Reader

Mathematical Analysis

Course: Mathematical Analysis and Probability Theory

Henk Norde

0 PRELIMINARIES

0.1 Number sets

The sequence 1, 2, 3, 4, 5, ... is the sequence of positive integer numbers. Mathematicians call these numbers *natural numbers*. The set of natural numbers is denoted by $\mathbb{N}$, so

$$\mathbb{N} = \{1, 2, 3, 4, 5, \ldots\}.$$

If you extend the set of natural numbers with the number 0 and with the negative integer numbers like -1, -12 and -2018 you get the set of *integer numbers*, which is denoted by $\mathbb{Z}$. So

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

These numbers do not cover of course all numbers you know. Missing numbers are for example fractions like $\frac{3}{2}(= 1\frac{1}{2})$, $-\frac{12}{7}$, $4\frac{1}{12}$, etcetera. These fractions are called *rational numbers*. The set of all rational numbers is denoted by $\mathbb{Q}$. Rational numbers are numbers obtained by dividing an integer number (positive, zero or negative) by a natural number. We therefore write

$$\mathbb{Q} = \{\frac{m}{n} \ : \ m \in \mathbb{Z}, n \in \mathbb{N}\}.$$

Note that the way we denoted the set $\mathbb{Q}$ is different from the way we denoted the sets $\mathbb{N}$ and $\mathbb{Z}$. For $\mathbb{N}$ and $\mathbb{Z}$ we simply wrote down a few elements, which suggested a certain regularity, and the dots indicated that the reader is supposed to be able to imagine what happens on the position of these dots. For $\mathbb{Q}$ this is not possible and we use a formula that generates (all elements of) the set. Read the line above as: $\mathbb{Q}$ consists of elements of the form $\frac{m}{n}$, where m represents an integer number ($m \in \mathbb{Z}$) and n a natural number ($n \in \mathbb{N}$).

Still, with $\mathbb{Q}$ we do not have all numbers. Missing numbers are for example $\sqrt{2}$, $\pi \approx 3.14$ and $e \approx 2.71$. Why is $\sqrt{2}$ not an element of $\mathbb{Q}$, by the way? Let us present this as the first mathematical theorem in this reader. Of course mathematical theorems need a solid mathematical proof.

THEOREM 0.1 $\sqrt{2} \notin \mathbb{Q}$.

($x \in A$ is the mathematical shorthand notation for "x is an element of the set A", whereas $x \notin A$ denotes "x is not an element of the set A".)

Proof. Suppose that $\sqrt{2} \in \mathbb{Q}$. Then there exist natural numbers m and n such that $\sqrt{2} = \frac{m}{n}$. Of course, we can assume that m and n are not both even (that is, they do not both belong to the set $\{2, 4, 6, 8, \ldots\} = \{2k : k \in \mathbb{N}\}$; why can we make this assumption?). So, we assume this. Now $2 = \frac{m^2}{n^2}$, so $m^2 = 2n^2$. So, m^2 is even, but then m has to be even as well (why?). So, we can write $m = 2\ell$ for some number $\ell \in \mathbb{N}$. But now $m^2 = (2\ell)(2\ell) = 4\ell^2$ and $m^2 = 2n^2$ so $n^2 = 2\ell^2$. So, n^2 is even and therefore n is even as well. So, m and n are both even, which contradicts the assumption that m and n are not both even.
(This proof is called a "proof by contradiction". We have shown that $\sqrt{2} \notin \mathbb{Q}$ is true by showing that the assumption that $\sqrt{2} \in \mathbb{Q}$ leads to a contradiction.) $\square$

EXERCISE 0.1 Show that $\sqrt[3]{3} \notin \mathbb{Q}$.

Adding numbers like $\sqrt{2}$, π and e to the set $\mathbb{Q}$ we get the set of *real numbers*. We denote this set as $\mathbb{R}$. Note that $\mathbb{Z}$ contains more numbers than $\mathbb{N}$, or, stated differently, that $\mathbb{N}$ is a subset of $\mathbb{Z}$. Moreover $\mathbb{Z}$ is a subset of $\mathbb{Q}$ and $\mathbb{Q}$ a subset of $\mathbb{R}$. We denote this as:

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

The symbol $\subset$ is a mathematical notation for "is a subset of".

Intervals form a frequently occurring type of subsets of the real numbers. The notation and the terminology for intervals are as follows (here a and b represent real numbers with $a < b$):

$$\begin{array}{lcll}
(a,b) & = & \{x \in \mathbb{R} : a < x < b\} & \text{(an open interval)}, \\
[a,b] & = & \{x \in \mathbb{R} : a \leq x \leq b\} & \text{(a closed interval)}, \\
(a,b] & = & \{x \in \mathbb{R} : a < x \leq b\} & \text{(a half open interval)}, \\
[a,b) & = & \{x \in \mathbb{R} : a \leq x < b\} & \text{(a half open interval)}.
\end{array}$$

The length of all these (bounded) intervals equals $b - a$.

Apart from these bounded intervals we also have the unbounded intervals:

$$\begin{array}{lcll}
(a,\infty) & = & \{x \in \mathbb{R} : a < x\} & \text{(an unbounded open interval),}\\
(-\infty,a) & = & \{x \in \mathbb{R} : x < a\} & \text{(an unbounded open interval),}\\
[a,\infty) & = & \{x \in \mathbb{R} : a \leq x\} & \text{(an unbounded closed interval),}\\
(-\infty,a] & = & \{x \in \mathbb{R} : x \leq a\} & \text{(an unbounded closed interval),}\\
(-\infty,\infty) & = & \mathbb{R} & \text{(an unbounded interval).}
\end{array}$$

For the notation of the unbounded intervals we have made use of the symbols $-\infty$ and ∞. We wish to emphasize that these symbols are not elements of $\mathbb{R}$, but merely notationally convenient symbols.

Note that the way of describing the set $(a,b) = \{x \in \mathbb{R} : a < x < b\}$ is different from the "dots"-method (for example $\mathbb{N} = \{1,2,3,4,5\ldots\}$) and the "generating"-method (for example $\mathbb{Q} = \{\frac{m}{n} : m \in \mathbb{Z}, n \in \mathbb{N}\}$). Read the line "$(a,b) = \{x \in \mathbb{R} : a < x < b\}$" as "$(a,b)$ consists of all numbers $x \in \mathbb{R}$ that satisfy the property $a < x < b$". This method is called the "selecting property"-method. Note that the set of even natural numbers can be described by using any of these three methods:

$$\{2,4,6,8\ldots\} = \{2n : n \in \mathbb{N}\} = \{n \in \mathbb{N} : n \text{ is even}\}.$$

0.2 Absolute value

If x is a real number unequal to 0, then either x is positive or $-x$ is positive. The absolute value of $x \neq 0$ is defined as the positive number of the numbers x and $-x$.

DEFINITION Let $x \in \mathbb{R}$. The *absolute value* of x, written as $|x|$, is given by

$$|x| = \begin{cases} -x & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ x & \text{if } x > 0. \end{cases}$$

The absolute value of a number can be seen as the distance between that number and the number zero, and the absolute value of the difference of two numbers can be seen as the distance between the numbers on the number line.

The absolute value meets a number of simple requirements in the form of properties which result directly from the definition of the absolute value.

THEOREM 0.2 *For every $x, y \in \mathbb{R}$ we have*

a) $|-x| = |x|$

b) $|xy| = |x|\,|y|$

c) $-|x| \leq x \leq |x|$. An inequality in combination with the absolute value can be replaced by a statement without the absolute value.

THEOREM 0.3 *For every $x \in \mathbb{R}$ and $c \geq 0$ the following holds:*

$$|x| \leq c \quad \textit{if and only if} \quad -c \leq x \leq c.$$

Proof. Let $x \in \mathbb{R}$ and $c \geq 0$.

($\Rightarrow$) Assume that $|x| \leq c$.

If $x = 0$, then it follows immediately that $-c \leq x \leq c$.

If $x > 0$, then $x = |x| \leq c$ and therefore $-c \leq x \leq c$ (since $-c \leq 0$ and $0 < x$).

If $x < 0$, then $-x = |x| \leq c$ and therefore $-c \leq x \leq c$ (since $x < 0$ and $0 \leq c$).

($\Leftarrow$) Assume that $-c \leq x \leq c$.

If $x = 0$, then it follows immediately that $|x| = 0 \leq c$.

If $x > 0$, then $|x| = x \leq c$.

If $x < 0$, then $|x| = -x \leq -(-c) = c$. $\square$

With a small adjustment of the arguments of the proof of Theorem 0.3 you can show that for any $c > 0$ and $x \in \mathbb{R}$ we have $|x| < c$ if and only if $-c < x < c$. The equivalence of the two statements in Theorem 0.3 is a fundamental property of the absolute value on which a series of results is based. One example of this is the Triangle Inequality for the absolute value.

THEOREM 0.4 (**Triangle Inequality**)

For every $x, y \in \mathbb{R}$ the following holds:

$$|x + y| \leq |x| + |y|.$$

Proof. Let $x, y \in \mathbb{R}$. In accordance with part c) of Theorem 0.2 it holds that $-|x| \leq x \leq |x|$ and $-|y| \leq y \leq |y|$.

We obtain subsequently:

$$\begin{array}{rcccl} -|x|-|y| & \leq & x+y & \leq & |x|+|y| \\ -(|x|+|y|) & \leq & x+y & \leq & |x|+|y| \end{array}$$

According to Theorem 0.3 we get $|x+y| \leq |x|+|y|$. □

The Triangle Inequality has several applications as is shown in the following exercise.

EXERCISE 0.2

a) Prove that $|x-y| \leq |x|+|y|$ for every $x, y \in \mathbb{R}$.

b) Prove that $|x+y+z| \leq |x|+|y|+|z|$ for every $x, y, z \in \mathbb{R}$.

Another application is known as the Inverse Triangle Inequality.

THEOREM 0.5 **(Inverse Triangle Inequality)**

For every $x, y \in \mathbb{R}$ the following holds:

$$|x-y| \geq \Big|\, |x|-|y| \,\Big|.$$

The Triangle Inequality (and sometimes the Inverse Triangle Inequality) can be used for making estimates.

From high school you definitely remember the concept of 'function'. A function is a rule f that assigns to any element x in the domain of f, denoted as D_f, an element $f(x)$. Often $D_f \subset \mathbb{R}$ and $f(x)$ is a real number for every element $x \in D_f$. An example is the function f with domain $D_f = [0, \infty)$, defined by $f(x) = \sqrt{x}$ for every $x \in D_f$.

A function f is called *bounded from above* if there exists a number m such that $f(x) \leq m$ for every $x \in D_f$. A function f is called *bounded from below* if there exists a number l such that $f(x) \geq l$ for every $x \in D_f$. A function is called *bounded* if it is both bounded from above and bounded from below. Equivalent to 'f is bounded' is the statement 'there is a number p such that $|f(x)| \leq p$ for every $x \in D_f$'.

EXAMPLE 0.1 The function f on $[2,3]$ is defined by

$$f(x) = \frac{2x^2-3x+1}{3x-1}.$$

Now determine a constant m such that $|f(x)| \leq m$ for every $x \in [2,3]$.

We estimate the numerator and denominator of $|f(x)|$ separately. From the Triangle Inequality (specifically from part b) of Exercise 0.2 it follows that

$$\begin{aligned} |2x^2 - 3x + 1| &\leq |2x^2| + |-3x| + |1| \\ &= 2|x|^2 + 3|x| + 1 \\ &\leq 2 \cdot 3^2 + 3 \cdot 3 + 1 = 28, \end{aligned}$$

since $|x| \leq 3$ if $x \in [2,3]$. Therefore 28 is an upper bound for the absolute value of the numerator.

In order to make a lower bound for the absolute value of the denominator we note that $3x - 1 \geq 0$ for $x \in [2,3]$. We have

$$|3x - 1| = 3x - 1 \geq 3 \cdot 2 - 1 = 5,$$

since $x \geq 2$ if $x \in [2,3]$. Consequently we have $\dfrac{1}{|3x-1|} \leq \frac{1}{5}$. So,

$$|f(x)| = \frac{|2x^2 - 3x + 1|}{|3x - 1|} = |2x^2 - 3x + 1| \cdot \frac{1}{|3x - 1|} \leq 28 \cdot \frac{1}{5} = \frac{28}{5}$$

for all $x \in [2,3]$. Hence $m = \frac{28}{5}$ can serve as the desired constant. It is obvious that $\frac{28}{5}$ is not the only value for m which does the job (for example $m = 1000000$ can be chosen as well), and neither do we exclude the possibility that much smaller values for m exist. $\triangleleft$

EXERCISE 0.3

a) The function f on $[-1,3]$ is defined by

$$f(x) = \frac{x^3 - 4x + 5}{2 + x^2}.$$

Determine a number m such that $|f(x)| \leq m$ for every $x \in [-1,3]$.

b) The function f on $[-3,1]$ is defined by

$$f(x) = \frac{x^3 - 4x + 5}{2 + x^2}.$$

Determine a number p such that $|f(x)| \leq p$ for every $x \in [-3,1]$.

0.3 Some important theorems

In the next chapters we need some important results from mathematical analysis, we are summarized in this subsection for future reference.

The first concept is the *Axiom of Completeness*:

AXIOM 0.1 **(axiom of completeness)** *every non-empty set S of real numbers which is bounded from above (i.e. there is a number $u \in \mathbb{R}$ such that $s \leq u$ for every $s \in S$) has a smallest upper bound. The smallest upper bound of a set S is also called the supremum of S and is written as* $\sup S$*;*
every non-empty set of real numbers which is bounded from below (i.e. there is a number $l \in \mathbb{R}$ such that $s \geq l$ for every $s \in S$) has a highest lower bound. The highest lower bound of a set S is also called the infimum of S and is written as $\inf S$.

Furthermore, we also need two theorems about convergence of sequences $x_1, x_2, x_3, \ldots$. The first theorem states that an increasing sequence that is bounded from above, or a decreasing sequence that is bounded from below has a limit.

THEOREM 0.6 **(convergence criterion for monotone sequences)**
a) Every increasing sequence $x_1, x_2, x_3, \ldots$ (i.e. $x_1 \leq x_2 \leq x_3 \leq \cdots$), which is bounded from above (i.e. there is a number $u \in \mathbb{R}$ such that $x_n \leq u$ for every $n \in \mathbb{N}$), is convergent and

$$\lim_{n \to \infty} x_n = \sup\{x_n \,:\, n \in \mathbb{N}\}.$$

b) Every decreasing sequence $x_1, x_2, x_3, \ldots$ (i.e. $x_1 \geq x_2 \geq x_3 \geq \cdots$), which is bounded from below (i.e. there is a number $l \in \mathbb{R}$ such that $x_n \geq l$ for every $n \in \mathbb{N}$), is convergent and

$$\lim_{n \to \infty} x_n = \inf\{x_n \,:\, n \in \mathbb{N}\}.$$

The second theorem states that a sequence that is bounded always has a convergent subsequence.

THEOREM 0.7 **(Theorem of Bolzano-Weierstrass)**
Every bounded sequence $x_1, x_2, x_3, \ldots$ (i.e. there exist numbers $l, u \in \mathbb{R}$ such that $l \leq x_n \leq u$ for every $n \in \mathbb{N}$) has at least one convergent subsequence (i.e. there exist natural numbers $n_1 < n_2 < n_3 < \cdots$ such that $x_{n_1}, x_{n_2}, x_{n_3}, \ldots$ is a convergent sequence).

1 THE PRINCIPLE OF INDUCTION

Fermat, a French mathematician from the 17-th century, was interested in the sequence of prime numbers:

$$2,\ 3,\ 5,\ 7,\ 11,\ 13,\ 17,\ldots$$

He was looking for some order in this sequence. He asked himself, for instance, whether it was possible to formulate an arithmetic rule f with the property that $f(1), f(2), f(3),\ldots$ would all be prime numbers. After searching for a long time he discovered that the numbers

$$\begin{array}{lcccl} 2^2 & + & 1 & = & 5 \\ 2^{2^2} & + & 1 & = & 17 \\ 2^{2^3} & + & 1 & = & 257 \\ 2^{2^4} & + & 1 & = & 65\,537 \end{array}$$

are prime numbers. He then expressed the conjecture that probably $2^{2^5} + 1, 2^{2^6} + 1,\ldots$ would also be prime numbers, but he could not manage the calculations. (If you know that $2^{2^5}+1$ equals $4\,294\,967\,297$ you can understand why). In a letter to Blaise Pascal he wrote: "I am certain that it is true; the proof is most unpleasant, and I must confess that I have not yet been capable of quite completing it."

About a hundred years later, however, Leonhard Euler proved that $2^{2^5}+1$ is divisible by 641, which caused the conjecture to collapse.

Moral: that the pitcher goes to the well four times without breaking, does not guarantee that it won't break the fifth time.

(From: *Analyse voor Beginners*, A. van Rooij)

In section 1.2 we discuss a tool for proving theorems which is known as the principle of induction; this tool is used if we need to prove that an infinite sequence of statements

$\mathcal{P}(1), \mathcal{P}(2), \mathcal{P}(3), \ldots$ is true. In section 1.3 we apply the principle of induction to Newton's binomium. In Newton's binomium we wish to eliminate the brackets in expressions of the form $(a+b)^2,\ (a+b)^3,\ (a+b)^4,$ etcetera. First, we discuss in section 1.1 the use of the summation and product sign.

1.1 The summation and product sign

By means of the summation sign Σ and the product sign Π we can denote the sum or product of a large or unknown collection of numbers. If you encounter the expression

$$\sum_{k=1}^{27} \ldots,$$

then you have to read this expression as follows: substitute in the expression on the position of the $\ldots$ subsequently $k = 1$, $k = 2$, ..., $k = 27$ and add the resulting 27 numbers. For example,

$$\sum_{k=1}^{5} (k^2 + k) = 2 + 6 + 12 + 20 + 30 = 70.$$

Note that $\displaystyle\sum_{k=1}^{5} (k^2 + k)$ is simply a number, that does not depend upon k. For the variable k we have substituted subsequently the values 1, 2, 3, 4 and 5 and the resulting numbers have been added. Hence, the variable k does not occur in the final result. We refer to k as the *summation index*. Note that $\displaystyle\sum_{j=1}^{5} (j^2 + j)$ represents exactly the same number.
Similarly, you have to read the expression

$$\prod_{k=1}^{27} \ldots$$

as follows: substitute in the expression on the position of the $\ldots$ subsequently $k = 1$, $k = 2$, ..., $k = 27$ and multiply the resulting 27 numbers. For example,

$$\prod_{k=1}^{5} (k^2 + k) = 2 \cdot 6 \cdot 12 \cdot 20 \cdot 30 = 86400.$$

Again, $\displaystyle\prod_{k=1}^{5} (k^2 + k)$ is simply a number, that does not depend upon k. The product sign is used in the following definition.

DEFINITION For every $n \in \mathbb{N}$, n *factorial*, written as $n!$, is given by

$$n! = \prod_{k=1}^{n} k = 1 \cdot 2 \cdot 3 \cdot \ \cdots \ \cdot (n-1) \cdot n.$$

We also define $0! = 1$.

