{4}\right)^2 - \frac{1}{3!}\frac{1}{\sqrt{2}}\left(x-\frac{\pi}{4}\right)^3 + \frac{1}{4!}\frac{1}{\sqrt{2}}\left(x-\frac{\pi}{4}\right)^4 + \ldots \\
&= \frac{1}{\sqrt{2}}\left[1 + \left(x-\frac{\pi}{4}\right) - \frac{1}{2}\left(x-\frac{\pi}{4}\right)^2 - \frac{1}{6}\left(x-\frac{\pi}{4}\right)^3 + \frac{1}{24}\left(x-\frac{\pi}{4}\right)^4 + \ldots\right]. \qquad \square
\end{aligned}$$

7.6 Approximation of a Function by the Taylor Polynomials

If we compute the first few terms of the Taylor series of f about x_0, then we have a polynomial and for x close to x_0 we may expect that this polynomial provides a reasonable approximation of f.

Definition 7.34 (Taylor polynomials)
Let $f : (a,b) \to \mathbb{R}$ be a smooth enough function, and let $x_0 \in (a,b)$. The polynomial $S_n(f,x_0)$ of degree n, obtained by taking the first $n+1$ terms of the Taylor series $T(f,x_0)$ of f about x_0, that is,

$$\begin{aligned}
S_n(f,x_0)(x) &= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k \\
&= f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \ldots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n, \quad (7.44)
\end{aligned}$$

*is called the **Taylor polynomial of degree** n **of** f **about** x_0.*

By definition, the Taylor polynomials of f about x_0 are just the **partial sums** of the Taylor series of f about x_0, that is,

$$T(f,x_0)(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}\,(x-x_0)^k = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}\,(x-x_0)^k = \lim_{n\to\infty} S_n(f,x_0).$$

For the special case of $x_0 = 0$, the Taylor polynomials are the partial sums of the Maclaurin series, and (7.44) reads

$$S_n(f,0)(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}\,x^k = f(0) + f'(0)\,x + \frac{f''(0)}{2!}\,x^2 + \ldots + \frac{f^{(n)}(0)}{n!}\,x^n. \quad (7.45)$$

Example 7.35 (Taylor polynomials for $\exp(x)$ **about** $x_0 = 0$**)**
Find the Taylor polynomials of the exponential function $f(x) = \exp(x) = e^x$ about $x_0 = 0$.

Solution: In Example 7.26, we have seen that the Maclaurin series of the exponential function is given by

$$T(t,0) = \sum_{k=0}^{\infty} \frac{1}{k!}\,x^k = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^k}{k!} + \ldots.$$

The Taylor polynomial $S_n(f,0)$ of $f(x) = e^x$ of degree n about $x_0 = 0$ is the partial sum up to $k = n$ of the Maclaurin series, and thus

$$S_n(f,0)(x) = \sum_{k=0}^{n} \frac{1}{k!}\,x^k = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!}.$$

We write down all Taylor polynomials up to degree 3 explicitly:

$$S_0(f,0)(x) = 1, \quad (7.46)$$

$$S_1(f,0)(x) = 1 + x, \quad (7.47)$$

$$S_2(f,0)(x) = 1 + x + \frac{x^2}{2!}, \quad (7.48)$$

$$S_3(f,0)(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}. \quad (7.49)$$

The Taylor polynomial $S_n(f,0)$ of degree $n = 0$ is a constant, of degree $n = 1$ is an affine linear function, of degree $n = 2$ is a quadratic function, and of degree $n = 3$ is a cubic function. □

Now we want to use the Taylor polynomial (7.49) of $f(x) = e^x$ about $x_0 = 0$ of degree $n = 3$ to **approximate** $f(x) = e^x$ **for** x **close to** $x_0 = 0$. We evaluate

$$S_3(f,0)(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6}$$

at $x = 1/2$ and $x = -1/2$ to get an approximation of the values $f(1/2) = e^{1/2}$ and $f(-1/2) = e^{-1/2}$, respectively. We have

$$\begin{aligned} S_3(f,0)(1/2) &= 1 + \frac{1}{2} + \frac{(1/2)^2}{2} + \frac{(1/2)^3}{6} = 1 + \frac{1}{2} + \frac{1}{8} + \frac{1}{48} \\ &= \frac{48 + 24 + 6 + 1}{48} = \frac{79}{48} \approx 1.6458, \\ S_3(f,0)(-1/2) &= 1 - \frac{1}{2} + \frac{(-1/2)^2}{2} + \frac{(-1/2)^3}{6} = 1 - \frac{1}{2} + \frac{1}{8} - \frac{1}{48} \\ &= \frac{48 - 24 + 6 - 1}{48} = \frac{29}{48} \approx 0.60417, \end{aligned}$$

and the exact values of the function are

$$f(1/2) = e^{1/2} \approx 1.6487 \qquad \text{and} \qquad f(-1/2) = e^{-1/2} \approx 0.60653.$$

We determine the errors

$$e^{1/2} - \frac{79}{48} \approx 0.0028879 \qquad \text{and} \qquad e^{-1/2} - \frac{29}{48} \approx 0.0023640,$$

and we see that the errors of the approximations are of the order 3×10^{-3}. Thus for $x = 1/2$ and $x = -1/2$ the Taylor polynomial of degree $n = 3$ of $f(x) = e^x$ about $x_0 = 0$ provides a good approximation of the values $f(1/2) = e^{1/2}$ and $f(-1/2) = e^{-1/2}$, respectively. This is illustrated in Figure 7.1.

We discuss some more examples where Taylor polynomials are used to approximate the functions close to the point of expansion x_0.

Example 7.36 (approximation of $\exp(\pm 1/2)$ **by Taylor polynomials)**
Use the Taylor polynomials of degree $n = 1$ and $n = 2$ of $f(x) = \exp(x) = e^x$ about $x_0 = 0$ to get an approximation of $f(1/2)$ and $f(-1/2)$.

Solution: For the Taylor polynomial of degree $n = 2$ of $f(x) = e^x$ about $x_0 = 0$,

$$S_2(f,0)(x) = 1 + x + \frac{x^2}{2!} = 1 + x + \frac{x^2}{2},$$

we find

$$S_2(f,0)(1/2) = 1 + \frac{1}{2} + \frac{(1/2)^2}{2} = 1 + \frac{1}{2} + \frac{1}{8} = \frac{8 + 4 + 1}{8} = \frac{13}{8} = 1.625,$$

$$S_2(f,0)(-1/2) = 1 - \frac{1}{2} + \frac{(1/2)^2}{2} = 1 - \frac{1}{2} + \frac{1}{8} = \frac{8-4+1}{8} = \frac{5}{8} = 0.625.$$

To compare with the true values

$$f(1/2) = e^{1/2} \approx 1.6487 \qquad \text{and} \qquad f(-1/2) = e^{-1/2} \approx 0.60653,$$

we work out the errors

$$e^{1/2} - \frac{13}{8} \approx 0.023721 \qquad \text{and} \qquad e^{-1/2} - \frac{5}{8} \approx -0.018469.$$

We see that the approximation is worse than for the Taylor polynomial of degree $n = 3$; the error is of the order 2×10^{-2}.

If we take the linear approximation given by the Taylor polynomial of degree $n = 1$,

$$S_1(f,0)(x) = 1 + x,$$

then

$$S_1(f,0)(1/2) = 1 + \frac{1}{2} = \frac{3}{2} = 1.5, \qquad S_2(f,0)(-1/2) = 1 - \frac{1}{2} = \frac{1}{2} = 0.5,$$

which is clearly an extremely crude approximation. The actual errors are

$$e^{1/2} - 1.5 \approx 0.14872 \qquad \text{and} \qquad e^{-1/2} - 0.5 \approx 0.10653,$$

and we see that the errors are of the order 1.5×10^{-1}. □

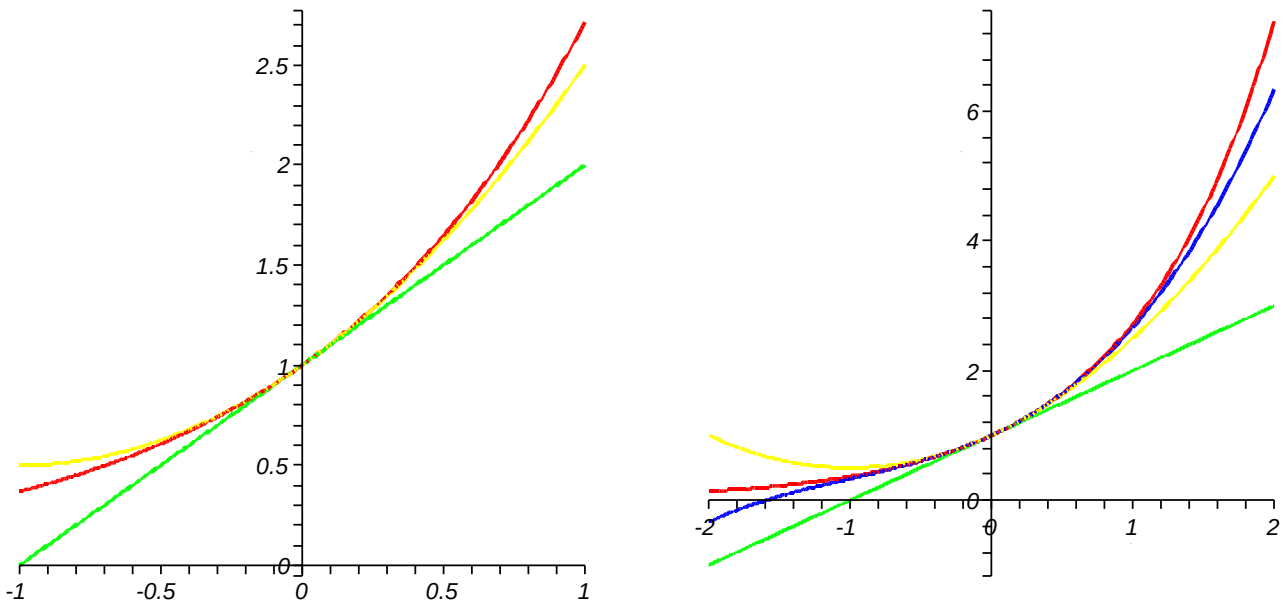


Figure 7.1: On the left, we see $f(x) = e^x$ (red), and its Taylor polynomials (about $x_0 = 0$) of degrees $n = 1$ (green) and $n = 2$ (yellow); and on the right, we see $f(x) = e^x$ (red), and its Taylor polynomials (about $x_0 = 0$) of degrees $n = 1$ (green), $n = 2$ (yellow), and $n = 3$ (blue).

Before we discuss another example, let us take a look at what we are doing: Assume that the function f has the power series expansion about x_0

$$T(f,x_0)(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}\,(x-x_0)^k = f(x_0) + f'(x_0)\,(x-x_0)$$
$$+\frac{f''(x_0)}{2!}\,(x-x_0)^2 + \frac{f'''(x_0)}{3!}\,(x-x_0)^3 + \frac{f^{(4)}(x_0)}{4!}\,(x-x_0)^4 + \ldots, \quad (7.50)$$

and that in addition

$$T(f,x_0)(x) = f(x) \qquad \text{for all } x \text{ close to } x_0.$$

For x very close to x_0, say $|x-x_0| \le 10^{-1}$, the terms $(x-x_0)^k$ with $k \ge 4$ will be very small. Indeed $|x-x_0|^4 \le (10^{-1})^4 = 10^{-4}$, and more generally

$$|x-x_0|^k \le (10^{-1})^k = 10^{-k}.$$

If we are satisfied with an accuracy up to the order 10^{-3}, then we can ignore the terms in the Taylor series with $k \ge 4$. Thus for x with $|x-x_0| \le 10^{-1}$, the Taylor polynomial of degree $n=3$,

$$S_3(f,x_0)(x) = f(x_0) + f'(x_0)\,(x-x_0) + \frac{f''(x_0)}{2!}\,(x-x_0)^2 + \frac{f'''(x_0)}{3!}\,(x-x_0)^3,$$

usually provides a **very good approximation of** $f(x)$ for x with $|x-x_0| \le 10^{-1}$. From our analysis we expect the approximation to be accurate up to the order 10^{-3}.

Approximations by the first two or three or four terms of a Taylor series are commonly used in physics, engineering, and mathematics. Towards the end of this section we will get some evidence that they are indeed a good approximation and we will even get a 'rule of thumb' for estimating the error.

Example 7.37 (approximation of $\sin(x)$ by its Taylor polynomials)
In Example 7.27 we saw that the Maclaurin series of $f(x) = \sin(x)$ is

$$\sin(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}\,x^{2k+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots.$$

Thus the Taylor polynomial of degree $n=3$ is given by

$$S_3(f,0) = x - \frac{x^3}{6}.$$

and we expect it to be a reasonable approximation of $f(x) = \sin(x)$ for **small** x.

Let's test this for some particular values of x. For $x = \pi/10$, a pocket calculator gives that

$$\sin(\pi/10) \approx 0.309017 \qquad \text{and} \qquad S_3(f,0)(\pi/10) = \frac{\pi}{10} - \frac{(\pi/10)^3}{6} \approx 0.308992.$$

The error of the approximation is of the order 2.5×10^{-5}, and so the approximation of $\sin(\pi/10)$ is very good.

However, for $x = \pi$

$$\sin(\pi) = 0 \qquad \text{and} \qquad S_3(f,0)(\pi) = \pi - \frac{\pi^3}{6} \approx -2.026120,$$

and so the approximation of $\sin(\pi)$ is very poor. This is not so surprising since π is comparatively far away from $x_0 = 0$. To overcome this, we could consider truncating after a greater number of terms. For example,

$$S_7(f,0)(\pi) = \pi - \frac{\pi^3}{3!} + \frac{\pi^5}{5!} - \frac{\pi^7}{7!} \approx -0.0752206,$$

which gives a better approximation of $\sin(\pi) = 0$. In Figure 7.2, we show $\sin(x)$ and its Taylor polynomials about $x_0 = 0$ of degree $n = 3$, $n = 5$, and $n = 7$. □

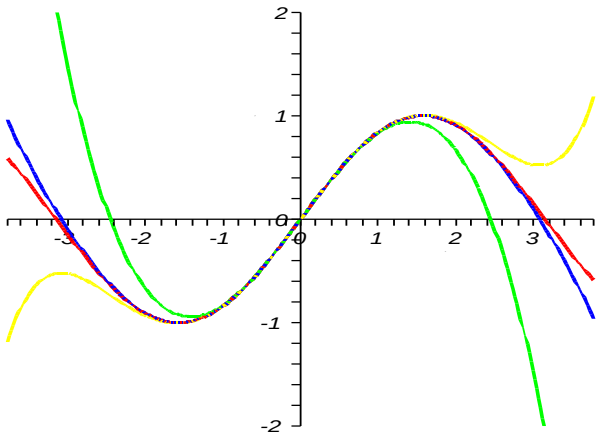


Figure 7.2: Graphs of $\sin(x)$ (red) and its Taylor polynomials about $x_0 = 0$ of degree $n = 3$ (green), $n = 5$ (yellow), and $n = 7$ (blue).

How large has the degree of the Taylor polynomial about x_0 to be, in order to get a **good approximation of the function close to** x_0**?** A rough answer to this question is given by the following **'rule of thumb'**, which you will apply on the corresponding exercise sheet.

Rule of Thumb: *The error of the Taylor polynomial*

$$S_n(f, x_0)(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

as an approximation of $f(x)$ for x close to x_0 has approximately the same order of magnitude as the first non-zero term that was omitted from the Taylor series expansion when we approximated $f(x)$ by $S_n(f, x_0)(x)$. In formulas,

$$\text{error} \approx \frac{f^{(n+1)}(x_0)}{(n+1)!} (x - x_0)^{n+1}$$

7.7 Differentiation and Integration of Power Series

In this very short section we discuss that power series may be **differentiated and integrated term by term**, and we give some examples.

Theorem 7.38 (differentiation/integration of power series)
Consider a power series about x_0,

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k. \tag{7.51}$$

*Then we may **differentiate the power series (7.51) term by term**, that is,*

$$\frac{d}{dx}\left[\sum_{k=0}^{\infty} a_k (x - x_0)^k\right] = \sum_{k=0}^{\infty} \frac{d}{dx}\left[a_k (x - x_0)^k\right] = \sum_{k=1}^{\infty} a_k\, k\, (x - x_0)^{k-1}. \tag{7.52}$$

*We may also **integrate (7.51) term by term**, and an **indefinite integral** of the power series (7.51) is*

$$\int_{x_0}^{x} \left[\sum_{k=0}^{\infty} a_k (t - x_0)^k\right] dt = \sum_{k=0}^{\infty} \int_{x_0}^{x} \left[a_k (t - x_0)^k\right] dt = \sum_{k=1}^{\infty} \frac{a_k}{k+1} (x - x_0)^{k+1}. \tag{7.53}$$

Usually, if (7.51) converges at some x then (7.52) and (7.53) will also converge at this x.

We can differentiate and integrate the power series of f arbitrarily often term by term.

Example 7.39 (derivative of $\sin(x)$)
For any real number $x \in \mathbb{R}$, we have

$$\sin(x) = \sum_{\ell=0}^{\infty} \frac{(-1)^\ell}{(2\ell+1)!}\, x^{2\ell+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots.$$

From Theorem 7.38, we have

$$\begin{aligned}\big[\sin(x)\big]' &= \frac{d}{dx}\left[\sum_{\ell=0}^{\infty} \frac{(-1)^\ell}{(2\ell+1)!}\, x^{2\ell+1}\right] = \sum_{\ell=0}^{\infty} \frac{d}{dx}\left[\frac{(-1)^\ell}{(2\ell+1)!}\, x^{2\ell+1}\right] \\ &= \sum_{\ell=0}^{\infty} \frac{(-1)^\ell}{(2\ell+1)!}\,(2\ell+1)\, x^{2\ell} = \sum_{\ell=0}^{\infty} \frac{(-1)^\ell}{(2\ell)!}\, x^{2\ell}.\end{aligned}$$

We recognize the last series as the Maclaurin series of $\cos(x)$. This is what we expected, since $\big[\sin(x)\big]' = \cos(x)$. □

Example 7.40 (primitive of $1/(1-x)$)
From Lemma 7.15 and Example 7.29, we know that the Maclaurin series of $f(x) = 1/(1-x)$ satisfies for $|x| < 1$

$$\frac{1}{1-x} = T(f,0)(x) = \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + x^3 + \dots.$$

We compute an indefinite integral on both sides.

$$\int_0^x \frac{1}{1-t}\, dt = -\left[\ln(1-t)\right]_0^x = -\ln(1-x) + \ln(1) = -\ln(1-x),$$

and integrating the series term by term gives

$$\int_0^x \left[\sum_{k=0}^{\infty} t^k\right] dt = \sum_{k=0}^{\infty} \int_0^x t^k\, dt = \sum_{k=0}^{\infty} \frac{t^{k+1}}{k+1}\bigg|_0^x = \sum_{k=0}^{\infty} \frac{x^{k+1}}{k+1} = \sum_{\ell=1}^{\infty} \frac{x^\ell}{\ell} = \sum_{k=1}^{\infty} \frac{x^k}{k},$$

where we have substituted the index $\ell = k+1$ in the second last step and have renamed $\ell = k$ in the last step. From Theorem 7.38, we know now that for x with $|x| < 1$,

$$-\ln(1-x) = \sum_{k=1}^{\infty} \frac{x^k}{k} = x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \dots. \tag{7.54}$$

We verify that this agrees with the Maclaurin of $\ln(1+x)$ series from Example 7.30: Multiplying (7.54) by (-1) yields

$$\ln(1-x) = -\sum_{k=1}^{\infty} \frac{x^k}{k} = \sum_{k=1}^{\infty} \frac{(-1)\,x^k}{k}.$$

If we know substitute $x = -y$, then we obtain for $|y| < 1$

$$\ln(1+y) = \sum_{k=1}^{\infty} \frac{(-1)\,(-y)^k}{k} = \sum_{k=1}^{\infty} \frac{(-1)\,(-1)^k\,y^k}{k} = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}\,y^k}{k} = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}\,y^k}{k}, \tag{7.55}$$

where we have removed one factor $(-1)^2 = 1$ from $(-1)^{k+1}$ in the last step. The series on the right-hand side is the Maclaurin series of $\ln(1+y)$ which we determined in Example 7.30. $\square$

Chapter 8

Complex Numbers

In this chapter we will get an introduction to **complex numbers**. From the mathematical point of view, the introduction of the complex numbers closes a 'gap' as now **every** quadratic equation has now two (complex) roots. The complex numbers are an **extension of the real numbers**, and real numbers are special cases of complex numbers. As we will learn, a complex number can be represented by a pair of real numbers and thus is geometrically a **point in the (complex) plane**. However, we will also learn two different representations of complex numbers: one involves trigonometric functions, and the other one is based on **Euler's formula** and is the so-called **polar form** of complex numbers. The ability to switch between the three representations of complex numbers facilitates adding, subtracting, multiplying, and dividing complex numbers enormously. We will learn how to take n**th roots** of complex numbers, and find that now any complex or real number has exactly n distinct complex nth roots. We will also introduce the **(natural) exponential function** and the **(natural) logarithm** for complex numbers. In physics and engineering complex numbers are used to describe **alternating current circuits**.

8.1 Introduction

The introduction of complex numbers is most natural starting with the equation

$$x^2 + 1 = 0 \qquad \Leftrightarrow \qquad x^2 = -1. \tag{8.1}$$

We know that this equation has no solution in the real numbers, but somehow this is not satisfactory: every quadratic equation should have two roots! To remedy this

we introduce the **imaginary unit** i as the (complex) number satisfying

$$\boxed{i^2 = -1.}$$

Then the equation (8.1) has two solutions $-i$ and $+i$. Indeed

$$(\pm i)^2 = i^2 = -1 \qquad \text{and} \qquad \pm\, i = \pm\sqrt{-1}.$$

We can multiply the imaginary unit with real numbers and add and subtract such multiples of i. For example

$$5\,i + 3\,i = 8\,i \quad \text{and} \quad 2\,i - 10\,i = -8\,i.$$

We can also take integer powers of (multiples of) the imaginary unit, where we simplify using $(\pm i)^2 = -1$. For example

$$(2\,i)^2 = 2^2\,i^2 = 4\,(-1) = -4 \quad \text{and} \quad (-3\,i)^3 = (-3)^3\,i^3 = -27\,i^2\,i = -27\,(-1)\,i = 27\,i.$$

Definition 8.1 (complex numbers)
*A **complex number** z is defined as a number of the form*

$$z = x + i\,y, \tag{8.2}$$

*where x and y are real numbers and i is the imaginary unit. In (8.2), the real number $\mathrm{Re}(z) = x$ is called the **real part** of z, and the real number $\mathrm{Im}(z) = y$ is called the **imaginary part** of z. The representation (8.2) of a complex number z is also referred to as the **Cartesian form**.*

Using $(\pm i)^2 = -1$ and otherwise performing multiplication as usual we can easily add, subtract and multiply complex numbers.

Lemma 8.2 (addition and subtraction of complex numbers)
Let $z_1 = x_1 + i\,y_1$ and $z_2 = x_2 + i\,y_2$ be two complex numbers.

(i) Addition:

$$\begin{aligned} z_1 + z_2 &= \big(x_1 + i\,y_1\big) + \big(x_2 + i\,y_2\big) \\ &= \big(x_1 + x_2\big) + i\,\big(y_1 + y_2\big). \end{aligned}$$

(ii) Substraction:

$$\begin{aligned} z_1 - z_2 &= \big(x_1 + i\,y_1\big) - \big(x_2 + i\,y_2\big) \\ &= \big(x_1 - x_2\big) + i\,\big(y_1 - y_2\big). \end{aligned}$$

Example 8.3 (addition and substraction of complex numbers)

$$\begin{aligned}(1+i\,2)+(4+i\,3) &= (1+4)+i\,(2+3) = 5+i\,5,\\ (2+i\,11)-(-1+i\,7) &= (2-(-1))+i\,(11-7) = 3+i\,4,\\ (3+i\,5)-(2-i\sqrt{3}) &= (3-2)+i\,\big(5-(-\sqrt{3})\big) = 1+i\,(5+\sqrt{3}).\end{aligned}$$

□

Lemma 8.4 (multiplication of complex numbers)
Let $z_1 = x_1 + i\,y_1$ and $z_2 = x_2 + i\,y_2$ be two complex numbers. Then the product of z_1 and z_2 is given by

$$\begin{aligned}z_1 \times z_2 = z_1\,z_2 &= (x_1+i\,y_1)\,(x_2+i\,y_2)\\ &= x_1\,x_2 + i\,x_1\,y_2 + i\,y_1\,x_2 + i^2\,y_1\,y_2\\ &= \big(x_1\,x_2 - y_1\,y_2\big) + i\,\big(x_1\,y_2 + y_1\,x_2\big).\end{aligned}$$

Example 8.5 (product of complex numbers)

$$\begin{aligned}(1+i\,2)\,(4+i\,3) &= 4+i\,3+i\,8+i^2\,6 = 4+i\,11+(-1)\,6 = -2+i\,11,\\ (2+i\sqrt{2})(2-i\sqrt{2}) &= 4-i\,2\sqrt{2}+i\,2\sqrt{2}-i^2\,2 = 4-(-1)\,2 = 6.\end{aligned}$$

□

8.2 The Complex Plane (Argand Diagram)

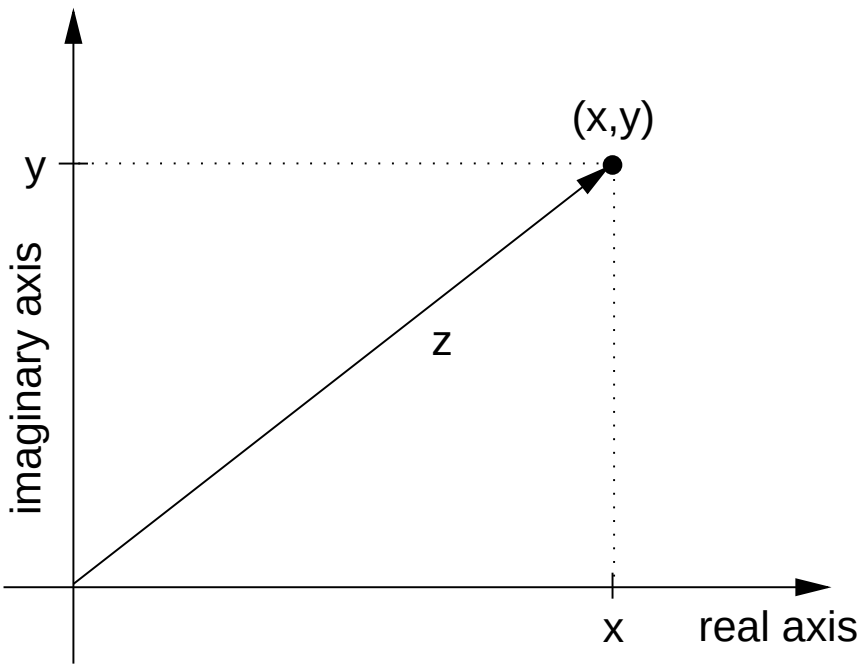


Figure 8.1: Visualization of $z = x + i\,y$ as the point (x, y) in the complex plane.

The so-called **complex plane** can be thought of as our usual (x, y)-plane, and it allows us to visualize complex numbers. More precisely, a complex number $z = x+i\,y$ can be represented as the **point** or **vector** (x, y) in the **complex plane** as illustrated in Figure 8.1. Diagrams like the one in Figure 8.1 (for visualizing complex numbers in the complex plane) are also called **Argand diagrams**.

With the help of this visualization of complex numbers the sum and difference of two complex numbers can be determined geometrically as illustrated in Figure 8.2. Let $z_1 = x_1 + i\,y_1$ and $z_2 = x_2 + i\,y_2$. The **sum** $w = z_1 + z_2$ is represented by the vertex of the parallelogram in the complex plane whose other three vertices are 0, z_1 and z_2. The **difference** $u = z_2 - z_1$ is represented by the vertex of the parallelogram whose other vertices are 0, $-z_1$ and z_2.

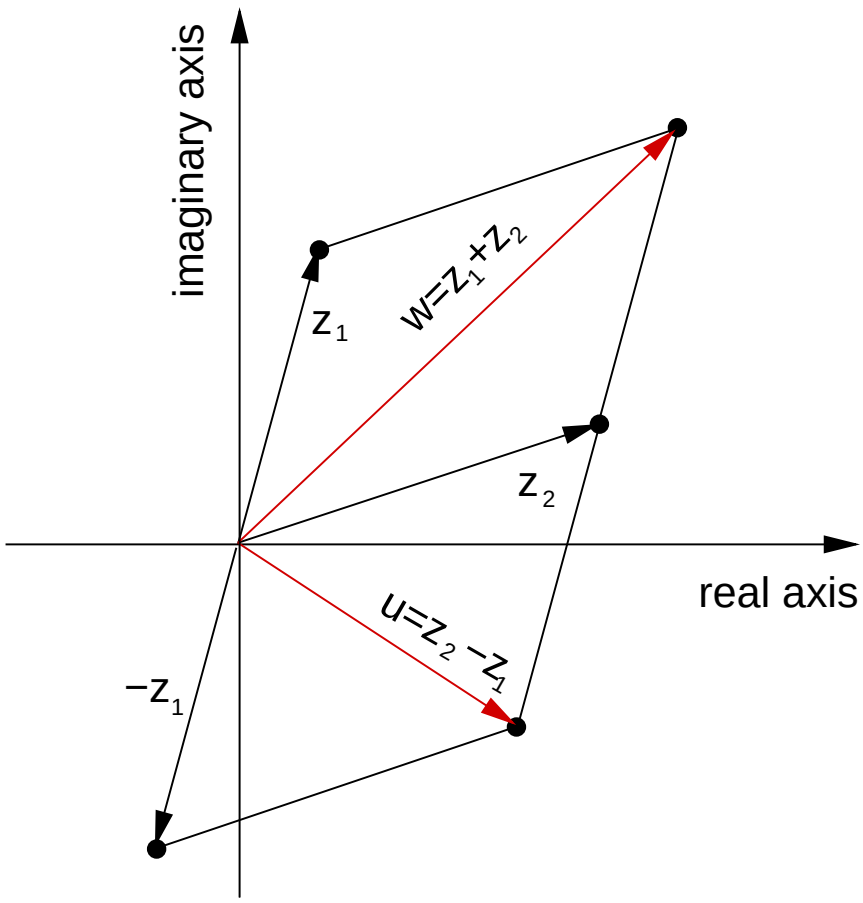


Figure 8.2: Visualization of the sum $w = z_1 + z_2$ and the difference $u = z_2 - z_1$ in the complex plane.

Definition 8.6 (complex conjugate of a complex number)
*The **complex conjugate** of the complex number $z = x + i\,y$ is defined by*

$$\bar{z} = x - i\,y.$$

The complex conjugate $\bar{z} = x - i\,y$ of the complex number $z = x + i\,y$ has the same real part as z, that is, $\mathrm{Re}(\bar{z}) = x = \mathrm{Re}(z)$, but the imaginary part $\mathrm{Im}(\bar{z}) = -y = -\mathrm{Im}(z)$ differs from $\mathrm{Im}(z) = y$ by the factor -1. Thus we see that in the complex plane the complex conjugate number can be obtained by **reflecting** the complex number z **on the real axis**.

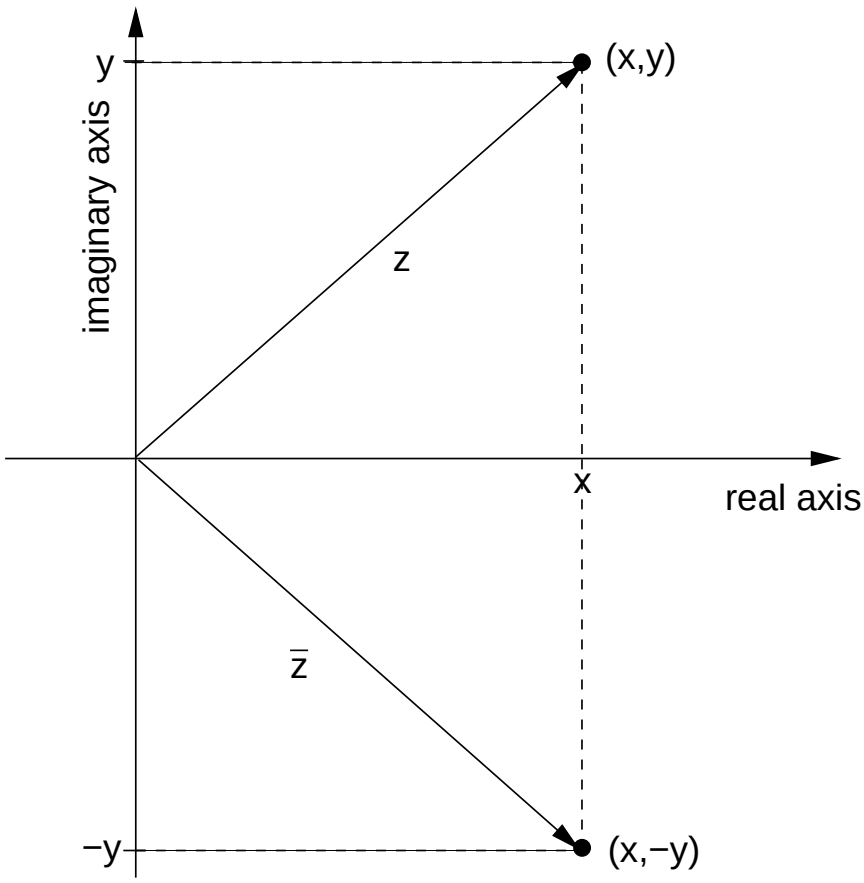


Figure 8.3: The complex conjugate $\bar{z} = x - i\,y$ of a complex number $z = x + i\,y$ is obtained by refection on the real axis.

Example 8.7 (complex conjugate number)

$$
\begin{aligned}
z &= 3 + i\,2, \qquad & \bar{z} &= 3 - i\,2;\\
w &= \sqrt{2} - i\,7, \qquad & \overline{w} &= \sqrt{2} + i\,7;\\
u &= -13 + i\,\frac{5}{3}, \qquad & \overline{u} &= -13 - i\,\frac{5}{3}.
\end{aligned}
\qquad \square
$$

We observe that for a complex number $z = x + i\,y$ the complex conjugate $\bar{\bar{z}}$ of the complex conjugate number $\bar{z} = x - i\,y$ of z is again z, that is,

$$\bar{\bar{z}} = \overline{x - i\,y} = \overline{x + i\,(-y)} = x - i\,(-y) = x + i\,y = z.$$

We will see that the complex conjugate $\bar{z} = x - i\,y$ of a complex number $z = x + i\,y$ is very useful for dividing by that complex number. We start with investigating what happens if we multiply a complex number $z = x + i\,y$ by its complex conjugate number $\bar{z} = x - i\,y$. Then

$$\begin{aligned} z\,\bar{z} &= \big(x + i\,y\big)\big(x - i\,y\big) \\ &= x^2 - i\,x\,y + i\,y\,x - i^2\,y^2 \\ &= x^2 - (-1)\,y^2 \;=\; x^2 + y^2. \end{aligned}$$

From $z\,\bar{z} = x^2 + y^2$ we see that the product of a complex number with its complex conjugate is **real** and **not negative**!

Definition 8.8 (modulus)
*The **modulus** of the complex number $z = x + iy$ is defined by*

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

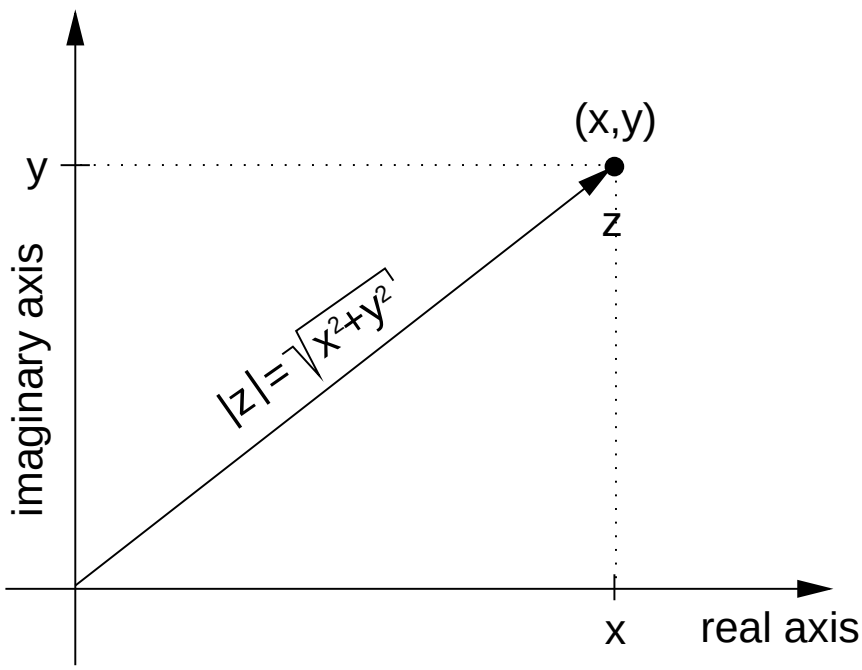


Figure 8.4: Visualization of the modulus $|z| = \sqrt{z\,\bar{z}} = \sqrt{x^2 + y^2}$ of the complex number $z = x + i\,y$.

From Pythagoras theorem, we see in Figure 8.4 that the modulus $|z| = \sqrt{z\,\bar{z}} = \sqrt{x^2 + y^2}$ is just the **length of the straight line from the origin** $(0, 0)$ **to the point** (x, y).

With the help of $z\,\bar{z} = x^2 + y^2$ we can now easily compute the **quotient** z_1/z_2 **of two complex numbers** $z_1 = x_1 + i\,y_1$ and $z_2 = x_2 + i\,y_2 \neq 0$. To do this we multiply both the numerator and the denominator of the quotient

$$\frac{z_1}{z_2} = \frac{x_1 + i\,y_1}{x_2 + i\,y_2}$$

by $\bar{z}_2 = x_2 - i\,y_2$ and use $z_2\,\bar{z}_2 = x_2^2 + y_2^2$

$$\frac{z_1}{z_2} = \frac{z_1\,\bar{z}_2}{z_2\,\bar{z}_2} = \frac{(x_1 + i\,y_1)(x_2 - i\,y_2)}{(x_2 + i\,y_2)(x_2 - i\,y_2)} = \frac{(x_1 + i\,y_1)(x_2 - i\,y_2)}{x_2^2 + y_2^2} = \frac{z_1\,\bar{z}_2}{|z_2|^2}.$$

We summarize this in a lemma:

Lemma 8.9 (division by complex numbers)
The division of $z_1 = x_1 + i\,y_1$ *by* $z_2 = x_2 + i\,y_2 \neq 0$ *is given by*

$$\frac{z_1}{z_2} = \frac{z_1\,\bar{z}_2}{z_2\,\bar{z}_2} = \frac{z_1\,\bar{z}_2}{|z_2|^2} = \frac{(x_1 + i\,y_1)(x_2 - i\,y_2)}{x_2^2 + y_2^2}. \tag{8.3}$$

We note that the two right-most expressions in (8.3) bring us back to the case of the product of two complex numbers since $|z_2|^2 = x_2^2 + y_2^2$ is a positive real number.

Example 8.10 (division of complex numbers)
Express the following quotients in Cartesian form $z = x + i\,y$.

$$\text{(a)}\quad \frac{2+i}{3-i}, \qquad \text{(b)}\quad \frac{1}{i^3}, \qquad \text{(c)}\quad \frac{1}{3+i\,4}.$$

Solution: We use Lemma 8.9 above.

(a) We multiply the numerator and denominator with $\overline{3-i} = 3+i$ and simplify

$$\frac{2+i}{3-i} = \frac{(2+i)(3+i)}{(3-i)(3+i)} = \frac{6+2\,i+3\,i+i^2}{9+3\,i-3\,i-i^2} = \frac{6+2\,i+3\,i-1}{9-(-1)} = \frac{5+5\,i}{10} = \frac{1}{2} + i\,\frac{1}{2}.$$

(b) We observe that $i^3 = i^2\,i = (-1)\,i = -i$. Now we multiply the numerator and denominator by $\overline{i^3} = \overline{-i} = i$. This yields

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

(c) We multiply the numerator and denominator with $\overline{3+i\,4} = 3 - i\,4$ and obtain (using $(3+i\,4)(3-i\,4) = 3^2 + 4^2 = 9 + 16 = 25$)

$$\frac{1}{3+i\,4} = \frac{3-i\,4}{(3+i\,4)(3-i\,4)} = \frac{3-i\,4}{9+16} = \frac{3-i\,4}{25} = \frac{3}{25} - i\,\frac{4}{25}.$$

Do not forget to simplify as far as possible! □

8.3 Polar Form of Complex Numbers

Another representation of complex numbers in the complex plane is given with the help of so-called **polar coordinates**. From Figure 8.5 we see that the complex number $z = x + i\,y$, or equivalently the point (x, y), can be described by giving the **angle** $\phi \in (-\pi, \pi]$ and the **length** r of the line from the origin $(0, 0)$ to the point (x, y).

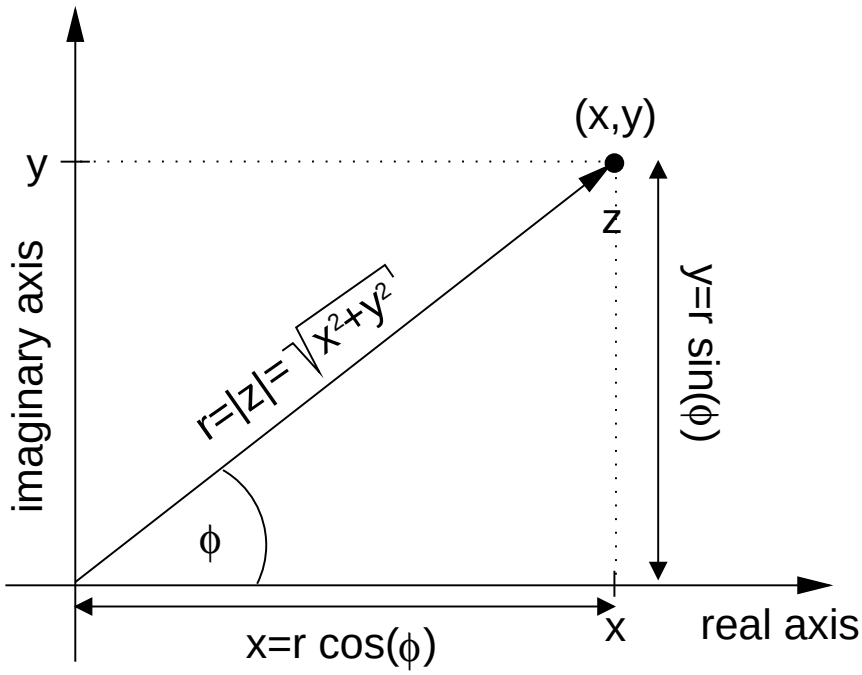


Figure 8.5: Visualization of $z = x + i\,y$ in the polar coordinate representation with angle $\phi \in (-\pi, \pi]$ and radius $r > 0$.

From Figure 8.5 we see that

$$\boxed{x = r\,\cos(\phi), \qquad y = r\,\sin(\phi), \qquad \text{and} \qquad r = |z| = \sqrt{x^2 + y^2},} \tag{8.4}$$

and obtain the following representation of $z = x + i\,y$ in terms of r and the angle ϕ

$$\boxed{z = r\,\cos(\phi) + i\,r\,\cos(\phi) = r\,\big(\cos(\phi) + i\,\sin(\phi)\big)} \tag{8.5}$$

We will say that (r, ϕ) is the **polar coordinate representation** of (8.5) We call $\phi \in (-\pi, \pi]$ the **angle** of z and $r > 0$ the **radius** of z.

If we have only $z = x + iy$ and do not know r and ϕ, how do we find them?

Clearly $r = |z| = \sqrt{x^2 + y^2}$ and from (8.4)

$$\boxed{\cos(\phi) = \frac{x}{r} \qquad \text{and} \qquad \sin(\phi) = \frac{y}{r}.} \tag{8.6}$$

We have also $\tan(\phi) = y/x$ but this equation is less useful than (8.6), since it does not allow the determination of ϕ without additional information. From the periodicity of the trigonometric functions $\sin(\phi)$ and $\cos(\phi)$, we know that

$$\sin(\phi) = \sin(\phi + k\,2\pi) \qquad \text{for all } -\pi < \phi \leq \pi \text{ and all } k \in \mathbb{Z},$$

$$\cos(\phi) = \cos(\phi + k\,2\pi) \qquad \text{for all } -\pi < \phi \leq \pi \text{ and all } k \in \mathbb{Z}.$$

Thus it is not possible to find a **uniquely determined** argument/angle ϕ that satisfies (8.6). However, for $-\pi < \phi \leq \pi$, the angle ϕ corresponds to a **uniquely determined point on the unit circle**, and we will explain below how to exploit this to find ϕ satisfying $-\pi < \phi \leq \pi$ and (8.6).

Definition 8.11 (argument and principal value of the argument)
Let $z = x + i\,y \neq 0$ be a complex number. Let $r \in \mathbb{R}$ with $r > 0$ and $\phi \in \mathbb{R}$ be such that

$$z = r\,\cos(\phi) + i\,r\,\cos(\phi) = r\,\big(\cos(\phi) + i\,\sin(\phi)\big).$$

*Then ϕ is called the **argument of** z, denoted by $\phi = \arg(z)$. The **uniquely determined argument** $\phi = \arg(z)$ **of** z **with** $-\pi < \phi \leq \pi$ is called the **principal value** of the argument of z. The principal value of the argument of z is denoted by $\mathrm{Arg}(z)$.*

Given $z = x + i\,y$, how do we find the principal value $\mathrm{Arg}(z)$ of the argument of z?

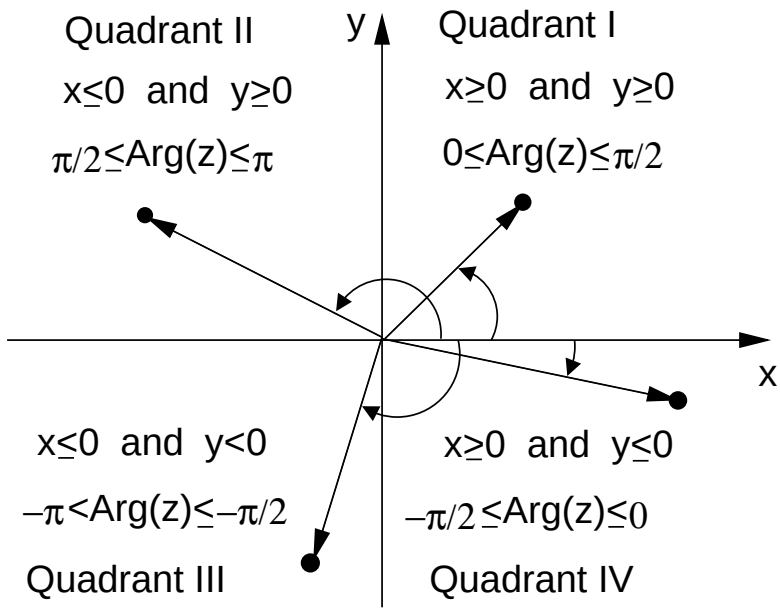


Figure 8.6: The signs of $x = r\,\cos(\phi)$ and $y = r\,\sin(\phi)$ determine the quadrant in which the principal value $\phi = \mathrm{Arg}(z)$ of the argument of $z = x + i\,y$ is located.

Once we have calculated $r = |z| = \sqrt{x^2 + y^2}$, we have from (8.6) that

$$\cos(\phi) = \frac{x}{r} \qquad \text{and} \qquad \sin(\phi) = \frac{y}{r}. \tag{8.7}$$

Since the principle value $\phi = \mathrm{Arg}(z)$ of the argument of z satisfies by definition $-\pi < \phi \leq \pi$, there exists **exactly one angle** $\phi = \mathrm{Arg}(z)$ for which (8.7) holds true. Interpreting $\big(\cos(\phi), \sin(\phi)\big)$ as a point on the unit circle, then the **signs of** x **and** y are the signs of $\cos(\phi)$ and $\sin(\phi)$, respectively, and they **tell us in which quadrant the angle** $\phi = \mathrm{Arg}(z)$ **is located**. This is illustrated in Figure 8.6. Once we have determined which quadrant contains $\phi = \mathrm{Arg}(z)$, we can easily determine the principle value $\phi = \mathrm{Arg}(z)$ from the value of $\sin(\phi) = y/r$ or $\cos(\phi) = x/r$.

In order to find the principle value $\phi = \mathrm{Arg}(z)$ of the argument of z with the help of (8.7), it is useful to know the following table.

angle ϕ in radians	0	$\pi/6$	$\phi/4$	$\phi/3$	$\pi/2$
angle ϕ in degree	0	30	45	60	90
$\sin(\phi)$	$\frac{\sqrt{0}}{2} = 0$	$\frac{\sqrt{1}}{2} = \frac{1}{2}$	$\frac{\sqrt{2}}{2} = \frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{4}}{2} = 1$
$\cos(\phi)$	$\frac{\sqrt{4}}{2} = 1$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2} = \frac{1}{\sqrt{2}}$	$\frac{\sqrt{1}}{2} = \frac{1}{2}$	$\frac{\sqrt{0}}{2} = 0$

Table 8.1: Important values of $\sin(\phi)$ and $\cos(\phi)$ in the first quadrant. Note that the values are all of the form $\sqrt{k}/2$ with $k \in \{0, 1, 2, 3, 4\}$.

Each complex number $z = x + i\,y$ has a **unique polar coordinate representation**

$$\boxed{(r, \phi), \qquad \text{where} \qquad r = |z| = \sqrt{x^2 + y^2} \quad \text{and} \quad \phi = \mathrm{Arg}(z).} \tag{8.8}$$

We demonstrate for some examples of complex numbers $z = x + i\,y$ how we find the polar coordinate representation (8.8).

When calculating $\mathrm{Arg}(z)$, it may be helpful to sketch an Argand diagram of z.

Example 8.12 (polar coordinate representation (r, ϕ) with $\phi = \mathrm{Arg}(z)$)
Find the polar coordinate representation (r, ϕ) of the complex number $z = 1 + i\sqrt{3}$.

Solution: We have $x = 1 > 0$ and $y = \sqrt{3} > 0$ and thus the principal value $\phi = \mathrm{Arg}(z)$ of the argument of z lies in the first quadrant, that is, $0 \le \phi \le \pi/2$ (see also left picture in Figure 8.7). We find that

$$r = \sqrt{x^2 + y^2} = \sqrt{1^2 + (\sqrt{3})^2} = \sqrt{4} = 2,$$

and, since $x = 1$ and $y = \sqrt{3}$, we know

$$\cos(\phi) = \frac{x}{r} = \frac{1}{2} \qquad \text{and} \qquad \sin(\phi) = \frac{y}{r} = \frac{\sqrt{3}}{2}.$$

Thus from Table 8.1, the principle value is $\phi = \mathrm{Arg}(z) = \pi/3$ and $z = 1 + i\sqrt{3}$ has the polar coordinate representation $(r, \phi) = (2, \pi/3)$. (It is always useful to test the result by converting back:

$$r\left(\cos(\phi) + i\,\sin(\phi)\right) = 2\left(\cos(\pi/3) + i\,\sin(\pi/3)\right) = 2\left(\frac{1}{2} + i\,\frac{\sqrt{3}}{2}\right) = 1 + i\sqrt{3},$$

as expected.) □

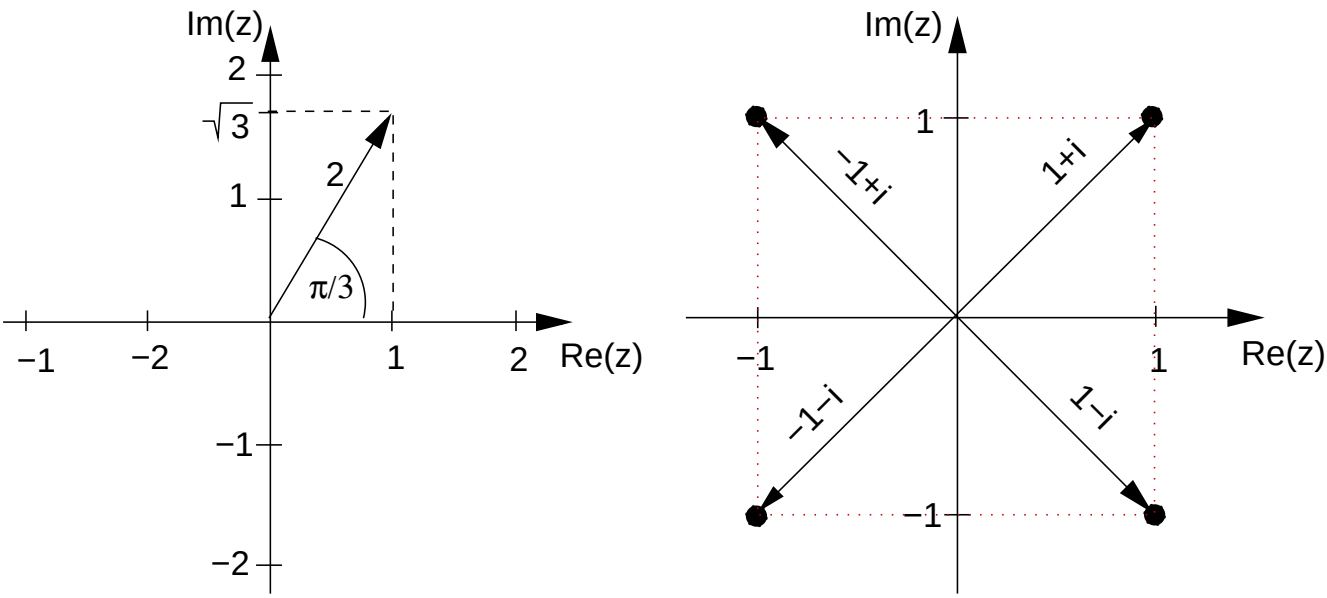


Figure 8.7: Argand diagram of $1 + i\sqrt{3}$ on the left, and on the right, an Argand diagram of the complex numbers $1 + i$, $-1 + i$, $1 - i$, and $-1 - i$.

Example 8.13 (polar coordinate representation (r, ϕ) with $\phi = \mathrm{Arg}(z)$)
Determine the polar coordinate representation (r, ϕ) of each of the following complex numbers:

(a) $z = 1 + i$, (b) $z = -1 + i$, (c) $z = 1 - i$, (d) $z = -1 - i$.

Solution: We have indicated the four complex numbers in the Argand diagram on the right-hand side in Figure 8.7.

(a) We have $x = 1$ and $y = 1$, and thus

$$r = |z| = \sqrt{x^2 + y^2} = \sqrt{1^2 + 1^2} = \sqrt{2}.$$

Thus the principal argument $\phi = \mathrm{Arg}(z)$ satisfies

$$\cos(\phi) = \frac{x}{r} = \frac{1}{\sqrt{2}} \qquad \text{and} \qquad \sin(\phi) = \frac{y}{r} = \frac{1}{\sqrt{2}}.$$

Since $\cos(\phi) > 0$ and $\sin(\phi) > 0$, we know that $\phi = \mathrm{Arg}(z)$ lies in the first quadrant, that is, $0 \le \phi \le \pi/2$. Thus, from Table 8.1, $\mathrm{Arg}(z) = \phi = \pi/4$. Thus the polar coordinate representation of $z = 1 + i$ is $(r, \phi) = (\sqrt{2}, \pi/4)$.

(b) We have $x = -1$ and $y = 1$, and thus

$$r = |z| = \sqrt{x^2 + y^2} = \sqrt{(-1)^2 + 1^2} = \sqrt{2}.$$

Thus the principal argument $\phi = \mathrm{Arg}(z)$ satisfies

$$\cos(\phi) = \frac{x}{r} = \frac{-1}{\sqrt{2}} \qquad \text{and} \qquad \sin(\phi) = \frac{y}{r} = \frac{1}{\sqrt{2}}.$$

Since $\cos(\phi) < 0$ and $\sin(\phi) > 0$, we know that $\phi = \mathrm{Arg}(z)$ lies in the second quadrant, that is, $\pi/2 \le \phi \le \pi$. Thus $-\pi/2 \ge -\phi \ge -\pi$ and from adding π we get $\pi/2 \ge \pi - \phi \ge 0$, that is, $\pi - \phi$ is in the first quadrant. From Lemma 2.10,

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

Thus we find, using Table 8.1, $\pi - \phi = \pi/4$, or equivalently $\phi = \pi - \pi/4 = 3\pi/4$. Thus $\mathrm{Arg}(z) = \phi = 3\pi/4$, and the polar coordinate representation of $z = -1 + i$ is $(r, \phi) = (\sqrt{2}, 3\pi/4)$.

(c) We have $x = 1$ and $y = -1$, and thus

$$r = |z| = \sqrt{x^2 + y^2} = \sqrt{1^2 + (-1)^2} = \sqrt{2}.$$

Thus the principal argument $\phi = \mathrm{Arg}(z)$ satisfies

$$\cos(\phi) = \frac{x}{r} = \frac{1}{\sqrt{2}} \qquad \text{and} \qquad \sin(\phi) = \frac{y}{r} = \frac{-1}{\sqrt{2}}.$$

Since $\cos(\phi) > 0$ and $\sin(\phi) < 0$, we know that $\phi = \mathrm{Arg}(z)$ lies in the fourth quadrant, that is, $-\pi/2 \le \phi \le 0$. Thus $-\phi$ satisfies $\pi/2 \ge -\phi \ge 0$, and hence $-\phi$ lies in the first quadrant. From $\cos(-\phi) = \cos(\phi) = 1/\sqrt{2}$ and Table 8.1, we find $-\phi = \pi/4$, and thus $\mathrm{Arg}(z) = \phi = -\pi/4$. Thus the polar coordinate representation of $z = 1 - i$ is $(r, \phi) = (\sqrt{2}, -\pi/4)$.