Now check yourself that

$$\begin{aligned}
\sum_{k=1}^{5}(2k-1) &= 25 \\
\prod_{k=1}^{5}(2k-1) &= 945 \\
\sum_{k=1}^{n} 2 &= 2n \\
\prod_{k=1}^{n} n &= n^n.
\end{aligned}$$

In a similar fashion we deal with expressions containing a double summation. For example,

$$\begin{aligned}
\sum_{k=1}^{3}\sum_{l=1}^{4}(k+l) &= \sum_{l=1}^{4}(1+l) + \sum_{l=1}^{4}(2+l) + \sum_{l=1}^{4}(3+l) \\
&= (2+3+4+5) + (3+4+5+6) + (4+5+6+7) \\
&= 54.
\end{aligned}$$

It often occurs that in a double summation the range of the second summation index depends upon the value of the first summation index. For example,

$$\begin{aligned}
\sum_{k=1}^{3}\sum_{l=k}^{4}(k+l) &= \sum_{l=1}^{4}(1+l) + \sum_{l=2}^{4}(2+l) + \sum_{l=3}^{4}(3+l) \\
&= (2+3+4+5) + (4+5+6) + (6+7) \\
&= 42.
\end{aligned}$$

Note however that it is not possible to allow for the range of the first summation index to be dependent upon the value of the second summation index (why?).

Of course, it does not stop here and one can consider triple summation or mixtures of summations and products. For example,

$$\begin{aligned}\prod_{k=1}^{3}\sum_{l=k}^{4}(k+l) &= \sum_{l=1}^{4}(1+l)\cdot\sum_{l=2}^{4}(2+l)\cdot\sum_{l=3}^{4}(3+l)\\ &= (2+3+4+5)\cdot(4+5+6)\cdot(6+7)\\ &= 2730.\end{aligned}$$

EXERCISE 1.1 Determine $\sum_{j=1}^{4}(3j+2)$ and $\prod_{j=2}^{3}(3j+2)$.

EXERCISE 1.2 Write $1+4+7+10+13+16$ and $1+4+7+\cdots+(3n+1)$ as an expression that uses the summation sign Σ.

EXERCISE 1.3 Write $\frac{1}{2}\cdot\frac{3}{4}\cdot\frac{5}{6}\cdot\frac{7}{8}$ and $\frac{1}{2}\cdot\frac{3}{4}\cdot\frac{5}{6}\cdot\cdots\cdot\frac{2n+1}{2n+2}$ as an expression that uses the product sign Π.

EXERCISE 1.4 Show that $\sum_{k=1}^{n}a_k=\sum_{k=0}^{n-1}a_{k+1}$.

EXERCISE 1.5 Compute $\sum_{k=1}^{1000000}(5k+7)-\sum_{k=2}^{1000000}(5k+2)$.

EXERCISE 1.6 Compute $\dfrac{\prod_{j=1}^{n}a_j}{\prod_{j=1}^{n-2}a_{j+1}}$.

EXERCISE 1.7 Show that $\sum_{i=1}^{n}\sum_{j=i}^{n}a_{ij}=\sum_{j=1}^{n}\sum_{i=1}^{j}a_{ij}$.

EXERCISE 1.8 Write $\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}-\sum_{i=2}^{n}\sum_{j=1}^{i-1}a_{ij}-\sum_{j=2}^{n}\sum_{i=1}^{j-1}a_{ij}$ as an expression that uses only one summation sign.

1.2 The principle of induction

As an introduction we study a numerical example.

EXAMPLE 1.1 Suppose that you want to know whether each of the numbers of the sequence

$$8^1 + 6,\ 8^2 + 6,\ 8^3 + 6,\ 8^4 + 6, \ldots$$

is divisible by 7. You start calculating and come up with the following results:

$$\begin{array}{lcccll} 8^1 & + & 6 & = & 14 & : \text{divisible by} \ \ 7 \\ 8^2 & + & 6 & = & 70 & : \text{divisible by} \ \ 7 \\ 8^3 & + & 6 & = & 518 & : \text{divisible by} \ \ 7 \\ 8^4 & + & 6 & = & 4102 & : \text{divisible by} \ \ 7. \end{array}$$

Now you are aware of the story of Fermat and therefore you feel some hesitation in concluding that this will work equally well for every next number. Besides, you do not really want to calculate $8^5 + 6$, $8^6 + 6$ and divide it by 7, for what is to be gained by knowing that $8^5 + 6$ and $8^6 + 6$ are divisible by 7? So let's take a more systematic approach:

$$\begin{array}{lclclclcl} 8^5 + 6 & = & 8 \cdot 8^4 + 6 & = & 7 \cdot 8^4 + (8^4 + 6) & = & \text{7-tuple} + \text{7-tuple} & = & \text{7-tuple} \\ 8^6 + 6 & = & 8 \cdot 8^5 + 6 & = & 7 \cdot 8^5 + (8^5 + 6) & = & \text{7-tuple} + \text{7-tuple} & = & \text{7-tuple} \\ 8^7 + 6 & = & 8 \cdot 8^6 + 6 & = & 7 \cdot 8^6 + (8^6 + 6) & = & \text{7-tuple} + \text{7-tuple} & = & \text{7-tuple} \\ 8^8 + 6 & = & 8 \cdot 8^7 + 6 & = & \ldots & & & & \end{array}$$

In the first line we used the result of the last calculation that we still liked doing, namely that $8^4 + 6$ is a 7-tuple; in the second line we used the result of the preceding line, namely that $8^5 + 6$ is a 7-tuple; etcetera.

You probably feel by now that the investigation of the divisibility by 7 of each next number of the sequence will proceed smoothly. You just continue doing the same calculations and plugging in the preceding line's result at the right moment. Nevertheless we are still not quite satisfied, because a proof ending with $\ldots$ is not a complete proof in the eyes of a mathematician. But in mathematics something has been invented to remedy this, namely the *principle of induction*. This is a technique for proving theorems which is tailored to the above situation: there is a(n open) statement which depends on a natural number n, and we wish to show that this statement is true for every natural number n.

The proof according to the principle of induction looks as follows when applied to our example. We denote the statement "$8^n + 6$ is divisible by 7" by $\mathcal{P}(n)$.

> **Step 1:** by means of a calculation we observe that $8^1 + 6$ is divisible by 7, in other words: statement $\mathcal{P}(1)$ holds.
>
> **Step 2:** assuming n to be a completely arbitrarily chosen natural number for which it is true that $8^n + 6$ is divisible by 7, we then show (in precisely the same way as has just been done) that $8^{n+1} + 6$ is divisible by 7:
>
> $$\begin{aligned} 8^{n+1} + 6 &= 8 \cdot 8^n + 6 = 7 \cdot 8^n + (8^n + 6) \\ &= \text{7-tuple} + \text{7-tuple} = \text{7-tuple}. \end{aligned}$$
>
> (For the last but one equality we have made use of the assumption that $8^n + 6$ is a 7-tuple.) In other words: it has been shown for every natural number n that from the assumption that $\mathcal{P}(n)$ holds, it follows that $\mathcal{P}(n+1)$ also holds.

It is intuitively clear that on the basis of the two above steps we are justified in concluding that for every $n \in \mathbb{N}$ the number $8^n + 6$ is divisible by 7: for according to the first step we know that $8^1 + 6$ is divisible by 7 and according to the second step we may then conclude that $8^2 + 6$ is also divisible by 7, so that in turn we can conclude, according to the second step, that $8^3 + 6$ is also divisible by 7, so that etcetera. $\triangleleft$

The principle of induction

A statement $\mathcal{P}(n)$ is proven to be true for every $n \in \mathbb{N}$ according to the *principle of induction*, if the following two conditions are satisfied:

condition 1: $\mathcal{P}(1)$ holds

condition 2: for every $n \in \mathbb{N}$ the implication $\mathcal{P}(n) \Rightarrow \mathcal{P}(n+1)$ holds.

The first condition is called the *induction basis* and the second condition the *induction step.*

We will now give a proof of Bernoulli's inequality, using the principle of induction.

THEOREM 1.1 (**Bernoulli's inequality**)
For every $x > -1$ and for every $n \in \mathbb{N}$ the following holds:

$$(1 + x)^n \geq 1 + nx.$$

Proof. Let $x > -1$. The statement "$(1 + x)^n \geq 1 + nx$" is denoted by $\mathcal{P}(n)$. We will prove that $\mathcal{P}(n)$ holds for every $n \in \mathbb{N}$ by means of the principle of induction.

> **Step 1:** ($\mathcal{P}(1)$ holds)
> $\mathcal{P}(1)$ is the statement "$(1 + x)^1 \geq 1 + 1 \cdot x$", which inequality certainly holds: in fact even the equality holds.
>
> **Step 2:** (for every $n \in \mathbb{N}$ the implication $\mathcal{P}(n) \Rightarrow \mathcal{P}(n + 1)$ holds)
> Let $n \in \mathbb{N}$. Assume that $\mathcal{P}(n)$ holds, in other words $(1 + x)^n \geq 1 + nx$. From this assumption and the fact that $1 + x > 0$ the following subsequent (in)equalities result:
>
> $$\begin{aligned} (1 + x)^{n+1} &= (1 + x)^n(1 + x) \geq (1 + nx)(1 + x) \\ &= 1 + (n + 1)x + nx^2 \geq 1 + (n + 1)x. \end{aligned}$$
>
> This proves the statement $\mathcal{P}(n + 1)$.

On the basis of the proof that $(1 + x)^n \geq 1 + nx$ holds for $n = 1$ and that for an arbitrary $n \in \mathbb{N}$ the implication $(1 + x)^n \geq 1 + nx \Rightarrow (1 + x)^{n+1} \geq 1 + (n + 1)x$ holds, the principle of induction states that the statement $(1 + x)^n \geq 1 + nx$ holds for every natural number n. $\square$

Before we continue, please note the following on Bernoulli's inequality and the principle of induction. Bernoulli's inequality, $(1 + x)^n \geq 1 + nx$, is a statement depending on x and n, which we have proved to be true for every $x > -1$ and for every natural number n. Our proof proceeded in the usual manner by initially choosing a completely arbitrary number $x > -1$ ("let $x > -1$"). In accordance with this we could also have chosen a completely arbitrary natural number for n. Why we have not in fact done so will become clear if we take a closer look at Bernoulli's inequality: how should we deal with the n-th power of $(1 + x)$? Elimination of brackets will work well for the case $n = 2$, and even for $n = 3$, but how will the elimination of brackets work in case n is arbitrary?

Therefore in this case it was more convenient to give the proof by means of the principle of induction. So it depends on the circumstances whether you should choose for the approach with “let $n \in \mathbb{N}$”, or for the principle of induction.
Note that a proof with induction to x was not possible, as x is allowed to take any *real* value in $(-1, \infty)$ and the principle of induction is only tailor-made for proving statements depending on *natural* numbers.

EXERCISE 1.9

a) Use the principle of induction to prove that for every $n \in \mathbb{N}$

i) $2^n \leq (n+1)!$

ii) $n < 2^n$.

b) Verify whether for every $n \in \mathbb{N}$

$$n \leq 3\sqrt{n} + 4.$$

EXERCISE 1.10

a) Prove that for every $n \in \mathbb{N}$ $1 + 2 + \cdots + n = \frac{1}{2}n(n+1)$.

b) (i)Prove that for every $n \in \mathbb{N}$ $1 + r + r^2 + \cdots + r^n = \frac{1-r^{n+1}}{1-r}$ where $r \in \mathbb{R}, r \neq 1$.

(ii) Let $-1 < r < 1$. Explain that $1 + r + r^2 + \cdots + r^n$ comes closer to $\frac{1}{1-r}$ the larger you take n.

(iii) Let $r \notin [-1, 1]$. What can you say about $1 + r + r^2 + \cdots + r^n$ if n tends to infinity?

EXERCISE 1.11

a) Prove that for every $n \in \mathbb{N}$

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

b) Prove that for every natural number $n \geq 2$

$$\sum_{r=1}^{n} \frac{1}{\sqrt{r}} > \sqrt{n}.$$

1.3 Newton's binomium

You know that $(a+b)^2 = a^2 + 2ab + b^2$ and perhaps that $(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$ (if not please work away the brackets in order to check this). We now wish to examine the question what this expansion looks like for higher powers, especially for the case $(a+b)^n$ where n is an arbitrary natural number. This expansion is known as *Newton's binomium* or as the *binomial expansion of* $(a+b)^n$. In order to be able to give this binomial expansion we need the following definition.

DEFINITION For every $n \in \mathbb{N}$ and for every $r \in \{0, 1, \ldots, n\}$, $\binom{n}{r}$ (pronounced as "n choose r") is given by

$$\binom{n}{r} = \frac{n!}{(n-r)!r!}.$$

In elementary probability theory, $\binom{n}{r}$ denotes the number of possible results when drawing r objects out of n without replacement, where the only thing that matters is *which* objects are drawn, *not* the order in which they are drawn. This is called the number of combinations of r from n.

EXERCISE 1.12 Calculate $\binom{n}{0}$, $\binom{n}{1}$, $\binom{n}{n-1}$ and $\binom{n}{n}$.

The numbers $\binom{n}{r}$ can be arranged in the so-called triangle of Pascal:

$$\begin{array}{ccccccccccccc}
 & & & & & \binom{1}{0} & & \binom{1}{1} & & & & & \\
 & & & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & & & \\
 & & & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & & & \\
 & & \binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} & & \\
 & \binom{5}{0} & & \binom{5}{1} & & \binom{5}{2} & & \binom{5}{3} & & \binom{5}{4} & & \binom{5}{5} & \\
\binom{6}{0} & & \binom{6}{1} & & \binom{6}{2} & & \binom{6}{3} & & \binom{6}{4} & & \binom{6}{5} & & \binom{6}{6} \\
\vdots & & & & & & & & & & & & \vdots
\end{array}$$

The actual calculation of the numbers $\binom{n}{r}$ produces the triangle of Pascal:

$$\begin{array}{ccccccccccccc}
 & & & & & 1 & & 1 & & & & & \\
 & & & & 1 & & 2 & & 1 & & & & \\
 & & & 1 & & 3 & & 3 & & 1 & & & \\
 & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
 & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 \\
\vdots & & & & & & & & & & & & \vdots
\end{array}$$

Please note that each number on the inside of this triangle is the sum of the two numbers placed above it to its left and right. In the following exercise you are asked to prove that this property is generally valid.

EXERCISE 1.13 Prove that for every $n \in \mathbb{N}$ and for every $r \in \{1, 2, \ldots, n\}$

$$\binom{n+1}{r} = \binom{n}{r-1} + \binom{n}{r}.$$

Now let us return to the purpose of this paragraph: the elimination of the brackets in the expression $(a+b)^n$. If we try our hand out at a few of the calculations, then we can see that

$$(a+b)^1 \quad = \quad a+b$$

$$\begin{aligned}(a+b)^2 &= a^2+2ab+b^2\\(a+b)^3 &= a^3+3a^2b+3ab^2+b^3\\(a+b)^4 &= a^4+4a^3b+6a^2b^2+4ab^3+b^4.\end{aligned}$$

We can conclude that the coefficients in the expansion of $(a+b)^1$, 1 and 1, precisely form the first row of Pascal's triangle, those in $(a+b)^2$, 1, 2 and 1, form the second row, those in $(a+b)^3$, 1, 3, 3 and 1, form the third row, and those in $(a+b)^4$, 1, 4, 6, 4 and 1, form the fourth row. By means of the principle of induction we will show that it is indeed generally true that the coefficients in the expansion of $(a+b)^n$ correspond to the numbers of the n-th row of Pascal's triangle.

THEOREM 1.2 (**Newton's binomium**)
For every $a, b \in \mathbb{R}$ and for every $n \in \mathbb{N}$

$$(a+b)^n = \binom{n}{0}a^n + \binom{n}{1}a^{n-1}b + \cdots + \binom{n}{n-1}ab^{n-1} + \binom{n}{n}b^n = \sum_{r=0}^{n}\binom{n}{r}a^{n-r}b^r.$$

Proof. Let $a, b \in \mathbb{R}$. We will prove Newton's binomium by means of the principle of induction. Define for $n \in \mathbb{N}$ the statement $\mathcal{P}(n)$ as follows:

$$(a+b)^n = \binom{n}{0}a^n + \binom{n}{1}a^{n-1}b + \cdots + \binom{n}{n-1}ab^{n-1} + \binom{n}{n}b^n.$$

Step 1: ($\mathcal{P}(1)$ is true)
For $n = 1$ the expressions on the left and right hand side of the binomial equation are equal, namely $a+b$, so that for $n = 1$ Newton's binomium is true.
Step 2: (for every $n \in \mathbb{N}$, $\mathcal{P}(n) \Rightarrow \mathcal{P}(n+1)$ is true)
Let $n \in \mathbb{N}$. Assume that $\mathcal{P}(n)$ is true, in other words

$$(a+b)^n = \binom{n}{0}a^n + \binom{n}{1}a^{n-1}b + \cdots + \binom{n}{n-1}ab^{n-1} + \binom{n}{n}b^n.$$

Then for $(a+b)^{n+1}$ the following subsequent equalities hold:

$$\begin{aligned}(a+b)^{n+1} &= (a+b)(a+b)^n = a(a+b)^n + b(a+b)^n\\ &= \binom{n}{0}a^{n+1} + \binom{n}{1}a^nb + \cdots + \binom{n}{n-1}a^2b^{n-1} + \binom{n}{n}ab^n\\ &\qquad + \binom{n}{0}a^nb + \binom{n}{1}a^{n-1}b^2 + \cdots + \binom{n}{n-1}ab^n + \binom{n}{n}b^{n+1}\end{aligned}$$

$$
\begin{aligned}
&= \binom{n}{0}a^{n+1} + \left[\binom{n}{1} + \binom{n}{0}\right]a^n b + \left[\binom{n}{2} + \binom{n}{1}\right]a^{n-1}b^2 \\
&\quad + \cdots + \left[\binom{n}{n} + \binom{n}{n-1}\right]ab^n + \binom{n}{n}b^{n+1} \\
&= \binom{n+1}{0}a^{n+1} + \binom{n+1}{1}a^n b + \cdots + \binom{n+1}{n}ab^n + \binom{n+1}{n+1}b^{n+1}.
\end{aligned}
$$

For the last equality we have made use of Exercise 1.13 and of the fact that $\binom{n}{0} = \binom{n+1}{0}$ and that $\binom{n}{n} = \binom{n+1}{n+1}$. Therefore $\mathcal{P}(n+1)$ is true. $\square$

So the conclusion for Newton's binomium is that the n-th row of Pascal's triangle precisely renders the coefficients in the expansion of $(a+b)^n$.

EXERCISE 1.14

a) Calculate the coefficient of x^2 in the binomial expansion of $(2+x)^5$.

b) Determine the coefficient of x^4 in $(3x+5)^{12}$.