(d) We have $x = -1$ and $y = -1$, and thus

$$r = |z| = \sqrt{x^2 + y^2} = \sqrt{(-1)^2 + (-1)^2} = \sqrt{2}.$$

Thus the principal argument $\phi = \mathrm{Arg}(z)$ satisfies

$$\cos(\phi) = \frac{x}{r} = \frac{-1}{\sqrt{2}} \qquad \text{and} \qquad \sin(\phi) = \frac{y}{r} = \frac{-1}{\sqrt{2}}.$$

Since $\cos(\phi) < 0$ and $\sin(\phi) < 0$, we know that $\phi = \mathrm{Arg}(z)$ lies in the third quadrant, that is, $-\pi < \phi \leq -\pi/2$. Thus, from adding π, $0 < \phi + \pi \leq \pi/2$, and hence $\phi + \pi$ lies in the first quadrant. From Lemma 2.10,

$$\sin(\phi + \pi) = \sin(\phi)\,\cos(\pi) + \sin(\pi)\,\cos(\phi) = \sin(\phi)\,(-1) + 0 = -\sin(\phi) = \frac{1}{\sqrt{2}}.$$

Using Table 8.1, we see that $\phi + \pi = \pi/4$. Thus $\mathrm{Arg}(z) = \phi = \pi/4 - \pi = -3\pi/4$, and the polar coordinate representation of $z = -1 - i$ is $(r, \phi) = (\sqrt{2}, -3\pi/4)$. □

The third and last representation of complex numbers follows from (8.5) with Euler's formula below.

Lemma 8.14 (Euler's formula)
Let $\phi \in \mathbb{R}$. Then

$$\exp(i\,\phi) = e^{i\,\phi} = \cos(\phi) + i\,\sin(\phi). \tag{8.9}$$

To understand (8.9), we have to say in which sense we interpret $\exp(i\,\phi) = e^{i\phi}$ in formula (8.9). Since we know how to compute sums and integer powers of complex numbers, we can define $\exp(i\,\phi)$ via the **Maclaurin series expansion of** $\exp(x)$.

The Maclaurin series of $\exp(x) = e^x$ converges for all $x \in \mathbb{R}$ to the function $\exp(x) = e^x$ (see Example 7.26), that is,

$$\exp(x) = e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots \quad \text{for all } x \in \mathbb{R}. \tag{8.10}$$

To defined $\exp(i\,\phi) = e^{i\phi}$ we now substitute $x = i\,\phi$ in the series expansion (8.10). Thus we find

$$\exp(i\,\phi) = e^{i\phi} = \sum_{k=0}^{\infty} \frac{(i\,\phi)^k}{k!} = 1 + i\,\phi + \frac{(i\,\phi)^2}{2!} + \frac{(i\,\phi)^3}{3!} + \frac{(i\,\phi)^4}{4!} + \ldots$$

$$= 1 + i\,\phi + \frac{(-1)\,\phi^2}{2!} + \frac{(-1)\,i\,\phi^3}{3!} + \frac{\phi^4}{4!} + \ldots = \sum_{\ell=0}^{\infty} \frac{(i\,\phi)^{2\ell}}{(2\ell)!} + \sum_{\ell=0}^{\infty} \frac{(i\,\phi)^{2\ell+1}}{(2\ell+1)!}$$

$$= \sum_{\ell=0}^{\infty} \frac{(-1)^\ell \, \phi^{2\ell}}{(2\ell)!} + i \sum_{\ell=0}^{\infty} \frac{(-1)^\ell \, \phi^{2\ell+1}}{(2\ell+1)!} = \left[1 - \frac{\phi^2}{2!} + \frac{\phi^4}{4!} - \dots\right] + i \left[\phi - \frac{\phi^3}{3!} + \dots\right], \tag{8.11}$$

where we have used $(i\,\phi)^{2\ell} = (i^2)^\ell \, \phi^{2\ell} = (-1)^\ell \, \phi^{2\ell}$ and $(i\,\phi)^{2\ell+1} = i\,(i^2)^\ell \, \phi^{2\ell+1} = i\,(-1)^\ell \, \phi^{2\ell+1}$. From the last chapter, we know that (see Examples 7.27 and 7.28)

$$\cos(x) = \sum_{\ell=0}^{\infty} \frac{(-1)^\ell \, x^{2\ell}}{(2\ell)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \qquad \text{for all } x \in \mathbb{R},$$

$$\sin(x) = \sum_{\ell=0}^{\infty} \frac{(-1)^\ell \, x^{2\ell+1}}{(2\ell+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \qquad \text{for all } x \in \mathbb{R}.$$

Thus we recognize the series in the last line of (8.11) as the Maclaurin series of $\cos(\phi)$ and $\sin(\phi)$, respectively. Replacing the series in the last line of (8.11) by $\cos(\phi)$ and $\sin(\phi)$, respectively, we obtain

$$\exp(i\,\phi) = e^{i\,\phi} = \cos(\phi) + i\,\sin(\phi),$$

and we have just proved Euler's formula.

Euler's formula (8.9) for $\phi = \pi$ gives the following interesting formula

$$e^{i\pi} = \cos(\pi) + i\,\sin(\pi) = -1 + 0 \qquad \Leftrightarrow \qquad e^{i\pi} + 1 = 0.$$

This equation links the five fundamental numbers 0, 1, e, π, and i.

If we use Euler's formula for a complex number $z = x + i\,y$, with polar coordinate representation (r, ϕ) with radius $r = |z|$ and $\phi = \mathrm{Arg}(z)$, then we obtain

$$z = r\,\big(\cos(\phi) + i\,\sin(\phi)\big) = r\,e^{i\,\phi}.$$

Definition 8.15 (polar form of a complex number)
*The **polar form** of a complex number $z = x + i\,y$ is*

$$z = r\,e^{i\,\phi}, \qquad \textit{where} \quad r = |z| \quad \textit{and} \quad \phi = \mathrm{Arg}(z). \tag{8.12}$$

Example 8.16 (polar form of complex numbers)
Find the polar form of the complex numbers

$$\text{(a)} \quad z = 1 + i\,\sqrt{3}, \qquad \text{(b)} \quad u = 1 + i, \qquad \text{(c)} \quad w = 1 - i.$$

Solution:

(a) In Example 8.12, we found that $z = 1 + i\sqrt{3}$ has the polar coordinate representation $(r, \phi) = (2, \pi/3)$. Thus from (8.12), $z = 1 + i\sqrt{3}$ has the polar form

$$z = 2\, e^{i\,\pi/3}.$$

(b) From Example 8.13 (a), we know that $u = 1 + i$ has the polar coordinate representation $(r, \phi) = (\sqrt{2}, \pi/4)$. Thus from (8.12), $u = 1 + i$ has the polar form

$$u = \sqrt{2}\, e^{i\,\pi/4}.$$

(c) From Example 8.13 (c), we know that $w = 1 - i$ has the polar coordinate representation $(r, \phi) = (\sqrt{2}, -\pi/4)$. Thus from (8.12), $w = 1 - i$ has the polar form

$$w = \sqrt{2}\, e^{-i\,\pi/4}. \qquad \square$$

Multiplication and division of complex numbers in polar form:

The polar form of complex numbers gives us a much easier way of multiplying and dividing complex numbers. If $z_1 = r_1\, e^{i\,\phi_1}$ and $z_2 = r_2\, e^{i\,\phi_2}$ are two complex numbers in polar form, then

$$\boxed{z_1\, z_2 = r_1\, e^{i\,\phi_1}\, r_2\, e^{i\,\phi_2} = r_1\, r_2\, e^{i\,(\phi_1+\phi_2)}.}$$

Thus we have

$$\boxed{|z_1\, z_2| = |z_1|\,|z_2| \qquad \text{and} \qquad \arg(z_1\, z_2) = \operatorname{Arg}(z_1) + \operatorname{Arg}(z_2).}$$

However, we may need to add $\pm 2\pi$ to $\operatorname{Arg}(z_1) + \operatorname{Arg}(z_2) = \phi_1 + \phi_2$ to obtain $\operatorname{Arg}(z_1\, z_2)$, since there is not guarantee that $\phi_1 + \phi_2 \in (-\pi, \pi]$.

Furthermore, the quotient of two complex number $z_1 = r_1\, e^{i\,\phi_1}$ and $z_2 = r_2\, e^{i\,\phi_2}$ in polar form, where $r_2 \neq 0$, is given by

$$\boxed{\frac{z_1}{z_2} = \frac{r_1\, e^{i\,\phi_1}}{r_2\, e^{i\,\phi_2}} = \frac{r_1}{r_2}\, e^{i\,(\phi_1-\phi_2)}.}$$

So, we find that

$$\boxed{\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|} \qquad \text{and} \qquad \arg\left(\frac{z_1}{z_2}\right) = \operatorname{Arg}(z_1) - \operatorname{Arg}(z_2).}$$

Again, we may need to add $\pm 2\pi$ to $\operatorname{Arg}(z_1) - \operatorname{Arg}(z_2) = \phi_1 - \phi_2$ to obtain $\operatorname{Arg}(z_1/z_2)$.

In particular, if $z = r\, e^{i\,\phi}$ and $r \neq 0$ then

$$\frac{1}{z} = \frac{1\, e^{i\,0}}{r\, e^{i\,\phi}} = \frac{1}{r}\, e^{-i\,\phi}.$$

Example 8.17 (multiplication and division in polar form)
Use the polar form of $z = 1 + i\sqrt{3}$ and $w = 1 + i$ to compute

$$\text{(a)} \qquad \big(1 + i\sqrt{3}\big)(1 + i), \qquad\qquad \text{(b)} \qquad \frac{1 + i\sqrt{3}}{1 + i}.$$

Solution: From Example 8.16 (a) and (b), we know that

$$z = 1 + i\sqrt{3} = 2\,e^{i\,\pi/3} \qquad \text{and} \qquad w = 1 + i = \sqrt{2}\,e^{i\,\pi/4}.$$

Thus we have that

$$\text{(a)} \qquad \big(1 + i\sqrt{3}\big)(1 + i) = \big(2\,e^{i\,\pi/3}\big)\big(\sqrt{2}\,e^{i\,\pi/4}\big) = 2\sqrt{2}\,e^{i\,(\frac{\pi}{3}+\frac{\pi}{4})} = 2\sqrt{2}\,e^{i\,7\pi/12},$$

$$\text{(b)} \qquad \frac{1 + i\sqrt{3}}{1 + i} = \frac{2\,e^{i\,\pi/3}}{\sqrt{2}\,e^{i\,\pi/4}} = \frac{2}{\sqrt{2}}\,e^{i\,(\frac{\pi}{3}-\frac{\pi}{4})} = \sqrt{2}\,e^{i\,\pi/12}. \qquad \square$$

Remark 8.18 (use Cartesian form for addition and subtraction)
While the multiplication and division of complex numbers is much easier in the polar form, the polar form does not allow an easy handling of addition and subtraction. Therefore you should ***use the Cartesian form of complex numbers for addition and subtraction!***

Example 8.19 (addition and subtraction of complex numbers)
Calculate

$$2\,e^{i\,\pi/3} + \sqrt{2}\,e^{i\,\pi/4} \qquad \text{and} \qquad 2\,e^{i\,\pi/3} - \sqrt{2}\,e^{i\,\pi/4}.$$

Solution: Let $z_1 = 2\,e^{i\,\pi/3}$ and $z_2 = \sqrt{2}\,e^{i\,\pi/4}$. Then $z_1 = r_1\,e^{i\,\phi_1}$ and $z_2 = r_2\,e^{i\,\phi_2}$ with

$$r_1 = 2, \qquad \phi_1 = \frac{\pi}{3}, \qquad r_2 = \sqrt{2}, \qquad \phi_2 = \frac{\pi}{4}.$$

We convert z_1 and z_2 into Cartesian form so that we can easily perform the addition and subtraction. We have $z_1 = x_1 + i\,y_1$ and $z_2 = x_2 + i\,y_2$ with

$$x_1 = r_1\cos(\phi_1) = 2\,\cos\left(\frac{\pi}{3}\right) = 2 \times \frac{1}{2} = 1,$$

$$y_1 = r_1\sin(\phi_1) = 2\,\sin\left(\frac{\pi}{3}\right) = 2 \times \frac{\sqrt{3}}{2} = \sqrt{3},$$

$$x_2 = r_2\cos(\phi_2) = \sqrt{2}\,\cos\left(\frac{\pi}{4}\right) = \sqrt{2} \times \frac{1}{\sqrt{2}} = 1,$$

$$y_2 = r_2\,\sin(\phi_2) = \sqrt{2}\,\sin\left(\frac{\pi}{4}\right) = \sqrt{2} \times \frac{1}{\sqrt{2}} = 1.$$

Thus $z_1 = x_1 + i\,y_1 = 1 + i\sqrt{3}$ and $z_2 = x_2 + i\,y_2 = 1 + i$, and so

$$2\,e^{i\,\pi/3} + \sqrt{2}\,e^{i\,\pi/4} = (1 + i\sqrt{3}) + (1 + i) = 2 + i\,(\sqrt{3} + 1),$$
$$2\,e^{i\,\pi/3} - \sqrt{2}\,e^{i\,\pi/4} = (1 + i\sqrt{3}) - (1 + i) = i\,(\sqrt{3} - 1). \qquad \square$$

Remark 8.20 (Complex conjugate in polar form)
Let $z = r\,e^{i\,\phi}$. Then from Euler's formula

$$z = r\,e^{i\,\phi} = r\,\big(\cos(\phi) + i\,\sin(\phi)\big) = r\,\cos(\phi) + i\,r\,\sin(\phi).$$

Thus we have that the complex conjugate of $z = r\,e^{i\,\phi}$ is given by

$$\bar{z} = r\,\cos(\phi) - i\,r\,\sin(\phi) = r\,\cos(-\phi) + i\,r\,\sin(-\phi) = r\,e^{-i\,\phi},$$

where we have used $\cos(\phi) = \cos(-\phi)$ and $-\sin(\phi) = \sin(-\phi)$ in the second last step. Thus we find

$$\boxed{\overline{r\,e^{i\,\phi}} = r\,e^{-i\,\phi}.} \tag{8.13}$$

Remark 8.21 (relation between trigonometric and hyperbolic functions)
From Euler's formula (8.9) and from (8.13), we know that for $\phi \in \mathbb{R}$

$$e^{i\,\phi} = \cos(\phi) + i\sin(\phi), \qquad e^{-i\,\phi} = \overline{e^{i\,\phi}} = \cos(\phi) - i\sin(\phi).$$

So from adding the two equations and subtracting the second from the first, respectively, we obtain

$$e^{i\,\phi} + e^{-i\,\phi} = 2\,\cos(\phi) \qquad \textit{and} \qquad e^{i\,\phi} - e^{-i\,\phi} = 2\,i\,\sin(\phi),$$

and thus we find

$$\cosh(i\,\phi) = \frac{e^{i\,\phi} + e^{-i\,\phi}}{2} = \cos(\phi) \qquad \textit{and} \qquad \frac{1}{i}\,\sinh(i\,\phi) = \frac{e^{i\,\phi} - e^{-i\,\phi}}{2i} = \sin(\phi).$$

Letting $\theta = i\,\phi$ gives that

$$\phi = \frac{\theta}{i} = \frac{-i\theta}{-i^2} = -i\,\theta, \tag{8.14}$$

and so

$$\cosh(\theta) = \cosh(i\,\phi) = \cos(\phi) = \cos(-i\,\theta), \tag{8.15}$$
$$\sinh(\theta) = \sinh(i\,\phi) = i\,\sin(\phi) = i\,\sin(-i\,\theta). \tag{8.16}$$

The formulas (8.15), and (8.16), with (8.14), establish an important relation between the hyperbolic and the trigonometric functions.

As an application of Euler's formula we will now prove the **addition theorems for** $\sin(\phi+\theta)$ **and** $\cos(\phi+\theta)$, which are (see Lemma 2.10)

$$\cos(\phi+\theta) = \cos(\phi)\cos(\theta) - \sin(\phi)\sin(\theta), \tag{8.17}$$

$$\sin(\phi+\theta) = \cos(\phi)\sin(\theta) + \sin(\phi)\cos(\theta). \tag{8.18}$$

Proof of (8.17) and (8.18): From the properties of the exponential function

$$e^{i\,\phi}\,e^{i\,\theta} = e^{i\,(\phi+\theta)},$$

and using Euler's formula we have that

$$\big[\cos(\phi) + i\,\sin(\phi)\big]\,\big[\cos(\theta) + i\,\sin(\theta)\big] = \cos(\phi+\theta) + i\,\sin(\phi+\theta).$$

Writing the left-hand side as a sum yields

$$\begin{aligned}
&\big[\cos(\phi) + i\,\sin(\phi)\big]\,\big[\cos(\theta) + i\,\sin(\theta)\big] \\
&\quad = \cos(\phi)\,\cos(\theta) + i\,\cos(\phi)\,\sin(\theta) + i\,\sin(\phi)\,\cos(\theta) + i^2\,\sin(\phi)\,\sin(\theta) \\
&\quad = \cos(\phi)\,\cos(\theta) - \sin(\phi)\,\sin(\theta) + i\,\big[\cos(\phi)\,\sin(\theta) + \sin(\phi)\,\cos(\theta)\big],
\end{aligned}$$

and substituting into the previous equation yields

$$\cos(\phi)\,\cos(\theta) - \sin(\phi)\,\sin(\theta) + i\big[\cos(\phi)\,\sin(\theta) + \sin(\phi)\,\cos(\theta)\big] = \cos(\phi+\theta) + i\,\sin(\phi+\theta).$$

Equating real and imaginary parts, respectively, gives

$$\begin{aligned}
\cos(\phi)\,\cos(\theta) - \sin(\phi)\,\sin(\theta) &= \cos(\phi+\theta) \\
\cos(\phi)\,\sin(\theta) + \sin(\phi)\,\cos(\theta) &= \sin(\phi+\theta),
\end{aligned}$$

as claimed. □

From the formula $\big(e^{i\,\phi}\big)^n = e^{i\,n\,\phi}$, we obtain with Euler's formula (8.9) De Moivre's theorem below.

Theorem 8.22 (De Moivre's Theorem)
For any integer n and any real number ϕ,

$$\big[\cos(\phi) + i\,\sin(\phi)\big]^n = \cos(n\,\phi) + i\,\sin(n\,\phi),$$

that is,

$$\big(e^{i\,\phi}\big)^n = e^{i\,n\,\phi}.$$

De Moivre's theorem is very useful for proving trigonometric identities. This is illustrated in the following two examples.

Example 8.23 (special case of Lemma 2.10)
Taking $n = 2$ in De Moivre's formula gives that

$$\big[\cos(\phi) + i\sin(\phi)\big]^2 = \big[e^{i\,\phi}\big]^2 = e^{i\,2\,\phi} = \cos(2\,\phi) + i\,\sin(2\,\phi).$$

Writing the left-hand side as a sum yields

$$\begin{aligned}\big[\cos(\phi) + i\sin(\phi)\big]^2 &= \big[\cos(\phi)\big]^2 + 2\,i\,\cos(\phi)\,\sin(\phi) + i^2\,\big[\sin(\phi)\big]^2 \\ &= \big[\cos(\phi)\big]^2 - \big[\sin(\phi)\big]^2 + 2\,i\,\cos(\phi)\,\sin(\phi).\end{aligned}$$

Thus

$$\big[\cos(\phi)\big]^2 - \big[\sin(\phi)\big]^2 + 2\,i\,\cos(\phi)\,\sin(\phi) = \cos(2\,\phi) + i\,\sin(2\,\phi).$$

Equating real and imaginary parts gives, respectively, gives

$$\big[\cos(\phi)\big]^2 - \big[\sin(\phi)\big]^2 = \cos(2\,\phi) \qquad \text{and} \qquad 2\,\cos(\phi)\,\sin(\phi) = \sin(2\,\phi).$$

These are the addition theorems (8.17) and (8.18) for the special case that $\theta = \phi$.
□

Example 8.24 (application of De Moivre's formula)
Prove that

$$\cos(3\,\phi) = 4\,\big[\cos(\phi)\big]^3 - 3\,\cos(\phi) \qquad \text{and} \qquad \sin(3\,\phi) = 3\,\sin(\phi) - 4\,\big[\sin(\phi)\big]^3.$$

Solution: Taking $n = 3$ in De Moivre's Theorem gives that

$$\cos(3\,\phi) + i\,\sin(3\,\phi) = e^{i\,3\,\phi} = \big[e^{i\,\phi}\big]^3 = \big[\cos(\phi) + i\,\sin(\phi)\big]^3. \tag{8.19}$$

We expand the right-hand side into a sum, using the binomial formula,

$$\begin{aligned}&\big[\cos(\phi) + i\,\sin(\phi)\big]^3 \\ &\quad= \big[\cos(\phi)\big]^3 + 3\,i\,\big[\cos(\phi)\big]^2\sin(\phi) + 3\,i^2\,\cos(\phi)\big[\sin(\phi)\big]^2 + i^3\,\big[\sin(\phi)\big]^3 \\ &\quad= \big[\cos(\phi)\big]^3 + 3\,i\big[\cos(\phi)\big]^2\sin(\phi) + 3\,(-1)\,\cos(\phi)\big[\sin(\phi)\big]^2 + i\,(-1)\,\big[\sin(\phi)\big]^3 \\ &\quad= \Big(\big[\cos(\phi)\big]^3 - 3\,\cos(\phi)\big[\sin(\phi)\big]^2\Big) + i\,\Big(3\,\big[\cos(\phi)\big]^2\sin(\phi) - \big[\sin(\phi)\big]^3\Big).\end{aligned}$$

From substituting into (8.19) we get

$$\begin{aligned}&\cos(3\,\phi) + i\,\sin(3\,\phi) \\ &\quad= \Big(\big[\cos(\phi)\big]^3 - 3\,\cos(\phi)\big[\sin(\phi)\big]^2\Big) + i\,\Big(3\,\big[\cos(\phi)\big]^2\,\sin(\phi) - \big[\sin(\phi)\big]^3\Big).\end{aligned}$$

Equating real and imaginary parts, respectively, and using $\big[\sin(\phi)\big]^2 + \big[\cos(\phi)\big]^2 = 1$ gives

$$\begin{aligned}
\cos(3\,\phi) &= \big[\cos(\phi)\big]^3 - 3\,\cos(\phi)\,\big[\sin(\phi)\big]^2 \\
&= \big[\cos(\phi)\big]^3 - 3\,\cos(\phi)\,\Big(1 - \big[\cos(\phi)\big]^2\Big) \\
&= 4\,\big[\cos(\phi)\big]^3 - 3\,\cos(\phi)
\end{aligned}$$

and

$$\begin{aligned}
\sin(3\,\phi) &= 3\,\big[\cos(\phi)\big]^2\,\sin(\phi) - \big[\sin(\phi)\big]^3 \\
&= 3\,\Big(1 - \big[\sin(\phi)\big]^2\Big)\,\sin(\phi) - \big[\sin(\phi)\big]^3 \\
&= 3\,\sin(\phi) - 4\,\big[\sin(\phi)\big]^3. \qquad \square
\end{aligned}$$

8.4 Roots of Complex Numbers

In this section we learn how to take n**th roots of complex numbers**. Unlike in the case of real numbers, **a complex number will now have n distinct nth roots**. We can easily find all nth roots by using the **polar form of complex numbers**.

Definition 8.25 (nth root(s) of a complex number)
*Given a complex number z, an n**th root** of z is any complex number w such that*

$$w^n = z.$$

An nth root is also denoted by $w = z^{1/n}$.

Remark 8.26 ($z = r\,e^{i\,\phi}$ with $r > 0$ has exactly n distinct nth roots)
Every non-zero complex number z has exactly n distinct nth roots. *It is easiest to find these by writing z in its polar form $z = r\,e^{i\,\phi}$. Then we have*

$$z = r\,e^{i\,\phi} = r\,e^{i\,(\phi + 2\pi m)} \quad \textit{for any integer } m.$$

Since $r > 0$, r has exactly one positive real nth root $r^{1/n}$. So from $(a\,b)^{1/n} = a^{1/n}\,b^{1/n}$ with $a = r$ and $b = e^{i(\phi+2\pi m)}$ (even though b is a complex number).

$$z^{1/n} = \Big(r\,e^{i\,(\phi+2\pi m)}\Big)^{1/n} = r^{1/n}\,e^{i\,\frac{\phi+2\pi m}{n}} = r^{1/n}\,e^{i\,(\frac{\phi}{n} + \frac{2\pi m}{n})} \qquad \textit{for any integer } m.$$

Since $2\pi m/n$ has values in $[0, 2\pi)$ for exactly $m = 0, 1, 2, 3, \ldots, n-1$, and since $2\pi m/n = 2\pi$ for $m = n$, we find that exactly n of these roots are distinct. The distinct roots are given by

$$\boxed{z^{1/n} = r^{1/n}\, e^{i(\frac{\phi}{n} + \frac{2\pi m}{n})} \qquad \textit{for} \quad m = 0, 1, 2, 3, \ldots, n-1.}$$

In the complex plane these roots lie all ***equally spaced on the circle with center in the origin*** $(0, 0)$ ***and radius*** $r^{1/n}$ *(as illustrated in Figure 8.8): For $m = 0$, we have the root $r^{1/n}\, e^{i\frac{\phi}{n}}$, and to get the next root we add $2\pi/n$ to the angle ϕ/n. Thus the roots are equally spaced, and for $m = n$ we have reached a full rotation and are again at the point on the circle with angle ϕ/n (since $e^{i\phi/n} = e^{i(\phi/n + 2\pi)}$).*

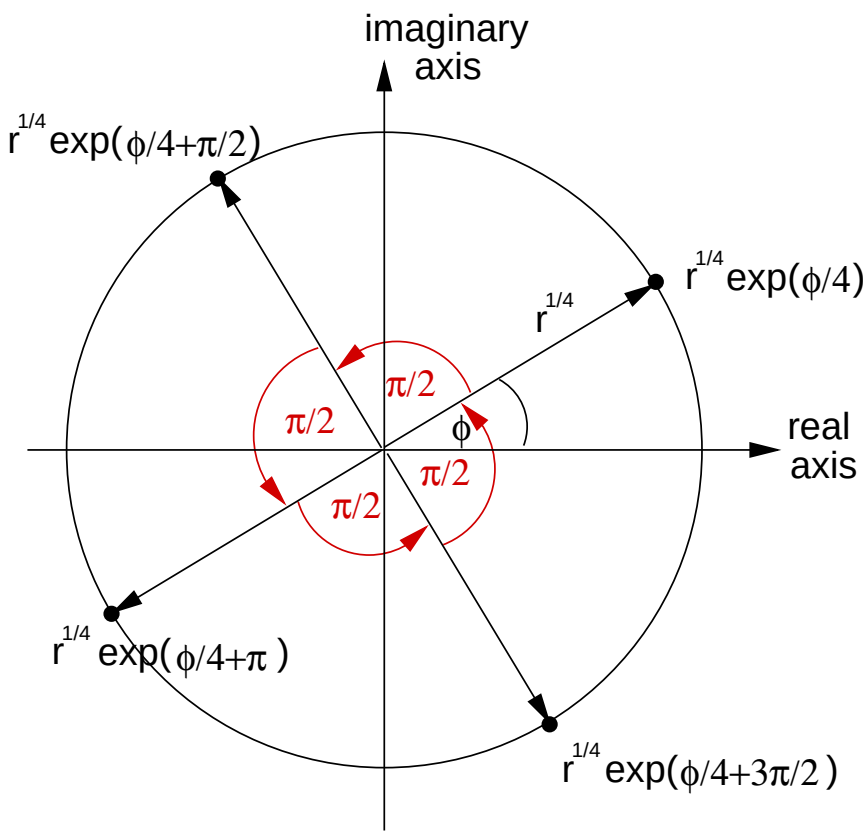


Figure 8.8: All 4th roots of $z = r\, e^{\phi}$. Here $2\pi/n = 2\pi/4 = \pi/2$.

We discuss some examples.

Example 8.27 (finding 2nd roots)
Find the square roots of $z = 1$.

Solution: We have that

$$z = 1 = 1\, e^{i\,0} = 1\, e^{i\,2\pi\, m} \qquad \text{for any integer } m.$$

Hence the roots are given by

$$z^{1/2} = 1^{1/2}\, e^{i\,2\pi m/2} = e^{i\,m\pi} \qquad \text{for any integer } m.$$

So the (distinct) square roots of 1 are $e^{i0\pi} = 1$ (for $m = 0$) band $e^{i\pi} = -1$ (for $m = 1$), as expected. This is illustrated in the left picture in Figure 8.9. □

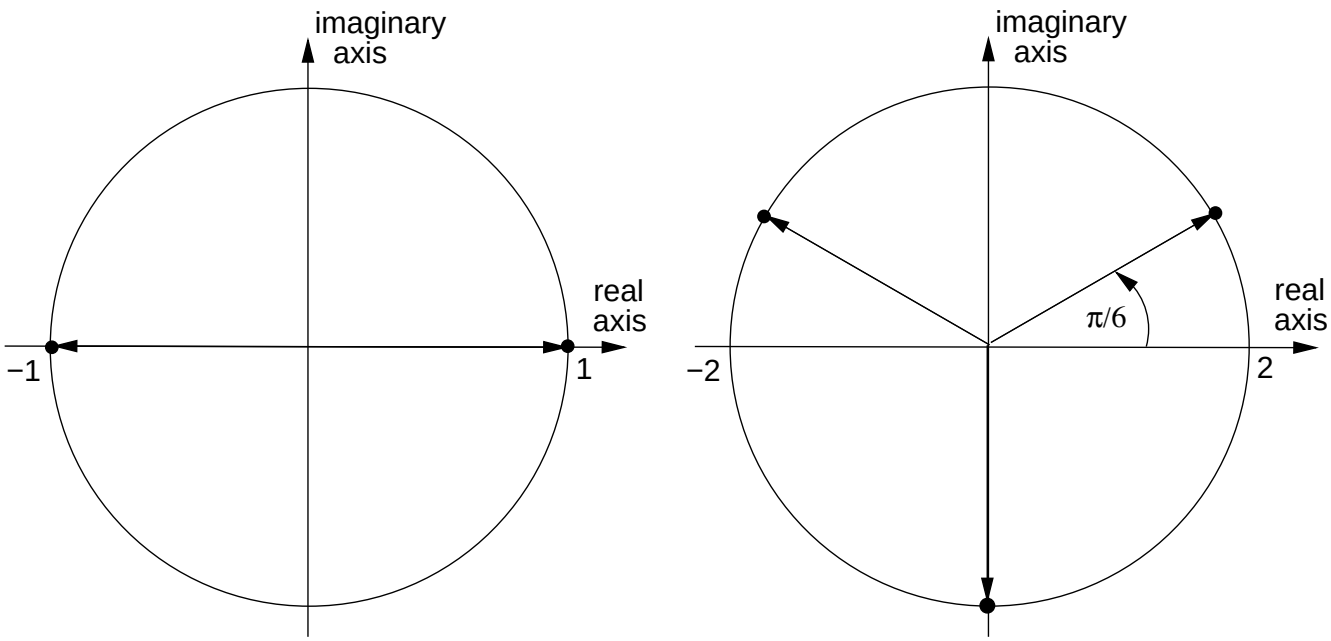


Figure 8.9: The 2nd roots of $z = 1$ on the left, and the 3rd roots of $z = 8\,i$ on the right.

Example 8.28 (finding 3rd roots)
Find the third roots of $z = 8\,i$.

Solution: We have that

$$z = 8\,i = 8\,e^{i\,\frac{\pi}{2}} = 8\,e^{i\,(\frac{\pi}{2}+2\pi m)} = 8\,e^{i\frac{\pi}{2}(1+4m)} \qquad \text{for any integer } m.$$

Since $2^3 = 8$, we have $8^{1/3} = 2$, and hence

$$z^{1/3} = 8^{1/3}\, e^{i\frac{\pi}{6}(1+4m)} = 2\,e^{i\frac{\pi}{6}(1+4m)} \qquad \text{for any integer } m.$$

So the third roots of $8\,i$ are (set $m = 0$, $m = 1$, and $m = 2$)

$$2\,e^{i\frac{\pi}{6}}, \qquad 2\,e^{i\frac{5\pi}{6}}, \qquad \text{and} \qquad 2\,e^{i\frac{9\pi}{6}} = 2\,e^{i(\frac{9\pi}{6}-2\pi)} = 2\,e^{-i\frac{3\pi}{6}} = 2\,e^{-i\frac{\pi}{2}} = -2\,i.$$

These three roots are equally spaced on a circle of radius 2 with center in the origin $(0, 0)$ of the complex plane, as illustrated in the right picture in Figure 8.9. □

Example 8.29 (fractional powers of complex numbers)
Find in polar form all complex numbers z such that $z = i^{2/5}$.

Solution: Using $i^2 = -1$, we can write $z = i^{2/5}$ as $z = i^{2/5} = (i^2)^{1/5} = (-1)^{1/5}$. Thus we are looking for all numbers z such that

$$z^5 = -1 = 1\,e^{i\pi} = 1\,e^{i\,(\pi+2\pi m)} = 1\,e^{i\pi(1+2m)} \qquad \text{for any integer } m.$$

Hence, from taking the 5th root

$$z = 1^{1/5}\,e^{i\frac{\pi(1+2m)}{5}} = e^{i\frac{\pi(1+2m)}{5}} \qquad \text{for any integer } m.$$

So the 5th roots of $i^2 = -1$, and thus the complex numbers z given by $z = i^{2/5}$, are

$$z = e^{i\frac{\pi}{5}}, \quad z = e^{i\frac{3\pi}{5}}, \quad e^{i\pi} = -1, \quad z = e^{i\frac{7\pi}{5}} = e^{-i\frac{3\pi}{5}}, \quad z = e^{i\frac{9\pi}{5}} = e^{i(\frac{9\pi}{5}-2\pi)} = e^{-i\frac{\pi}{5}}.$$

These five roots are equally spaced on a circle of radius 1 with center in the origin $(0,0)$ of the complex plane. □

8.5 Exponential Function and Logarithm of Complex Numbers

Now we finally learn how to calculate the **exponential function** and the **natural logarithm** of a complex number. As in the case of the nth root, we will see that the natural logarithm is not unique.

While we do not justify and explain the definition of the exponential function and the logarithm for a complex variable mathematically thoroughly, we give some comments that help with the intuitive understanding.

Let us assume that the usual properties of the exponential function do also hold for complex numbers, which is indeed the case. For every complex number $z = x + i\,y$ (where $x, y \in \mathbb{R}$) there exist exactly one $\phi \in (-\pi, \pi]$ and one $n \in \mathbb{Z}$ such that $y = \phi + 2\pi\,n$. Then we have (from $e^{a+b} = e^a\,e^b$)

$$e^z = e^{x+i\,y} = e^x\,e^{i\,y} = r\,e^{i\,(\phi+2\pi n)} = r\,e^{i\,\phi} \qquad \text{with } r = e^x,\ \phi = \mathrm{Arg}(z) = y - 2\pi\,n. \tag{8.20}$$

The right-most expression in (8.20) is in polar form $r\,e^{i\,\phi}$ with the radius $r = e^x$ and the principle value $\phi = \mathrm{Arg}(z) = y - 2\pi\,n$ of the argument of z. Thus we can use

$$\boxed{e^z = e^{x+i\,y} = e^x\,e^{i\,y}, \qquad z = x + i\,y.} \tag{8.21}$$

to define the **exponential function for all complex numbers** $z = x + i\,y$.

Example 8.30 (exponential function of complex numbers)

$$e^{2+i\,3\pi} = e^2\, e^{i\,3\pi} = e^2\, e^{i\,\pi} = e^2\,\big[\cos(\pi) + i\,\sin(\pi)\big] = -e^2,$$

$$e^{1+i\,\pi/3} = e^1\, e^{i\,\pi/3} = e\,\big[\cos(\pi/3) + i\,\sin(\pi/3)\big] = e\,\left[\frac{1}{2} + i\,\frac{\sqrt{3}}{2}\right] = \frac{e}{2} + i\,\frac{\sqrt{3}\,e}{2},$$

where we have used Table 8.1 □

The logarithm for real numbers was defined as the inverse function of the exponential function, that is, we have

$$e^{\ln(a)} = a \quad \text{for all } a > 0 \qquad \text{and} \qquad \ln\big(e^b\big) = b \quad \text{for all } b \in \mathbb{R}. \tag{8.22}$$

When we introduce the complex logarithm, then we want that (8.22) holds for complex a and b (with suitable restrictions on on a and b). To introduce the complex logarithm, we start with a complex number in polar form $z = r\,e^{i\,\phi}$. Then

$$z = r\,e^{i\,\phi} = r\,e^{i\,(\phi+2\pi n)} \qquad \text{for any integer } n.$$

Since our logarithm with a complex variable shall also satisfy $\ln(a\,b) = \ln(a) + \ln(b)$ for complex a and b, we write

$$\ln(z) = \ln\big(r\,e^{i\,\phi}\big) = \ln\big(r\,e^{i\,(\phi+2\pi n)}\big) = \ln(r) + \ln\big(e^{i\,(\phi+2\pi n)}\big) = \ln(r) + i\,(\phi + 2\pi n), \tag{8.23}$$

where $n \in \mathbb{Z}$ and where we have used in the last step that we want the logarithm to be the inverse function of the exponential function, and thus $\ln(e^a) = a$. Here $\ln(r)$ is the usual logarithm of the positive real number r. Thus we define the **logarithm for complex numbers** by

$$\boxed{\ln(z) = \ln\big(r\,e^{i\,\phi}\big) = \ln(r) + i\,(\phi + 2\pi n), \qquad n \in \mathbb{Z}, \quad z = r\,e^{i\,\phi}.} \tag{8.24}$$

Since n in (8.24) can be any integer, $\ln(z)$ has **infinitely many values**: **all values of** $\ln(z)$ **have the same real part and their imaginary parts differ by integer multiples of** 2π.

Remark 8.31 (logarithm as inverse function of $\exp(x)$**)**
From (8.24) and (8.21), we have now for any complex number $z = r\,e^{i\phi}$ *in polar form and any integer* n*,*

$$\exp\big(\ln(r\,e^{i\phi})\big) = \exp\big(\ln(r) + i\,(\phi + 2\pi n)\big) = \exp\big(\ln(r)\big)\,\exp(i\,(\phi + 2\pi n)) = r\,e^{i\phi},$$

where we have used in the last step $e^{i\,(\phi+2\pi n)} = e^{i\,\phi}\,e^{i\,2\pi n} = e^{i\,\phi}$. *For any complex number* z, *written in Cartesian form* $z = x + i\,y$, *we have from (8.24) and (8.21)*

$$\ln(e^z) = \ln\left(e^{x+i\,y}\right) = \ln\left(e^x\,e^{iy}\right) = \ln(e^x) + i\,(y + 2\pi n) = x + i\,(y + 2\pi n), \quad n \in \mathbb{Z}. \tag{8.25}$$

This is essentially what we expected for the logarithm as the inverse function of the exponential function; only we do ***not*** *have* ***one but infinitely many values of*** $\ln(e^z)$. *More precisely, we have a different value of* $\ln(e^z)$ *for each integer* $n \in \mathbb{Z}$. *Strictly speaking, the exponential function has to be one-to-one in order to have an inverse function. For example, the exponential function is one-to-one for* $z = r\,e^{i\phi}$ *with* $-\pi < \phi < \pi$ *and* $r > 0$. *Then we have from the definition of the inverse function*

$$\ln\left(r\,e^{i\,\phi}\right) = \ln(r) + \ln(e^{i\,\phi}) = \ln(r) + i\,\phi. \tag{8.26}$$

In (8.26) we only consider the value of the logarithm (8.25) for $n = 0$.

We give some examples.

Example 8.32 (logarithm of complex numbers)
Find the logarithm of the complex numbers

$$\text{(a)} \quad z = 1 + i\,\sqrt{3} \qquad \text{and} \qquad \text{(b)} \quad z = i.$$

Solution: In both cases we have to first find the polar form of the given complex number.

(a) For $z = 1 + i\sqrt{3} = r\,e^{i\,\phi}$, we have $x = 1 > 0$ and $y = \sqrt{3} > 0$. Thus the principal argument $\phi = \mathrm{Arg}(z)$ lies in the first quadrant, that is, $0 \le \phi \le \pi/2$. From $r = |z| = \left(1^2 + \sqrt{3}^2\right)^{1/2} = \sqrt{4} = 2$, we have

$$\cos(\phi) = \frac{x}{r} = \frac{1}{2}, \qquad \sin(\phi) = \frac{y}{r} = \frac{\sqrt{3}}{2}.$$

Thus from the values of sine or cosine (see Table 8.1) we know that $\sin(\pi/3) = \sqrt{3}/2$, and thus the principal argument is $\phi = \mathrm{Arg}(z) = \pi/3$. Thus $z = 1 + i\,\sqrt{3}$ has the polar form

$$z = 2\,e^{i\,\pi/3} = 2\,e^{i\,(\frac{\pi}{3}+2\pi n)} = 2\,e^{i\,\pi\,(\frac{1}{3}+2n)} \qquad \text{for any integer } n.$$

Thus we find

$$\ln(z) = \ln(2) + i\,\pi\left(\frac{1}{3} + 2n\right) \qquad \text{for any integer } n.$$

(b) If $z = i$, then the polar form is $z = e^{i\,\pi/2}$. Thus

$$z = 1\,e^{i\,\pi/2} = 1\,e^{i\,(\frac{\pi}{2}+2\pi n)} = 1\,e^{i\,\frac{\pi}{2}(1+4n)} \qquad \text{for any integer } n,$$

and so

$$\ln(z) = \ln(1)+i\,\frac{\pi}{2}\,(1+4n) = i\,\frac{\pi}{2}\,(1+4n) \qquad \text{for any integer } n. \qquad \square$$

Example 8.33 (logarithm of complex numbers)
Find the logarithm of the complex numbers

(a) $z = x > 0$, where $x \in \mathbb{R}$, and (b) $z = x < 0$, where $x \in \mathbb{R}$.

Solution:

(a) If $z = x$, with x a positive real number, then

$$z = x = x\,e^{i0} = x\,e^{i\,2\pi n} \qquad \text{for any integer } n,$$

and so

$$\ln(z) = \ln(x) + i\,2\pi n \qquad \text{for any integer } n.$$

(b) If $z = x$, with x a negative real number, then $-x > 0$ and

$$z = (-x)\,(-1) = (-x)\,e^{i\,\pi} = (-x)\,e^{i\,(\pi+2\pi n)} = (-x)\,e^{i\,\pi(1+2n)} \qquad \text{for any integer } n.$$

Thus

$$\ln(z) = \ln(-x)+i\,\pi(1+2n) \qquad \text{for any integer } n. \qquad \square$$

Chapter 9

Vectors

Forces and velocities are **vector** quantities, that is, quantities which have both **direction** and **magnitude**. In this chapter, we will encounter vectors in three-dimensional Euclidean space $\mathbb{R}^3$, that is, triples $\mathbf{x} = (x, y, z)$ or $\mathbf{a} = (a_1, a_2, a_3)$, where $x, y, z \in \mathbb{R}$ and $a_1, a_2, a_3 \in \mathbb{R}$, respectively, are the components of the vectors. In Section 9.2, we will learn basic vector operations, that is, how to **add** and **subtract vectors** and how to **multiply vectors with scalars**. These elementary vector operations have geometric interpretations, and we will discuss some properties of vector addition/substraction and scalar multiplication. In Section 9.3, we will introduce the **scalar product** or **dot product of two vectors** which results in a scalar (a real number), and in Section 9.4 we will discuss the **vector product** or **cross product of two vectors** which results in another vector. Both the scalar product and the vector product will be defined geometrically, using trigonometric functions, and this **geometric interpretation** of the scalar product and vector product is often very useful. We will also derive **another formula** for computing the scalar product and the vector product, respectively, **in terms of the components of the two vectors**.

9.1 Introduction

You will have encountered vectors in school. Here we will discuss **vectors** in the so-called **three dimensional Euclidean space** $\mathbb{R}^3$ as triples of real numbers.

Definition 9.1 (vector in $\mathbb{R}^3$)
*A **vector** in $\mathbb{R}^3$ is a triple*

$$\mathbf{x} = (x, y, z) \qquad \text{or} \qquad \mathbf{a} = (a_1, a_2, a_3), \qquad \text{where } x, y, z \in \mathbb{R} \text{ and } a_1, a_2, a_3 \in \mathbb{R},$$

*respectively. The real numbers x, y, z and a_1, a_2, a_3 are called the **components** of the vector $\mathbf{x}$ and $\mathbf{a}$, respectively. The order of the components is important, that is, $(x, y, z) \neq (y, x, z)$, etc..*

Example 9.2 (vectors)
Two examples of vectors in $\mathbb{R}^3$ are $\mathbf{a} = (4, 2, 3)$ and $\mathbf{b} = (-1, 0, 6)$. For $\mathbf{a} = (a_1, a_2, a_3) = (4, 2, 3)$ we have $a_1 = 4$, $a_2 = 2$, and $a_3 = 3$. For $\mathbf{b} = (-1, 0, 6) = (b_1, b_2, b_3)$ we have $b_1 = -1$, $b_2 = 0$, and $b_3 = 6$. □

Definition 9.3 (zero vector)
*The **zero vector** is the unique vector whose **components are all zero**. The zero vector is denoted by*

$$\mathbf{0} = (0, 0, 0).$$

A vector $\mathbf{a} = (a_1, a_2, a_3)$ in $\mathbb{R}^3$ is **visualized** by plotting the point with coordinates (a_1, a_2, a_3) and drawing an arrow that connects the origin $(0, 0, 0)$ with (a_1, a_2, a_3) and points from $(0, 0, 0)$ to (a_1, a_2, a_3). This is indicated in Figure 9.1 below.

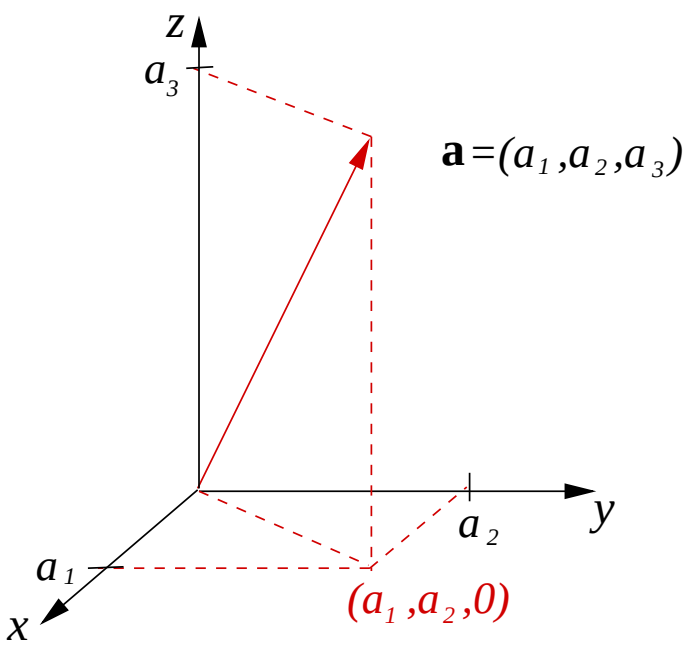


Figure 9.1: Plotting a vector $\mathbf{a} = (a_1, a_2, a_3)$ in $\mathbb{R}^3$.

The vector $\mathbf{a} = (a_1, a_2, a_3)$, represented by the arrow with 'arrow head' at the point (a_1, a_2, a_3) in Figure 9.1 has a **direction** and has a **length**. Under length we understand here the length of the arrow, or in words, the **distance of the point** (a_1, a_2, a_3) **from the origin** $(0, 0, 0)$. The direction is intuitively the **direction into which the arrow points**. We will now discuss these quantities.

Definition 9.4 (length/magnitude of a vector)
*The **length** or **magnitude** of the vector* $\mathbf{a} = (a_1, a_2, a_2)$ *is*

$$|\mathbf{a}| = |(a_1, a_2, a_3)| = \sqrt{a_1^2 + a_2^2 + a_3^2}. \tag{9.1}$$

From **Pythagoras' theorem**, we can see that the length/magnitude (9.1) is indeed the length of the arrow, that is, the distance from $(0, 0, 0)$ to the point (a_1, a_2, a_3), as follows: Look again at Figure 9.1 and consider the rectangle in the (x, y)-plane with corners $(0, 0, 0)$, $(a_1, 0, 0)$, $(a_1, a_2, 0,)$, and $(0, a_2, 0)$. The diagonal of this rectangle (indicated as the line from $(0, 0, 0)$ to $(a_1, a_2, 0)$) has from Pythagoras' theorem the length $\sqrt{a_1^2 + a_2^2}$. Now we consider the rectangle whose corners are $(0, 0, 0)$, $(a_1, a_2, 0)$, (a_1, a_2, a_3), and $(0, 0, a_3)$. Then the 'bottom side' and 'top side' of this rectangle have length $\sqrt{a_1^2 + a_2^2}$ (as we worked out previously), and the other two sides have length $|a_3|$, since the distance from $(0, 0, 0)$ to $(0, 0, a_3)$ is just $|a_3|$. The diagonal of this rectangle from $(0, 0, 0)$ to (a_1, a_2, a_3) is just the vector $\mathbf{a} = (a_1, a_2, a_3)$. From Pythagoras' theorem the length of this diagonal is

$$\sqrt{\left(\sqrt{a_1^2 + a_2^2}\right)^2 + a_3^3} = \sqrt{a_1^2 + a_2^2 + a_3^3}.$$

Thus we see that formula (9.1) gives indeed the **geometric length** of the vector $\mathbf{a} = (a_1, a_2, a_3)$.

Example 9.5 (length of vectors)
If $\mathbf{a} = (4, 2, 3)$ and $\mathbf{b} = (-1, 0, 6)$, then

$$|\mathbf{a}| = |(4, 2, 3)| = \sqrt{4^2 + 2^2 + 3^2} = \sqrt{16 + 4 + 9} = \sqrt{29},$$
$$|\mathbf{b}| = |(-1, 0, 6)| = \sqrt{(-1)^2 + 0^2 + 6^2} = \sqrt{1 + 36} = \sqrt{37}. \qquad \square$$

Remark 9.6 (only the zero vector has length zero)
The zero vector $\mathbf{0} = (0, 0, 0)$ *satisfies*

$$|\mathbf{0}| = |(0, 0, 0)| = \sqrt{0^2 + 0^2 + 0^2} = \sqrt{0} = 0,$$

and has thus length/magnitude $|\mathbf{0}| = 0$. *For any other vector* $\mathbf{a} = (a_1, a_2, a_3)$, *we have that at least one of the components* a_1, a_2, a_3 *is different from zero. For example, assume that* $a_1 \neq 0$. *Then*

$$a_1^2 > 0, \quad a_2^2 \geq 0, \quad \textit{and} \quad a_3^3 \geq 0 \qquad \Rightarrow \qquad a_1^2 + a_2^2 + a_3^3 > 0.$$

Thus we can conclude that

$$|\mathbf{a}| = |(a_1, a_2, a_3)| = \sqrt{a_1^2 + a_2^2 + a_3^2} > 0.$$

We have thus verified that $\boxed{\textbf{\textit{if}}\ \mathbf{a} \neq \mathbf{0}\ \textbf{\textit{then}}\ |\mathbf{a}| > 0\textbf{\textit{; and if}}\ |\mathbf{a}| = 0\ \textbf{\textit{then}}\ \mathbf{a} = \mathbf{0}.}$

Now we come back to the second quantity that defines a vector, namely its **direction**. Once we have the length of a vector $\mathbf{x} = (x, y, z)$, its direction can be indicated by any vector (or arrow) pointing in the same direction as $\mathbf{x} = (x, y, z)$. It is customary to describe the direction of a vector by a vector that points in the same direction and has length one.

Definition 9.7 (unit vector)
A vector $\mathbf{x} = (x, y, z)$ *is a* ***unit vector*** *if it has* ***length/magnitude one****, that is, if*

$$|\mathbf{x}| = \sqrt{x^2 + y^2 + z^2} = 1.$$

Example 9.8 (unit vectors)
Show that the vectors $\mathbf{a} = (1, 0, 0)$ and $\mathbf{b} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$ are unit vectors.

Solution: We compute the length/magnitude of each vector

$$|\mathbf{a}| = \sqrt{1^2 + 0^2 + 0^2} = \sqrt{1} = 1$$

and

$$|\mathbf{b}| = \sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 + 0^2} = \sqrt{\frac{1}{2} + \frac{1}{2}} = \sqrt{1} = 1.$$

Since both vectors have length one, they are indeed unit vectors. □

Once we have introduced the scalar multiplication of vectors by real numbers (and thus the division of vectors by real numbers), we will see that the unit vector pointing in the same direction as a given vector $\mathbf{x}$ can be obtained by dividing $\mathbf{x}$ through its length/magnitude $|\mathbf{x}|$.

9.2 Basic Vector Operations

In this section we learn how to **add and subtract vectors** and how to **multiply vectors with scalars** (real numbers). We will also discuss the **geometric interpretation** of vector addition/substraction and scalar multiplication.

Definition 9.9 (addition/subtraction of vectors)
Let $\mathbf{a} = (a_1, a_2, a_3)$ *and* $\mathbf{b} = (b_1, b_2, b_3)$ *be two vectors in* $\mathbb{R}^3$. *Then the* ***sum*** $\mathbf{a}+\mathbf{b}$ *is defined by*

$$\mathbf{a} + \mathbf{b} = (a_1, a_2, a_3) + (b_1, b_2, b_3) = \big(a_1 + b_1, a_2 + b_2, a_3 + b_3\big),$$

and the ***difference*** $\mathbf{a} - \mathbf{b}$ *is defined by*

$$\mathbf{a} - \mathbf{b} = (a_1, a_2, a_3) - (b_1, b_2, b_3) = \big(a_1 - b_1, a_2 - b_2, a_3 - b_3\big).$$

That is, vectors are added or subtracted by adding or subtracting the components of the two vectors, respectively. We say, the vectors are added or subtracted ***component-wise****.*

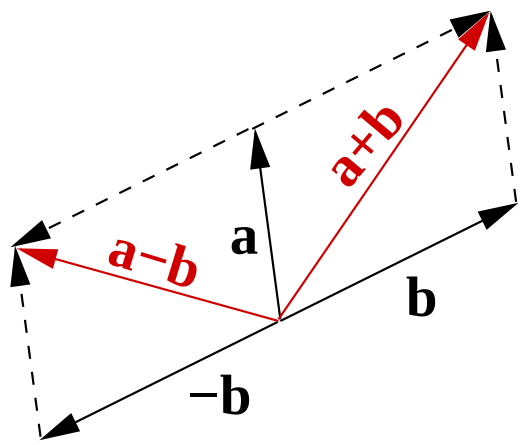


Figure 9.2: Geometric determination of the sum $\mathbf{a} + \mathbf{b}$ and the difference $\mathbf{a} - \mathbf{b}$ of the vectors $\mathbf{a}$ and $\mathbf{b}$.

Vector **addition** and **substraction** can be **geometrically visualized** as indicated in Figure 9.2: Consider the vectors $\mathbf{a}$ and $\mathbf{b}$ as plotted in Figure 9.2. Then $\mathbf{a}+\mathbf{b}$ is the vector given by the diagonal of the parallelogram spanned by the vectors $\mathbf{a}$ and $\mathbf{b}$ with the 'head' of the vector as indicated. In other words, we can obtain $\mathbf{a}+\mathbf{b}$ by taking a copy of $\mathbf{b}$ and moving it (without changing its direction) so that its 'tail' touches the 'head' of the vector $\mathbf{a}$. The 'head' of this copy of the vector $\mathbf{b}$ gives then

the 'head' of the vector $\mathbf{a}+\mathbf{b}$, and the 'tail' of $\mathbf{a}+\mathbf{b}$ is located where the 'tail' of $\mathbf{a}$ is. The difference $\mathbf{a}-\mathbf{b}$ is given as the sum $\mathbf{a}+(-\mathbf{b})$, where we interpret the vector $-\mathbf{b}$ as the vector with the same length as $\mathbf{b}$ but pointing in the opposite direction. In formulas, $-\mathbf{b} = (-b_1, -b_2, -b_3)$.

Example 9.10 (addition/substraction of vectors)
If $\mathbf{a} = (4, 2, 3)$ and $\mathbf{b} = (-1, 0, 6)$, then

$$\mathbf{a}+\mathbf{b} = (4,2,3)+(-1,0,6) = \big(4+(-1), 2+0, 3+6\big) = (3,2,9),$$
$$\mathbf{a}-\mathbf{b} = (4,2,3)-(-1,0,6) = \big(4-(-1), 2-0, 3-6\big) = (5,2,-3). \qquad \square$$

Definition 9.11 (scalars and scalar multiplication)
In the context of vector spaces real numbers are denoted as ***scalars****. Let $\lambda \in \mathbb{R}$ be a scalar and let $\mathbf{x} = (x, y, z)$ be a vector in $\mathbb{R}^3$. Then the* ***scalar multiplication*** *of the vector $\mathbf{x} = (x, y, z)$ by the scalar λ is defined by*

$$\lambda\,\mathbf{x} = \lambda\,(x, y, z) = \big(\lambda\,x, \lambda\,y, \lambda\,z\big).$$

In other words, the vector $\mathbf{x}$ is multiplied by the scalar λ ***component-wise****, that is, each component is multiplied with λ.*

The **scalar multiplication** of a vector $\mathbf{x}$ by a scalar $\lambda \in \mathbb{R}$ can also be **geometrically visualized** which is illustrated in Figure 9.3: For $\lambda > 0$, the vector $\lambda\,\mathbf{x}$ is given by the vector with the same direction as $\mathbf{x}$ and the length $\lambda\,|\mathbf{x}|$. For $\lambda < 0$, the vector $\lambda\,\mathbf{x}$ is given by the vector with the opposite direction as $\mathbf{x}$ and the length $|\lambda|\,|\mathbf{x}|$.

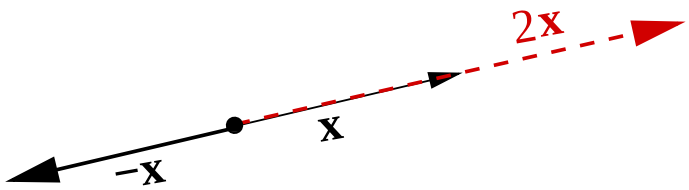


Figure 9.3: The vectors $\mathbf{x}$, $2\,\mathbf{x}$, and $-\mathbf{x}$.