1.4 Mixed exercises

EXERCISE 1.15 Prove by means of the principle of induction that for every $n \in \mathbb{N}$

a) $\displaystyle\sum_{k=1}^{n} \frac{1}{k(k+1)} = \frac{n}{n+1}$

b) $\displaystyle\sum_{j=1}^{n} j^3 = \Big(\sum_{j=1}^{n} j\Big)^2$

(Clue: use the formula for the sum of the arithmetic sequence in Exercise 1.10 a).)

c) $\displaystyle\sum_{k=1}^{n} k^2 = \frac{1}{6}n(n+1)(2n+1).$

EXERCISE 1.16 Prove by means of induction that for every $n \in \mathbb{N}$:

$$\sum_{k=1}^{n} \frac{1}{k!} \leq 2 - \frac{1}{2^{n-1}}.$$

(Clue: use Exercise 1.9.)

EXERCISE 1.17 Determine the coefficient of x^4 in $(2x^2 + \frac{1}{3x})^5$.

EXERCISE 1.18 Prove that for every $n \in \mathbb{N}, n \geq 2$ and for every $k \in \{1, \ldots, n\}$

$$k\binom{n}{k} = n\binom{n-1}{k-1}.$$

EXERCISE 1.19 Prove by means of the principle of induction that for every $n \in \mathbb{N}$ the number $n^3 + 2n$ is divisible by 3.

EXERCISE 1.20 Write down the binomial expansion of $e_n = \left(1 + \frac{1}{n}\right)^n$ and show that this can be rewritten as

$$\begin{aligned} e_n \quad &= 1 + 1 + \tfrac{1}{2!}\left(1 - \tfrac{1}{n}\right) + \tfrac{1}{3!}\left(1 - \tfrac{1}{n}\right)\left(1 - \tfrac{2}{n}\right) + \cdots \\ &\qquad \cdots + \tfrac{1}{n!}\left(1 - \tfrac{1}{n}\right)\left(1 - \tfrac{2}{n}\right)\cdots\left(1 - \tfrac{n-1}{n}\right). \end{aligned}$$

EXERCISE 1.21 Prove by means of the principle of induction that for every $n \in \mathbb{N}$

$$\prod_{k=1}^{n} \frac{2k-1}{2k} < \frac{1}{\sqrt{2n+1}}.$$

EXERCISE 1.22 Prove by means of the principle of induction that for every $n \in \mathbb{N}$ and $x_1, \ldots, x_n \in \mathbb{R}$ we have

$$|\sum_{k=1}^{n} x_k| \leq \sum_{k=1}^{n} |x_k|.$$

2 LIMITS OF FUNCTIONS

In this chapter we will discuss the concept of 'limit of a function'. In section 2.1 we will give the definition of limit of a function in infinity. The starting-point for the limit of a function in infinity is an arbitrarily chosen positive number ε. For this ε a number H must then be determined that meets certain requirements. For this reason the definition of a limit of a function in infinity is also commonly referred to as the (ε, H)-definition. In section 2.2 we will formulate a number of arithmetic rules for limits of functions in infinity. In section 2.3 we consider limits of functions in a point and in section 2.4 the corresponding arithmetic rules. Finally, in section 2.5 other variants of the concept of limit are studied.

2.1 The limit of a function in infinity

In Figure 2.1 the graph is sketched of the function f on $(0, \infty)$, defined by $f(x) = \frac{1}{\sqrt{x}}$ for every $x \in (0, \infty)$. We observe that the values $f(x)$ of the function get closer and closer to the number 0: the greater the value of x, the smaller the distance of $f(x)$ to the number 0. Better still, we can indicate for any positive number, no matter how small, from which point in the domain onwards the distance of *each* function value to the number 0 is less than that positive number. If, for instance, we take for this positive number $\frac{1}{10}$, then for *every* $x > 100$ the distance of $f(x)$ to the number 0 is less than $\frac{1}{10}$ (please verify). If we take for this positive number $\frac{1}{100}$ then for *every* $x > 10000$ the distance of the function value to the number 0 is less than $\frac{1}{100}$. A 'tolerance' of $\frac{1}{100\,000}$ is reached for every $x > 10^{10}$.

To recapitulate, the function $f(x) = \frac{1}{\sqrt{x}}$ has the property that for *each* positive number, whether that be $\frac{1}{10}$ or $\frac{1}{100}$ or $\frac{1}{100\,000}$ or any other positive number, a number H can be determined such that for *each* $x > H$ the corresponding function value $f(x)$ has a distance to the number 0 less than that positive number. In our example (but also more generally) you will find that the smaller you take this positive number, the greater you will have to take H.

A formal way of expressing this type of behavior of the function is to say that the function converges or that it has a limit in infinity. The arbitrarily selected positive number for which a number H has to be determined is indicated in the definition by the variable ε.

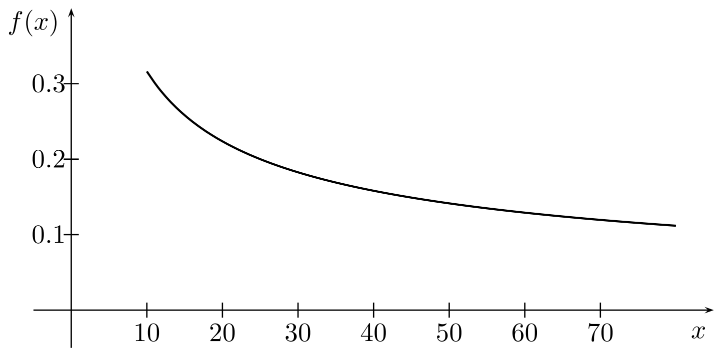

Figure 2.1: The graph of the function $f(x) = \frac{1}{\sqrt{x}}$.

DEFINITION Let f be a function such that an $a \in \mathbb{R}$ exists with $(a, \infty) \subset D_f$.
A real number l is called the limit of the function f in infinity, if for every $\varepsilon > 0$ a real number H exists such that

$$|f(x) - l| < \varepsilon \quad \text{for every} \quad x \in D_f, \quad \text{for which} \quad x > H.$$

If l is the limit of the function f in infinity, then we say that f *converges to l in infinity.* We write this as

$$\lim_{x\to\infty} f(x) = l \qquad \text{or} \qquad f(x) \to l \quad \text{as} \quad x \to \infty.$$

We call a function *convergent in infinity* if the function has a limit in infinity.

Now we will make some remarks on the definition of the limit of a function in infinity.

1) In quantifier-language the statement 'the function f has limit l in infinity' reads as follows:

$$\forall \varepsilon > 0 \exists H \in \mathbb{R} \forall x \in D_f, x > H : |f(x) - l| < \varepsilon.$$

2) The number H as described in the definition of the limit usually depends on ε.

3) The number l is the limit of the function f if for *every* $\varepsilon > 0$ a number H can be found

such that for all terms $x \in D_f$ with $x > H$ the following holds:

$$|f(x)-l| < \varepsilon \quad \text{or equivalently} \quad l-\varepsilon < f(x) < l+\varepsilon \quad \text{or equivalently} \quad f(x) \in (l-\varepsilon, l+\varepsilon).$$

Graphically we can illustrate the convergence of a function in infinity as follows. For an arbitrarily selected positive number ε we draw two horizontal lines in the coordinate system, one at a height of $l-\varepsilon$ and one at a height of $l+\varepsilon$. The function is convergent in infinity with limit l if for all $\varepsilon > 0$ it is possible to indicate a point H such that from H onwards the graph of f is fully located between the horizontal lines at heights of $l-\varepsilon$ and $l+\varepsilon$ (see Figure 2.2).

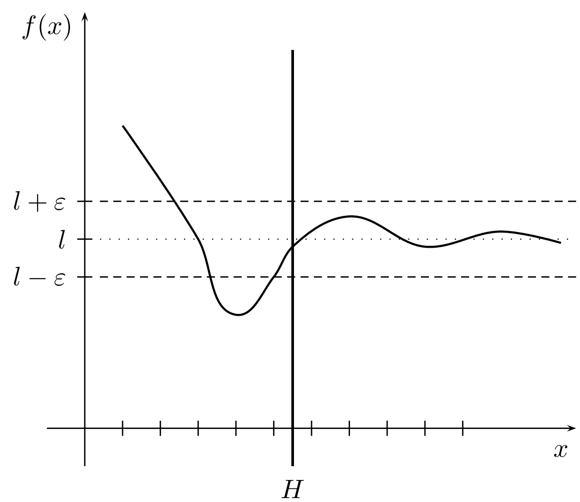

Figure 2.2: Graphical illustration of the limit of a function in infinity.

It is of the utmost importance that you try to fully fathom the concept of limit, since it is the foundation of mathematical analysis.

EXAMPLE 2.1 We consider the function f on $(0,\infty)$, defined by $f(x) = \frac{1+x}{x}$ for every $x \in (0,\infty)$, and prove by means of the definition that the function is convergent in infinity with limit 1: $\lim\limits_{x\to\infty} \frac{1+x}{x} = 1$.
If we wish to prove that 1 is the limit of the function $\frac{1+x}{x}$ then we must prove that for *every* positive number ε, whether that be $\frac{1}{10}$ or $\frac{3}{184}$ or $\frac{\sqrt{5}}{163\,584}$, a number H can be found such that for every natural number $x > H$ the distance of $f(x)$ to 1 is less than ε, in other words

$|\frac{1+x}{x} - 1| < \varepsilon$.

Let $\varepsilon > 0$. Take $H = \frac{1}{\varepsilon}$ (see 'draft'). Next let $x \in D_f$ with $x > H = \frac{1}{\varepsilon}$. The following holds:

$$\left|f(x) - 1\right| = \left|\frac{1+x}{x} - 1\right| = \left|\frac{1}{x} + 1 - 1\right| = \left|\frac{1}{x}\right| = \frac{1}{x} < \frac{1}{\frac{1}{\varepsilon}} = \varepsilon.$$

This proves that $\lim\limits_{n\to\infty} \frac{1+x}{x} = 1$.

The proof looks brief and to the point. Still you may wonder how we came to the choice of $H = \frac{1}{\varepsilon}$. Well now, this is the so-called preparatory work that you always must do before you can put together a really substantial proof. This preparatory work is your draft; I luckily still preserved mine.

Draft:

We must determine for $\varepsilon > 0$ a number H such that for all $x \in D_f$ with $x > H$

$$\left|\frac{1+x}{x} - 1\right| < \varepsilon.$$

Now

$$\left|\frac{1+x}{x} - 1\right| = \left|\frac{1}{x} + 1 - 1\right| = \left|\frac{1}{x}\right| = \frac{1}{x}.$$

Now if $\frac{1}{x} < \varepsilon$, then $\left|\frac{1+x}{x} - 1\right| < \varepsilon$,

in other words

if $x > \frac{1}{\varepsilon}$, then $\left|\frac{1+x}{x} - 1\right| < \varepsilon$.

This solves the problem: take $N = \frac{1}{\varepsilon}$.

If for example you have the value $\frac{3}{184}$ in mind for ε, then for every $x > H = \frac{184}{3} = 61\frac{1}{3}$ we have $|\frac{1+x}{x} - 1| < \frac{3}{184}$. But the inequality also holds of course for every $x > 62$ or every $x > 2018$. ◁

EXERCISE 2.1 We consider the function f on $(0, \infty)$, defined by $f(x) = \dfrac{1}{x^2}$, for every $x \in (0, \infty)$.

a) Determine for each of the inequalities below a number H such that for every $x \in D_f$ with $x > H$

$$\frac{1}{x^2} < \frac{1}{100}, \quad \frac{1}{x^2} < \frac{1}{105}, \quad \frac{1}{x^2} < \frac{1}{123} \text{ and } \frac{1}{x^2} < \frac{1}{10^5}.$$

b) Let $\varepsilon > 0$. Determine a number H such that $\dfrac{1}{x^2} < \varepsilon$ for every $x \in D_f$ with $x > H$.

c) Does $f(x) = \dfrac{1}{x^2}$ have a limit in infinity?

EXERCISE 2.2 Prove by means of the definition of a limit that

a) $\lim_{x\to\infty} \dfrac{1}{x} = 0$

b) $\lim_{x\to\infty} \dfrac{3}{\sqrt{x}} = 0$

EXAMPLE 2.2 By means of the definition of a limit we prove that $\lim_{x\to\infty} \dfrac{2x^2+3x+4}{2x^2+3} = 1$. Let $\varepsilon > 0$. Take $H = \max\{1, \frac{2}{\varepsilon}\}$. Next let $x \in D_f(= \mathbb{R})$ with $x > H = \frac{2}{\varepsilon}$. We then have subsequently

$$\begin{aligned}
\left|\frac{2x^2+3x+4}{2x^2+3} - 1\right| &= \left|\frac{2x^2+3x+4-(2x^2+3)}{2x^2+3}\right| = \left|\frac{3x+1}{2x^2+3}\right| \\
&= \frac{3x+1}{2x^2+3} \\
&< \frac{3x+1}{2x^2} \\
&\le \frac{3x+x}{2x^2} = \frac{2}{x} \\
&< \frac{2}{H} \\
&\le \frac{2}{\frac{2}{\varepsilon}} \\
&= \varepsilon.
\end{aligned}$$

This proves that $\lim_{x\to\infty} \frac{2x^2+3x+4}{2x^2+3} = 1$.
Because it is not at all obvious in advance that $\max\{1, \frac{2}{\varepsilon}\}$ would be a good choice for H, it is a good thing that the preparatory work for this proof was preserved.

Draft:

The question is whether for $\varepsilon > 0$ a number H exists such that for every $x > H$ the following holds:

$$\left|\frac{2x^2+3x+4}{2x^2+3} - 1\right| < \varepsilon.$$

Now we can see that

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

provided we take at least $x > -\frac{1}{3}$ because only then we are allowed to leave out the absolute value signs in the last step. This is not a problem as we are letting $x \to \infty$. If $\frac{3x+1}{2x^2+3} < \varepsilon$, then $|\frac{2x^2+3x+4}{2x^2+3} - 1| < \varepsilon$. In view of the way in which x occurs in the inequality $\frac{3x+1}{2x^2+3} < \varepsilon$, we are not capable of solving x from it in terms of ε. If we assume that $x > 1$ as well, i.e. if we take at least $H \geq 1$, we can proceed as follows.

For every $x > 1$ it holds that

$$\frac{3x+1}{2x^2+3} < \frac{3x+1}{2x^2} \leq \frac{3x+x}{2x^2} = \frac{4x}{2x^2} = \frac{2}{x}.$$

Therefore

$$\text{if} \quad \frac{2}{x} < \varepsilon, \quad \text{then} \quad \frac{3x+1}{2x^2+3} < \varepsilon,$$

in other words

$$\text{if} \quad x > \frac{2}{\varepsilon}, \quad \text{then} \quad \frac{3x+1}{2x^2+3} < \varepsilon.$$

So, the problem has been solved if we take $x > 1$ *and* $x > \dfrac{2}{\varepsilon}$. This is the reason why we chose $H = \max\{1, \frac{2}{\varepsilon}\}$.

By increasing the numerator in $\frac{3x+1}{2x^2+3}$ and decreasing the denominator we have obtained an expression which 'majors' the original expression, thus allowing x to be solved in terms of ε. Finding expressions which 'major' original expressions can be done in several ways. ◁

Hopefully this example made clear to you that you can never perform mathematics without a scratch pad.

EXERCISE 2.3 Use the definition of a limit to prove that

a) $\lim\limits_{x\to\infty} \dfrac{2x}{x+2} = 2$

b) $\lim\limits_{x\to\infty} \dfrac{1-x^2}{2x^2+3} = -\frac{1}{2}$.

2.2 Arithmetic rules for limits of functions in infinity

We can formulate some arithmetic rules for limits of functions in infinity. For example, the sum of two functions with a limit in infinity has itself a limit in infinity, and the limit of the sum of the functions is the sum of the limits. Similar statements can be made about the product and quotient of functions with a limit in infinity. As a result of this, functions with a limit in infinity produce new functions with a limit in infinity. But conversely, the arithmetic rules also allow us to show how we can reduce the study of the existence of a limit in infinity for a complicated function to a series of studies of the existence of a limit in infinity for less complicated functions.

EXAMPLE 2.3 We want to prove that the function $\frac{x+\sqrt{x}}{x\sqrt{x}}$ converges to 0 in infinity. We can do this by means of the definition of the limit in infinity, but we can also break down the function into convergent functions. To this purpose we rewrite the function as follows:

$$\frac{x+\sqrt{x}}{x\sqrt{x}} = \frac{1}{\sqrt{x}} + \frac{1}{x}.$$

So the function $\frac{x+\sqrt{x}}{x\sqrt{x}}$ can be seen as the sum of the function $\frac{1}{\sqrt{x}}$ and the function $\frac{1}{x}$.
In section 2.1 it was proved that the function $\frac{1}{\sqrt{x}}$ converges to 0 in infinity and in Exercise 2.2 that the function $\frac{1}{x}$ also converges to 0 in infinity. Now if we would know that the sum of two convergent functions in infinity is itself a convergent function in infinity, and that the limit of the sum of the functions is the sum of the limits of the functions, then the function $\frac{x+\sqrt{x}}{x\sqrt{x}}$ is convergent in infinity with limit 0. The following theorem provides a solution. $\triangleleft$

THEOREM 2.1 **(Arithmetic rules for limits of functions in infinity)**
For every pair of functions f and g and every pair of numbers $l, m \in \mathbb{R}$ the following holds: if $\lim\limits_{x\to\infty} f(x) = l$ and $\lim\limits_{x\to\infty} g(x) = m$, then
a) $\lim\limits_{x\to\infty} (f+g)(x) = l+m$,
b) $\lim\limits_{x\to\infty} (f\cdot g)(x) = l\cdot m$,
c) $\lim\limits_{x\to\infty} \dfrac{f}{g}(x) = \dfrac{l}{m}$ *(provided that $m \neq 0$ and $g(x) \neq 0$ for every $x \in D_g$).*

PROOF Suppose that $\lim\limits_{x\to\infty} f(x) = l$ and $\lim\limits_{x\to\infty} g(x) = m$.
a) Let $\varepsilon > 0$.
Since $\lim\limits_{x\to\infty} f(x) = l$, for $\frac{\varepsilon}{2} > 0$ a number H' exists such that $|f(x)-l| < \frac{\varepsilon}{2}$ for all $x > H'$.
Since $\lim\limits_{x\to\infty} g(x) = m$, for $\frac{\varepsilon}{2} > 0$ a number H'' exists such that $|g(x)-m| < \frac{\varepsilon}{2}$ for all $x > H''$.
Choose $H = \max\{H', H''\}$. Then for $x > H$

$$\begin{aligned} |(f(x)+g(x))-(l+m)| &= |(f(x)-l)+(g(x)-m)| \\ &\leq |f(x)-l|+|g(x)-m| \\ &< \frac{\varepsilon}{2}+\frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

This proves part a).

This proof too was prepared by us in a draft version.