Example 9.12 (multiplication of a vector by a scalar)
If $\mathbf{a} = (4, 2, 3)$ and $\lambda = 3$, then

$$3\,\mathbf{a} = 3\,(4,2,3) = \big(3\times 4, 3\times 2, 3\times 3\big) = (12,6,9).$$

The vector $3\,\mathbf{a}$ points in the same direction as $\mathbf{a}$, but is three times as long.

If $\mathbf{b} = (-1, 0, 6)$ and $\lambda = -2$, then

$$-2\,\mathbf{b} = (-2)\,(-1, 0, 6) = \big((-2) \times (-1), (-2) \times 0, (-2) \times 6\big) = (2, 0, -12).$$

The vector $-2\,\mathbf{b}$ is twice as long as $\mathbf{b}$, and points in the opposite direction to $\mathbf{b}$. □

Definition 9.13 (division of a vector by a scalar)
Let $\lambda \in \mathbb{R}$ be a non-zero scalar, and let $\mathbf{x} = (x, y, z) \in \mathbb{R}^3$. Then we define

$$\frac{\mathbf{x}}{\lambda} = \frac{1}{\lambda}\,\mathbf{x} = \frac{1}{\lambda}\,(x, y, z) = \left(\frac{x}{\lambda}, \frac{y}{\lambda}, \frac{z}{\lambda}\right).$$

*In words, $\mathbf{x}/\lambda$ is defined to be the vector obtained from scalar multiplication of $\mathbf{x}$ by $1/\lambda$, that is, $\mathbf{x}$ is divided **component-wise** by $\lambda \neq 0$.*

Remark 9.14 (length of the 'scalar multiple' of a vector)
If $\mathbf{x} \in \mathbb{R}^3$ is a vector and $\lambda \in \mathbb{R}$ is a scalar, then

$$\boxed{|\lambda\,\mathbf{x}| = |\lambda|\,|\mathbf{x}|.} \tag{9.2}$$

*So $\lambda\,\mathbf{x}$ is $|\lambda|$ times as long as $\mathbf{x}$. If $\lambda > 0$, then $\lambda\,\mathbf{x}$ points in the **same direction** as $\mathbf{x}$; and if $\lambda < 0$, then $\lambda\,\mathbf{x}$ points in the **opposite direction** to $\mathbf{x}$. We can also prove (9.2) formally: for $\mathbf{x} = (x, y, z)$, we have*

$$\begin{aligned}
|\lambda\,\mathbf{x}| = |\lambda\,(x, y, z)| = |(\lambda\,x, \lambda\,y, \lambda\,z)| &= \sqrt{(\lambda\,x)^2 + (\lambda\,y)^2 + (\lambda\,z)^2} \\
&= \sqrt{|\lambda|^2\,(x^2 + y^2 + z^2)} \\
&= |\lambda|\,\sqrt{x^2 + y^2 + z^2} = |\lambda|\,|\mathbf{x}|,
\end{aligned}$$

which indeed proves (9.2). □

Definition 9.15 (parallel vectors)
*Two non-zero vectors $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^3$ are called **parallel**, if there exists a non-zero scalar $\lambda \in \mathbb{R}$ such that*

$$\mathbf{a} = \lambda\,\mathbf{b}.$$

Remark 9.16 (unit vector in the same direction as a given vector)
If $\mathbf{x} \neq \mathbf{0}$ then $|\mathbf{x}| > 0$, and hence from (9.2)

$$\left|\frac{\mathbf{x}}{|\mathbf{x}|}\right| = \left|\frac{1}{|\mathbf{x}|}\,\mathbf{x}\right| = \left|\frac{1}{|\mathbf{x}|}\right|\,|\mathbf{x}| = \frac{1}{|\mathbf{x}|}\,|\mathbf{x}| = \frac{|\mathbf{x}|}{|\mathbf{x}|} = 1.$$

So the vector

$$\hat{\mathbf{x}} = \frac{\mathbf{x}}{|\mathbf{x}|} = \frac{1}{|\mathbf{x}|}\,\mathbf{x} \tag{9.3}$$

points in the same direction as $\mathbf{x}$ *and has length* 1*, that is,* $\hat{\mathbf{x}}$ *is the* ***unit vector in the direction of*** $\mathbf{x}$*. From (9.3) we obtain by multiplying by* $|\mathbf{x}|$

$$|\mathbf{x}|\,\hat{\mathbf{x}} = \mathbf{x} \qquad \textit{or} \qquad \mathbf{x} = |\mathbf{x}|\,\hat{\mathbf{x}} \tag{9.4}$$

From (9.4) we see that we can describe a vector ***uniquely*** *by giving the length/magnitude* $|\mathbf{x}|$ *and the unit vector* $\hat{\mathbf{x}} = \mathbf{x}/|\mathbf{x}|$ *pointing in the same direction.*

Example 9.17 (unit vectors)
If $\mathbf{a} = (4, 2, 3)$ and $\mathbf{b} = (-1, 0, 6)$, then, from Example 9.5, $|\mathbf{a}| = \sqrt{29}$ and $|\mathbf{b}| = \sqrt{37}$. Thus the unit vectors in the direction of $\mathbf{a}$ and $\mathbf{b}$ are

$$\hat{\mathbf{a}} = \frac{\mathbf{a}}{|\mathbf{a}|} = \frac{1}{\sqrt{29}}\,(4, 2, 3) = \left(\frac{4}{\sqrt{29}}, \frac{2}{\sqrt{29}}, \frac{3}{\sqrt{29}}\right),$$

$$\hat{\mathbf{b}} = \frac{\mathbf{b}}{|\mathbf{b}|} = \frac{1}{\sqrt{37}}\,(-1, 0, 6) = \left(\frac{-1}{\sqrt{37}}, \frac{0}{\sqrt{37}}, \frac{6}{\sqrt{37}}\right),$$

respectively. □

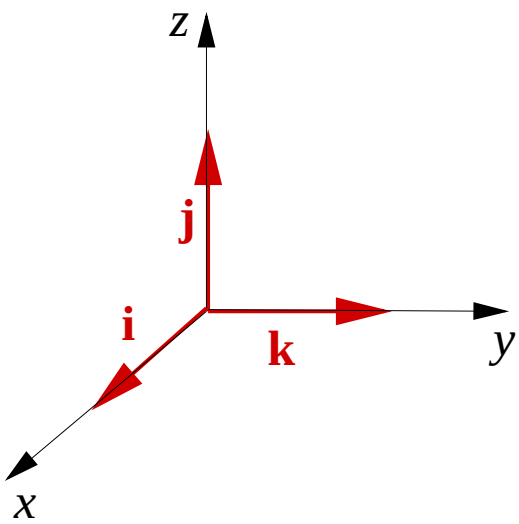


Figure 9.4: The cartesian unit vectors $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, and $\mathbf{k} = (0, 0, 1)$.

Definition 9.18 (cartesian unit vectors)
The unit vectors

$$\mathbf{i} = (1, 0, 0), \qquad \mathbf{j} = (0, 1, 0), \qquad \textit{and} \qquad \mathbf{k} = (0, 0, 1),$$

are called the ***Cartesian unit vectors****.*

The Cartesian unit vectors are indicated in Figure 9.4.

Remark 9.19 (representation of vector with Cartesian unit vectors)
If $\mathbf{x} = (x, y, z)$*, then*

$$\boxed{\mathbf{x} = (x, y, z) = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}.}$$

Example 9.20 (representation of vector with Cartesian unit vectors)
If $\mathbf{a} = (4, 2, 3)$ and $\mathbf{b} = (-1, 0, 6)$, then

$$\begin{aligned}\mathbf{a} &= (4, 2, 3) = 4\,\mathbf{i} + 2\,\mathbf{j} + 3\,\mathbf{k},\\ \mathbf{b} &= (-1, 0, 6) = -1\,\mathbf{i} + 0\,\mathbf{j} + 6\,\mathbf{k} = -\mathbf{i} + 6\,\mathbf{k}.\end{aligned}$$

□

9.3 Scalar Product (or Dot Product)

The **scalar product** or **dot product of two vectors** gives a scalar. It is **not** to be mixed up with the scalar multiplication $\lambda\,\mathbf{x}$ of a vector $\mathbf{x} \in \mathbb{R}^3$ by a scalar $\lambda \in \mathbb{R}$ which gives a vector.

Definition 9.21 (scalar product/dot product)
Let $\mathbf{a}$ *and* $\mathbf{b}$ *be two vectors in* $\mathbb{R}^3$*. We define the* ***scalar product*** *or* ***dot product*** $\mathbf{a} \cdot \mathbf{b}$ *of the non-zero vectors* $\mathbf{a}$ *and* $\mathbf{b}$ *to be the scalar (real number)*

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}|\,|\mathbf{b}|\,\cos(\theta), \tag{9.5}$$

where $\theta \in [0, \pi]$ *is the angle between* $\mathbf{a}$ *and* $\mathbf{b}$*. If* $\mathbf{a} = \mathbf{0}$ *or* $\mathbf{b} = \mathbf{0}$*, then we define* $\mathbf{a} \cdot \mathbf{b} = 0$*.*

The scalar product $\mathbf{a} \cdot \mathbf{b}$ has a **geometric interpretation**: Writing the scalar product $\mathbf{a} \cdot \mathbf{b}$ as

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}|\,\big(|\mathbf{b}|\,\cos(\theta)\big),$$

it is the product of the length $|\mathbf{a}|$ of $\mathbf{a}$ and the value $|\mathbf{b}|\,\cos(\theta)$. The value $|\mathbf{b}|\,\cos(\theta)$ may be interpreted as the **length of the projection of b onto the vector a**, as indicated in Figure 9.5.

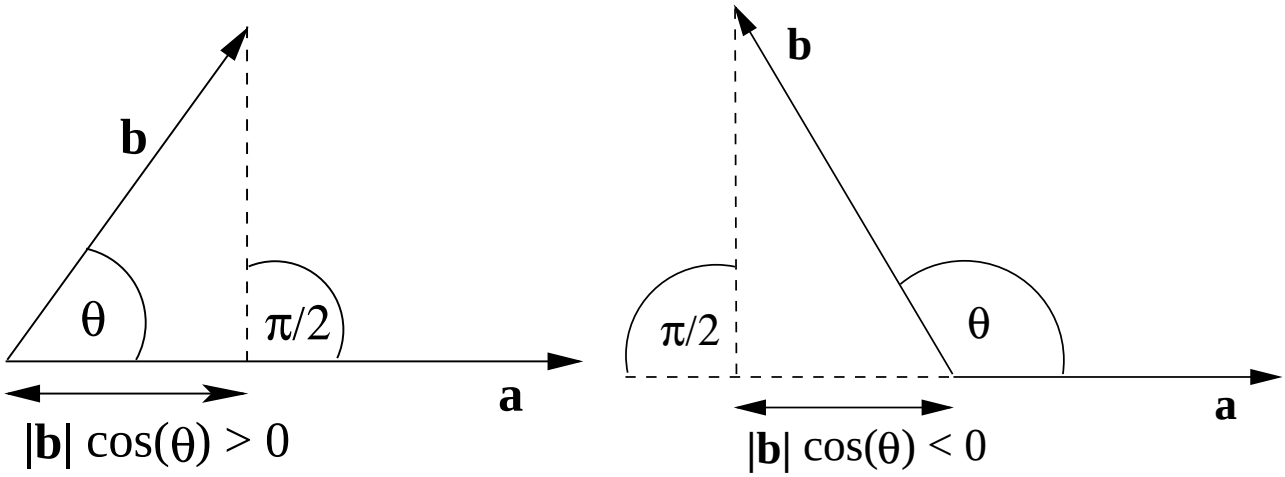


Figure 9.5: The value of the scalar product $\mathbf{a} \cdot \mathbf{b}$ is the product of the length $|\mathbf{a}|$ of $\mathbf{a}$ and the length $|\mathbf{b}| \cos(\theta)$ of the projection of $\mathbf{b}$ onto $\mathbf{a}$.

Example 9.22 (scalar product)
Compute the scalar product of $\mathbf{a} = (2, 2, 0)$ and $\mathbf{b} = (0, 1, 0)$.

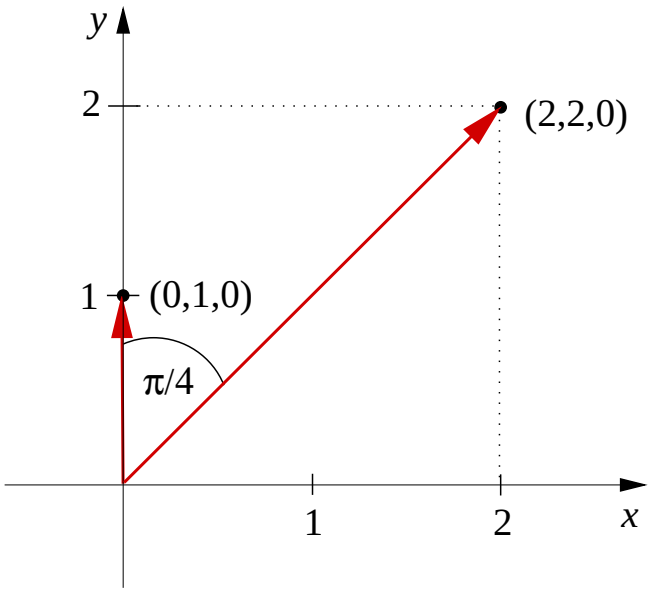


Figure 9.6: Angle between the vectors $\mathbf{a} = (2, 2, 0)$ and $\mathbf{b} = (0, 1, 0)$.

Solution: Since the vectors have both the third component zero, we can consider them as vectors in the (x, y)-plane. From the plot in Figure 9.6 below, we see that the angle between the two vectors is $\theta = \pi/4$. Now we compute the lengths of the two vectors $\mathbf{a}$ and $\mathbf{b}$, and we find

$$|\mathbf{a}| = |(2, 2, 0)| = \sqrt{2^2 + 2^2 + 0^2} = \sqrt{8} = 2\sqrt{2},$$
$$|\mathbf{b}| = |(0, 1, 0)| = \sqrt{0^2 + 1^2 + 0^1} = \sqrt{1} = 1.$$

Thus the scalar product of $\mathbf{a}$ and $\mathbf{b}$ is given by

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}|\,|\mathbf{b}|\,\cos(\theta) = 2\sqrt{2} \times 1 \times \cos(\pi/4) = 2\sqrt{2} \times \frac{1}{\sqrt{2}} = 2. \qquad \square$$

Remark 9.23 (special cases of the scalar product)

(a) For any vector $\mathbf{a} \in \mathbb{R}^3$*, the angle between* $\mathbf{a}$ *and itself is* $\theta = 0$*. So we have* $\cos(\theta) = \cos(0) = 1$*, giving that*

$$\boxed{\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|\,|\mathbf{a}|\,\cos(0) = |\mathbf{a}|^2.}$$

(b) If the vectors $\mathbf{a} \neq \mathbf{0}$ *and* $\mathbf{b} \neq \mathbf{0}$ *are* ***perpendicular to each other****, that is, the angle between the* $\mathbf{a}$ *and* $\mathbf{b}$ *is* $\pi/2$*, then*

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}|\,|\mathbf{b}|\,\cos(\pi/2) = |\mathbf{a}|\,|\mathbf{b}| \times 0 = 0.$$

(c) For any two vectors $\mathbf{a}$ *and* $\mathbf{b}$*,*

$$\boxed{\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}.} \tag{9.6}$$

Remark 9.23 (b) leads to the following definition and lemma.

Definition 9.24 (orthogonal/perpendicular)
Two non-zero vectors $\mathbf{a}$ *and* $\mathbf{b}$ *in* $\mathbb{R}^3$ *are* ***orthogonal*** *or* ***perpendicular (to each other)*** *if the angle between the two vectors is* $\pi/2$*.*

Lemma 9.25 (criterion for orthogonality)
Two non-zero vectors $\mathbf{a}$ *and* $\mathbf{b}$ *in* $\mathbb{R}^3$ *are* ***orthogonal*** *or* ***perpendicular to each other*** *if and only if*

$$\mathbf{a} \cdot \mathbf{b} = 0.$$

Example 9.26 (scalar products of the Cartesian unit vectors)
Any two different Cartesian unit vectors have the angle $\pi/2$, since they lie on the axes of the coordinate system (see Figure 9.4). Thus any two different Cartesian unit vectors are orthogonal/perpendicular to each other. Thus

$$\mathbf{i} \cdot \mathbf{j} = \mathbf{j} \cdot \mathbf{i} = 0, \qquad \mathbf{i} \cdot \mathbf{k} = \mathbf{k} \cdot \mathbf{i} = 0, \qquad \text{and} \qquad \mathbf{j} \cdot \mathbf{k} = \mathbf{k} \cdot \mathbf{j} = 0.$$

Furthermore, since the Cartesian unit vectors are unit vectors, we have

$$\mathbf{i} \cdot \mathbf{i} = |\mathbf{i}|^2 = 1, \qquad \mathbf{j} \cdot \mathbf{j} = |\mathbf{j}|^2 = 1, \qquad \text{and} \qquad \mathbf{k} \cdot \mathbf{k} = |\mathbf{k}|^2 = 1. \qquad \square$$

Theorem 9.27 (distributive law for scalar product)
For any three vectors $\mathbf{a}$, $\mathbf{b}$ *and* $\mathbf{c}$, *we have that*

$$\mathbf{a}\cdot(\mathbf{b}+\mathbf{c}) = \mathbf{a}\cdot\mathbf{b}+\mathbf{a}\cdot\mathbf{c} \qquad \textit{and} \qquad (\mathbf{b}+\mathbf{c})\cdot\mathbf{a} = \mathbf{b}\cdot\mathbf{a}+\mathbf{c}\cdot\mathbf{a}. \tag{9.7}$$

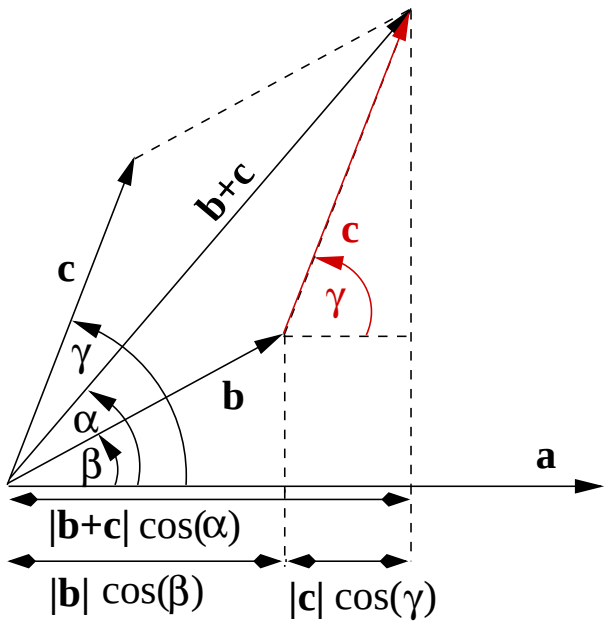


Figure 9.7: Geometric proof of the distributive law of scalar multiplication.

Proof of Theorem 9.27: We only need to prove the first formula in (9.7), since the second formula in (9.7) follows immediately from the first formula and (9.6). The vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, and $\mathbf{b}+\mathbf{c}$ are plotted in Figure 9.7, and the angles between the vector $\mathbf{a}$ and the vectors $\mathbf{b}$, $\mathbf{c}$, and $\mathbf{b}+\mathbf{c}$ are denoted by β, γ, and α, respectively. We have that

$$\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}|\,|\mathbf{b}|\,\cos(\beta), \tag{9.8}$$

$$\mathbf{a}\cdot\mathbf{c} = |\mathbf{a}|\,|\mathbf{c}|\,\cos(\gamma), \tag{9.9}$$

$$\mathbf{a}\cdot(\mathbf{b}+\mathbf{c}) = |\mathbf{a}|\,|\mathbf{b}+\mathbf{c}|\,\cos(\alpha). \tag{9.10}$$

From Figure 9.7, we see that

$$|\mathbf{b}|\,\cos(\beta) + |\mathbf{c}|\,\cos(\gamma) = |\mathbf{b}+\mathbf{c}|\,\cos(\alpha). \tag{9.11}$$

Multiplying (9.11) with $|\mathbf{a}|$ yields

$$|\mathbf{a}|\,|\mathbf{b}|\,\cos(\beta) + |\mathbf{a}|\,|\mathbf{c}|\,\cos(\gamma) = |\mathbf{a}|\,|\mathbf{b}+\mathbf{c}|\,\cos(\alpha),$$

which, using (9.8), (9.9), and (9.10), is just

$$\mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} + \mathbf{c}),$$

and we have proved the distributive law. □

Remark 9.28 (scalars can be pulled out of the scalar product)
Let $\mathbf{a}$ *and* $\mathbf{b}$ *be two vectors in* $\mathbb{R}^3$*, and let* $\lambda \in \mathbb{R}$ *be a scalar. Then*

$$\boxed{\lambda\,(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot (\lambda\,\mathbf{b}) = (\lambda\,\mathbf{a}) \cdot \mathbf{b}.} \tag{9.12}$$

Proof: *If one of the vectors is the zero vector or if* $\lambda = 0$*, then all three terms in (9.12) are zero and (9.12) holds true. Now let us assume that* $\mathbf{a} \neq \mathbf{0}$*,* $\mathbf{b} \neq \mathbf{0}$*, and* $\lambda \neq 0$*. Let* θ *be the angle between* $\mathbf{a}$ *and* $\mathbf{b}$*.*

If $\lambda > 0$*, then the angle between* $\mathbf{a}$ *and* $\lambda\,\mathbf{b}$ *is also* θ*, and hence*

$$\mathbf{a} \cdot (\lambda\,\mathbf{b}) = |\mathbf{a}|\,|\lambda\,\mathbf{b}|\,\cos(\theta) = |\mathbf{a}|\,|\lambda|\,|\mathbf{b}|\,\cos(\theta) = \lambda\,|\mathbf{a}|\,|\mathbf{b}|\,\cos(\theta) = \lambda\,(\mathbf{a} \cdot \mathbf{b}),$$

where we have used $|\lambda| = \lambda$ *since* $\lambda > 0$*.*

If $\lambda < 0$*, then the angle between* $\mathbf{a}$ *and* $\lambda\,\mathbf{b}$ *is* $\pi - \theta$*, and so from the addition theorem for the cosine (see Lemma 2.10)*

$$\cos(\pi - \theta) = \cos(\pi)\,\cos(\theta) + \sin(\pi)\,\sin(\theta) = (-1)\,\cos(\theta) + 0 = -\cos(\theta).$$

Thus we have, using $|\lambda| = -\lambda$ *(since* $\lambda < 0$*),*

$$\begin{aligned}\mathbf{a} \cdot (\lambda\,\mathbf{b}) &= |\mathbf{a}|\,|\lambda\,\mathbf{b}|\,\cos(\pi - \theta) = |\mathbf{a}|\,|\lambda|\,|\mathbf{b}| \times [-\cos(\theta)] \\ &= |\mathbf{a}| \times (-\lambda) \times |\mathbf{b}| \times [-\cos(\theta)] = \lambda\,|\mathbf{a}|\,|\mathbf{b}|\,\cos(\theta) = \lambda\,(\mathbf{a} \cdot \mathbf{b}).\end{aligned}$$

Thus we have proved the first equality in (9.12). The second equality in (9.12) follows from the case already proved. Indeed, we have

$$(\lambda\,\mathbf{a}) \cdot \mathbf{b} = \mathbf{b} \cdot (\lambda\,\mathbf{a}) = \lambda\,(\mathbf{b} \cdot \mathbf{a}) = \lambda\,(\mathbf{a} \cdot \mathbf{b}),$$

where the second step follows from the first equality in (9.12). In the first and third step we have used (9.6). □

Lemma 9.29 (formula for the scalar product in terms of components)
Let $\mathbf{a} = (a_1, a_2, a_3)$ *and let* $\mathbf{b} = (b_1, b_2, b_3)$ *be vectors in* $\mathbb{R}^3$*. Then*

$$\mathbf{a} \cdot \mathbf{b} = a_1\,b_1 + a_2\,b_2 + a_3\,b_3. \tag{9.13}$$

Proof of Lemma 9.29: If $\mathbf{a} = (a_1,\, a_2,\, a_3)$ and $\mathbf{b} = (b_1,\, b_2,\, b_3)$, then

$$\mathbf{a} = a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k} \qquad \text{and} \qquad \mathbf{b} = b_1\,\mathbf{i} + b_2\,\mathbf{j} + b_3\,\mathbf{k}.$$

Using the distributive law (see Theorem 9.27) and (9.12) gives that

$$\begin{aligned}
\mathbf{a}\cdot\mathbf{b} &= (a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k})\cdot(b_1\,\mathbf{i} + b_2\,\mathbf{j} + b_3\,\mathbf{k}) \\
&= (a_1\,\mathbf{i})\cdot(b_1\,\mathbf{i} + b_2\,\mathbf{j} + b_3\,\mathbf{k}) + (a_2\,\mathbf{j})\cdot(b_1\,\mathbf{i} + b_2\,\mathbf{j} + b_3\,\mathbf{k}) \\
&\quad + (a_3\,\mathbf{k})\cdot(b_1\,\mathbf{i} + b_2\,\mathbf{j} + b_3\,\mathbf{k}) \\
&= (a_1\,\mathbf{i})\cdot(b_1\,\mathbf{i}) + (a_1\,\mathbf{i})\cdot(b_2\,\mathbf{j}) + (a_1\,\mathbf{i})\cdot(b_3\,\mathbf{k}) \\
&\quad + (a_2\,\mathbf{j})\cdot(b_1\,\mathbf{i}) + (a_2\,\mathbf{j})\cdot(b_2\,\mathbf{j}) + (a_2\,\mathbf{j})\cdot(b_3\,\mathbf{k}) \\
&\quad + (a_3\,\mathbf{k})\cdot(b_1\,\mathbf{i}) + (a_3\,\mathbf{k})\cdot(b_2\,\mathbf{j}) + (a_3\,\mathbf{k})\cdot(b_3\,\mathbf{k}) \\
&= (a_1\,b_1)\,\mathbf{i}\cdot\mathbf{i} + (a_1\,b_2)\,\mathbf{i}\cdot\mathbf{j} + (a_1\,b_3)\,\mathbf{i}\cdot\mathbf{k} + (a_2\,b_1)\,\mathbf{j}\cdot\mathbf{i} + (a_2\,b_2)\,\mathbf{j}\cdot\mathbf{j} \\
&\quad + (a_2\,b_3)\,\mathbf{j}\cdot\mathbf{k} + (a_3\,b_1)\,\mathbf{k}\cdot\mathbf{i} + (a_3\,b_2)\,\mathbf{k}\cdot\mathbf{j} + (a_3\,b_3)\,\mathbf{k}\cdot\mathbf{k} \\
&= a_1\,b_1 + a_2\,b_2 + a_3\,b_3,
\end{aligned}$$

where we have used $\mathbf{i}\cdot\mathbf{j} = \mathbf{i}\cdot\mathbf{k} = \mathbf{j}\cdot\mathbf{k} = 0$ and $\mathbf{i}\cdot\mathbf{i} = |\mathbf{i}|^2 = 1$, $\mathbf{j}\cdot\mathbf{j} = |\mathbf{j}|^2 = 1$, and $\mathbf{k}\cdot\mathbf{k} = |\mathbf{k}|^2 = 1$ in the last step (see Example 9.26). □

Example 9.30 (scalar product)
If $\mathbf{a} = (4, 2, 3)$ and $\mathbf{b} = (-1, 0, 6)$, then

$$\mathbf{a}\cdot\mathbf{b} = (4,2,3)\cdot(-1,0,6) = 4\times(-1)+2\times 0+3\times 6 = -4+0+18 = 14. \qquad \square$$

Example 9.31 (scalar product)
Compute the scalar products $\mathbf{a}\cdot\mathbf{b}$ and $\mathbf{a}\cdot\mathbf{c}$ of the vectors $\mathbf{a} = (2, 2, 0)$, $\mathbf{b} = (0, 1, 0)$, and $\mathbf{c} = (-1, 3, 5)$.