> **Draft:**
>
> For $\varepsilon > 0$ we must find an H such that $|(f(x)+g(x))-(l+m)| < \varepsilon$ for every $x > H$. According to the Triangle Inequality the following is true for every $x \in D_f \cap D_g$:
>
> $$|(f(x)+g(x))-(l+m)| \leq |f(x)-l| + |g(x)-m|.$$
>
> If $|f(x)-l| < \frac{\varepsilon}{2}$ and $|g(x)-m| < \frac{\varepsilon}{2}$, then $|(f(x)+g(x))-(l+m)| < \varepsilon$. The functions f and g are known to converge respectively to l and m in infinity, so that for sufficiently large x it indeed holds that $|f(x)-l| < \frac{\varepsilon}{2}$ and $|g(x)-m| < \frac{\varepsilon}{2}$.

b) Let $\varepsilon > 0$.

Since $\lim\limits_{x\to\infty} f(x) = l$, for $\varepsilon = 1$ a number H' exists such that $|f(x)-l| < 1$ for all $x > H'$. In particular we have $|f(x)| = |(f(x)-l)+l| \leq |f(x)-l| + |l| < 1+|l|$ for all $x > H'$.

Since $\lim\limits_{x\to\infty} f(x) = l$, a number H'' exists for $\frac{\varepsilon}{2|m|+1}$, such that $|f(x)-l| < \frac{\varepsilon}{2|m|+1}$ for all $x > H''$.

Since $\lim\limits_{x\to\infty} g(x) = m$, a number H''' exists for $\frac{\varepsilon}{2(1+|l|)}$, such that $|g(x)-m| < \frac{\varepsilon}{2(1+|l|)}$ for all $x > H'''$.

Choose $H = \max\{H', H'', H'''\}$ and let $x > H$. Then

$$\begin{aligned}
|f(x)g(x)-lm| &= |f(x)g(x)-f(x)m+f(x)m-lm| \\
&= |f(x)(g(x)-m)+m(f(x)-l)| \\
&\leq |f(x)|\ \ |g(x)-m| + |m|\ \ |f(x)-l| \\
&< (1+|l|)\frac{\varepsilon}{2(1+|l|)} + \frac{\varepsilon|m|}{2|m|+1} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

The product of the functions is therefore convergent in infinity and the limit is the product of the limits.

In order to understand such a proof, the scratch pad is once again indispensable.

Draft:

For $\varepsilon > 0$ we must determine an H such that $|f(x)g(x) - lm| < \varepsilon$ for all $x > H$. Since f and g converge respectively to l and m in infinity, we can make $|f(x) - l|$ and $|g(x) - m|$ as small as we like, simply by taking x large enough. These two expressions can be integrated into our calculations by means of a frequently used method called the 'Telescope Trick': we add the number 0 in the form $0 = -a + a$. For every $x \in D_f \cap D_g$ (here a equals $f(x)m$) the following holds:

$$\begin{aligned} |f(x)g(x) - lm| &= |f(x)g(x) - f(x)m + f(x)m - lm| \\ &= |f(x)(g(x) - m) + m(f(x) - l)|. \end{aligned}$$

Next we apply the Triangle Inequality,

$$|f(x)(g(x) - m) + m(f(x) - l)| \leq |f(x)|\ |g(x) - m| + |m|\ |f(x) - l|.$$

If $|f(x)|\ |g(x) - m| < \frac{\varepsilon}{2}$ and $|m|\ |f(x) - l| < \frac{\varepsilon}{2}$, then $|f(x)g(x) - lm| < \varepsilon$.
Since $|f(x) - l|$ can be made arbitrarily small by taking x large enough, we can manage that $|f(x) - l| < \frac{\varepsilon}{2|m|}$. But this only goes well if $|m| \neq 0$. To avoid this problem we see to it that $|f(x) - l| < \frac{\varepsilon}{2|m|+1}$. (Question: why does $|m| = 0$ no longer present a problem now?)
Now we would like to see that similarly $|g(x) - m| < \frac{\varepsilon}{2|f(x)|}$, but that does not work because $f(x)$ has a value which varies with x and is therefore not constant. However, for large x the function f is bounded: we can find a number H' such that $|f(x)| \leq 1 + |l|$ for every $x > H'$. Then, if we can arrange (by taking x large enough) that $|g(x) - m| < \frac{\varepsilon}{2(1+|l|)}$, then we also have that $|f(x)|\ \ |g(x) - m| < \frac{\varepsilon}{2}$. This concludes the preparatory 'calculating/thinking behind the screens', known as draft.

The proof of part c) is omitted. $\triangleleft$

EXERCISE 2.4 Let f be convergent in infinity with limit l and g convergent in infinity with limit m.
Prove that for every $\lambda, \mu \in \mathbb{R}$ the function $\lambda f + \mu g$ is convergent in infinity with limit $\lambda l + \mu m$.
(Clue: interpret the function λf as the product of the constant function λ and the function f and apply Theorem 2.1.)

EXAMPLE 2.4 By means of the arithmetic rules for limits of functions in infinity we prove that the function

$$\frac{2x+4x^2}{x^2+1}$$

is convergent in infinity with limit 4.

If we divide the numerator and the denominator of the function by x^2 (for example for any $x > 0$), then we can rewrite the function as

$$\frac{2x+4x^2}{x^2+1} = \frac{\frac{2}{x}+4}{1+\frac{1}{x^2}}.$$

The function $\frac{2x+4x^2}{x^2+1} = \frac{\frac{2}{x}+4}{1+\frac{1}{x^2}}$ can be decomposed in the following manner for the purpose of the study of its convergence in infinity:

According to Exercise 2.2 the function $\frac{1}{x}$ has limit 0 in infinity. As a consequence, according to part b) of Theorem 2.1, we have that $\frac{1}{x^2} = \frac{1}{x} \cdot \frac{1}{x}$ has limit $0 \cdot 0 = 0$ in infinity. Of course the constant functions 1, 2 and 4 have limits 1, 2 and 4 respectively in infinity. Again according to part b) of Theorem 2.1 we have that $\frac{2}{x} = 2 \cdot \frac{1}{x}$ has limit $2 \cdot 0 = 0$ in infinity. Now according to part a) of Theorem 2.1 the function $\frac{2}{x}+4$ converges to $0+4$ in infinity and the function $1+\frac{1}{x^2}$ to $1+0=1$. Finally, according to part c) of Theorem 2.1 the function $\frac{2x+4}{1+\frac{1}{x^2}}$ converges to $\frac{4}{1} = 4$. ◁

EXERCISE 2.5 Prove that the following functions are convergent in infinity and determine their limits:

a) $f(x) = \dfrac{2x-1}{x+1}$

b) $g(x) = \dfrac{2x}{3x^2+1}$

c) $h(x) = -\dfrac{1}{3x+1}$.

EXERCISE 2.6 Let f be a convergent function in infinity with limit l. Prove that the function g, defined by $g(x) = f(x) + \frac{f(x)}{x}$, converges to l in infinity.

There are more operations on functions under which the convergence of the function is preserved, for example taking the square root or the absolute value.

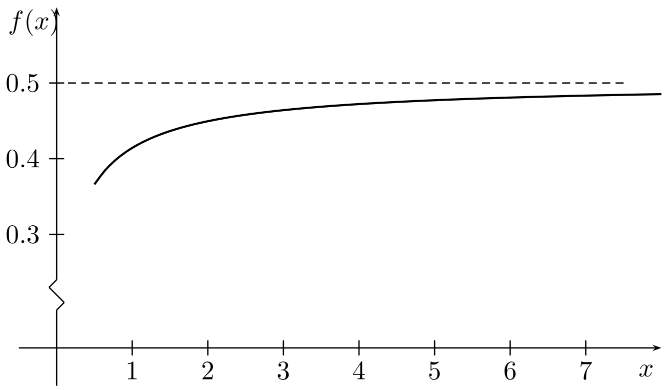

Figure 2.3: The graph of the function $f(x) = \sqrt{x^2 + x} - x$

EXERCISE 2.7 Prove the following statements.

a) If f is a non-negative function converging to $m \geq 0$ in infinity, then the function g, defined by $g(x) = \sqrt{f(x)}$ for every $x \in D_f$, converges to $\sqrt{m}$ in infinity.

b) If f is a convergent function in infinity with limit l, then the function g, defined by $g(x) = |f(x)|$ for every $x \in D_f$, converges to $|l|$ in infinity.

EXAMPLE 2.5 In this example we investigate the convergence in infinity of the function f on $(0, \infty)$, defined by

$$f(x) = \sqrt{x^2 + x} - x,$$

for every $x \in (0, \infty)$, and in the case of convergence we wish to know what its limit is.
To get a rough idea of the behaviour of the function we first calculate a few of its values:

$$f(1) = 0.414\ldots,\ f(2) = 0.449\ldots,\ f(3) = 0.464\ldots,\ f(4) = 0.472\ldots, f(5) = 0.477\ldots\ .$$

In Figure 2.3 the initial part of the graph has been sketched. The diminishing steepness suggests that the function reaches a 'ceiling', but such observations may be deceptive. That is why we calculate some more values, with higher arguments:

$$f(10) = 0.488\ldots,\ f(50) = 0.497\ldots,\ f(100) = 0.498\ldots.$$

Now it seems as though for large values of x, $f(x)$ is just a little less than $\frac{1}{2}$. We start to conjecture that $\lim\limits_{x\to\infty} f(x) = \frac{1}{2}$. We can prove this conjecture by means of the so-called *square-root-trick*. If we simultaneously multiply and divide $f(x) = \sqrt{x^2+x} - x$ by $\sqrt{x^2+x} + x$, then in the numerator we obtain a product of the form $(a-b)(a+b)$ which equals $a^2 - b^2$. Let's see where this will bring us:

$$f(x) = \left(\sqrt{x^2+x} - x\right) \cdot \left(\frac{\sqrt{x^2+x}+x}{\sqrt{x^2+x}+x}\right) = \frac{x^2+x-x^2}{\sqrt{x^2+x}+x} = \frac{x}{\sqrt{x^2+x}+x}.$$

Finally, if we divide numerator and denominator by x then we get

$$f(x) = \frac{1}{\sqrt{1+\frac{1}{x}}+1}.$$

Using the latter expression for $f(x)$, the function $f(x)$ can be decomposed in the following manner in order to study its convergence: the function $\frac{1}{x}$ converges in infinity with limit 0 according to Exercise 2.2 a), the function $1 + \frac{1}{x}$ converges in infinity with limit $1 + 0 = 1$ according to part a) of Theorem 2.1, the function $\sqrt{1+\frac{1}{x}}$ converges in infinity with limit $\sqrt{1} = 1$ according to part a) of Exercise 2.7), the function $\sqrt{1+\frac{1}{x}}+1$ converges in infinity with limit $1 + 1 = 2$ according to part a) of Theorem 2.1, and finally the function $\frac{1}{\sqrt{1+\frac{1}{x}}+1}$ converges in infinity with limit $\frac{1}{2}$ according to part c) of Theorem 2.1.

EXERCISE 2.8 Prove that the function f, defined by $f(x) = \dfrac{x}{\sqrt{|2x^2-20x+49|}}$, converges in infinity and determine the limit.

2.3 Limit of a function in a point

A function which is defined in the neighborhood of $x = 1$ by $\dfrac{3x-3}{\sqrt{x}-1}$, cannot be defined by this expression at $x = 1$. For, at $x = 1$ this expression is of the type $\frac{0}{0}$ and as such indefinite. Now the question arises how the function 'behaves' in the neighborhood of the point $x = 1$.

The behavior of a function in the neighborhood of a point c can be described by means of the concept of limit of a function in a point. The idea behind the assertion 'a function f has the limit l in the point c' is, that the values $f(x)$ lie arbitrarily close to l for *every* point x which is close enough to c.

DEFINITION Let f be a function and let $c \in \mathbb{R}$ be such that an $a < c$ exists with $(a, c) \subset D_f$ or a $b > c$ exists with $(c, b) \subset D_f$ (remember that D_f is the domain of f).
A real number l is called the *limit* of the function f in c if for every $\varepsilon > 0$ a $\delta > 0$ exists such that

$$\text{for every } x \in D_f, \quad \text{for which} \quad c - \delta < x < c + \delta \quad \text{and} \quad x \neq c, \quad \text{we have} \quad |f(x) - l| < \varepsilon.$$

If l is the limit of a function f in c, then we say that f *approaches* l *as* x *approaches* c or that f *converges to* l *in* c. This is written as

$$\lim_{x \to c} f(x) = l \qquad \text{or} \qquad f(x) \to l \;\text{ as }\; x \to c.$$

We call a function *convergent in* c, if the function has a limit in c.

We want to make some remarks regarding this definition.

1. In quantifier-language the statement '$\lim_{x \to c} f(x) = l$' reads as follows:

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in D_f \text{ with } c - \delta < x < c + \delta \text{ and } x \neq c : |f(x) - l| < \varepsilon.$$

2. According to the definition of the limit of a function in a point, we must determine for *every* $\varepsilon > 0$ a $\delta > 0$, such that $|f(x) - l| < \varepsilon$ for every $x \in D_f$, for which $c - \delta < x < c + \delta$ and $x \neq c$.
Obviously in general this δ depends on the choice of ε.
In Figure 2.4 for some value of ε it has been indicated which δ satisfies the requirements which are stated in the definition of the limit of a function in a point. In general it holds that the smaller values we take for ε, the smaller δ must be chosen.

3. Of course the concept of 'limit of a function in a point' only has a meaning if in every neighborhood of c there are points of the domain of f. In order to guarantee this it is stated at the beginning of the definition that f (a function) and $c \in \mathbb{R}$ are such that an $a < c$ exists with $(a, c) \subset D_f$ or a $b > c$ exists with $(c, b) \subset D_f$. In this way in every neighborhood of c there will be points $x \in D_f$ to the left of c or to the right of c or possibly to both sides of c.

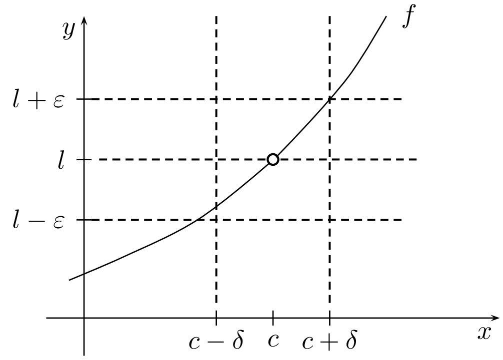

Figure 2.4: The limit of f in c equals l.

If in the following a limit of a function f in a point c is mentioned, it is silently assumed that an $a < c$ or $b > c$ exists such that $(a, c) \subset D_f$ or $(c, b) \subset D_f$. The examples and exercises of this reader are such that this condition is always fulfilled:
if $[0, \infty)\backslash\{1\}$ is the domain of definition of the function f defined by $f(x) = \frac{3x-3}{\sqrt{x}-1}$, then a $b > 1$ exists such that $(1, b) \subset D_f$.
4. If we examine the definition of the limit of a function in a point carefully, then we can see that it is not necessary for c to belong to the domain of f. Moreover, if c does belong to the domain of f, then still $f(c)$ plays no role whatsoever in the limit's definition.
5. Instead of writing '$x \in D_f$, for which $c-\delta < x < c+\delta$ and $x \neq c$,' we may also write '$x \in D_f$, for which $0 < |x-c| < \delta$'.

EXAMPLE 2.6 We will show that the function f on $[0, \infty)\backslash\{1\}$, defined by

$$f(x) = \frac{3x-3}{\sqrt{x}-1},$$

has the limit 6 at 1.
Let $\varepsilon > 0$. Take $\delta = \frac{1}{3}\varepsilon$ (see 'draft') and let $x \in D_f$, for which $0 < |x-1| < \delta$. We then have subsequently

$$\begin{aligned}\Big|\frac{3x-3}{\sqrt{x}-1} - 6\Big| &= \Big|\frac{3(x-1)}{\sqrt{x}-1} - 6\Big| \\ &= \Big|\frac{3\left(\sqrt{x}-1\right)\left(\sqrt{x}+1\right)}{\sqrt{x}-1} - 6\Big|\end{aligned}$$

$$
\begin{aligned}
&= \left|3\left(\sqrt{x}+1\right)-6\right| \\
&= \left|3\left(\sqrt{x}-1\right)\right| \\
&= \left|3\frac{x-1}{\sqrt{x}+1}\right| \\
&\leq 3|x-1| < 3\delta = \varepsilon.
\end{aligned}
$$

Thus we have proved that $\lim\limits_{x\to 1} \dfrac{3x-3}{\sqrt{x}-1} = 6$.

Of course this proof has been carefully prepared before it was put to paper in its present form. These considerations can be found in the draft which we also give here for the sake of completeness. It contains the reasoning which teaches us what could be an adequate choice for δ once ε is chosen.

Draft:

For $\varepsilon > 0$ we must determine a $\delta > 0$ such that the following holds for all $x \in D_f$, for which $1-\delta < x < 1+\delta$ and $x \neq 1$, in other words for all x with $0 < |x-1| < \delta$:

$$\left|\frac{3x-3}{\sqrt{x}-1} - 6\right| < \varepsilon.$$

Now we have for every $x \geq 0, x \neq 1$,

$$
\begin{aligned}
\left|\frac{3x-3}{\sqrt{x}-1} - 6\right| &= \left|\frac{3(x-1)}{\sqrt{x}-1} - 6\right| \\
&= \left|\frac{3\left(\sqrt{x}-1\right)\left(\sqrt{x}+1\right)}{\sqrt{x}-1} - 6\right| \\
&= \left|3\left(\sqrt{x}+1\right)-6\right| \\
&= \left|3\left(\sqrt{x}-1\right)\right|.
\end{aligned}
$$

Now if $|3\left(\sqrt{x}-1\right)| < \varepsilon$, then $\left|\frac{3x-3}{\sqrt{x}-1} - 6\right| < \varepsilon$. It is immediately clear that, if x is sufficiently close to 1, $|3(\sqrt{x}-1)|$ is approximately equal to 0. The only question is how close x must be to 1, so that $|3(\sqrt{x}-1)| < \varepsilon$. In order to answer that question we conveniently discover the presence of the factor $|x-1|$ in the expression $|3(\sqrt{x}-1)|$. For this is the factor which measures the distance separating x from 1. We proceed as follows:

$$\left|3\left(\sqrt{x}-1\right)\right| = \left|3\frac{\left(\sqrt{x}-1\right)\left(\sqrt{x}+1\right)}{\left(\sqrt{x}+1\right)}\right| = 3\frac{|x-1|}{\sqrt{x}+1}.$$

Now $\sqrt{x}+1 \geq 1$ for all $x \geq 0$, so that $3\frac{|x-1|}{\sqrt{x}+1} \leq 3|x-1|$. So if $3|x-1| < \varepsilon$, then $3\frac{|x-1|}{\sqrt{x}+1} < \varepsilon$, and therefore $|3(\sqrt{x}-1)| < \varepsilon$ and also $\left|\frac{3x-3}{\sqrt{x}-1} - 6\right| < \varepsilon$. This shows us that by choosing $\delta = \frac{1}{3}\varepsilon$ we have for all $x \in D_f$, for which $0 < |x-1| < \delta$, that $3|x-1| < \varepsilon$, and that this choice therefore also satisfies the original inequality. $\triangleleft$

EXAMPLE 2.7 We will prove that the function f, defined by $f(x) = x^3$, has the limit c^3 at c for every $c \in \mathbb{R}$.