Solution: From (9.13), we have

$$\mathbf{a}\cdot\mathbf{b} = (2,2,0)\cdot(0,1,0) = 2\times 0 + 2\times 1 + 0\times 0 = 0 + 2 + 0 = 2.$$

$$\mathbf{a}\cdot\mathbf{c} = (2,2,0)\cdot(-1,3,5) = 2\times(-1) + 2\times 3 + 0\times 5 = -2 + 6 + 0 = 4. \qquad \square$$

Remark 9.32 (finding the angle between two vectors)
To find the angle θ ***between*** $\mathbf{a}$ ***and*** $\mathbf{b}$*, we can use that, from (9.5) and (9.13),*

$$|\mathbf{a}|\,|\mathbf{b}|\,\cos(\theta) = \mathbf{a}\cdot\mathbf{b} = a_1\,b_1 + a_2\,b_2 + a_3\,b_3.$$

Indeed, if $\mathbf{a} \neq \mathbf{0}$ *and* $\mathbf{b} \neq \mathbf{0}$*, then* $|\mathbf{a}| \neq 0$ *and* $|\mathbf{b}| \neq 0$*, and it follows that*

$$\boxed{\cos(\theta) = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|} = \frac{a_1\,b_1 + a_2\,b_2 + a_3\,b_3}{|\mathbf{a}|\,|\mathbf{b}|}.} \tag{9.14}$$

Now we can determine $\theta \in [0, \pi]$ *by taking the* arccos *of (9.14).*

Example 9.33 (determination of the angle between two vectors)
The angle θ between $\mathbf{a} = (4, 2, 3)$ and $\mathbf{b} = (-1, 0, 6)$ is given by

$$\cos(\theta) = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|} = \frac{14}{\sqrt{29}\sqrt{37}} = \frac{14}{\sqrt{29 \times 37}} \approx 0.42739,$$

where we have used $|\mathbf{a}| = \sqrt{29}$, $|\mathbf{b}| = \sqrt{37}$, and $\mathbf{a} \cdot \mathbf{b} = 14$ from Examples 9.5 and 9.30. So $\theta = \arccos(14/\sqrt{29 \times 37}) \approx 1.1292$ radians, that is, $\theta \approx 64.698°$. □

Application 9.34 (a force drags an object horizontally at an angle)
One place where scalar product/dot products are useful is finding the work done. Suppose a ***force*** $\mathbf{F}$ ***drags an object of mass*** m ***along a horizontal rough surface*** *giving rise to a displacement* $\mathbf{r}$*, and assume that the* ***force acts at an angle*** θ ***to the motion****. Think, for example, of using a rope to pull a sledge, where the angle between the rope and the horizontal ground is* θ *(see Figure 9.8). Then the total* ***work done by the force in overcoming friction*** *is*

$$\text{work} = \text{horizontal component of force} \times \text{distance moved} = |\mathbf{F}|\cos(\theta)\cdot|\mathbf{r}| = \mathbf{F}\cdot\mathbf{r}.$$

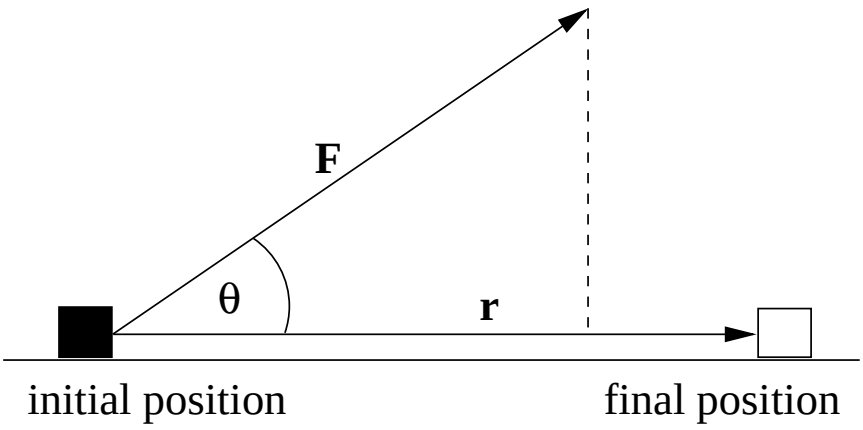


Figure 9.8: An object is dragged horizontally by a force acting at an angle θ.

9.4 Vector Product or Cross Product

In this section we discuss the **vector product** or **cross product of two vectors** which gives another vector.

Definition 9.35 (vector product or cross product)
The ***vector product*** *or* ***cross product*** *of two non-zero vectors* $\mathbf{a}$ *and* $\mathbf{b}$ *is the vector* $\mathbf{a} \times \mathbf{b}$ *whose* ***magnitude/length*** *is*

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}|\,\sin(\theta), \tag{9.15}$$

where $\theta \in [0, \pi]$ *is the angle between* $\mathbf{a}$ *and* $\mathbf{b}$*, and whose* ***direction is perpendicular to both*** $\mathbf{a}$ ***and*** $\mathbf{b}$ ***when taken in the right-handed sense from*** $\mathbf{a}$ ***to*** $\mathbf{b}$*. If* $\mathbf{a} = \mathbf{0}$ *or* $\mathbf{b} = \mathbf{0}$*, then we define* $\mathbf{a} \times \mathbf{b} = \mathbf{0}$*.*

What it means that the direction of $\mathbf{a} \times \mathbf{b}$ is 'perpendicular to both $\mathbf{a}$ and $\mathbf{b}$ when taken in the **right-handed sense** from $\mathbf{a}$ to $\mathbf{b}$' is illustrated in the left picture in Figure 9.9. Right-handed sense from $\mathbf{a}$ to $\mathbf{b}$ means that when determining the direction of $\mathbf{a} \times \mathbf{b}$, we hold thumb, index finger, and middle finger of the right hand such that they are perpendicular to each other. If we take $\mathbf{a}$ to be the thumb and $\mathbf{b}$ to be the index finger, then the direction of the middle finger gives the direction of $\mathbf{a} \times \mathbf{b}$.

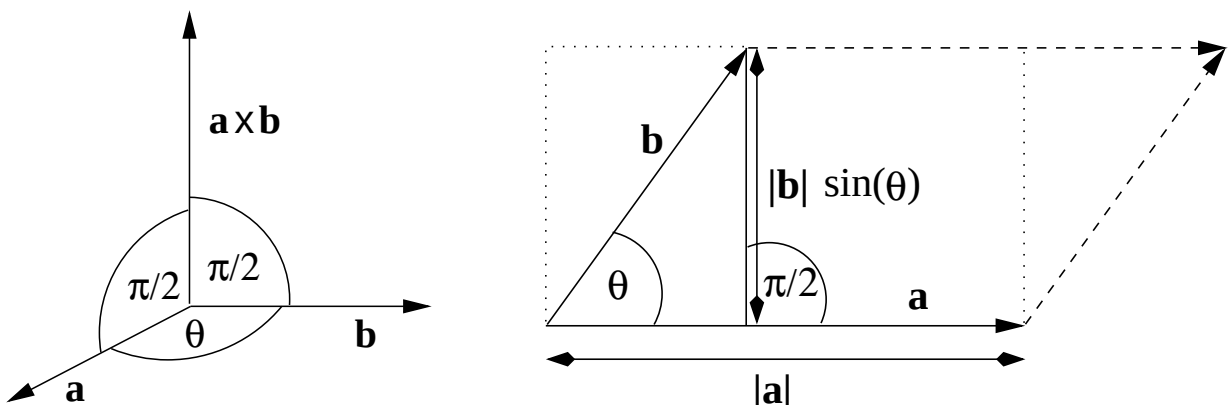


Figure 9.9: On the left, the vector product/cross product $\mathbf{a} \times \mathbf{b}$ is perpendicular to $\mathbf{a}$ and $\mathbf{b}$ in the right-handed sense from $\mathbf{a}$ to $\mathbf{b}$. On the right, the length $|\mathbf{a} \times \mathbf{b}|$ of the vector product $\mathbf{a} \times \mathbf{b}$ is the area of the parallelogram.

The **length** $|\mathbf{a} \times \mathbf{b}|$ **of the vector product/cross product** has a **geometric interpretation** as illustrated in the right picture in Figure 9.9: In (9.15)

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}|\,\sin(\theta) = |\mathbf{a}|\,\big(|\mathbf{b}|\,\sin(\theta)\big), \tag{9.16}$$

and the term $|\mathbf{b}|\,\sin(\theta)$ is the height of the parallelogram with corner points given by the vectors $\mathbf{0}$, $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{a} + \mathbf{b}$ as indicated in Figure 9.9. Thus (9.16) is the area of the parallelogram in Figure 9.16.

Example 9.36 (vector product/cross product)
Let $\mathbf{a} = (1, 0, 0)$ and $\mathbf{b} = (1, -1, 0)$. Then $|\mathbf{a}| = \sqrt{1^2 + 0^2 + 0^2} = 1$ and $|\mathbf{b}| = \sqrt{1^2 + (-1)^2 + 0^2} = \sqrt{2}$. From $\mathbf{a} \cdot \mathbf{b} = 1 \times 1 + 0 \times (-1) + 0 \times 0 = 1$, we can determine the angle θ between $\mathbf{a}$ and $\mathbf{b}$ as follows (see (9.14))

$$\cos(\theta) = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|} = \frac{1}{1 \times \sqrt{2}} = \frac{1}{\sqrt{2}} \qquad \Rightarrow \qquad \theta = \frac{\pi}{4}.$$

Thus the length/magnitude $|\mathbf{a} \times \mathbf{b}|$ of $\mathbf{a} \times \mathbf{b}$ is therefore

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}|\,\sin(\theta) = 1 \times \sqrt{2} \times \sin(\pi/4) = \sqrt{2} \times \frac{1}{\sqrt{2}} = 1.$$

It remains to determine the direction of $\mathbf{a} \times \mathbf{b}$. From the fact that $\mathbf{a}$ and $\mathbf{b}$ lie in the (x, y)-plane and the fact that $\mathbf{a} \times \mathbf{b}$ is perpendicular to $\mathbf{a}$ and $\mathbf{b}$ in the right-handed sense from $\mathbf{a}$ to $\mathbf{b}$, we can determine that the direction of $\mathbf{a} \times \mathbf{b}$ is $-\mathbf{k} = (0, 0, -1)$. Since $|\mathbf{a} \times \mathbf{b}| = 1$ and $|-\mathbf{k}| = 1$, we have

$$\mathbf{a} \times \mathbf{b} = -\mathbf{k} = (0, 0, -1). \qquad \square$$

Remark 9.37 (properties of the vector product)

(a) For any two vectors $\mathbf{a}$ *and* $\mathbf{b}$*,*

$$\boxed{\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a},} \tag{9.17}$$

which can easily be seen from the fact that the direction of the vector product/cross product is perpendicular to $\mathbf{a}$ *and* $\mathbf{b}$ *in the right-handed sense from* $\mathbf{a}$ *to* $\mathbf{b}$*.*

(b) For any vector $\mathbf{a} \neq \mathbf{0}$*, we have that* $\mathbf{a}$ *has the angle* $\theta = 0$ *to itself. Thus we have* $|\mathbf{a} \times \mathbf{a}| = |\mathbf{a}|\,|\mathbf{a}|\,\sin(0) = 0$*, and so*

$$\boxed{\mathbf{a} \times \mathbf{a} = \mathbf{0}.} \tag{9.18}$$

Example 9.38 (vector product of the Cartesian unit vectors)
Since any two different Cartesian unit vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ have the angle $\pi/2$ to each other, the vector product/cross product of any two different Cartesian unit vectors satisfies

$$|\mathbf{i} \times \mathbf{j}| = |\mathbf{i}|\,|\mathbf{j}|\,\sin(\pi/2) = 1 \times 1 \times 1 = 1,$$
$$|\mathbf{i} \times \mathbf{k}| = |\mathbf{i}|\,|\mathbf{k}|\,\sin(\pi/2) = 1 \times 1 \times 1 = 1,$$
$$|\mathbf{j} \times \mathbf{k}| = |\mathbf{j}|\,|\mathbf{k}|\,\sin(\pi/2) = 1 \times 1 \times 1 = 1.$$

Since $\mathbf{i} \times \mathbf{j}$ is perpendicular to both $\mathbf{i}$ and $\mathbf{j}$ in the right-handed sense from $\mathbf{i}$ to $\mathbf{j}$, we have that the direction of $\mathbf{i} \times \mathbf{j}$ is $\mathbf{k}$. Since $|\mathbf{i} \times \mathbf{j}| = 1$ and $|\mathbf{k}| = 1$, we have $\mathbf{i} \times \mathbf{j} = \mathbf{k}$. Analogously, we can determine all the following vector products

$$\mathbf{i} \times \mathbf{j} = -\mathbf{j} \times \mathbf{i} = \mathbf{k}, \qquad \mathbf{j} \times \mathbf{k} = -\mathbf{k} \times \mathbf{j} = \mathbf{i}, \qquad \mathbf{k} \times \mathbf{i} = -\mathbf{i} \times \mathbf{k} = \mathbf{j}.$$

Furthermore, from (9.18), we have

$$\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = \mathbf{0}. \qquad \square$$

Definition 9.39 (scalar triple product)
Given three vectors $\mathbf{a}$, $\mathbf{b}$ *and* $\mathbf{c}$, *the scalar*

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$$

is called the ***scalar triple product*** *of* $\mathbf{a}$, $\mathbf{b}$ *and* $\mathbf{c}$.

The next lemma states that the scalar triple product is the volume of the parallelepiped spanned by the three vectors. This lemma will be proved on the exercise sheet for this chapter.

Lemma 9.40 (interpretation of scalar triple product)
We have that for any three vectors $\mathbf{a}$, $\mathbf{b}$ *and* $\mathbf{c}$,

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}). \tag{9.19}$$

If the vectors $\mathbf{a}$, $\mathbf{b}$, *and* $\mathbf{c}$ *are not* ***coplanar*** *(that is, if they do* ***not*** *lie all in one plane), then the scalar triple product is the* ***volume of the parallelepiped*** *spanned by* $\mathbf{a}$, $\mathbf{b}$, *and* $\mathbf{c}$, *as illustrated in Figure 9.10.*

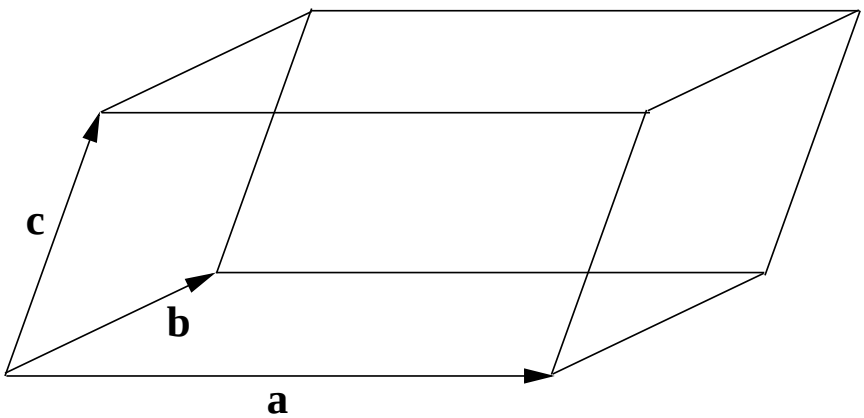


Figure 9.10: The volume of the parallelepiped spanned by $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ is the scalar triple product $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$.

For the vector/cross product, we have the following distributive law.

Theorem 9.41 (distributive law for vector product)
For any three vectors $\mathbf{a}$, $\mathbf{b}$ *and* $\mathbf{c}$ *in* $\mathbb{R}^3$, *we have that*

$$\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) + (\mathbf{a} \times \mathbf{c}) \quad \textit{and} \quad (\mathbf{b} + \mathbf{c}) \times \mathbf{a} = (\mathbf{b} \times \mathbf{a}) + (\mathbf{c} \times \mathbf{a}). \tag{9.20}$$

Proof of Theorem 9.41: We only need to prove the first equation in (9.20), since the second equation follows immediately from the first and (9.17). Let

$$\mathbf{d} = \mathbf{a} \times (\mathbf{b} + \mathbf{c}) - (\mathbf{a} \times \mathbf{b}) - (\mathbf{a} \times \mathbf{c}),$$

that is, we have arranged all terms in (9.20) on the left-hand side of (9.20). In order to prove (9.20), we have to show that $\mathbf{d} = \mathbf{0}$. To do this, we show that $|\mathbf{d}|^2 = \mathbf{d} \cdot \mathbf{d}$ is zero, and then we know from Remark 9.6 that $\mathbf{d} = \mathbf{0}$. Indeed, using the distributive law for the scalar product (see Theorem 9.27) gives that

$$\begin{aligned} \mathbf{d} \cdot \mathbf{d} &= \mathbf{d} \cdot \big[\mathbf{a} \times (\mathbf{b} + \mathbf{c}) - (\mathbf{a} \times \mathbf{b}) - (\mathbf{a} \times \mathbf{c})\big] \\ &= \mathbf{d} \cdot \big[\mathbf{a} \times (\mathbf{b} + \mathbf{c})\big] - \mathbf{d} \cdot \big[\mathbf{a} \times \mathbf{b}\big] - \mathbf{d} \cdot \big[\mathbf{a} \times \mathbf{c}\big]. \end{aligned}$$

Using the property (9.19) of the triple scalar product gives that

$$\mathbf{d} \cdot \mathbf{d} = (\mathbf{b} + \mathbf{c}) \cdot [\mathbf{d} \times \mathbf{a}] - \mathbf{b} \cdot [\mathbf{d} \times \mathbf{a}] - \mathbf{c} \cdot [\mathbf{d} \times \mathbf{a}].$$

Using the distributive law for the scalar product (see Theorem 9.27) again gives

$$\mathbf{d} \cdot \mathbf{d} = \big[(\mathbf{b} + \mathbf{c}) - \mathbf{b} - \mathbf{c}\big] \cdot [\mathbf{d} \times \mathbf{a}] = \mathbf{0} \cdot [\mathbf{d} \times \mathbf{a}] = 0. \qquad \square$$

Remark 9.42 (scalar factors can be pulled out of the vector product)
For any two vectors $\mathbf{a}$ *and* $\mathbf{b}$ *in* $\mathbb{R}^3$ *and any scalar* $\lambda \in \mathbb{R}$,

$$\boxed{\mathbf{a} \times (\lambda\, \mathbf{b}) = (\lambda\, \mathbf{a}) \times \mathbf{b} = \lambda\, (\mathbf{a} \times \mathbf{b}).} \tag{9.21}$$

Proof: *If* $\mathbf{a} = \mathbf{0}$, $\mathbf{b} = \mathbf{0}$, *or* $\lambda = 0$, *then all three terms in (9.21) are equal to the zero vector* $\mathbf{0}$ *and the statement is true. Now assume that* $\mathbf{a} \neq \mathbf{0}$, $\mathbf{b} \neq \mathbf{0}$, *and* $\lambda \neq 0$. *We will show that the term on the left-hand side of (9.21) equals the term on the right-hand side of (9.21). The proof that the term in the middle of (9.21) equals the term on the right-hand side of (9.21) is analogous.*

Let θ *be the angle between* $\mathbf{a}$ *and* $\mathbf{b}$. *If* $\lambda > 0$, *then the angle between* $\mathbf{a}$ *and* $\lambda\, \mathbf{b}$ *is also* θ, *and hence the magnitude/length of* $\mathbf{a} \times (\lambda\, \mathbf{b})$ *is*

$$|\mathbf{a} \times (\lambda\, \mathbf{b})| = |\mathbf{a}|\, |\lambda\, \mathbf{b}| \sin(\theta) = |\mathbf{a}| \big(|\lambda|\, |\mathbf{b}|\big) \sin(\theta)$$

$$= |\lambda| \big(|\mathbf{a}|\,|\mathbf{b}|\,\sin(\theta)\big) = |\lambda|\,|\mathbf{a}\times\mathbf{b}| = |\lambda\,(\mathbf{a}\times\mathbf{b})|. \tag{9.22}$$

Furthermore, $\mathbf{a}\times(\lambda\,\mathbf{b})$ *is perpendicular to both* $\mathbf{a}$ *and* $\lambda\,\mathbf{b}$ *in the right-handed sense from* $\mathbf{a}$ *to* $\lambda\,\mathbf{b}$*, that is, to both* $\mathbf{a}$ *and* $\mathbf{b}$ *in the right-handed sense from* $\mathbf{a}$ *to* $\mathbf{b}$*. So* $\mathbf{a}\times(\lambda\,\mathbf{b})$ *points in the same direction as* $\mathbf{a}\times\mathbf{b}$*, and hence in the same direction as* $\lambda\,(\mathbf{a}\times\mathbf{b})$*. From this and (9.22), it follows that*

$$\mathbf{a}\times(\lambda\,\mathbf{b}) = \lambda\,(\mathbf{a}\times\mathbf{b}).$$

If $\lambda<0$*, then the angle between* $\mathbf{a}$ *and* $\lambda\,\mathbf{b}$ *is* $\pi-\theta$*, and so the length of* $\mathbf{a}\times(\lambda\,\mathbf{b})$ *is*

$$\begin{aligned}|\mathbf{a}\times(\lambda\,\mathbf{b})| &= |\mathbf{a}|\,|\lambda\,\mathbf{b}|\,\sin(\pi-\theta) = |\mathbf{a}|\,\big(|\lambda|\,|\mathbf{b}|\big)\,\big(\sin(\pi)\,\cos(\theta)-\sin(\theta)\,\cos(\pi)\big)\\ &= |\mathbf{a}|\,|\lambda|\,|\mathbf{b}|\,\big(0-\sin(\theta)\times(-1)\big)\\ &= |\lambda|\,\big(|\mathbf{a}|\,|\mathbf{b}|\,\sin(\theta)\big) = |\lambda|\,|\mathbf{a}\times\mathbf{b}| = |\lambda\,(\mathbf{a}\times\mathbf{b})|. \end{aligned}\tag{9.23}$$

Furthermore, $\mathbf{a}\times(\lambda\mathbf{b})$ *is perpendicular to both* $\mathbf{a}$ *and* $\lambda\,\mathbf{b}$ *in the right-handed sense from* $\mathbf{a}$ *to* $\lambda\,\mathbf{b}$*, that is, perpendicular to both* $\mathbf{a}$ *and* $\mathbf{b}$ *in the left-handed sense from* $\mathbf{a}$ *to* $\mathbf{b}$*. So* $\mathbf{a}\times(\lambda\,\mathbf{b})$ *points in the opposite direction to* $\mathbf{a}\times\mathbf{b}$*, and hence in the same direction as* $\lambda\,(\mathbf{a}\times\mathbf{b})$*. From this and (9.23), it follows that*

$$\mathbf{a}\times(\lambda\,\mathbf{b}) = \lambda\,(\mathbf{a}\times\mathbf{b}).$$

So we have proved that for any two vectors $\mathbf{a}$ *and* $\mathbf{b}$ *and any scalar* $\lambda\in\mathbb{R}$*,*

$$\mathbf{a}\times(\lambda\,\mathbf{b}) = \lambda\,(\mathbf{a}\times\mathbf{b}). \qquad \square$$

The next theorem gives us a very useful representation of the vector product/cross product $\mathbf{a}\times\mathbf{b}$ in terms of the components of the vectors $\mathbf{a}$ and $\mathbf{b}$.

Theorem 9.43 (formula for cross product in terms of components)
Let $\mathbf{a}=(a_1,a_2,a_3)$ *and* $\mathbf{b}=(b_1,b_2,b_3)$ *be two vectors in* $\mathbb{R}^3$*. Then*

$$\begin{aligned}\mathbf{a}\times\mathbf{b} &= (a_2\,b_3-a_3\,b_2)\,\mathbf{i}-(a_1\,b_3-a_3\,b_1)\,\mathbf{j}+(a_1\,b_2-a_2\,b_1)\,\mathbf{k}\\ &= \big(a_2\,b_3-a_3\,b_2,\,-(a_1\,b_3-a_3\,b_1),\,a_1\,b_2-a_2\,b_1\big).\end{aligned}\tag{9.24}$$

Proof of Theorem 9.43: To prove (9.24), we start by writing the vectors $\mathbf{a}$ and $\mathbf{b}$ in terms of the Cartesian unit vectors, that is, $\mathbf{a}=(a_1,a_2,a_3)=a_1\,\mathbf{i}+a_2\,\mathbf{j}+a_3\,\mathbf{k}$ and $\mathbf{b}=(b_1,b_2,b_3)=b_1\,\mathbf{i}+b_2\,\mathbf{j}+b_3\,\mathbf{k}$, and substitute this into the vector product/cross product $\mathbf{a}\times\mathbf{b}$. Thus

$$\mathbf{a}\times\mathbf{b} = (a_1\,\mathbf{i}+a_2\,\mathbf{j}+a_3\,\mathbf{k})\times(b_1\,\mathbf{i}+b_2\,\mathbf{j}+b_3\,\mathbf{k}).$$

Using the distributive law (9.20) and (9.21) to transform the right-hand side gives

$$\begin{aligned}
\mathbf{a}\times\mathbf{b} &= (a_1\,\mathbf{i}+a_2\,\mathbf{j}+a_3\,\mathbf{k})\times(b_1\,\mathbf{i}+b_2\,\mathbf{j}+b_3\,\mathbf{k})\\
&= (a_1\,\mathbf{i})\times(b_1\,\mathbf{i}+b_2\,\mathbf{j}+b_3\,\mathbf{k})+(a_2\,\mathbf{j})\times(b_1\,\mathbf{i}+b_2\,\mathbf{j}+b_3\,\mathbf{k})\\
&\quad+(a_3\,\mathbf{k})\times(b_1\,\mathbf{i}+b_2\,\mathbf{j}+b_3\,\mathbf{k})\\
&= (a_1\,\mathbf{i})\times(b_1\,\mathbf{i})+(a_1\,\mathbf{i})\times(b_2\,\mathbf{j})+(a_1\,\mathbf{i})\times(b_3\,\mathbf{k})\\
&\quad+(a_2\,\mathbf{j})\times(b_1\,\mathbf{i})+(a_2\,\mathbf{j})\times(b_2\,\mathbf{j})+(a_2\,\mathbf{j})\times(b_3\,\mathbf{k})\\
&\quad+(a_3\,\mathbf{k})\times(b_1\,\mathbf{i})+(a_3\,\mathbf{k})\times(b_2\,\mathbf{j})+(a_3\,\mathbf{k})\times(b_3\,\mathbf{k})\\
&= (a_1\,b_1)\,(\mathbf{i}\times\mathbf{i})+(a_1\,b_2)\,(\mathbf{i}\times\mathbf{j})+(a_1\,b_3)\,(\mathbf{i}\times\mathbf{k})\\
&\quad+(a_2\,b_1)\,(\mathbf{j}\times\mathbf{i})+(a_2\,b_2)\,(\mathbf{j}\times\mathbf{j})+(a_2\,b_3)\,(\mathbf{j}\times\mathbf{k})\\
&\quad+(a_3\,b_1)\,(\mathbf{k}\times\mathbf{i})+(a_3\,b_2)\,(\mathbf{k}\times\mathbf{j})+(a_3\,b_3)\,(\mathbf{k}\times\mathbf{k})\\
&= (a_1\,b_2)\,\mathbf{k}-(a_1\,b_3)\,\mathbf{j}-(a_2\,b_1)\,\mathbf{k}+(a_2\,b_3)\,\mathbf{i}+(a_3\,b_1)\,\mathbf{j}-(a_3\,b_2)\,\mathbf{i}\\
&= (a_2\,b_3-a_3\,b_2)\,\mathbf{i}-(a_1\,b_3-a_3\,b_1)\,\mathbf{j}+(a_1\,b_2-a_2\,b_1)\,\mathbf{k},
\end{aligned}$$

where we have used $\mathbf{i}\times\mathbf{i}=\mathbf{j}\times\mathbf{j}=\mathbf{k}\times\mathbf{k}=\mathbf{0}$ and

$$\mathbf{i}\times\mathbf{j}=-\mathbf{j}\times\mathbf{i}=\mathbf{k},\qquad \mathbf{j}\times\mathbf{k}=-\mathbf{k}\times\mathbf{j}=\mathbf{i}\qquad\text{and}\qquad \mathbf{k}\times\mathbf{i}=-\mathbf{i}\times\mathbf{k}=\mathbf{j}$$

in the last step (see Example 9.38). Note the alternating signs, $+$, $-$, $+$, of the terms involving $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$. □

Remark 9.44 (alternative representation of (9.24))
In school some of you may have used the formula (9.25) below for the vector product which also leads to (9.24). This formula uses the notation of the ***determinant*** *which we will encounter in the next chapter. If you have not seen the notation below, do not worry about it for the moment; we will come back to it in the next chapter.*

$$\begin{aligned}
\mathbf{a}\times\mathbf{b} &= \begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\end{vmatrix} = \begin{vmatrix}a_2 & a_3\\ b_2 & b_3\end{vmatrix}\mathbf{i}-\begin{vmatrix}a_1 & a_3\\ b_1 & b_3\end{vmatrix}\mathbf{j}+\begin{vmatrix}a_1 & a_2\\ b_1 & b_2\end{vmatrix}\mathbf{k}\\
&= (a_2\,b_3-a_3\,b_2)\,\mathbf{i}-(a_1\,b_3-a_3\,b_1)\,\mathbf{j}+(a_1\,b_2-a_2\,b_1)\,\mathbf{k}. \qquad (9.25)
\end{aligned}$$

Example 9.45 (vector product/cross product)
Compute the vector product/cross product $\mathbf{a}\times\mathbf{b}$ for the following pairs of vectors:

(a) $\mathbf{a}=(4,2,3)$, $\mathbf{b}=(-1,0,6)$; (b) $\mathbf{a}=(1,2,1)$, $\mathbf{b}=(3,1,2)$.