Let $c \in \mathbb{R}$ and $\varepsilon > 0$. Call $H = (1+|c|)^2 + (1+|c|)|c| + |c^2|$. Take $\delta = \min\{1, \frac{\varepsilon}{H}\}$ and let $x \in \mathbb{R}$, for which $0 < |x-c| < \delta$. We then have subsequently

$$\begin{array}{rcl} |x^3 - c^3| & = & |x-c|\,|x^2+xc+c^2| \\ & \leq^* & |x-c|\,\left(|x^2|+|x||c|+|c^2|\right) \\ & \leq^{**} & |x-c|\,\left((1+|c|)^2+(1+|c|)|c|+|c^2|\right) \\ & < & \varepsilon, \end{array}$$

where in * the Triangle Inequality was applied to the second factor and in ** the inequality $|x| = |x-c+c| \leq |x-c|+|c| < 1+|c|$ was used.

For the sake of completeness we will add to the proof the draft that was made in preparation for this proof.

Draft:

Once more we need to determine for $\varepsilon > 0$ a $\delta > 0$ such that $|x^3 - c^3| < \varepsilon$ for all $x \in \mathbb{R}$, for which $0 < |x-c| < \delta$. It is at once clear that if x lies sufficiently close to c, then $|x^3-c^3|$ is approximately equal to 0. But the question is, how close must x lie to c in order for $|x^3-c^3|$ to be smaller than the arbitrarily chosen positive number ε. Therefore we wish to recognize the factor $|x-c|$ in $|x^3-c^3|$. For this reason we break down $|x^3-c^3|$ into two factors

$$|x^3 - c^3| = |x-c||x^2+xc+c^2|.$$

Now the important thing to note is that the second factor $|x^2+xc+c^2|$ is bounded if x lies sufficiently close to c. For if we take an $x \in \mathbb{R}$ such that $0 < |x-c| < \delta$

with for instance $\delta \le 1$*, then* $|x| = |x-c+c| \le |x-c| + |c| \le 1 + |c|$ *so that*

$$|x^2 + xc + c^2| \le |x^2| + |x||c| + |c^2| \le (1+|c|)^2 + (1+|c|)|c| + |c^2|.$$

Therefore we definitely take a δ *no bigger than* 1*.*

Now if $|x-c|\left((1+|c|)^2 + (1+|c|)|c| + |c^2|\right) < \varepsilon$*, then* $|x^3 - c^3| < \varepsilon$*, in other words, if* $|x-c| < \frac{\varepsilon}{(1+|c|)^2+(1+|c|)|c|+|c^2|}$*, then* $|x^3 - c^3| < \varepsilon$*.*

This explains the choice $\delta = \min\{1, \frac{\varepsilon}{(1+|c|)^2+(1+|c|)|c|+|c^2|}\}$*.* ◁

EXERCISE 2.9 Use the definition of limit to prove that

a) $\lim\limits_{x\to 0}(6x^2 + 4x^3 + x^4) = 0$

b) $\lim\limits_{x\to 0} \dfrac{(2+3x)^2 - 4}{2x} = 6$

c) $\lim\limits_{x\to c} \dfrac{1}{x} = \dfrac{1}{c} \quad (c > 0)$.

EXERCISE 2.10 Let $\lim\limits_{x\to c} f(x) = l > 0$. Prove that an $\alpha > 0$ exists such that $f(x) > 0$ for every $x \in D_f$ for which $0 < |x-c| < \alpha$.

As was stated at the beginning of the definition of the limit of a function in a point c, f must be defined on an interval to the left of c or on an interval to the right of c. In the example below we will consider a function that is only defined on an interval to the right of c.

EXAMPLE 2.8 We will prove that the function f on $(0,\infty)$, defined by

$$f(x) = \frac{x\sqrt{x}}{x+\sqrt{x}},$$

has the limit 0 in 0.

Let $\varepsilon > 0$. Take $\delta = \varepsilon$ and let $x \in D_f$, for which $0 < x < \delta$. Then the following holds:

$$\left|\frac{x\sqrt{x}}{x+\sqrt{x}} - 0\right| = \frac{x\sqrt{x}}{x+\sqrt{x}} = \frac{x}{\sqrt{x}+1} < x < \delta = \varepsilon.$$

This proves that $\lim\limits_{x\to 0} \dfrac{x\sqrt{x}}{x+\sqrt{x}} = 0$.

The corresponding draft looks as follows.

Draft:

For $\varepsilon > 0$ *we must determine a* $\delta > 0$ *such that* $\left|\frac{x\sqrt{x}}{x+\sqrt{x}} - 0\right| < \varepsilon$ *for all* $x \in D_f$*, for which* $0 < x < \delta$*. Now we have* $\left|\frac{x\sqrt{x}}{x+\sqrt{x}} - 0\right| = \left|\frac{x}{\sqrt{x}+1} - 0\right| = \frac{x}{\sqrt{x}+1}$*. It is* *at once clear that if x lies sufficiently close to 0, then $\frac{x}{\sqrt{x}+1}$ is approximately equal to 0. The only question is how close must x lie to 0 so that $\frac{x}{\sqrt{x}+1} < \varepsilon$. In the expression $\frac{x}{\sqrt{x}+1}$ we can recognize the factor x*

$$\frac{x}{\sqrt{x}+1} = x\frac{1}{\sqrt{x}+1}.$$

Now the important thing to note is that the second factor $\frac{1}{\sqrt{x}+1}$ is bounded from above by 1 for every $x \in D_f$, so that $x\frac{1}{\sqrt{x}+1} < x$ for every $x \in D_f$. If now $0 < x < \varepsilon$, then we certainly have that $\frac{x}{\sqrt{x}+x} < \varepsilon$ and therefore that $|f(x) - 0| < \varepsilon$. This explains the choice $\delta = \varepsilon$. ◁

EXERCISE 2.11

a) The function f on $\mathbb{R}$ is defined by $f(x) = \max\{-5x, 3x\}$. Prove by means of the definition of limit of a function that $\lim_{x\to 0} f(x) = 0$.

b) The function f on $\mathbb{R}$ is defined by

$$f(x) = \begin{cases} 3x+2 & \text{if } x \leq 1 \\ x^2+3x+1 & \text{if } x > 1. \end{cases}$$

Prove that $\lim_{x\to 1} f(x)$ exists.

2.4 Arithmetic rules for limits of functions in a point

Limits of a function in a point have similar properties as limits of functions in infinity. For example, similar arithmetic rules (see Theorem 2.1) are valid. In this way we are able to avoid the sometimes laborious method of using the (ε, δ)-definition in order to show that a function is having a limit in some point. The proof of the theorem below is similar to the proof of Theorem 2.1 and therefore omitted.

THEOREM 2.2 (Arithmetic rules for limits of functions)

For every pair of functions f and g the following holds:

if $\lim_{x\to c} f(x) = l$ and $\lim_{x\to c} g(x) = m$, then

a) $\lim_{x\to c}(f+g)(x) = l + m$,

b) $\lim_{x\to c}(f \cdot g)(x) = l \cdot m$,

c) $\lim_{x\to c} \dfrac{f}{g}(x) = \dfrac{l}{m}$ *(provided that $m \neq 0$ and $g(x) \neq 0$ for every $x \in D_g$).*

EXAMPLE 2.9 We will show that

$$\lim_{x\to 2}\frac{x+3}{x^2+1}=1.$$

Using the (ε,δ)-definition of limit in a point it is straightforward to see that $\lim_{x\to 2} x = 2$, $\lim_{x\to 2} 1 = 1$ and $\lim_{x\to 2} 3 = 3$ (please check this). Using part a) of Theorem 2.2 we get $\lim_{x\to 2}(x+3) = 2+3 = 5$. Using part b) of Theorem 2.2 we get $\lim_{x\to 2} x^2 = \lim_{x\to 2} x\cdot x = 2\cdot 2 = 4$. Again, using part a) of Theorem 2.2 we get $\lim_{x\to 2}(x^2+1) = 4+1 = 5$. Finally, using part c) of Theorem 2.2 we get $\lim_{x\to 2}\frac{x+3}{x^2+1} = \frac{5}{5} = 1$. $\triangleleft$

EXERCISE 2.12 Formulate statements for limits of functions in a point similar to the statements for limits of functions in infinity in Exercise 2.7.

EXERCISE 2.13 Prove that the following limits exist and determine their values:
a) $\lim_{x\to 2}\dfrac{2x^2-3}{4x^2+3}$
b) $\lim_{x\to 1}(x^2-1)\sqrt{x+1}$.

2.5 Extensions of the concept of limit

In this chapter we have considered two types of limits: limits in infinity and limits in a point. There are more types of asymptotic behavior, for example 'diverging to infinity'.

DEFINITION Let f be a function such that an $a\in\mathbb{R}$ exists with $(a,\infty)\subset D_f$.
The function f is *divergent to infinity if x tends to infinity* if for every $u\in\mathbb{R}$ a real number H exists such that

$$f(x) > u \quad \text{for every} \quad x\in D_f, \quad \text{for which} \quad x > H.$$

We write this as

$$\lim_{x\to\infty} f(x) = \infty \qquad \text{or} \qquad f(x)\to\infty \quad \text{if} \quad x\to\infty.$$

EXAMPLE 2.10 We will show that

$$\lim_{x\to\infty}\frac{x^2+1}{x+3}=\infty.$$

Let $u \in \mathbb{R}$. Choose $H = \max\{3, 2u\}$. For every $x \in \mathbb{R}$, $x > H$ we have $x > 3$, so

$$\frac{x^2+1}{x+3} > \frac{x^2}{x+3} > \frac{x^2}{x+x} = \frac{x^2}{2x} = \frac{1}{2}x > \frac{1}{2}H \geq \frac{1}{2} \cdot 2u = u.$$

$\triangleleft$

EXERCISE 2.14 Show that $\lim_{x\to\infty} \frac{\sqrt{x^4+1}}{x} = \infty$.

Up to now we have considered three types of limits. All these types have the format

$$\lim_{x\to c} f(x) = l.$$

In section 2.1 we considered 'limits in infinity', i.e. we considered the situation where $c = \infty$ and $l \in \mathbb{R}$, in section 2.3 we considered 'limits in a point', i.e. $c \in \mathbb{R}$ and $l \in \mathbb{R}$, and in this section we considered 'diverging to infinity', i.e. $c = l = \infty$. Any choice for c (either $c = -\infty$ or $c \in \mathbb{R}$ or $c = \infty$) and any choice for l (either $l = -\infty$ or $l \in \mathbb{R}$ or $l = \infty$) corresponds to a different type of asymptotic behavior.

EXERCISE 2.15 Give the definition of $\lim_{x\to-\infty} f(x) = l$ (where $l \in \mathbb{R}$) and the definition of $\lim_{x\to\infty} f(x) = -\infty$.

EXERCISE 2.16

a) Use the definition to prove that

i) $\lim_{x\to\infty} \dfrac{x^2+1}{x^2+2} = 1$ ii) $\lim_{x\to\infty} \dfrac{x}{x^2+1} = 0$ iii) $\lim_{x\to-\infty} \dfrac{\sqrt{1-x}}{x} = 0$.

b) Determine, if possible, the following limits:

i) $\lim_{x\to\infty} \dfrac{3+x}{\sqrt{x}}$ ii) $\lim_{x\to\infty} \dfrac{\sqrt{x}-x}{\sqrt{x}+x}$.

2.6 Mixed exercises

EXERCISE 2.17 Show, by using the $(\varepsilon, H_\varepsilon)$-definition, that

$$\lim_{x\to\infty} \frac{3x+18}{x+1}$$

exists.

EXERCISE 2.18 The function f on $\mathbb{R}$ is defined by

$$f(x) = \begin{cases} x & \text{if } x > 1 \\ 2x^2 - 1 & \text{if } x \leq 1. \end{cases}$$

Use the definition to prove that $\lim_{x \to 1} f(x) = 1$.

EXERCISE 2.19 Determine, if possible, $\lim_{x \to \infty} \dfrac{x^2 - 4}{2(x-1)(x-2)}$.

3 CONTINUITY

The concept of 'limit of a function' is logically succeeded by the concept of 'continuity of a function'. In section 3.1 we start with the definition of the concept of continuity of a function in a point. Next we define what is meant by a continuous function. In section 3.2 we discuss some arithmetic rules for continuous functions and in section 3.3 we deduce properties of continuous functions on a compact interval (which is a bounded and closed interval). One of the results is known as the Theorem of Weierstrass and concerns the existence of a maximal and a minimal value of the function. It goes without saying that this theorem is widely applied, for instance in decision theory. In section 3.4 we show that the inverse of a continuous and invertible function is itself continuous.

3.1 Continuous functions

First we provide the definition of 'continuity of a function in a point' and of a 'continuous function'.

DEFINITION Let f be a function on an interval I and $c \in I$.
The function f is called *continuous in* c if for every $\varepsilon > 0$ a $\delta > 0$ exists such that

$$\text{for every } x \in I, \text{ for which } |x - c| < \delta, \text{ we have } |f(x) - f(c)| < \varepsilon.$$

The function f is called *continuous* if f is continuous in every point of I.

According to the definition of continuity of a function f in a point c it is required that f is defined in c (which is not the case for the limit of f in c, where f need not be defined in c). We therefore assume that f is a function on an interval I with $c \in I$. With the interval I as the domain of f we are also certain that an $a < c$ exists such that $(a, c) \subset D_f = I$, or a $b > c$ such that $(c, b) \subset D_f = I$. Therefore, we can examine the limit of f in c. It turns out that there is a link between the concepts of 'continuity of a function in a point' and of 'limit of a function in a point':

$$f \text{ is continuous in } c \text{ if and only if } \lim_{x \to c} f(x) = f(c).$$

For if $\lim_{x\to c} f(x) = f(c)$, then according to the definition of limit of a function in a point, for every $\varepsilon > 0$ a $\delta > 0$ exists such that $|f(x) - f(c)| < \varepsilon$ for every $x \in I$, for which $0 < |x - c| < \delta$. But since f is now also defined in c and $|f(x) - f(c)| = 0$ for $x = c$, we now have that $|f(x) - f(c)| < \varepsilon$ for every $x \in I$, for which $|x - c| < \delta$. So, f is continuous in c. If conversely f is continuous in c, then for every $\varepsilon > 0$ a $\delta > 0$ exists such that $|f(x)-f(c)| < \varepsilon$ for all $x \in I$, for which $|x - c| < \delta$. Then certainly $|f(x) - f(c)| < \varepsilon$ for all $x \in I$, for which $|x-c| < \delta$ and $x \neq c$ (in other words, for which $0 < |x-c| < \delta$). Therefore $\lim_{x\to c} f(x) = f(c)$.

EXAMPLE 3.1 The function f on $\mathbb{R}$, defined by $f(x) = x^3$, is continuous in 2 because $f(2) = 2^3 = 8 = \lim_{x\to 2} x^3$ (refer to Example 2.7). ◁

EXAMPLE 3.2 The function f on $[0, \infty)$, defined by

$$f(x) = \begin{cases} \dfrac{3x-3}{\sqrt{x}-1} & \text{if } x \neq 1 \\ 6 & \text{if } x = 1, \end{cases}$$

is continuous in 1, since $f(1) = 6 = \lim_{x\to 1} \dfrac{3x-3}{\sqrt{x}-1}$ (refer to Example 2.6). ◁

EXERCISE 3.1

a) The function f on $\mathbb{R}$ is defined by

$$f(x) = \begin{cases} \dfrac{(1+x)^2-1}{x} & \text{if } x \neq 0 \\ 2 & \text{if } x = 0. \end{cases}$$

Prove that f is continuous in 0.

b) The function g on $\mathbb{R}$ is defined by

$$g(x) = \begin{cases} x^3 & \text{if } x < 0 \\ \sqrt{x} & \text{if } x \geq 0. \end{cases}$$

Prove that g is continuous in 0.

The following example provides an example of a function that is not continuous.

EXAMPLE 3.3 We prove that the function f on $\mathbb{R}$, defined by

$$f(x) = \begin{cases} x^2+1 & \text{if } x \leq 2 \\ x+4 & \text{if } x > 2, \end{cases}$$

is not continuous in 2.

Note that $f(2) = 5$. Suppose that f is continuous in 2. Then for every $\varepsilon > 0$ there is a $\delta > 0$ such that $|f(x) - f(2)| < \varepsilon$ for every $x \in \mathbb{R}$ with $|x - 2| < \delta$. Stated differently, for every $\varepsilon > 0$ there is a $\delta > 0$ such that $f(x) \in (f(2) - \varepsilon, f(2) + \varepsilon)$ for every $x \in (2 - \delta, 2 + \delta)$. In particular this is true for, e.g., $\varepsilon = \frac{1}{2}$. So there is a $\delta > 0$ such that $f(x) \in (4.5, 5.5)$ for every $x \in (2-\delta, 2+\delta)$. However, for every $x \in (2, 2+\delta) \subset (2-\delta, 2+\delta)$ we have $f(x) = x+4 > 6$, so $f(x) \notin (4.5, 5.5)$. Contradiction. Therefore f is not continuous in 2. $\triangleleft$

EXERCISE 3.2 Prove that the function f on $[0, \infty)$, defined by

$$f(x) = \begin{cases} \dfrac{1}{x} & \text{if } x > 0 \\ 0 & \text{if } x = 0, \end{cases}$$

is not continuous in 0.

Because the concepts of 'continuity of a function in a point' and of 'limit of a function in a point' are linked, we are able to use our knowledge about limits of functions from chapter 2 in order to prove continuity of some functions.

EXAMPLE 3.4 We prove that the function f on $[0, \infty)$, defined by $f(x) = \sqrt{x}$, is a continuous function.

Let $c \geq 0$. Note that $\lim\limits_{x \to c} x = c$. According to Exercise 2.12 we have $\lim\limits_{x \to c} \sqrt{x} = \sqrt{c}$, or $\lim\limits_{x \to c} f(x) = f(c)$. So f is continuous in c. Since this is true for every $c \in [0, \infty)$ we can conclude that f is a continuous function. $\triangleleft$

3.2 Arithmetic rules for continuous functions

For continuous functions we can formulate arithmetic rules that are similar to those for limits of functions.