Solution: We use (9.24) to compute the cross products:

$$\begin{aligned}\mathbf{a}\times\mathbf{b} &= (4,2,3)\times(-1,0,6)\\ &= \big(2\times 6-3\times 0\big)\,\mathbf{i}-\big(4\times 6-3\times(-1)\big)\,\mathbf{j}+\big(4\times 0-2\times(-1)\big)\,\mathbf{k}\\ &= 12\,\mathbf{i}-27\,\mathbf{j}+2\,\mathbf{k} \;=\; (12,-27,2),\end{aligned}$$

and

$$\begin{aligned}\mathbf{a}\times\mathbf{b} &= (1,2,1)\times(3,1,2)\\ &= \big(2\times 2-1\times 1\big)\,\mathbf{i}-\big(1\times 2-1\times 3\big)\,\mathbf{j}+\big(1\times 1-2\times 3\big)\,\mathbf{k}\\ &= 3\,\mathbf{i}-(-1)\,\mathbf{j}+(-5)\,\mathbf{k} \;=\; 3\,\mathbf{i}+\mathbf{j}-5\,\mathbf{k} \;=\; (3,1,-5).\end{aligned}$$

Remark 9.46 (unit vectors perpendicular to two non-parallel vectors) *For any two non-zero, non-parallel vectors* $\mathbf{a}$ *and* $\mathbf{b}$*, the vectors* $\pm(\mathbf{a}\times\mathbf{b})$ *are non-zero vectors which are perpendicular to both* $\mathbf{a}$ *and* $\mathbf{b}$*. So the* ***unit vectors which are perpendicular to both*** $\mathbf{a}$ ***and*** $\mathbf{b}$ *are*

$$\boxed{\text{unit vectors perpendicular to } \mathbf{a} \text{ and } \mathbf{b} \;=\; \pm\,\frac{\mathbf{a}\times\mathbf{b}}{|\mathbf{a}\times\mathbf{b}|}.}$$

Example 9.47 (unit vectors perpendicular to two non-parallel vectors) Find the unit vectors which are perpendicular to both vectors $\mathbf{a}=(3,6,9)$ and $\mathbf{b}=(-2,3,1)$.

Solution: We have that

$$\begin{aligned}\mathbf{a}\times\mathbf{b} &= (3,6,9)\times(-2,3,1)\\ &= \big(6\times 1-9\times 3\big)\,\mathbf{i}-\big(3\times 1-9\times(-2)\big)\,\mathbf{j}+\big(3\times 3-6\times(-2)\big)\,\mathbf{k}\\ &= -21\,\mathbf{i}-21\,\mathbf{j}+21\,\mathbf{k}=21\,(-\mathbf{i}-\mathbf{j}+\mathbf{k})=21\,(-1,-1,1).\end{aligned}$$

Now we work out the length of $\mathbf{a}\times\mathbf{b}$

$$|\mathbf{a}\times\mathbf{b}|=|21\,(-1,-1,1)|=21|(-1,-1,1)|=21\,\sqrt{(-1)^2+(-1)^2+1^2}=21\,\sqrt{3},$$

and hence the unit vectors which are perpendicular to both $\mathbf{a}$ and $\mathbf{b}$ are

$$\pm\,\frac{\mathbf{a}\times\mathbf{b}}{|\mathbf{a}\times\mathbf{b}|}=\pm\,\frac{21\,(-1,-1,1)}{21\,\sqrt{3}}=\frac{\pm 1}{\sqrt{3}}\,(-1,-1,1). \qquad \square$$

Remark 9.48 (Warning: $(\mathbf{a}\times\mathbf{b})\times\mathbf{c}\neq\mathbf{a}\times(\mathbf{b}\times\mathbf{c})$) *In general,*

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{c}\neq\mathbf{a}\times(\mathbf{b}\times\mathbf{c}).$$

Indeed, the vector on the left-hand side is perpendicular to $\mathbf{c}$ *and to the vector* $\mathbf{a} \times \mathbf{b}$ *(which itself is perpendicular to* $\mathbf{a}$ *and* $\mathbf{b}$*). The vector on the right-hand side is perpendicular* $\mathbf{a}$ *and to the vector* $\mathbf{b} \times \mathbf{c}$ *(which itself is perpendicular to* $\mathbf{b}$ *and* $\mathbf{c}$*).*

Application 9.49 (torque)
If, for example, you use a box spanner to undo a wheel nut, you apply a force $\mathbf{F}$ *at some point* $\mathbf{r}$*. It is assumed that the centre of the nut is at* $\mathbf{0}$*. If the angle between* $\mathbf{F}$ *and* $\mathbf{r}$ *is* θ*, then the* ***torque*** *is a vector with magnitude*

$$|\mathbf{r}|\,|\mathbf{F}|\,\sin(\theta) = |\mathbf{r} \times \mathbf{F}|.$$

The direction of the torque is perpendicular to the plane containing $\mathbf{F}$ *and* $\mathbf{r}$*. In fact, the torque is given by*

$$\text{torque} \;=\; \mathbf{r} \times \mathbf{F},$$

the ***vector product/cross product*** *of* $\mathbf{r}$ *and* $\mathbf{F}$*.*

Chapter 10

Matrices

In Section 10.1 we introduce $m \times n$ **matrices** which are rectangular arrays of real numbers consisting of m rows and n columns. For example,

$$A = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 3 & 1 \end{pmatrix}$$

is a 2×3 matrix with two rows and three columns. Vectors which we have encountered in Chapter 9, are a special case of matrices. In Section 10.2, we learn how to **add** and **subtract matrices** of the same size and how to **multiply matrices with scalars**. We also learn how to **multiply matrices** with each other. In Section 10.3, we introduce the **determinant** of a square matrix, which is a scalar. In Section 10.4, we learn how to compute the **inverse matrix** of a matrix with determinant different from zero. Finally we will use matrices and their inverse matrices to **solve linear systems of equations** in Section 10.5.

10.1 Introduction

We start by introducing matrices as rectangular arrays of real numbers. In the previous chapter we have already encountered vectors which are special cases of matrices.

Definition 10.1 (matrix)
*An $m \times n$ (read 'm by n') **matrix** $A = (a_{i,j}) = (a_{i,j})_{j=1,2,\ldots,n}^{i=1,2,\ldots,m}$ is a rectangular array of real numbers containing m **rows** and n **columns***

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & a_{m,3} & \cdots & a_{m,n} \end{pmatrix}$$

*The entry in the intersection of the ith row and jth column is denoted by $a_{i,j}$ and is called the (i,j)**th entry** of the matrix. That is, the **first index** i **of** $a_{i,j}$ **denotes the row** and the **second index** j **of** $a_{i,j}$ **denotes the column**.*

Example 10.2 (matrices)

(a) The matrix

$$A = (a_{i,j}) = \begin{pmatrix} 2 & 3 & 1 & -4 \\ 2 & 1 & 0 & 5 \end{pmatrix}$$

has $m = 2$ rows and $n = 4$ columns, and hence is a 2×4 matrix. Here we have $a_{1,1} = 2$, $a_{1,2} = 3$, $a_{1,3} = 1$, $a_{1,4} = -4$, and $a_{2,1} = 2$, $a_{2,2} = 1$, $a_{2,3} = 0$, $a_{2,4} = 5$.

(b) The matrix

$$B = (b_{i,j}) = \begin{pmatrix} 1 & -2 & 3 \\ -2 & 1 & 4 \\ 7 & 5 & 9 \\ -3 & -4 & -8 \end{pmatrix}$$

has $m = 4$ rows and $n = 3$ columns, and hence is a 4×3 matrix.

(c) The **row vector** $\mathbf{a} = (1, 4, -5)$ has $m = 1$ rows and $n = 3$ columns, and hence is a 1×3 matrix.

(d) The **column vector**

$$\mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

has $m = 3$ rows and $n = 1$ columns, and hence is a 3×1 matrix. □

Definition 10.3 (equal matrices)
Two matrices are said to be ***equal*** *if they have the* ***same size****, that is, the same numbers of rows and the same number of columns, and if* ***all corresponding entries are the same****. In formulas, an $m \times n$ matrix $A = (a_{i,j})$ and an $\ell \times k$ matrix $B = (b_{i,j})$ are* ***equal*** *if $m = \ell$ and $n = k$ and*

$$a_{i,j} = b_{i,j} \qquad \text{for all } i = 1, 2, \ldots, m \text{ and all } j = 1, 2, \ldots, n.$$

If two matrices A and B are ***equal*** *we may write $A = B$, and if two matrices are* ***not equal*** *we may write $A \neq B$.*

Example 10.4 (equal and unequal matrices)
(a) Consider the matrices

$$A = \begin{pmatrix} 2 & 3 & 1 & -4 \\ 2 & 1 & 0 & 5 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 2 & 3 & 1 \\ 2 & 1 & 0 \end{pmatrix}.$$

The matrix A is a 2×4 matrix, and B is an 2×3 matrix. Since the matrices have not the same dimension, we have $A \neq B$.

(b) Are the matrices

$$A = (a_{i,j}) = \begin{pmatrix} 2 & 3 & 1 & -4 \\ 2 & 1 & 0 & 5 \end{pmatrix} \qquad \text{and} \qquad B = (b_{i,j}) = \begin{pmatrix} 2 & 1 & 0 & 5 \\ 2 & 3 & 1 & -4 \end{pmatrix}$$

equal? *Solution:* The matrices A and B are both 2×4 matrices, but we have $a_{1,2} = 3 \neq b_{1,2} = 1$. Hence the matrices are not equal, that is, $A \neq B$.

(c) The matrices

$$A = (a_{i,j}) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad \text{and} \qquad B = (b_{i,j}) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

are both 2×2 matrices. In addition,

$$a_{1,1} = b_{1,1} = 1, \qquad a_{1,2} = b_{1,2} = 2, \qquad a_{2,1} = b_{2,1} = 3, \qquad \text{and} \qquad a_{2,2} = b_{2,2} = 4.$$

Thus the two matrices are equal, that is, $A = B$. □

Definition 10.5 *A matrix which has the same number of rows and columns is called* ***square*** *or a* ***square matrix****. In formulas, an $m \times n$ matrix $A = (a_{i,j})$ is square if $m = n$.*

Example 10.6 (square matrices)
Consider the matrices

$$A = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 2 & 1 & 0 \\ 2 & 3 & 1 \end{pmatrix}.$$

The matrix A is a 2×2 matrix and is square. The matric B is a 2×3 matrix and is not square. □

Definition 10.7 (zero matrix)
The $m \times n$ matrix $\mathcal{O}$ which has all entries zero

$$\mathcal{O} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

is called the $m \times n$ ***zero matrix****.*

10.2 Elementary Matrix Operations

In this section we learn **how to multiply a matrix by a scalar** and how to **add** and **subtract matrices**. Finally, we learn how to **multiply an $m \times n$ matrix by an $n \times \ell$ matrix**.

Definition 10.8 (scalar multiplication of a matrix by a scalar)
Let $A = (a_{i,j})$ be an $m \times n$ matrix, and let $\lambda \in \mathbb{R}$ be a scalar (real number). Then the ***scalar multiplication*** *of A by λ is the $m \times n$ matrix*

$$\lambda\, A = \lambda\, (a_{i,j}) = (\lambda\, a_{i,j}) = \begin{pmatrix} \lambda\, a_{1,1} & \lambda\, a_{1,2} & \cdots & \lambda\, a_{1,n} \\ \lambda\, a_{2,1} & \lambda\, a_{2,2} & \cdots & \lambda\, a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda\, a_{m,1} & \lambda\, a_{m,2} & \cdots & \lambda\, a_{m,n} \end{pmatrix}.$$

In shorter notation, the (i,j)th entry of $\lambda\, A$ is

$$[\lambda\, A]_{i,j} = \lambda\, a_{i,j}.$$

In words, the matrix A is multiplied with the scalar λ by ***multiplying every entry of the matrix by*** *λ.*

Example 10.9 (multiplication of a matrix by a scalar)
Consider the matrix

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

(a) If $\lambda = 2$, then $\lambda\, A$ is given by

$$2\, A = \begin{pmatrix} 2 \times 3 & 2 \times 1 \\ 2 \times 2 & 2 \times 5 \end{pmatrix} = \begin{pmatrix} 6 & 2 \\ 4 & 10 \end{pmatrix}.$$

(b) If $\lambda = 1$, then $\lambda\, A$ is given by

$$1\, A = \begin{pmatrix} 1 \times 3 & 1 \times 1 \\ 1 \times 2 & 1 \times 5 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix} = A.$$

(c) If $\lambda = 0$, then $\lambda\, A$ is given by

$$0\, A = \begin{pmatrix} 0 \times 3 & 0 \times 1 \\ 0 \times 2 & 0 \times 5 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \mathcal{O}. \qquad \square$$

We can also **add** and **subtract** matrices A and B if they have the **same size**, that is, the same numbers of rows and columns.

Definition 10.10 (addition of matrices)
*Let $A = (a_{i,j})$ be an $m \times n$ matrix, and let $B = (b_{i,j})$ be also an $m \times n$ matrix. Then the **sum** $A + B$ is the $m \times n$ matrix defined by*

$$A + B = (a_{i,j} + b_{i,j}) = \begin{pmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} & \cdots & a_{1,n} + b_{1,n} \\ a_{2,1} + b_{2,1} & a_{2,2} + b_{2,2} & \cdots & a_{2,n} + b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} + b_{m,1} & a_{m,2} + b_{m,2} & \cdots & a_{m,n} + b_{m,n} \end{pmatrix}.$$

*In words, to add two $m \times n$ matrices A and B, we just **add the corresponding elements of the matrices**, that is,*

$$[A + B]_{i,j} = a_{i,j} + b_{i,j}.$$

*Note that we only can add matrices if they have the **same size**, that is, the **same number of rows** and the **same number of columns**.*

Remark 10.11 (substraction of matrices)
The ***difference*** $A - B$ *of two* $m \times n$ *matrices* $A = (a_{i,j})$ *and* $B = (b_{i,j})$ *can now easily be defined with the help of the scalar multiplication of a matrix with a scalar and the addition of two matrices by*

$$\boxed{A - B = A + (-1)\,B = (a_{i,j}) + (-b_{i,j}) = (a_{i,j} - b_{i,j}),}$$

or equivalently, the (i,j)*th entry of* $A - B$ *is*

$$\boxed{[A - B]_{i,j} = a_{i,j} - b_{i,j}.}$$

Example 10.12 (addition and subtraction of matrices)
(a) The sum $A + B$ of the two 2×3 matrices

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 2 & 1 & 0 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 2 & 2 & 6 \\ 3 & 1 & 7 \end{pmatrix},$$

is given by

$$A + B = \begin{pmatrix} 2 & 3 & 1 \\ 2 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 2 & 2 & 6 \\ 3 & 1 & 7 \end{pmatrix} = \begin{pmatrix} 2+2 & 3+2 & 1+6 \\ 2+3 & 1+1 & 0+7 \end{pmatrix} = \begin{pmatrix} 4 & 5 & 7 \\ 5 & 2 & 7 \end{pmatrix}.$$

(b) The difference $A - B$ of the two 2×2 matrices

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

is given by

$$A - B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1-1 & 2-1 \\ 3-1 & 4-1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix},$$

and the sum $A + B$ is given by

$$A + B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1+1 & 2+1 \\ 3+1 & 4+1 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix}.$$

(c) Let A be an arbitrary $m \times n$ matrix and let $\mathcal{O}$ be the $m \times n$ zero matrix. Then

$$A + \mathcal{O} = \mathcal{O} + A = A.$$

Indeed

$$[A+\mathcal{O}]_{i,j} = a_{i,j}+0 = [\mathcal{O}+A]_{i,j} = 0+a_{i,j} = a_{i,j} \qquad \text{for all } 1 \le i \le m \text{ and all } 1 \le j \le n.$$

(d) If A is a matrix and k is a positive integer, then

$$\underbrace{A + A + \ldots + A}_{k\text{-times}} = k\,A. \qquad \square$$

Next we learn that we can multiply an $m \times n$ matrix A with and $n \times \ell$ matrix B, and the result will be an $m \times \ell$ matrix. The **matrix multiplication** $A\,B$ is rather more complicated than scalar multiplication or matrix addition, since matrix multiplication is not performed by multiplying corresponding entries but by taking the **scalar products of rows the matrix A and columns of the matrix B**.

Definition 10.13 (multiplication of an $m \times n$ by an $n \times \ell$ matrix)
*Let $A = (a_{i,j})$ be an $m \times n$ matrix, and let $B = (b_{j,k})$ be an $n \times \ell$ matrix. Then the **product** $A\,B$ **of the** $m \times n$ **matrix** A **with the** $n \times \ell$ **matrix** B is defined to be the $m \times \ell$ matrix $A = ([A\,B]_{i,j})$ with the entries*

$$\begin{aligned}
(i,j)\text{th entry of } A\,B \;&=\; [A\,B]_{i,j} = \sum_{k=1}^{n} a_{i,k}\,b_{k,j} \\
&=\; \text{scalar product of } i\text{th row vector of } A \text{ and } j\text{th column vector of } B \\
&=\; \big(a_{i,1},\, a_{i,2},\, \ldots,\, a_{i,n}\big) \cdot \begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{n,j} \end{pmatrix}.
\end{aligned}$$

Note that the **product $A\,B$ only exists** if the **number of columns in the first matrix A equals the number of rows in the second matrix B**.

Note that if A and B are both $n \times n$ matrices, then both $A\,B$ and $B\,A$ exist but in general $A\,B$ and $B\,A$ are **not necessarily equal**.

Example 10.14 (matrix multiplication)
The matrices A and B, given by

$$A = \begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} -2 & 6 \\ 4 & 7 \end{pmatrix}.$$

They are both 2×2 matrices, and we can compute both $A\,B$ and $B\,A$. We find

$$A\,B \;=\; \begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} -2 & 6 \\ 4 & 7 \end{pmatrix}$$

$$= \begin{pmatrix} 3\times(-2)+1\times 4 & 3\times 6+1\times 7 \\ 2\times(-2)+5\times 4 & 2\times 6+5\times 7 \end{pmatrix}$$

$$= \begin{pmatrix} -2 & 25 \\ 16 & 47 \end{pmatrix},$$

and

$$\begin{aligned} B\,A &= \begin{pmatrix} -2 & 6 \\ 4 & 7 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix} \\ &= \begin{pmatrix} (-2)\times 3+6\times 2 & (-2)\times 1+6\times 5 \\ 4\times 3+7\times 2 & 4\times 1+7\times 5 \end{pmatrix} \\ &= \begin{pmatrix} 6 & 28 \\ 26 & 39 \end{pmatrix}. \end{aligned}$$

We see that $A\,B \neq B\,A$. □

Example 10.15 (matrix multiplication)
Consider the two matrices

$$A = (1 \;\; 2) \qquad \text{and} \qquad B = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Then $A\,B$ does **not** exist, since A is a 1×2 matrix and B is a 3×1 matrix. The product $B\,A$ does exist, since B is a 3×1 and A is a 1×2 matrix. We find

$$B\,A = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}(1 \;\; 2) = \begin{pmatrix} 1\times 1 & 1\times 2 \\ 2\times 1 & 2\times 2 \\ 3\times 1 & 3\times 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{pmatrix}.$$ □

Definition 10.16 (identity matrix)
*The $n\times n$ **identity matrix** I_n is the $n\times n$ matrix which has ones on the leading diagonal and zeros everywhere else, that is,*

$$[I_n]_{i,i} = 1 \quad \textit{for } 1\le i\le n \qquad \textit{and} \qquad [I_n]_{i,j} = 0 \quad \textit{if } i\neq j.$$

*Explicitly the **identity matrix** I_n is*

$$I_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Example 10.17 (identity matrix)
The 2×2 identity matrix and the 3×3 identity matrix are given by

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

respectively. □

Example 10.18 (multiplication with the identity matrix)
Let the 2×2 matrix A be given by

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

Compute the products $I_2\, A$ and $A\, I_2$, where I_2 is the 2×2 identity matrix.

Solution:

$$I_2\, A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} 1 \times 3 + 0 \times 2 & 1 \times 1 + 0 \times 5 \\ 0 \times 3 + 1 \times 2 & 0 \times 1 + 1 \times 5 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix},$$

and

$$I_2\, A = \begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 \times 1 + 1 \times 0 & 3 \times 0 + 1 \times 1 \\ 2 \times 1 + 5 \times 0 & 2 \times 0 + 5 \times 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix}.$$

We observe that $I_2\, A = A\, I_2 = A$. □

The identity $I_2\, A = A\, I_2 = A$ which we observed in the last example is true for every 2×2 matrix A. Moreover, an analogous property holds for the $n \times n$ identity matrix.

Lemma 10.19 (multiplication by the identity matrix)
Let A be an $m \times n$ matrix, let B be a $n \times \ell$ matrix, and let I_n be the $n \times n$ identity matrix. Then we have

$$A\, I_n = A \qquad \text{and} \qquad I_n\, B = B. \tag{10.1}$$

In particular, if A is an $n \times n$ square matrix, then (10.1) implies $A\, I_n = I_n\, A = A$.

We observe that I_n plays the role in the multiplication of $n \times n$ matrices that the real number 1 plays in the multiplication of real numbers, that is,

$$x\, 1 = 1\, x = x \qquad \text{for any real number } x \in \mathbb{R}.$$

Theorem 10.20 (associative law of matrix multiplication)
Let A be an $m \times n$ matrix, B an $n \times \ell$ matrix, and C an $\ell \times k$ matrix. Then

$$(A\,B)\,C = A\,(B\,C).$$

Example 10.21 (associative law of matrix multiplication)
Verify the associative law for the three matrices

$$A = \begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Solution: We have have

$$A\,B = \begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 3-1 & -3+1 \\ 2-5 & -2+5 \end{pmatrix} = \begin{pmatrix} 2 & -2 \\ -3 & 3 \end{pmatrix}$$

and

$$(A\,B)\,C = \begin{pmatrix} 2 & -2 \\ -3 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 2-4 & 4-2 \\ -3+6 & -6+3 \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ 3 & -3 \end{pmatrix}.$$

On the other hand

$$B\,C = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1-2 & 2-1 \\ -1+2 & -2+1 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

and

$$A\,(B\,C) = \begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} -3+1 & 3-1 \\ -2+5 & 2-5 \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ 3 & -3 \end{pmatrix},$$

and we find indeed $(A\,B)\,C = A\,(B\,C)$. □

Definition 10.22 (k-fold product of a matrix)
*For any integer $k \in \mathbb{N}$, we denote by A^k the k-**fold product of** A, that is,*

$$A^k = \underbrace{A\,A\,\ldots\,A}_{k\text{-times}}$$

Finally we introduce the transpose of a matrix.

Definition 10.23 (transpose of a matrix)
The ***transpose*** A^T *of an* $m \times n$ *matrix* A *is the* $n \times m$ *matrix* A^T *obtained by interchanging the rows and columns of* A*. In formulas,*

$$[A^T]_{i,j} = a_{j,i} \qquad \textit{for all } 1 \le i \le n \textit{ and } 1 \le j \le m. \tag{10.2}$$

Example 10.24 (transpose of a matrix)

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 2 & 1 & 0 \end{pmatrix} \qquad \text{and} \qquad A^T = \begin{pmatrix} 2 & 2 \\ 3 & 1 \\ 1 & 0 \end{pmatrix}.$$

Remark 10.25 (transpose of transpose gives original matrix)
For any matrix A*, we have*

$$\boxed{(A^T)^T = A.}$$

Proof: *Indeed, if* A *is an* $m \times n$ *matrix, then* A^T *is an* $n \times m$ *matrix, and so* $(A^T)^T$ *is an* $m \times n$ *matrix. Furthermore, we have from (10.2) that*

$$\left[(A^T)^T\right]_{i,j} = \left[A^T\right]_{j,i} = a_{i,j} \qquad \textit{for all } 1 \le i \le m \textit{ and all } 1 \le j \le n. \qquad \square$$

10.3 The Determinant of a Square Matrix

The determinant associates a **scalar** with a matrix and can only be defined for **square matrices**. We will start by discussing the determinant for the special cases of 2×2 matrices and 3×3 matrices. After that we will define the determinant recursively for $n \times n$ matrices. We will learn several **properties of determinants** that can be used to **simplify the computation of determinants**.

Definition 10.26 (determinant of 2×2 matrix)
The ***determinant*** $\det(A)$ *of a* 2×2 *matrix*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is defined by

$$\det(A) \;=\; \begin{vmatrix} a & b \\ c & d \end{vmatrix} \;=\; a\,d - b\,c. \tag{10.3}$$

Example 10.27 (determinants of 2×2 matrices)
Compute the determinants of the 2×2 matrices

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 2 & 5 \\ 4 & 10 \end{pmatrix}.$$

Solution: From (10.3), we have that

$$\det(A) = \begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix} = 1 \times 4 - 3 \times 2 = 4 - 6 = -2,$$

$$\det(B) = \begin{vmatrix} 2 & 5 \\ 4 & 10 \end{vmatrix} = 2 \times 10 - 5 \times 4 = 20 - 20 = 0. \qquad \square$$

Definition 10.28 (determinant of a 3×3 matrix)
*The **determinant** of a 3×3 matrix*

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}$$

is defined by

$$\det(A) = \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} = a_{1,1}\, C_{1,1} - a_{1,2}\, C_{1,2} + a_{1,3}\, C_{1,3}, \tag{10.4}$$

*where $C_{1,1}$, $-C_{1,2}$, and $C_{1,3}$ are the so-called **cofactors** of $a_{1,1}$, $a_{1,2}$, and $a_{1,3}$, respectively, and are defined by*

$$C_{1,1} = \begin{vmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix}, \qquad C_{1,2} = \begin{vmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{vmatrix}, \qquad \text{and} \qquad C_{1,3} = \begin{vmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{vmatrix}.$$

We observe that $C_{1,j}$ is the determinant of the 2×2 submatrix of A that is obtained by deleting the 1st row and jth column of A.

Example 10.29 (determinant of a 3×3 matrix)
Find the determinant of the matrix

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 2 & 0 & 0 \end{pmatrix}.$$

Solution: We apply formula (10.4) to compute the determinant:

$$\begin{aligned}
\det(A) &= \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 2 & 0 & 0 \end{vmatrix} = 1 \times \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} - 0 \times \begin{vmatrix} 0 & 0 \\ 2 & 0 \end{vmatrix} + 1 \times \begin{vmatrix} 0 & 1 \\ 2 & 0 \end{vmatrix} \\
&= 1 \times (1 \times 0 - 0 \times 0) - 0 + 1 \times (0 \times 0 - 1 \times 2) \\
&= 1 \times 0 + 1 \times (-2) = -2.
\end{aligned}$$

□

Example 10.30 (determinant of a 3×3 matrix)
Determine the determinant of the 3×3 matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 6 & 5 \\ 9 & 8 & 7 \end{pmatrix}.$$

Solution: We have from (10.4) that

$$\begin{aligned}
\det(A) &= \begin{vmatrix} 1 & 2 & 3 \\ 4 & 6 & 5 \\ 9 & 8 & 7 \end{vmatrix} = 1 \times \begin{vmatrix} 6 & 5 \\ 8 & 7 \end{vmatrix} - 2 \times \begin{vmatrix} 4 & 5 \\ 9 & 7 \end{vmatrix} + 3 \times \begin{vmatrix} 4 & 6 \\ 9 & 8 \end{vmatrix} \\
&= (6 \times 7 - 5 \times 8) - 2 \times (4 \times 7 - 5 \times 9) + 3 \times (4 \times 8 - 6 \times 9) \\
&= (42 - 40) - 2 \times (28 - 45) + 3 \times (32 - 54) \\
&= 2 - 2 \times (-17) + 3 \times (-22) = 2 + 34 - 66 = -30.
\end{aligned}$$

□

Remark 10.31 (use 3×3 matrix to determine the vector product)
In Remark 9.44 we mentioned the following formula for the ***vector product***

$$\begin{aligned}
\mathbf{a} \times \mathbf{b} &= (a_1, a_2, a_3) \times (b_1, b_2, b_3) \\
&= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \mathbf{k} \\
&= (a_2\, b_3 - a_3\, b_2)\, \mathbf{i} - (a_1\, b_3 - a_3\, b_1)\, \mathbf{j} + (a_1\, b_2 - a_2\, b_1)\, \mathbf{k},
\end{aligned}$$

which we can now understand with our knowledge of the determinant notation.

Now we learn the formula for the determinant of an $n \times n$ matrix. This formula is a generalization of the formula (10.4) for the determinant of a 3×3 matrix.

Definition 10.32 (determinant of an $n \times n$ matrix)
Let $A = (a_{i,j})$ be an $n \times n$ matrix whose entry in ith row and jth column is denoted by $a_{i,j}$. Then the ***determinant*** *of the $n \times n$ matrix A is given by*

$$\det(A) = \sum_{j=1}^{n} a_{i,j}\,(-1)^{i+j}\,C_{i,j}, \tag{10.5}$$

for any $i \in \{1, 2, \ldots, n\}$, where $C_{i,j}$ is the determinant of the $(n-1) \times (n-1)$ matrix obtained by deleting the ith row and jth column from A. The term $(-1)^{i+j}\,C_{i,j}$ is called the ***cofactor*** *of $a_{i,j}$.*

Remark 10.33 (comments on the determinant)
(a) The notion 'for any $i \in \{1, 2, \ldots, n\}$' in Definition 10.32 implies that for any choice of i the value of the determinant $\det(A)$ is the same.