THEOREM 3.1 **(Arithmetic rules for continuous functions)**
For every pair of functions f and g on an interval I containing c the following holds: if f and g are continuous in c, then

a) *the sum function $f+g$ is continuous in c,*

b) *the product function $f \cdot g$ is continuous in c,*

c) *the quotient function $\frac{f}{g}$ is continuous in c (provided that $g(x) \neq 0$ for every $x \in I$).*

PROOF We only give the proof of part a) and leave the proofs of parts b) and c) to the reader.
Assume that f and g are continuous in $c \in I$. Then $\lim_{x \to c} f(x) = f(c)$ and $\lim_{x \to c} g(x) = g(c)$. So according to the arithmetic rules for limits of functions in a point (refer to Theorem 2.2)

$$\begin{aligned} \lim_{x \to c} (f+g)(x) &= \lim_{x \to c} (f(x) + g(x)) \\ &= \lim_{x \to c} f(x) + \lim_{x \to c} g(x) \\ &= f(c) + g(c) = (f+g)(c). \end{aligned}$$

Therefore the sum function $f+g$ is continuous in c. $\triangleleft$

In the following theorem it is claimed that the composite function of two continuous functions is itself continuous. We provide this theorem without proof.

THEOREM 3.2 *Let f be a function on an interval I, and let g be a function on an interval J, such that $g(J) \subseteq I$ and $c \in J$. Then the following holds: if g is continuous in c and f is continuous in $g(c)$, then the composite function $f \circ g$ is continuous in c.*

> **EXERCISE 3.3** The functions f and g are defined on $[0,1]$ respectively by $f(x) = \dfrac{1-x}{1+x}$ and $g(x) = 4x(1-x)$. Prove that $\lim_{x \to \frac{1}{2}} f \circ g(x) = 0$.

EXAMPLE 3.5 Let g be a continuous function on an interval J, which is non-negative, in other words, $g(x) \geq 0$ for every $x \in J$. We prove that the function $\sqrt{g}$ on J, defined by $(\sqrt{g})(x) = \sqrt{g(x)}$, is also continuous.
The root function f on $[0, \infty)$, defined by $f(x) = \sqrt{x}$, is continuous, and $g(J) \subseteq [0, \infty)$. According to Theorem 3.2, applied to f and g, the composite function $f \circ g$ on J, given by

$(f \circ g)(x) = f(g(x)) = \sqrt{g(x)}$, is continuous. ◁

EXERCISE 3.4 Let g be a continuous function on an interval J. Define the function $|g|$ on J by $|g|(x) = |g(x)|$.

Prove that $|g|$ is continuous.

3.3 Continuous functions on an interval

A *compact interval* is an interval that is both bounded and closed, so an interval of the form $[a, b]$ with $a, b \in \mathbb{R}$. Continuous functions on a compact interval have some important properties:

a) a continuous function on a compact interval is bounded

b) a continuous function on a compact interval has a minimum and a maximum

c) every number between the minimal and the maximal value of a continuous function on a compact interval is a function value.

The first property is the property of boundedness of a continuous function on a compact interval, the second property is known as the 'extreme value property' or the theorem of Weierstrass, and the third as the 'intermediate value property'. For the proof of the first two theorems we make use of the Theorem of Bolzano-Weierstrass (Theorem 0.3): 'every bounded sequence has a convergent subsequence'.

First we again provide the definition of 'bounded function'.

DEFINITION A function f is *bounded on an interval* $I \subseteq D_f$ if a number p (> 0) exists such that

$$|f(x)| \leq p \quad \text{for every} \quad x \in I.$$

If I is the domain D_f of f, then we call f a *bounded function.*

A function f is therefore bounded on I (respectively bounded) if the range $f(I) = \{f(x) : x \in I\}$ (respectively R_f) is bounded.

THEOREM 3.3 *Every continuous function on a compact interval is bounded.*

PROOF Let f be a continuous function on the compact interval $I = [a, b]$. By means of a proof by contradiction we prove that f is bounded.
Assume that f is *not* bounded. On the basis of the definition of boundedness it then holds that

$$\neg\ \exists m\ (>0)\ \ \forall x \in I\ :\ |f(x)| \le m.$$

Now we have the following series of equivalent statements:

$$\begin{array}{lccc} & \neg\ \exists m > 0 & \forall x \in I\ : & |f(x)| \le m \\ \Leftrightarrow & \forall m > 0 & \neg\ \forall x \in I\ : & |f(x)| \le m \\ \Leftrightarrow & \forall m > 0 & \exists x \in I\ : & \neg\ (|f(x)| \le m) \\ \Leftrightarrow & \forall m > 0 & \exists x \in I\ : & |f(x)| > m. \end{array}$$

In other words, for every number $m > 0$ an $x \in I = [a, b]$ exists such that $|f(x)| > m$. Now we choose subsequently $m = 1, 2, 3, \ldots$. Then we can find points $x_1, x_2, x_3, \ldots$, with $|f(x_1)| > 1$, $|f(x_2)| > 2$, $|f(x_3)| > 3, \ldots$. In this way a sequence $x_1, x_2, x_3, \ldots$ in $I = [a, b]$ is obtained (in other words, $a \le x_n \le b$ for every $n \in \mathbb{N}$) such that the sequence $f(x_1), f(x_2), f(x_3), \ldots$ is an unbounded sequence ($|f(x_n)| \ge n$ for every $n \in \mathbb{N}$).
According to Theorem 0.3, the bounded sequence $x_1, x_2, x_3, \ldots$ has a convergent subsequence $x_{n_1}, x_{n_2}, x_{n_3}, \ldots$, say with limit $c \in [a, b]$. Since f is continuous in c we have $\lim\limits_{k\to\infty} f(x_{n_k}) = f(c)$. This implies that $\lim\limits_{k\to\infty} |f(x_{n_k})| = |f(c)|$. However, this is inconsistent with the fact that $|f(x_{n_k})| > n_k \ge k$ for every $k \in \mathbb{N}$.
Therefore the assumed claim is not true, in other words, the function f is bounded. $\triangleleft$

In an example we show that in Theorem 3.3 neither the interval's property of being closed, nor it's property of being bounded can be omitted.

EXAMPLE 3.6 The function f on the (non-closed) interval $(0, 1]$, defined by $f(x) = \dfrac{1}{x}$, is continuous but not bounded.
The function g on the (non-bounded) interval $[0, \infty)$, defined by $g(x) = x$, is continuous but not bounded. $\triangleleft$

Now that we know that for a continuous function f on a compact interval $I = [a, b]$, the set of values $f(I)$ is bounded, $f(I)$ must have an infimum and a supremum according to the

Axiom of Completeness. If the infimum and supremum of $f(I)$ belong to the set $f(I)$ as well (i.e. they are function values themselves), then f has a minimum and a maximum on I.

DEFINITION Let f be a function on an interval I and let $c \in I$.
We call the value $f(c)$ the *minimum* of f if $f(x) \geq f(c)$ for every $x \in I$. The point c is called a *minimum location* of f.
We call the value $f(c)$ the *maximum* of f if $f(x) \leq f(c)$ for every $x \in I$. The point c is called a *maximum location* of f.
We call $f(c)$ an *extreme value* of f if $f(c)$ is the minimum or the maximum of f.

Since it must hold that $f(x) \geq f(c)$ or that $f(x) \leq f(c)$ *for every* $x \in I$, we also speak of *global extrema*.

THEOREM 3.4 **(Theorem of Weierstrass)**
For every continuous function f on a compact interval I, points $c, d \in I$ exist such that $f(c) \leq f(x) \leq f(d)$ for every $x \in I$.
In other words, every continuous function on a compact interval has a minimum and a maximum.

PROOF Let f be a continuous function on a compact interval $I = [a, b]$. Then according to Theorem 3.3, f is bounded. So we can define

$$\inf f(I) = \inf\{f(x) : x \in I\} \quad \text{and} \quad \sup f(I) = \sup\{f(x) : x \in I\}.$$

We must show that a $c \in I$ exists such that $f(c) = \inf f(I)$ and that a $d \in I$ exists such that $f(d) = \sup f(I)$. We only prove the existence of the maximum and leave the proof of the existence of the minimum to the reader.
Let $u = \sup f(I)$, i.e. u is the lowest upper bound of the set $f(I)$. For every $\varepsilon > 0$ we have that $u - \varepsilon < u$ so $u - \varepsilon$ is not an upper bound of $f(I)$. So for every $\varepsilon > 0$ an $x \in I$ exists such that $u - \varepsilon < f(x)$. For the consecutive values $\varepsilon = 1, \frac{1}{2}, \frac{1}{3}, \ldots$ therefore, points $x_1, x_2, x_3, \ldots$ exist with the property that $a \leq x_n \leq b$ and $u - \frac{1}{n} < f(x_n)$ $(\leq u)$ for every $n \in \mathbb{N}$.
It can easily be proved from the inequality $u - \frac{1}{n} < f(x_n) \leq u$ for every $n \in \mathbb{N}$, that the sequence of function values $f(x_1), f(x_2), f(x_3), \ldots$ converges to u.

Because $a \leq x_n \leq b$ for every $n \in \mathbb{N}$, according to Theorem 0.3 the sequence $x_1, x_2, x_3, \ldots$ has a convergent subsequence $x_{n_1}, x_{n_2}, x_{n_3}, \ldots$, say with limit $d \in [a, b]$. Since f is continuous in d, the sequence of function values $f(x_{n_1}), f(x_{n_2}), f(x_{n_3}), \ldots$ converges to $f(d)$.
But the sequence $f(x_{n_1}), f(x_{n_2}), f(x_{n_3}), \ldots$ also converges to u, since every subsequence of the convergent sequence $f(x_1), f(x_2), f(x_3), \ldots$ has the limit u.
It therefore follows that $f(d) = u$. This concludes the proof of the existence of the maximum of f. $\triangleleft$

EXERCISE 3.5 Prove that a continuous function on a compact interval I has a minimum.

EXAMPLE 3.7 The function f on $[0, 1]$, defined by $f(x) = x^3 - x + 1$ is continuous and therefore has a minimal and maximal value. Later we will discuss a method how to determine these values. $\triangleleft$

EXERCISE 3.6 This exercise shows that the condition in Theorem 3.4, that I must be a bounded and closed interval, cannot be omitted.
a) Verify that the function h on $[0, 1)$, defined by $h(x) = x^2 + 1$, does have a minimum, but no maximum.
b) Verify that the function k on $[0, \infty)$, defined by $k(x) = -x$, does have a maximum, but no minimum.

EXERCISE 3.7 Let f be a continuous function on a compact interval $I = [a, b]$, which is positive, in other words, $f(x) > 0$ for every $x \in I$. Prove that a number $\alpha > 0$ exists such that $f(x) \geq \alpha$ for every $x \in I$.

The third important theorem that we will prove is the Intermediate Value Theorem. In combination with the Theorem of Weierstrass, this theorem enables us to prove that the image of a compact interval under a continuous function is itself a compact interval (see Theorem 3.7).

THEOREM 3.5 **(Intermediate Value Theorem)**

For every continuous function f on a compact interval $I = [a, b]$ the following holds: for every number t between $f(a)$ and $f(b)$ a $c \in [a, b]$ exists with $f(c) = t$.
(that t lies between $f(a)$ and $f(b)$ means that $t \in [f(a), f(b)]$ if $f(a) \leq f(b)$ and $t \in [f(b), f(a)]$ if $f(a) > f(b)$.)

PROOF Let f be a continuous function on a compact interval $I = [a, b]$. Assume that $f(a) \leq f(b)$ and let $t \in [f(a), f(b)]$. If $t = f(a)$ or $t = f(b)$ then the proof is complete. Therefore assume that $f(a) < t < f(b)$. By means of the so-called method of interval halving we construct a sequence of intervals, all of which contain the point c we are looking for.

[step 1:]

Consider the point $x_1 = \frac{a+b}{2}$, which is the middle of the interval $I = [a, b]$.
If $f(x_1) = t$ then the proof of the theorem is complete with $c = x_1$.
If $f(x_1) \neq t$ then we define an interval $I_1 = [a_1, b_1]$ in the following way:
if $t < f(x_1)$ (so that $f(a) < t < f(x_1)$), then choose $a_1 = a$ and $b_1 = x_1$ (so that $I_1 = [a_1, b_1] = [a, x_1]$, the subinterval which is the left half of I) and
if $f(x_1) < t$ (so that $f(x_1) < t < f(b)$), then choose $a_1 = x_1$ and $b_1 = b$ (so that $I_1 = [a_1, b_1] = [x_1, b]$, the subinterval which is the right half of I).
In each of these cases it is true that $a \leq a_1 < b_1 \leq b$, $b_1 - a_1 = \frac{1}{2}(b-a)$ and $f(a_1) < t < f(b_1)$.
See Figure 3.1, where $a_1 = x_1$ and $b_1 = b$.

[step 2:]

Consider the point $x_2 = \frac{a_1+b_1}{2}$, which is the middle of the interval $I_1 = [a_1, b_1]$.
If $f(x_2) = t$ then the proof of the theorem is complete with $c = x_2$.
If $f(x_2) \neq t$ then we define the interval $I_2 = [a_2, b_2]$ in the following way:
if $t < f(x_2)$ (so that $f(a_1) < t < f(x_2)$), then choose $a_2 = a_1$ and $b_2 = x_2$ (so that $I_2 = [a_2, b_2] = [a_1, x_2]$, the subinterval which is the left half of I_1) and
if $f(x_2) < t$ (so that $f(x_2) < t < f(b_1)$), then choose $a_2 = x_2$ and $b_2 = b_1$ (so that $I_2 = [a_2, b_2] = [x_2, b_1]$, the subinterval which is the right half of I_1).
In each of these cases it is true that $a_1 \leq a_2 < b_2 \leq b_1$, $b_2 - a_2 = (\frac{1}{2})^2(b - a)$ and $f(a_2) < t < f(b_2)$.
See Figure 3.1 where $a_2 = x_1 = a_1$ and $b_2 = x_2$.

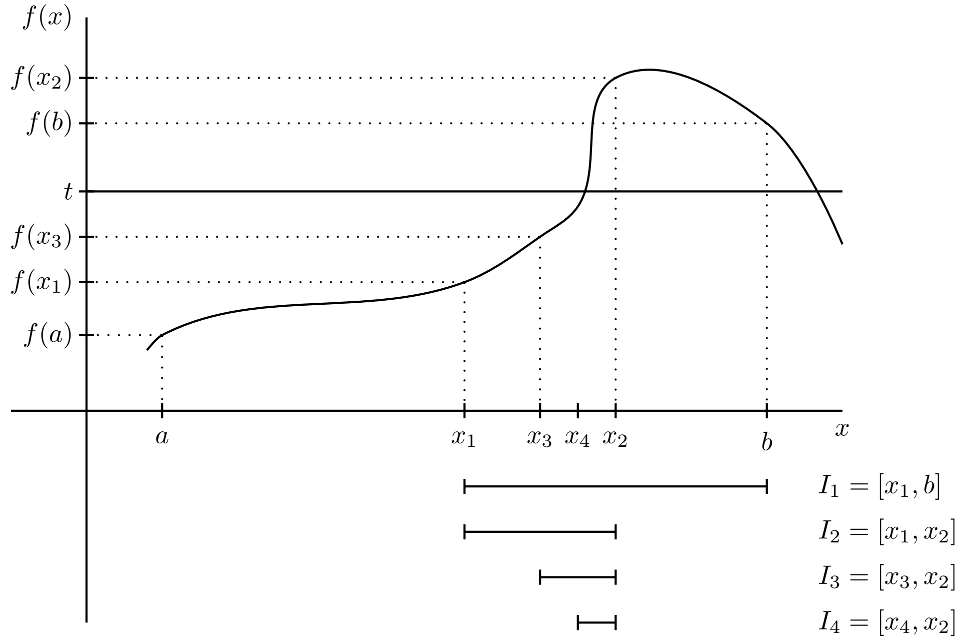

Figure 3.1: The method of interval halving graphically explained.

By continuing the method of interval halving we either find an $n \in \mathbb{N}$ with $f(x_n) = t$ (in which case the proof is complete), or we obtain a sequence $a_1, a_2, a_3, \ldots$ and a sequence $b_1, b_2, b_3, \ldots$ with the property that for every $n \in \mathbb{N}$ it is true that $a_n \leq a_{n+1} < b_{n+1} \leq b_n$, $\;b_n - a_n = (\frac{1}{2})^n(b-a)$ and $f(a_n) < t < f(b_n)$.

Since the sequence $a_1, a_2, a_3, \ldots$ is increasing and bounded from above (for instance by b), and the sequence $b_1, b_2, b_3, \ldots$ is decreasing and bounded from below, both sequences are convergent according to Theorem 0.2. Call their limits c and d respectively. Then, we have

$$c = \lim_{n\to\infty} a_n = \sup\{a_n : n \in \mathbb{N}\} \quad \text{and} \quad d = \lim_{n\to\infty} b_n = \inf\{b_n : n \in \mathbb{N}\}.$$

According to the arithmetic rules for a limit of a sequence, we have

$$d - c = \lim_{n\to\infty} b_n - \lim_{n\to\infty} a_n = \lim_{n\to\infty}(b_n - a_n) = \lim_{n\to\infty}(\tfrac{1}{2})^n(b-a) = 0,$$

so that $c = d$. In addition, we also have that $c \in [a_1, b_1] \subset [a, b]$.

Because f is continuous in c and $c = \lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n$ we get

$$f(c) = \lim_{n\to\infty} f(a_n) \quad \text{and} \quad f(c) = \lim_{n\to\infty} f(b_n).$$

Moreover, since $f(a_n) \leq t \leq f(b_n)$ for every $n \in \mathbb{N}$,

$$f(c) = \lim_{n\to\infty} f(a_n) \leq t \leq \lim_{n\to\infty} f(b_n) = f(c),$$

so that $f(c) = t$. The point c is therefore a point that we were looking for in I. $\triangleleft$

EXERCISE 3.8 How does the method of interval halving work if $f(a) > t > f(b)$?

The proof of the Intermediate Value Theorem is a so-called constructive proof: it does not only show the existence of the point c, but also provides an algorithm for the determination of this point. This interval halving technique can easily be transformed into a computer program.

EXERCISE 3.9

a) Let f be defined on $\mathbb{R}$ by $f(x) = x^3 + x^2 - 17x + 16$. Prove that f has at least one zero on (0,2), at least one on (2,4) and at least one on $(-\infty, 0)$.

b) Let f be a continuous function on the interval $I = [a, b]$ with $f(a)f(b) < 0$. Show that f has a zero on (a, b), in other words, that there is a $c \in (a, b)$ with $f(c) = 0$.

The Intermediate Value Theorem has an interesting consequence, known as Brouwer's theorem.

THEOREM 3.6 (Brouwer's theorem)

Let f be a continuous function on an interval $[a, b]$ such that $f(x) \in [a, b]$ for every $x \in [a, b]$. Then an $x^* \in [a, b]$ exists such that $f(x^*) = x^*$. (x^* is called a fixed point of f.)

PROOF Let f be a continuous function on $[a, b]$. Define the auxiliary function g on the interval $[a, b]$, by $g(x) = f(x) - x$. The function g is continuous, $g(a) = f(a) - a \geq 0$ and $g(b) = f(b) - b \leq 0$. According to the Intermediate Value Theorem, applied to g and the value 0, a point $x^* \in [a, b]$ exists such that $g(x^*) = 0$, in other words, $f(x^*) = x^*$.
(We got the idea of introducing the auxiliary function g by sketching in one and the same coordinate system both the graph of f and the line $y = x$.) $\triangleleft$

The meaning of the condition $f([a, b]) \subseteq [a, b]$ is that f maps the interval $[a, b]$ into itself: the graph of f lies inside the square $[a, b] \times [a, b]$, see Figure 3.2.

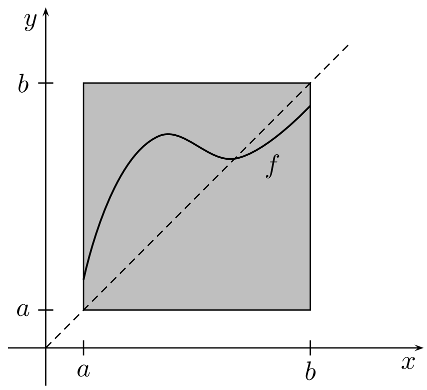

Figure 3.2: f maps $[a, b]$ in itself.

EXERCISE 3.10 Give an example of a continuous function f on $(0, 1)$ with $f((0, 1)) \subseteq (0, 1)$, for which no $c \in (0, 1)$ exists with $f(c) = c$.

The results of the theorem of Weierstrass and the Intermediate Value Theorem can be summarized in a single theorem.

THEOREM 3.7 *For every continuous function f on a compact interval $I = [a, b]$ the following holds:*
f has a minimum m and a maximum M and for every y for which $m \leq y \leq M$ an $x \in I$ exists such that $f(x) = y$, in other words, $f(I) = [m, M]$.
To put it differently, the image of a compact interval under a continuous function is a compact interval.

The above theorem has a wider validity: the image of an interval (not necessarily compact) under a continuous function is itself an interval. For example, the image of the interval $(0, \infty)$ under the continuous function $f(x) = \frac{1}{1+x^2}$ is the interval $(0, 1)$ and the image of the interval $(0, 1]$ under the continuous function $f(x) = \frac{1}{\sqrt{x}}$ is the interval $[1, \infty)$.

3.4 Continuity and the inverse function

We begin this paragraph by discussing the relation between monotony and invertibility of a function.

DEFINITION Let f be a function and let $I \subseteq D_f$.

The function f is called *increasing* on I if $f(x) \leq f(y)$ for every $x, y \in I$ with $x < y$ (and *strictly increasing* on I if $f(x) < f(y)$).

The function f is called *decreasing* on I if $f(x) \geq f(y)$ for every $x, y \in I$ with $x < y$ (and *strictly decreasing* on I if $f(x) > f(y)$).

If the function f is (strictly) increasing or (strictly) decreasing on I, then f is called *(strictly) monotone* on I. If in addition $I = D_f$, then we call f a *(strictly) monotone function.*

EXERCISE 3.11 Determine by means of their graphs which of the functions below are (strictly) increasing or (strictly) decreasing.

a) The function f on $\mathbb{R}$, defined by

$$f(x) = \begin{cases} k & \text{if } 2k-1 \leq x < 2k, k \in \mathbb{Z} \\ x-k & \text{if } 2k \leq x < 2k+1, k \in \mathbb{Z}. \end{cases}$$

b) The function g on $[0,1]$, defined by

$$g(x) = \begin{cases} 0 & \text{if } x = 0 \\ 2^{1-n} & \text{if } 2^{-n} < x \leq 2^{1-n}, n \in \mathbb{N}. \end{cases}$$

DEFINITION Let f be a function and let $I \subseteq D_f$.

The function f is *invertible* (on I) if for every $y \in f(I)$ precisely one $x \in I$ exists such that $f(x) = y$.

If f is invertible on I, then the *inverse* of f (with respect to the interval I) is the function f^{-1} on $f(I)$, which is defined by $f^{-1}(y) = x$, where x is *the* element from I with $f(x) = y$.

EXERCISE 3.12 The function f on $\mathbb{R}$ is defined by $f(x) = x^2 + 1$. Verify that f is not invertible on $\mathbb{R}$ but is invertible on $[0, \infty)$.

THEOREM 3.8 *Every function which is strictly monotone on I is invertible on I.*

PROOF Let f be a function which is strictly monotone on I.
We must show that for every $y \in f(I)$ precisely one $x \in I$ exists with $f(x) = y$. We give a proof by contradiction.
Assume that a $y \in f(I)$ exists such that $f(x_1) = f(x_2) = y$ for a certain $x_1, x_2 \in I$ with $x_1 < x_2$. Then $f(x_1) < f(x_2)$ if f is strictly increasing on I and $f(x_1) > f(x_2)$ if f is strictly decreasing on I. In both cases we have $f(x_1) \neq f(x_2)$. This contradicts $f(x_1) = f(x_2) = y$.
Therefore f is invertible on I (the graph of f satisfies the horizontal line test). $\triangleleft$

> **EXERCISE 3.13** Prove that the inverse of a strictly monotone function is itself strictly monotone.

The following theorem states that if a continuous function f has an inverse, then f^{-1} is also continuous. Although this statement is intuitively obvious, its proof is not so obvious. We will not provide the proof here.

THEOREM 3.9 *For every invertible function f on an interval I the following holds: if f is a continuous function, then the inverse f^{-1} is a continuous function on the interval $f(I)$.*

EXAMPLE 3.8 In this example we prove that the function f on $[0, 3]$, defined by

$$f(x) = x^5 + x,$$

has a continuous inverse on $[0, 246]$.
It is obvious that f is strictly increasing: if $x, y \in [0, 3]$ and $x < y$ then $x^5 < y^5$ and therefore $f(x) = x^5 + x < y^5 + y = f(y)$. Thus according to Theorem 3.8 f is invertible.
Because f is continuous in the interval $[0, 3]$, according to Theorem 3.9, f^{-1} is a continuous function on the interval $f([0, 3])$. From $f(0) = 0$ and $f(3) = 246$ and the fact that f is strictly increasing, it now follows that 0 is the minimum and 246 is the maximum of f. By means of Theorem 3.7 we find that $f([0, 3]) = [0, 246]$.

Comment on function rules

We now know that the inverse f^{-1} of the function f from the above example is a continuous, strictly increasing (Exercise 3.13) function on $[0, 246]$, but we have no function rule, other than

$$f^{-1}(y) = x \text{ where } x \text{ is } \textit{the} \text{ element from } D_f \text{ with } f(x) = y.$$

By 'function rule, other than ...', we mean a rule that makes use of polynomial functions, root functions, sine and cosine functions, exponential functions, etcetera, and that makes use of the operations of addition, subtraction, multiplication, division and composition. It can be proved, after using a lot of difficult mathematics, that for this f it is impossible to give such a rule, and we will have to live with that. However, it is possible to determine $f^{-1}(y)$ for certain values of y. For example, $f^{-1}(2) = 1$ because $f(1) = 2$, and $f^{-1}(34) = 2$ because $f(2) = 34$. But for instance the exact determination of $f^{-1}(1)$ is not possible, although it can be approximated.

Since $f(0) = 0$ and $f(1) = 2$ it follows from the Intermediate Value Theorem that the value 1 is taken by f on the interval $(0, 1)$, in other words, $0 < f^{-1}(1) < 1$.

Since $f(0.7) = 0.86807$ and $f(0.8) = 1.12768$ it follows from the Intermediate Value Theorem that the value 1 is taken by f on the interval $(0.7, 0.8)$, in other words, $0.7 < f^{-1}(1) < 0.8$ or $f^{-1}(1) = 0.7\cdots$.

Since $f(0.75) = 0.987\cdots$ and $f(0.76) = 1.013\cdots$ it follows from the Intermediate Value Theorem that the value 1 is taken by f on the interval $(0.75, 0.76)$, in other words, $0.75 < f^{-1}(1) < 0.76$ or $f^{-1}(1) = 0.75\cdots$.

By continuing in this way we can determine $f^{-1}(1)$ to a degree of accuracy which corresponds to as many decimal places as we wish. $\triangleleft$

The following exercise proves the continuity of the power function.

EXERCISE 3.14 Let $n \in \mathbb{N}$. The function f on $(0, \infty)$ is defined by $f(x) = x^n$.

a) Prove that the function f is invertible and determine $D_{f^{-1}}$.

b) Prove that the inverse f^{-1} of f is continuous.

For every $y > 0$ the positive solution of the equation $x^n = y$ is denoted as $x = y^{\frac{1}{n}}$. Therefore the inverse f^{-1} of f is given by $f^{-1}(y) = y^{\frac{1}{n}}$.

c) Prove that for every $k, l \in \mathbb{Z}$ the function g on $(0, \infty)$, defined by $g(x) = x^{\frac{k}{l}}$, is continuous.

3.5 Mixed exercises

EXERCISE 3.15 The function f on $(2, \infty)$ is defined by $f(x) = \dfrac{1-x}{1+x}$. Prove that f is a continuous function.

EXERCISE 3.16 Determine the value(s) of $a \in \mathbb{R}$, for which the functions below are continuous.

a) $f(x) = \begin{cases} 1 + a & \text{if } 0 \leq x \leq 1 \\ x^2 & \text{if } x < 0 \text{ or } x > 1. \end{cases}$

b) $g(x) = \begin{cases} (x - a)^3 & \text{if } 0 \leq x \leq 1 \\ \frac{1}{2}x^2 - a^3 & \text{if } x < 0 \text{ or } x > 1. \end{cases}$

EXERCISE 3.17 Let f be a continuous function on the interval $I = [a, b]$ with $f(x) \neq 0$ for every $x \in [a, b]$. Prove that the following holds:

Either $f(x) > 0$ for every $x \in [a, b]$ or $f(x) < 0$ for every $x \in [a, b]$.

EXERCISE 3.18 Let f be a function on a closed interval I. Assume that a constant $\alpha > 0$ exists such that for every $x, y \in I$ it is true that $\mid f(x) - f(y) \mid \leq \alpha \mid x - y \mid$. Prove that the function f is continuous.

EXERCISE 3.19 Let f be a continuous function on $\mathbb{R}$ and let $c \in \mathbb{R}$ such that $f(c) > 0$. Prove that a $\delta > 0$ exists such that $f(x) > 0$ for every $x \in (c - \delta, c + \delta)$.

EXERCISE 3.20 Let f be a continuous function on $\mathbb{R}$ for which

$$\lim_{x \to -\infty} f(x) = 0 \quad \text{and} \quad \lim_{x \to \infty} f(x) = 0.$$

Prove that f has a maximum or a minimum.

EXERCISE 3.21 Consider a continuous function f on $[0, +\infty)$, such that

$$f(x) \leq 10 - x$$

for every $x \in [0, +\infty)$. Show that f has a maximum.

EXERCISE 3.22 Let f and g be functions on $\mathbb{R}$.

Prove the following statements or give a counter-example.

a) If f and g are increasing, then the composite function $f \circ g$ is increasing.

b) If f and g are increasing, then the product $f \cdot g$ is increasing.

c) If f is both increasing and decreasing, then f is a constant function.

d) If f is increasing and g is decreasing, then the composite function $f \circ g$ is decreasing.

EXERCISE 3.23 Let f be a continuous function on $\mathbb{R}$ which is periodical with period 2, which means that $f(x+2) = f(x)$ for every $x \in \mathbb{R}$.

Prove that f has a minimum and maximum.

EXERCISE 3.24 Let f be a continuous function on the interval $[0, 1]$, for which $f(0) = f(1)$. Prove that a $c \in [0, \frac{1}{2}]$ exists such that $f(c) = f(c + \frac{1}{2})$.

EXERCISE 3.25 Let f and g be functions on $\mathbb{R}$ and let f_+ be the function on $\mathbb{R}$, defined by

$$f_+(x) = \max\{f(x), 0\}.$$

The (auxiliary) function h on $\mathbb{R}$ is defined by

$$h(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0. \end{cases}$$

a) Show that $f_+(x) = (h \circ f)(x)$ for every $x \in \mathbb{R}$.

b) Prove that h is continuous.

c) Prove the statement: if f is continuous, then f_+ is also continuous.

d) Show that $\max\{f(x), g(x)\} = f(x) + (g - f)_+(x)$ for every $x \in \mathbb{R}$.

e) Prove the statement: if f and g are continuous, then the function k on $\mathbb{R}$, defined by $k(x) = \max\{f(x), g(x)\}$, is also continuous.

4 DIFFERENTIABLE FUNCTIONS

Most likely you remember the concept of derivative from secondary school. In section 4.1 we introduce this concept once again, using the notion of difference quotient and limit of a function. In section 4.2 we deduce some arithmetic rules for differentiating, among which the chain rule for composite functions. In section 4.3 we pay specific attention to the approximation of a function, especially of a differentiable function. The geometrical interpretation of this is that the graph of a differentiable function has a tangent line at every point. Rolle's Theorem and the Mean Value Theorem form the content of section 4.4. In sections 4.5 and 4.6 we make use of the Mean Value Theorem to study a function's behavior if its derivative has a fixed sign. In section 4.7 we formulate a version of Taylor's Theorem, which can be seen as an extension of the Mean Value Theorem to higher derivatives. Finally in section 4.8 we discuss a limit rule for functions, which is known as the rule of de l'Hôpital.

4.1 Differentiable functions

We want to investigate the change of the values of a function f on an interval I in the neighborhood of a point $c \in I$. For every $x \in I$ the difference Δx,

$$\Delta x = x - c$$

can be seen as the *change* with respect to c. Consequently we can write x as $x = c + \Delta x$. Analogously the difference Δf,

$$\Delta f = f(x) - f(c) \quad \text{or} \quad \Delta f = f(c + \Delta x) - f(c)$$

can be seen as the change of the function with respect to $f(c)$.
If the value of the change is positive, then the change is called an increase, if the value of the change is negative, then we speak of a decrease.

DEFINITION Let f be a function on an interval I and let $c \in I$.

For every $x \in I$ with $x \neq c$ (in other words, for every $\Delta x \neq 0$ with $c + \Delta x \in I$) the quotient of the changes Δf and Δx is called the *difference quotient* of the function f in c for the change Δx,

$$\frac{\Delta f}{\Delta x} = \frac{f(x) - f(c)}{x - c} \quad \text{or} \quad \frac{\Delta f}{\Delta x} = \frac{f(c + \Delta x) - f(c)}{\Delta x}.$$

Geometrically the difference quotient represents the slope of the line through the point with coordinates $(c, f(c))$ and the point with coordinates $(c + \Delta x, f(c + \Delta x))$, see Figure 4.1.

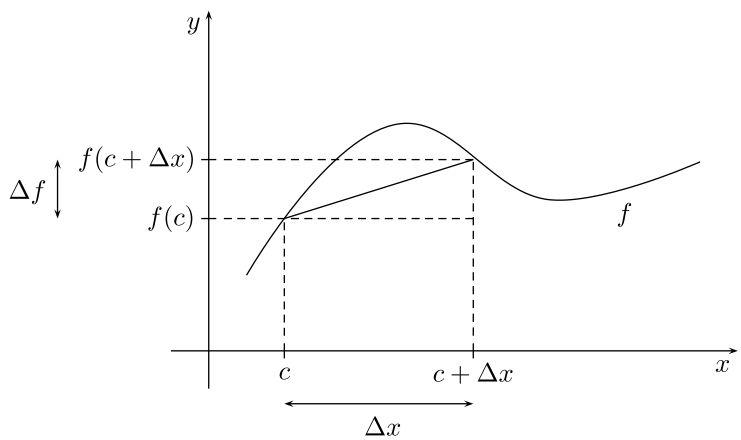

Figure 4.1: The slope of the line.

In terms of changes, the difference quotient is the average change of the values of a function, measured over a change Δx (with respect to c). In the following example we determine the average change in the costs if the level of production changes. We do this for a number of different values of the change in the production level.

EXAMPLE 4.1 Consider the cost function k on $[0, \infty)$, defined by

$$k(q) = 0.001q^3 - 0.06q^2 + 1.3q + 10,$$

where q represents the production level. For a production level of 30 units, the average change in the costs, calculated over a change in production of Δq units, equals

$$\frac{\Delta k}{\Delta q} = \frac{k(30+\Delta q) - k(30)}{\Delta q}.$$

In Table 4.1 you can find the average change in the costs if the production level of $q = 30$ changes by $20,\ 10,\ -5,\ 1$ and 0.1 units ($\Delta q = 20,\ 10,\ -5,\ 1,\ 0.1$).

Δq	Δk	$\frac{\Delta k}{\Delta q}$
20	28.0	1.4
10	8.0	0.8
-5	-1.375	0.275
1	0.431	0.431
0.1	0.040301	0.40301

Table 4.1 The average change in the costs.

It appears as though the average change in the costs tends to approach the value 0.4 if the change in production diminishes. We can prove this, since

$$\begin{aligned}
\frac{\Delta k}{\Delta q} &= \frac{k(30+\Delta q) - k(30)}{\Delta q} \\
&= \frac{1}{\Delta q}\big[0.001(30+\Delta q)^3 - 0.06(30+\Delta q)^2 + 1.3(30+\Delta q) + 10 \\
&\qquad -0.001(30)^3 + 0.06(30)^2 - 1.3(30) - 10\big] \\
&= 3(0.001)(30)^2 - 2(0.06)(30) + 1.3 + [3(0.001)30 - 0.06]\,\Delta q + 0.001(\Delta q)^2 \\
&= 0.4 + 0.03\Delta q + 0.001(\Delta q)^2 \to 0.4 \ \text{ as } \ \Delta q \to 0,
\end{aligned}$$

in other words, $\lim\limits_{\Delta q \to 0} \dfrac{\Delta k}{\Delta q} = 0.4.$ ◁

EXERCISE 4.1 The function f on $\mathbb{R}$ is defined by $f(x) = x^2$. Verify whether the difference quotient $\frac{f(1+\Delta x) - f(1)}{\Delta x}$ has a limit as $\Delta x \to 0$.

Example 4.1 shows that there is sufficient reason to check whether or not a difference quotient of a function in a point has a limit. If the limit of a difference quotient of a function in a point exists, we call the function differentiable in that point.

DEFINITION Let f be a function on an interval I and let $c \in I$.
We call the function f *differentiable* in c if $\lim_{x \to c} \dfrac{f(x) - f(c)}{x - c}$ exists. In that case the value of the limit is called the *derivative* of f in c and is written as $f'(c)$,

$$f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} \quad \text{or} \quad f'(c) = \lim_{\Delta x \to 0} \frac{f(c + \Delta x) - f(c)}{\Delta x}.$$

The function f is called *differentiable* if f is differentiable in every point of I. In that case we call the function on I which adds to every point $x \in I$ the derivative of f in x the *derivative function* of f, and we write it as f'.

Other notations for the derivative $f'(c)$ of f in c are $\dfrac{\mathrm{d}f}{\mathrm{d}x}(c)$ and $Df(c)$. (In economic literature you can also find the notations $\dfrac{\partial f}{\partial x}(c)$ and $\partial_x f(c)$. We will not use these notations.)

EXAMPLE 4.2 We prove that the function f on $\mathbb{R}$, defined by $f(x) = x^5 + x$, is differentiable in 1.