*(b) In the above definition, we have '**expanded about the ith row**'. We can also '**expand about the jth column**'. Indeed, the determinant of A is also given by*

$$\boxed{\det(A) = \sum_{i=1}^{n} a_{i,j}\,(-1)^{i+j}\,C_{i,j},} \tag{10.6}$$

for any $j \in \{1, 2, \ldots, n\}$.

(c) For a 3×3 matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix},$$

expanding about the first row *(that is, taking $i = 1$ in Definition 10.32) gives that the determinant of A is*

$$\begin{aligned} \det(A) &= \sum_{j=1}^{3} a_{1,j}\,(-1)^{1+j}\,C_{1,j} \\ &= a_{1,1}\,(-1)^{1+1}\,C_{1,1} + a_{1,2}\,(-1)^{1+2}\,C_{1,2} + a_{1,3}\,(-1)^{1+3}\,C_{1,3} \\ &= a_{1,1}\,C_{1,1} - a_{1,2}\,C_{1,2} + a_{1,3}\,C_{1,3}, \end{aligned}$$

where

$$C_{1,1} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \qquad C_{1,2} = \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}, \qquad \text{and} \qquad C_{1,3} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

This is the formula (10.4) given in Definition 10.28, and we see that Definition 10.28 is just a special case of Definition 10.32.

We now come back to the matrix from Example 10.30, and we compute its determinant in two different ways.

Example 10.34 (determinant of 3×3 matrix)
Consider again the matrix A from Example 10.30, given by

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 6 & 5 \\ 9 & 8 & 7 \end{pmatrix}.$$

(a) Expanding about the second row (that is, taking (10.5) with $i = 2$) gives that the determinant of A is

$$\begin{aligned} \det(A) &= \sum_{j=1}^{3} a_{2,j}\,(-1)^{2+j}\,C_{2,j} \\ &= a_{2,1}\,(-1)^{2+1}\,C_{2,1} + a_{2,2}\,(-1)^{2+2}\,C_{2,2} + a_{2,3}\,(-1)^{2+3}\,C_{2,3} \\ &= -a_{2,1}\,C_{2,1} + a_{2,2}\,C_{2\,2} - a_{2,3}\,C_{2,3} \\ &= -4\,C_{2,1} + 6\,C_{2,2} - 5\,C_{2,3}, \end{aligned}$$

where

$$C_{2,1} = \begin{vmatrix} 2 & 3 \\ 8 & 7 \end{vmatrix} = 14 - 24 = -10,$$

$$C_{2,2} = \begin{vmatrix} 1 & 3 \\ 9 & 7 \end{vmatrix} = 7 - 27 = -20,$$

$$C_{2,3} = \begin{vmatrix} 1 & 2 \\ 9 & 8 \end{vmatrix} = 8 - 18 = -10.$$

Hence

$$\begin{aligned} \det(A) &= -4\,C_{2,1} + 6\,C_{2,2} - 5\,C_{2,3} \\ &= -4 \times (-10) + 6 \times (-20) - 5 \times (-10) \\ &= 40 - 120 + 50 \;=\; -30, \end{aligned}$$

agreeing with the answer we obtained previously in Example 10.30.

(b) Expanding about the first column (that is, taking (10.6) with $j = 1$) gives that the determinant of A is

$$\det(A) = \sum_{i=1}^{3} a_{i,1}\,(-1)^{i+1}\,C_{i,1}$$

$$
\begin{aligned}
&= a_{1,1}\,(-1)^{1+1}\,C_{1,1} + a_{2,1}\,(-1)^{2+1}\,C_{2,1} + a_{3,1}\,(-1)^{3+1}\,C_{3,1} \\
&= a_{1,1}\,C_{11} - a_{2,1}\,C_{2,1} + a_{3,1}\,C_{3,1} \\
&= C_{1,1} - 4\,C_{2,1} + 9\,C_{3,1},
\end{aligned}
$$

where

$$C_{1,1} = \begin{vmatrix} 6 & 5 \\ 8 & 7 \end{vmatrix} = 42 - 40 = 2,$$

$$C_{2,1} = \begin{vmatrix} 2 & 3 \\ 8 & 7 \end{vmatrix} = 14 - 24 = -10,$$

$$C_{3,1} = \begin{vmatrix} 2 & 3 \\ 6 & 5 \end{vmatrix} = 10 - 18 = -8.$$

Hence

$$\det(A) = C_{1,1} - 4\,C_{2,1} + 9\,C_{3,1} = 2 - 4 \times (-10) + 9 \times (-8) = 2 + 40 - 72 = -30,$$

again agreeing with the answer we obtained previously in Example 10.30. □

In the next lemma we get information about how the determinant changes if we perform certain operations on the matrix, such as interchanging rows (or columns) of the matrix, multiplying a row (or column) with a scalar, or even adding a multiple of one row (or column) to another row (or column). Such operations can be used to modify the the matrix (whose determinant we want to compute) and thus may allow a much easier computation of the determinant of the modified matrix.

Lemma 10.35 (properties of the determinant)
We can use the following ***properties of the determinant*** *when finding the determinant of an $n \times n$ matrix:*

(i) ***multiplying all the elements of a single row or column*** *of a matrix* ***by a real number*** *λ results in the determinant of the matrix being multiplied by λ;*

(ii) ***interchanging any two rows or any two columns*** *of a matrix changes the sign of the determinant of the matrix;*

(iii) ***adding a multiple of one row (or column)*** *of a matrix* ***to another row (or column)*** *does not change the determinant of the matrix;*

(iv) if any ***row or column*** *of a matrix* ***consists entirely of zeros****, then the determinant of the matrix is zero;*

(v) if any ***two rows*** *or any* ***two columns*** *of a matrix* ***are the same****, then the determinant of the matrix is zero.*

Example 10.36 (determinant of a 4×4 matrix)
Find the determinant of the matrix

$$A = \begin{pmatrix} 4 & 3 & 0 & 1 \\ 9 & 7 & 2 & 3 \\ 4 & 0 & 2 & 1 \\ 3 & -1 & 4 & 0 \end{pmatrix}.$$

Solution: Rule (i) allows us to take out the common factor 2 from column 3:

$$\det(A) = 2 \times \begin{vmatrix} 4 & 3 & 0 & 1 \\ 9 & 7 & 1 & 3 \\ 4 & 0 & 1 & 1 \\ 3 & -1 & 2 & 0 \end{vmatrix}.$$

Next, we use rule (ii) to interchange columns 1 and 4 to make the entry in the top left corner 1:

$$\det(A) = 2 \times (-1) \times \begin{vmatrix} 1 & 3 & 0 & 4 \\ 3 & 7 & 1 & 9 \\ 1 & 0 & 1 & 4 \\ 0 & -1 & 2 & 3 \end{vmatrix}.$$

We now apply rule (iii) twice. Firstly, we subtract three times row 1 from row 2 (that is, $r_2^{\text{new}} = r_2 - 3\,r_1$):

$$\det(A) = -2 \times \begin{vmatrix} 1 & 3 & 0 & 4 \\ 0 & -2 & 1 & -3 \\ 1 & 0 & 1 & 4 \\ 0 & -1 & 2 & 3 \end{vmatrix}.$$

Now, we subtract row 1 from row 3 (that is, $r_3^{\text{new}} = r_3 - r_1$):

$$\det(A) = -2 \times \begin{vmatrix} 1 & 3 & 0 & 4 \\ 0 & -2 & 1 & -3 \\ 0 & -3 & 1 & 0 \\ 0 & -1 & 2 & 3 \end{vmatrix}.$$

Now, we have only one non-zero entry in the column 1, namely, the number 1 at the top. Expanding about column 1 gives that

$$\det(A) = -2 \times 1 \times \begin{vmatrix} -2 & 1 & -3 \\ -3 & 1 & 0 \\ -1 & 2 & 3 \end{vmatrix}.$$

So we have reduced the problem to that of finding the determinant of a 3×3 matrix.

We now apply the same procedure to find this determinant. Indeed, rule (i) allows us to take out the common factors -1 and 3 from columns 1 and 3, respectively:

$$\det(A) = -2 \times (-1) \times 3 \times \begin{vmatrix} 2 & 1 & -1 \\ 3 & 1 & 0 \\ 1 & 2 & 1 \end{vmatrix} = 6 \times \begin{vmatrix} 2 & 1 & -1 \\ 3 & 1 & 0 \\ 1 & 2 & 1 \end{vmatrix}.$$

Next, we use rule (ii) to interchange columns 1 and 2 to make the entry in the top left corner 1:

$$\det(A) = 6 \times (-1) \times \begin{vmatrix} 1 & 2 & -1 \\ 1 & 3 & 0 \\ 2 & 1 & 1 \end{vmatrix}.$$

We now apply rule (iii) twice. Subtracting row 1 from row 2 (that is, $r_2^{\text{new}} = r_2 - r_1$) and subtracting twice row 1 from row 3 (that is, $r_3^{\text{new}} = r_3 - 2r_1$) gives that

$$\det(A) = -6 \times \begin{vmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 0 & -3 & 3 \end{vmatrix}.$$

Now, we have only one non-zero entry in the column 1, namely, the number 1 at the top. Expanding about column 1 gives that

$$\det(A) = -6 \times \begin{vmatrix} 1 & 1 \\ -3 & 3 \end{vmatrix}.$$

Rule (i) allows us to take out the common factor 3 from row 2:

$$\det(A) = -18 \times \begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix}.$$

Hence

$$\det(A) = -18\times[1\times 1 - 1\times(-1)] = -18\times[1+1] = -18\times 2 = -36. \qquad \square$$

Example 10.37 (determinant of the identity matrix)
The $n \times n$ identity matrix I_n has determinant

$$\boxed{\det(I_n) = 1.} \tag{10.7}$$

Proof: Let us start with the 2×2 matrix I_2. We have

$$\det(I_2) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \times 1 - 0 \times 0 = 1. \tag{10.8}$$

For the 3×3 identity matrix I_3, we find by expanding the determinant $\det(I_3)$ about the first row

$$\det(I_3) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1 \times \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \times \det(I_2) = 1 \times 1 = 1,$$

where we have used (10.8) in the second last step.

Now we use induction and assume that we have already verified that $\det(I_{n-1}) = 1$. Then we want to show that $\det(I_n) = 1$. By expanding the determinant $\det(I_n)$ of the $n \times n$ identity matrix I_n about the first row, we find (in analogy to above)

$$\det(I_n) = 1 \times \det(I_{n-1}) = 1 \times 1 = 1,$$

where have used the assumption that $\det(I_{n-1}) = 1$ in the second last step. □

The next lemma gives information about the relation between the determinant $\det(A\,B)$ of the product $A\,B$ of two square matrices A and B and the determinants of the individual matrices $\det(A)$ and $\det(B)$.

Lemma 10.38 (determinant of the product of matrices)
If A and B are $n \times n$ matrices, then the $n \times n$ matrix $A\,B$ has the determinant

$$\det(A\,B) = \det(A) \times \det(B) = \det(A)\,\det(B). \tag{10.9}$$

Note that, in general, for two $n \times n$ matrices

$$\det(A + B) \neq \det(A) + \det(B).$$

Example 10.39 (determinant of product of matrices)
Determine the determinant $\det(A\,B)$ where

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 3 & 3 & 3 \end{pmatrix}.$$

Solution: We will determine the determinant in two ways: (a) by using (10.9); and (b) by computing $A\,B$ first and then computing the determinant of $A\,B$.

(a) We have, from expanding $\det(A)$ about the first column and from expanding $\det(B)$ about the first row,

$$\det(A) = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} = 1 \times \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1 \times (1 \times 1 - 1 \times 0) = 1,$$

$$\det(B) = \begin{vmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 3 & 3 & 3 \end{vmatrix} = 1 \times \begin{vmatrix} 2 & 0 \\ 3 & 3 \end{vmatrix} = 1 \times (2 \times 3 - 0 \times 3) = 6.$$

From (10.9), we find

$$\det(A\,B) = \det(A) \times \det(B) = 1 \times 6 = 6.$$

(b) Computing the matrix $A\,B$ yields

$$A\,B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 3 & 3 & 3 \end{pmatrix} = \begin{pmatrix} 6 & 5 & 3 \\ 5 & 5 & 3 \\ 3 & 3 & 3 \end{pmatrix}.$$

To compute the determinant $\det(A\,B)$, we use that, from Lemma 10.35 (iii), we may subtract the second row of the matrix from the first row without changing the value of the determinant (that is, $r_1^{\text{new}} = r_1 - r_2$). Thus

$$\det(A\,B) = \begin{vmatrix} 6 & 5 & 3 \\ 5 & 5 & 3 \\ 3 & 3 & 3 \end{vmatrix} = \begin{vmatrix} 6-5 & 5-5 & 3-3 \\ 5 & 5 & 3 \\ 3 & 3 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 5 & 5 & 3 \\ 3 & 3 & 3 \end{vmatrix}.$$

Now we expand about the first row and get

$$\det(A\,B) = \begin{vmatrix} 1 & 0 & 0 \\ 5 & 5 & 3 \\ 3 & 3 & 3 \end{vmatrix} = 1 \times \begin{vmatrix} 5 & 3 \\ 3 & 3 \end{vmatrix} = 1 \times (5 \times 3 - 3 \times 3) = 15 - 9 = 6.$$

We get indeed the same result. □

10.4 Invertible Matrices

In this section, we introduce the **inverse matrix** of an (invertible) square $n \times n$ matrix A. The inverse matrix A^{-1} of A satisfies $A\,A^{-1} = A^{-1}\,A = I_n$. If we compare

invertible matrices A to real numbers $x \neq 0$, then the inverse matrix A^{-1} corresponds to the number $1/x$. We learn that a **matrix A has an inverse matrix** if and only if $\det(A) \neq 0$. We introduce the **adjoint matrix**, and we learn a **formula for the inverse matrix** which involves the adjoint matrix and the determinant.

Definition 10.40 (invertible and inverse matrix)
Let A be an $n \times n$ matrix. If there exists an $n \times n$ matrix B such that

$$A\,B = I_n \qquad \text{and} \qquad B\,A = I_n, \tag{10.10}$$

*then A is **invertible**. The matrix B satisfying (10.10) is called the **inverse (matrix) of** A, and usually the inverse matrix of A is denoted by A^{-1}.*

In Definition 10.40 we would prefer that, if there exists a matrix B satisfying (10.10), then this matrix is the **only matrix satisfying (10.10)**. In other words, we would like the inverse matrix to be **unique**. To show that this is indeed true, assume that there exits another $n \times n$ matrix C satisfying

$$A\,C = I_n \qquad \text{and} \qquad C\,A = I_n. \tag{10.11}$$

Multiplying the first formula in (10.11) from the left by the matrix B yields

$$B\,(A\,C) = B\,I_n \quad \Rightarrow \quad (B\,A)\,C = B \quad \Rightarrow \quad I_n\,C = B \quad \Rightarrow \quad C = B,$$

where we have used the associative law of matrix multiplication (see Theorem 10.20) and (10.1). Since $C = B$, we see that the inverse matrix is indeed uniquely determined.

Example 10.41 (matrix and its inverse)
Consider the matrices A and B, given by

$$A = \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} -4 & 3 \\ 3 & -2 \end{pmatrix}.$$

Show that the matrix B is the inverse matrix to the matrix A.

Solution: We compute $A\,B$ and $B\,A$, and find

$$\begin{aligned} A\,B &= \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} -4 & 3 \\ 3 & -2 \end{pmatrix} \\ &= \begin{pmatrix} 2 \times (-4) + 3 \times 3 & 2 \times 3 + 3 \times (-2) \\ 3 \times (-4) + 4 \times 3 & 3 \times 3 + 4 \times (-2) \end{pmatrix} \end{aligned}$$

$$
\begin{aligned}
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \\
B\,A &= \begin{pmatrix} -4 & 3 \\ 3 & -2 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix} \\
&= \begin{pmatrix} (-4) \times 2 + 3 \times 3 & (-4) \times 3 + 3 \times 4 \\ 3 \times 2 + (-2) \times 3 & 3 \times 3 + (-2) \times 4 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.
\end{aligned}
$$

Since $A\,B = B\,A = I_2$, we know that B is the inverse matrix to A. We also see that A is the inverse matrix of B. □

Remark 10.42 (not every matrix is invertible)
It is easy to see that the zero matrix

$$\mathcal{O} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

does not have an inverse and is thus not invertible. Indeed, if B is any other 2×2 matrix, then

$$\mathcal{O}\,B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \textit{and} \qquad B\,\mathcal{O} = B \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

We see that ***not every square matrix is invertible****.*

If an $n \times n$ matrix A has an inverse A^{-1}, then taking the determinant on both sides of the equation

$$I_n = A^{-1}\,A$$

and using (10.9) yields

$$\det(I_n) = \det(A^{-1}\,A) = \det(A^{-1})\,\det(A).$$

From (10.7), we know that $\det(I_n) = 1$, and thus

$$1 = \det(A^{-1})\,\det(A). \tag{10.12}$$

From (10.12), we know that **neither** $\det(A)$ **nor** $\det(A^{-1})$ **can be zero**, since zero multiplied with any other numbers is always zero. Thus we know that **any matrix**

which has an inverse matrix satisfies $\det(A) \neq 0$. If A has in inverse matrix then we have from (10.12) that

$$\boxed{\det(A^{-1}) = \frac{1}{\det(A)}.} \tag{10.13}$$

In fact, for any square matrix A, the condition $\det(A) \neq 0$ **guarantees that** A **has an inverse matrix**.

Theorem 10.43 (criterion for having an inverse matrix)
*A square matrix A is **invertible**, and thus has an inverse matrix, **if and only if its determinant** $\det(A)$ **is non-zero**, that is,*

$$\det(A) \neq 0.$$

Now that we have a criterion for checking whether a matrix is invertible, that is, has an inverse, it remains to learn how to **compute the inverse matrix**. This leads to the definition and the theorem below.

Definition 10.44 (adjoint matrix)
*The **adjoint** of an $n \times n$ square matrix A is that $n \times n$ square matrix $\operatorname{adj}(A)$ whose **entry in the ith row and jth column** is the **cofactor** $(-1)^{j+i}\, C_{j,i}$ of $a_{j,i}$ (note the changing order of the subscripts!). In formulas,*

$$[\operatorname{adj}(A)]_{i,j} = (-1)^{j+i}\, C_{j,i} \qquad \text{for } 1 \leq i,\, j \leq n.$$

With the help of the adjoint matrix, we can now easily give a formula for the inverse matrix of a square matrix.

Theorem 10.45 (formula for inverse matrix)
*Suppose that A is an $n \times n$ square matrix with $\det(A) \neq 0$. Then A is invertible and the **inverse matrix** A^{-1} of A is given by*

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A).$$

Example 10.46 (special case: inverse of a 2×2 matrix)
Suppose that we have a 2×2 matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Then $\det(A) = a\,d - b\,c$. The matrix of cofactors of A is

$$\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}.$$

So the adjoint of A is

$$\operatorname{adj}(A) = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}^T = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

So if $a\,d \neq b\,c$, then the inverse matrix A^{-1} of A is

$$\boxed{\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A) = \frac{1}{a\,d - b\,c} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.} \tag{10.14}$$

The formula (10.14) for the inverse of a 2×2 matrix with $\det(A) = a\,d - b\,c \neq 0$ is easy to remember and worth knowing from memory. □

Example 10.47 (inverse of a 2×2 matrix)
We use the formula (10.14) to determine the inverse of the matrix

$$A = \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix}$$

from Example 10.41. The determinant is $\det(A) = 2 \times 4 - 3 \times 3 = 8 - 9 = -1$. Thus from (10.14)

$$A^{-1} = \frac{1}{(-1)} \begin{pmatrix} 4 & -3 \\ -3 & 2 \end{pmatrix} = \begin{pmatrix} -4 & 3 \\ 3 & -2 \end{pmatrix},$$

and we have the matrix B from Example 10.41. □

Example 10.48 (inverse of a 2×2 matrix)
Determine whether the 2×2

$$A = \begin{pmatrix} 3 & 2 \\ 5 & 3 \end{pmatrix}$$

is invertible, and if yes find its inverse matrix A^{-1}.

Solution: Since $\det(A) = 3 \times 3 - 2 \times 5 = 9 - 10 = -1 \neq 0$, we know that the matrix A is invertible. From (10.14), its inverse is given by

$$A^{-1} = \frac{1}{(-1)} \begin{pmatrix} 3 & -2 \\ -5 & 3 \end{pmatrix} = \begin{pmatrix} -3 & 2 \\ 5 & -3 \end{pmatrix}.$$ □

Example 10.49 Determine whether the matrix

$$A = \begin{pmatrix} 2 & 1 & -3 \\ 3 & 1 & 0 \\ 1 & 2 & 3 \end{pmatrix}.$$

is invertible, and if yes, find its inverse matrix.

Solution: From the computations in Example 10.36, we can deduce that this matrix has the determinant

$$\det(A) = -18.$$

Thus A is invertible. Furthermore, the matrix $(C_{i,j})$ of cofactors of A is

$$\begin{aligned}(C_{i,j}) &= \begin{pmatrix} 1\times 3 - 0\times 2 & -(3\times 3 - 0\times 1) & 3\times 2 - 1\times 1 \\ -(1\times 3 - (-3)\times 2) & 2\times 3 - (-3)\times 1 & -(2\times 2 - 1\times 1) \\ 1\times 0 - (-3)\times 1 & -(2\times 0 - (-3)\times 3) & 2\times 1 - 1\times 3 \end{pmatrix} \\ &= \begin{pmatrix} 3 & -9 & 5 \\ -9 & 9 & -3 \\ 3 & -9 & -1 \end{pmatrix}.\end{aligned}$$

Hence

$$\operatorname{adj}(A) = \begin{pmatrix} 3 & -9 & 5 \\ -9 & 9 & -3 \\ 3 & -9 & -1 \end{pmatrix}^T = \begin{pmatrix} 3 & -9 & 3 \\ -9 & 9 & -9 \\ 5 & -3 & -1 \end{pmatrix}.$$

It follows that

$$A^{-1} = \frac{1}{\det(A)}\operatorname{adj}(A) = -\frac{1}{18}\begin{pmatrix} 3 & -9 & 3 \\ -9 & 9 & -9 \\ 5 & -3 & -1 \end{pmatrix} = \frac{1}{18}\begin{pmatrix} -3 & 9 & -3 \\ 9 & -9 & 9 \\ -5 & 3 & 1 \end{pmatrix}. \qquad \square$$

10.5 Solving Linear Systems of Equations

Finally we show how we can use inverse matrices to solve **linear systems of equations**. We will explain the idea first and then we apply it in two examples.

Let us start by looking at an example of a linear system of 2 equations in 2 unknowns.

$$\begin{aligned} 1\,x + 2\,y &= 5, \\ 3\,x + 4\,y &= 6. \end{aligned} \tag{10.15}$$

We can rewrite this linear system in the form

$$A\,\mathbf{x} = \mathbf{b}$$

with the 2×2 matrix

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

and the vectors

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} \qquad \text{and} \qquad \mathbf{b} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}.$$

That is, (10.15) is equivalent to the linear system in matrix form

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}.$$

After this motivation we will think from now on of **square linear systems** with n equations and n unknowns as an expression of the form

$$\boxed{A\,\mathbf{x} = \mathbf{b},} \tag{10.16}$$

where A is an $n \times n$ matrix, and $\mathbf{b}$ is a given $n \times 1$ matrix (an n-vector in $\mathbb{R}^n$) and $\mathbf{x}$ is an $n \times 1$ matrix (an n-vector in $\mathbb{R}^n$) that contains the unknowns. Without further information about A and $\mathbf{b}$, we **cannot know** whether an $\mathbf{x}$ exists that satisfies (10.16) and if such an $\mathbf{x}$ **exists** whether it is **unique**.

However, **if** $\det(A) \neq 0$, then there exists a **uniquely determined** n**-vector** $\mathbf{x}$ that satisfies (10.16). 'Uniquely determined' means that $\mathbf{x}$ is the only vector that satisfies (10.16). Now we explain how we can use the inverse matrix A^{-1} of A to solve (10.16) if $\det(A) \neq 0$.

If $\det(A) \neq 0$, then A is invertible and has an inverse matrix A^{-1} (see Theorem 10.43). We multiply (10.16) from the left with the inverse matrix A^{-1} and obtain

$$A^{-1}\,(A\,\mathbf{x}) = A^{-1}\,\mathbf{b}. \tag{10.17}$$

Using the associative law of matrix multiplication (see Theorem 10.20) and $A^{-1}\,A = I_n$ and $I_n\,\mathbf{x} = \mathbf{x}$ (from (10.1)), we then transform the left-hand side

$$A^{-1}\,(A\,\mathbf{x}) = (A^{-1}\,A)\,\mathbf{x} = I_n\,\mathbf{x} = \mathbf{x}. \tag{10.18}$$

Combing (10.17) and (10.18) yields that the **solution** $\mathbf{x}$ **of (10.16)** is given by

$$\boxed{\mathbf{x} = A^{-1}\,\mathbf{b}.} \tag{10.19}$$

If it is not too demanding to determine the inverse matrix A^{-1}, then (10.19) provides an easy way for computing the solution $\mathbf{x}$ of (10.16).

We illustrate this for two examples.

Example 10.50 (solve linear system of equations)
Solve the linear system of equations

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

Solution: We start by writing the linear system of equations in matrix form as

$$\begin{pmatrix} 3 & 2 \\ 5 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ 7 \end{pmatrix}.$$

Hence $A\,\mathbf{x} = \mathbf{b}$, with

$$A = \begin{pmatrix} 3 & 2 \\ 5 & 3 \end{pmatrix}, \qquad \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} \qquad \text{and} \qquad \mathbf{b} = \begin{pmatrix} 4 \\ 7 \end{pmatrix}.$$

Since $\det(A) = 3 \times 3 - 2 \times 5 = 9 - 10 = -1 \neq 0$, the matrix A is invertible. Its inverse matrix A^{-1} exists and is given by (use (10.14))

$$A^{-1} = \frac{1}{(-1)} \begin{pmatrix} 3 & -2 \\ -5 & 3 \end{pmatrix} = \begin{pmatrix} -3 & 2 \\ 5 & -3 \end{pmatrix}.$$

Thus we know from (10.19) that $\mathbf{x}$ is given by $\mathbf{x} = A^{-1}\,\mathbf{b}$, that is,

$$\begin{pmatrix} x \\ y \end{pmatrix} = A^{-1}\,\mathbf{b} = \begin{pmatrix} -3 & 2 \\ 5 & -3 \end{pmatrix} \begin{pmatrix} 4 \\ 7 \end{pmatrix} = \begin{pmatrix} (-3) \times 4 + 2 \times 7 \\ 5 \times 4 + (-3) \times 7 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Thus $x = 2$ and $y = -1$. Indeed, testing of the solution shows

$$3\,x + 2\,y = 3 \times 2 + 2 \times (-1) = 6 - 2 = 4 \quad \text{and} \quad 5\,x + 3\,y = 5 \times 2 + 3 \times (-1) = 10 - 3 = 7,$$

as required. □

Example 10.51 (solve linear system of equations)
Solve the linear system of equations

$$\begin{aligned} 2\,x + y - 3\,z &= 3, \\ 3\,x + y \phantom{{}- 3\,z} &= 0, \\ x + 2\,y + 3\,z &= -3. \end{aligned}$$

Solution: We start by writing the linear system of equations in matrix form

$$\begin{pmatrix} 2 & 1 & -3 \\ 3 & 1 & 0 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ -3 \end{pmatrix}.$$

Hence $A\,\mathbf{x} = \mathbf{b}$, with

$$A = \begin{pmatrix} 2 & 1 & -3 \\ 3 & 1 & 0 \\ 1 & 2 & 3 \end{pmatrix}, \qquad \mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \qquad \text{and} \qquad \mathbf{b} = \begin{pmatrix} 3 \\ 0 \\ -3 \end{pmatrix}.$$

From Example 10.49, A^{-1} exists and is given by

$$A^{-1} = -\frac{1}{18} \begin{pmatrix} 3 & -9 & 3 \\ -9 & 9 & -9 \\ 5 & -3 & -1 \end{pmatrix}.$$

So from (10.19), we have

$$\begin{aligned} \mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= A^{-1}\,\mathbf{b} = -\frac{1}{18} \begin{pmatrix} 3 & -9 & 3 \\ -9 & 9 & -9 \\ 5 & -3 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \\ -3 \end{pmatrix} \\ &= -\frac{1}{18} \begin{pmatrix} 3 \times 3 + (-9) \times 0 + 3 \times (-3) \\ (-9) \times 3 + 9 \times 0 + (-9) \times (-3) \\ 5 \times 3 + (-3) \times 0 + (-1) \times (-3) \end{pmatrix} \\ &= -\frac{1}{18} \begin{pmatrix} 0 \\ 0 \\ 18 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \end{aligned}$$

that is, $x = 0$, $y = 0$ and $z = -1$. Testing of the solution by substituting back into the equations shows that indeed

$$\begin{aligned} &2\,x + y - 3\,z = 2 \times 0 + 0 - 3 \times (-1) = 3, \\ &3\,x + y = 3 \times 0 + 0 = 0, \\ &x + 2\,y + 3\,z = 0 + 2 \times 0 + 3 \times (-1) = -3, \end{aligned}$$

as required. □