We have the following series of subsequent statements for $\Delta x \neq 0$:

$$\begin{aligned}
\frac{f(1 + \Delta x) - f(1)}{\Delta x} &= \frac{(1 + \Delta x)^5 + (1 + \Delta x) - 2}{\Delta x} \\
&= \frac{1 + 5\Delta x + 10(\Delta x)^2 + 10(\Delta x)^3 + 5(\Delta x)^4 + (\Delta x)^5 + 1 + \Delta x - 2}{\Delta x} \\
&= 6 + 10\Delta x + 10(\Delta x)^2 + 5(\Delta x)^3 + (\Delta x)^4,
\end{aligned}$$

so that $\lim_{\Delta x \to 0} \dfrac{f(1 + \Delta x) - f(1)}{\Delta x} = 6$. Therefore the function f is differentiable in 1 and $f'(1) = 6$. ◁

EXAMPLE 4.3 The function f on $[0, \infty)$, defined by $f(x) = \sqrt{x^3}$, is differentiable in 0, since for $x > 0$ it holds that

$$\frac{f(x) - f(0)}{x - 0} = \frac{\sqrt{x^3}}{x} = \sqrt{x},$$

so that $\lim_{x \to 0} \frac{f(x) - f(0)}{x - 0} = 0$. Therefore the function f is differentiable in 0 and $f'(0) = 0$. ◁

EXERCISE 4.2 Show that the function f on $\mathbb{R}$, defined by $f(x) = x^2$, is differentiable in c for every $c \in \mathbb{R}$ and that $f'(c) = 2c$.

EXERCISE 4.3 The function f on $[0, \infty)$ is defined by $f(x) = \sqrt{x}$.

a) Show that for every $c \in (0, \infty)$ the function f is differentiable in c and that $f'(c) = \frac{1}{2\sqrt{c}}$.

b) Show that f is not differentiable in 0.

An important observation is the fact that a function, which is differentiable in a point, is continuous in that point.

THEOREM 4.1 *A function which is differentiable in a point is continuous in that point.*

PROOF Let f be a function on an interval I and let $c \in I$. Assume that f is differentiable in c. Then $\lim\limits_{x \to c} \frac{f(x)-f(c)}{x-c}$ exists and equals $f'(c)$. Then for $x \in I$, $x \neq c$, it holds that

$$f(x) - f(c) = \left(\frac{f(x) - f(c)}{x - c} \right) \cdot (x - c).$$

By means of the arithmetic rules for limits of functions (Theorem 2.2) it follows that

$$\lim_{x \to c} (f(x) - f(c)) = f'(c) \cdot 0 = 0,$$

in other words, f is continuous in c. Therefore the function f is continuous. $\triangleleft$

The following example shows that a function, which is continuous in a point, need not be differentiable in that point. So it is not permitted to reverse the 'if ..., then ...' statement of Theorem 4.1.

EXAMPLE 4.4 The (absolute value) function f on $\mathbb{R}$, defined by $f(x) = |x|$, is continuous in 0, but not differentiable in 0. For, the difference quotient of f in 0, $\frac{|x|-0}{x-0}$, has no limit in 0. $\triangleleft$

EXERCISE 4.4 The function f on $\mathbb{R}$ is defined by

$$f(x) = \begin{cases} x^2 & \text{if } x \leq 1 \\ x^2 + 1 & \text{if } x > 1. \end{cases}$$

a) Show that, for every $x \neq 1$, f is differentiable in x and that f' on $\mathbb{R} \backslash \{1\}$ is given by $f'(x) = 2x$.

b) Show that $\lim\limits_{x \to 1} f'(x)$ exists (and equals 2).

c) Show that f is not differentiable in 1.

EXERCISE 4.5 The function f on $\mathbb{R}$ is defined by

$$f(x) = \begin{cases} \frac{4}{3+x^2} & \text{if } x \leq 1 \\ \frac{3-x}{2} & \text{if } x > 1. \end{cases}$$

a) Calculate for $x \neq 1$ the difference quotient $\frac{f(x)-f(1)}{x-1}$ of f.

b) Use part a) to show that f is differentiable in 1 and determine $f'(1)$.

DEFINITION A function is called *continuously differentiable* if the derivative function is continuous. A function f on I is twice differentiable, if beside f itself, the derivative function f' is differentiable as well. We call the derivative of f' the *second order derivative* of f. We write the second derivative of f as f'' or as $f^{(2)}$. In a similar way we define the *third (order) derivative* f''' or $f^{(3)}$ and, in general, the n^{th} *(order) derivative* $f^{(n)}$.

We note that for the existence of the n^{th} derivative of f it is silently assumed that the derivative of the $(n-1)^{th}$ order exists.

4.2 Arithmetic rules for differentiable functions

The following arithmetic rules apply to differentiable functions.

THEOREM 4.2 *For functions f and g on an interval I the following holds: if f and g are differentiable in $c \in I$, then the sum function $f+g$, the product function $f \cdot g$ and the quotient function $\frac{f}{g}$ (provided that $g(x) \neq 0$ for every $x \in I$) are differentiable in c and*

a) $(f+g)'(c) = f'(c) + g'(c)$,

b) $(f \cdot g)'(c) = f'(c) \cdot g(c) + f(c) \cdot g'(c)$,

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

PROOF We only prove parts b) and c), leaving the proof of part a) to the reader.
b) Let f and g be functions on I and let $c \in I$. Assume that f and g are differentiable in c. Then, for every $x \in I$, $x \neq c$,

$$\frac{(f \cdot g)(x) - (f \cdot g)(c)}{x-c} = \frac{f(x)g(x) - f(c)g(x) + f(c)g(x) - f(c)g(c)}{x-c}$$

$$= \quad \frac{f(x)-f(c)}{x-c}g(x) + f(c)\frac{g(x)-g(c)}{x-c}.$$

According to Theorem 4.1 a differentiable function is continuous. Therefore the function g is continuous in c, so that $g(x) \to g(c)$ as $x \to c$. Because f and g are differentiable in c it holds that $\frac{f(x)-f(c)}{x-c} \to f'(c)$ and $\frac{g(x)-g(c)}{x-c} \to g'(c)$ as $x \to c$. So according to the arithmetic rules for limits

$$\lim_{x\to c} \frac{(f \cdot g)(x) - (f \cdot g)(c)}{x-c} = f'(c)g(c) + f(c)g'(c).$$

Therefore $f \cdot g$ is differentiable in c and $(f \cdot g)'(c) = f'(c)g(c) + f(c)g'(c)$.

c) Assume that $g(x) \neq 0$ for every $x \in I$ and that f and g are differentiable in c.

First we prove that the function $\frac{1}{g}$ is differentiable in c and that it holds that

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

For every $x \in I$, $x \neq c$, we have

$$\frac{\frac{1}{g}(x) - \frac{1}{g}(c)}{x-c} = \frac{\frac{1}{g(x)} - \frac{1}{g(c)}}{x-c} = -\frac{1}{g(x)g(c)} \cdot \frac{g(x)-g(c)}{x-c}.$$

Because g is differentiable in c, this function is therefore also continuous in c, $g(x) \to g(c)$ and $\frac{g(x)-g(c)}{x-c} \to g'(c)$ as $x \to c$, so that

$$\lim_{x\to c} \frac{\frac{1}{g}(x) - \frac{1}{g}(c)}{x-c} = -\lim_{x\to c} \frac{1}{g(x)g(c)} \cdot \lim_{x\to c} \frac{g(x)-g(c)}{x-c} = -\frac{g'(c)}{(g(c))^2}.$$

Therefore $\dfrac{1}{g}$ is differentiable in c and $\left(\dfrac{1}{g}\right)'(c) = -\dfrac{g'(c)}{(g(c))^2}$.

Since $\dfrac{f}{g} = f \cdot \dfrac{1}{g}$, $\dfrac{f}{g}$ is now also differentiable according to the product rule of part b), and

$$\left(\frac{f}{g}\right)'(c) = f'(c) \cdot \frac{1}{g}(c) + f(c) \cdot \left(\frac{1}{g}\right)'(c) = \frac{f'(c) \cdot g(c) - f(c) \cdot g'(c)}{(g(c))^2}. \qquad \triangleleft$$

EXAMPLE 4.5 We prove that for every $n \in \mathbb{N}$ the function f on $\mathbb{R}$, defined by $f(x) = x^n$, is differentiable and that $f'(x) = nx^{n-1}$ for every $x \in \mathbb{R}$. We make use of the principle of induction.

By $\mathcal{P}(n)$ we denote the statement 'the function f on $\mathbb{R}$, defined by $f(x) = x^n$, is differentiable and $f'(x) = nx^{n-1}$ for every $x \in \mathbb{R}$'.

Step 1: ($\mathcal{P}(1)$ is true)

$\mathcal{P}(1)$ is the statement 'the function f on $\mathbb{R}$, defined by $f(x) = x$, is differentiable and $f'(x) = 1$ for every $x \in \mathbb{R}$' and is therefore certainly true.

Step 2: (for every $n \in \mathbb{N}$ it holds that $\mathcal{P}(n) \Rightarrow \mathcal{P}(n+1)$)

Let $n \in \mathbb{N}$. Assume that the function f on $\mathbb{R}$, defined by $f(x) = x^n$, is differentiable and that $f'(x) = nx^{n-1}$ for every $x \in \mathbb{R}$.

The function g on $\mathbb{R}$, defined by $g(x) = x^{n+1}$, can be written as the product of the function h on $\mathbb{R}$, defined by $h(x) = x$, and of f. According to the product rule for differentiation, g, being the product of the two differentiable functions h and f, is differentiable, and it holds for every $x \in \mathbb{R}$ that

$$g'(x) = h'(x) \cdot f(x) + h(x) \cdot f'(x) = x^n + x \cdot nx^{n-1} = (n+1)x^n.$$

This proves that the statement $\mathcal{P}(n+1)$ is true. $\triangleleft$

EXERCISE 4.6 Determine, by means of the arithmetic rules for differentiable functions, the derivative of the function f on $\mathbb{R}$, defined by

a) $f(x) = \dfrac{x}{1+x^2}$

b) $f(x) = (x-2)(x-7)(x-4)$.

EXERCISE 4.7 Prove that for every $n \in \mathbb{N}$ the function g on $(0, \infty)$, defined by $g(x) = \dfrac{1}{x^n}$, is differentiable and that $g'(x) = (-n)\dfrac{1}{x^{n+1}}$ for every $x > 0$.

One of the most frequently applied arithmetic rules for differentiation is the chain rule.

THEOREM 4.3 **(Chain rule)**

Let f be a function on an interval I and g a function on an interval J, such that $g(J) \subseteq I$ and $c \in J$. If g is differentiable in c and f is differentiable in $g(c)$, then the composite function $f \circ g$ is differentiable in c and

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

PROOF Assume that g is differentiable in c and that f is differentiable in $g(c)$ and in addition assume that g satisfies the extra condition that $g(x) - g(c) \neq 0$ for all $x \in J, x \neq c$, in a neighborhood of c. (Here neighborhood of c is short for an ε-neighborhood of c.) Then, for this $x \in J$, $x \neq c$,

$$\begin{aligned}\frac{(f \circ g)(x) - (f \circ g)(c)}{x - c} &= \frac{f(g(x)) - f(g(c))}{x - c} \\ &= \frac{f(g(x)) - f(g(c))}{g(x) - g(c)} \cdot \frac{g(x) - g(c)}{x - c}.\end{aligned}$$

Since g is differentiable in c, g is continuous in c, so that $g(x) \to g(c)$ as $x \to c$. And since f is differentiable at $g(c)$, it subsequently holds that $\frac{f(g(x))-f(g(c))}{g(x)-g(c)} \to f'(g(c))$ as $x \to c$. Moreover it holds that $\frac{g(x)-g(c)}{x-c} \to g'(c)$ as $x \to c$. According to the product rule for limits of functions therefore, the limit of the difference quotient exists and

$$\lim_{x \to c} \frac{(f \circ g)(x) - (f \circ g)(c)}{x - c} = f'(g(c)) \cdot g'(c).$$

Under the extra condition that $g(x) - g(c) \neq 0$ for every $x \in J, x \neq c$ in a neighborhood of c, the composite function $f \circ g$ is therefore differentiable in $c \in J$.
If g is a function that does not satisfy the extra condition, then the proof for the differentiability of the composite function runs less smoothly than above. $\triangleleft$

EXERCISE 4.8 Show that the function f on $\mathbb{R}$, defined by $f(x) = \sqrt{x^2 + 1}$, is differentiable and determine its derivative.

4.3 Linear approximation

Geometrically a difference quotient represents the slope of a line segment. For this reason it is obvious that we can interpret the derivative $f'(c)$ geometrically as the slope of the tangent line to the graph of f in $(c, f(c))$. We call the function, of which the graph is this tangent line, the linear approximation of f in c.

DEFINITION Let f be a function on an interval I, which is differentiable in $c \in I$. The function g on $\mathbb{R}$, defined by

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

is called the *linear approximation* of f in c. The graph of g,

$$\{(x, y) \in \mathbb{R}^2 |\ y = f(c) + f'(c)(x - c)\},$$

is called the *tangent line* to the graph of f at $(c, f(c))$.

(Please note that the function g has the property that $g(c) = f(c)$ and $g'(c) = f'(c)$.) The name 'tangent line' for the graph of the function g may be intuitively clear, but the name 'linear approximation' seems to have fallen more or less from the blue sky. First we will explain what we mean by 'approximating a function' and then we will explain the concept of 'linear approximation'.

We call a function h an *approximation* of f in c if the *remainder* r, defined by $r(x) = f(x) - h(x)$, has the property that

$$\lim_{x \to c} r(x) = 0.$$

What interests us is the question: under which condition(s) does a function have an approximation?

Let's assume that the function f is continuous in c. Then the function h on $\mathbb{R}$, defined by $h(x) = f(c)$, is an approximation of f in c. For it follows from the continuity of f in c that $\lim_{x \to c} f(x) = f(c)$, so that for the remainder $r(x) = f(x) - h(x)$ it holds that

$$\lim_{x \to c} r(x) = \lim_{x \to c} (f(x) - f(c)) = 0.$$

Sometimes we also call the function h the *constant approximation* of f in c.

If we subsequently assume that the function f is differentiable in c, then the function g on $\mathbb{R}$, defined by $g(x) = f(c) + f'(c)(x - c)$, is also an approximation of f in c, since the following series of subsequent statements holds for $r(x) = f(x) - g(x)$:

$$\begin{aligned}\lim_{x \to c} r(x) &= \lim_{x \to c} [f(x) - f(c) - f'(c)(x - c)] \\ &= \lim_{x \to c} [f(x) - f(c)] - \lim_{x \to c} f'(c)(x - c) = 0 - 0 = 0.\end{aligned}$$

The linear approximation g is a 'better' approximation than the constant approximation h of f in c. To make it clear what we mean by 'better' approximation we consider the quotient $\frac{r(x)}{x-c}$. The following holds:

$$\begin{aligned}\lim_{x\to c}\frac{r(x)}{x-c} &= \lim_{x\to c}\frac{f(x)-f(c)-f'(c)(x-c)}{x-c}\\ &= \lim_{x\to c}\left(\frac{f(x)-f(c)}{x-c}-f'(c)\right)=f'(c)-f'(c)=0.\end{aligned}$$

Here we have made use of the differentiability of f in c. So the quotient $\frac{r(x)}{x-c}$ is 'something small' if x approaches c. Because the denominator $x-c$ of the quotient is also 'something small' if x approaches c, the numerator $r(x)$ is therefore the product of 'something small' times 'something small' if x approaches c. Because generally we do not know more about the remainder for the constant approximation h than that is 'something small' if x approaches c, the remainder for the linear approximation g will therefore be less than the remainder for the constant approximation. For this reason the linear approximation g is a 'better' approximation than the constant approximation h of f in c.

We can summarize what we know about the linear approximation g of f in c as follows.

The remainder r, defined by $r(x)=f(x)-g(x)$, for the *linear approximation* g of f in c has the property that

$$\lim_{x\to c}\frac{r(x)}{x-c}=0.$$

In Figure 4.2 we can see the graphical representation of the constant vs. the linear approximation of a function f in c. We can also see in the figure what the graphical interpretation is of the statement 'the linear approximation is better than the constant approximation'.

EXAMPLE 4.6 For the continuous function f on $\mathbb{R}$, defined by $f(x)=x^5+x$, it holds that $f(1)=2$. Therefore the constant approximation of f in 1 is the function h on $\mathbb{R}$, defined by $h(x)=2$.

In Example 4.2 it was proved that f is differentiable in 1 and that $f'(1)=6$. The linear approximation of f in 1 is the linear function g on $\mathbb{R}$, defined by $g(x)=2+6(x-1)$. $\triangleleft$

EXERCISE 4.9 The function f on $\mathbb{R}$ is defined by $f(x)=x^2+2$.

a) Determine the constant approximation h and the linear approximation g of f in 2 and sketch the graphs of f, g and h in one single figure.

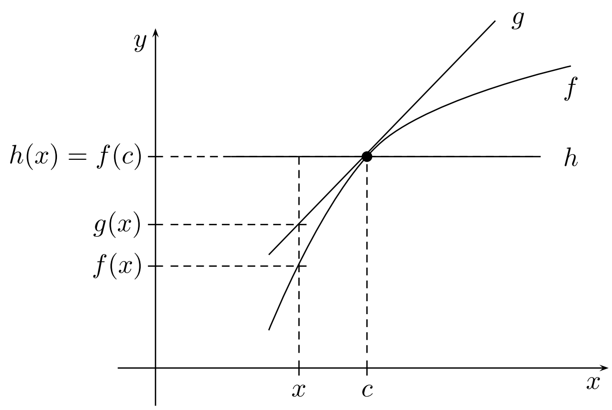

Figure 4.2: The constant and the linear approximation of f in c.

b) Investigate the existence of the following limits and determine their values if they exist:

i) $\lim\limits_{x\to 2}(f(x) - g(x))$ ii) $\lim\limits_{x\to 2} \frac{f(x)-g(x)}{x-2}$ iii) $\lim\limits_{x\to 2} \frac{f(x)-g(x)}{(x-2)^2}$.

4.4 Mean Value Theorem

Besides a function's global extreme values, we can also define local extreme values.

DEFINITION Let f be a function on an interval I and let $c \in I$.

We call $f(c)$ a *local minimum* of f if an ε-neighborhood $(c - \varepsilon, c + \varepsilon)$ of c exists such $f(x) \geq f(c)$ that for every $x \in I \cap (c - \varepsilon, c + \varepsilon)$.

We call $f(c)$ a *local maximum* of f if an ε-neighborhood $(c - \varepsilon, c + \varepsilon)$ of c exists such that $f(x) \leq f(c)$ for every $x \in I \cap (c - \varepsilon, c + \varepsilon)$.

We call $f(c)$ a *local extreme value* of f if $f(c)$ is a local minimum or a local maximum of f on I.

Every global extreme value is of course also a local extreme value (see the definition of global extreme value in section 3.3). The interval $(c - \varepsilon, c + \varepsilon)$ is called the ε-neighborhood of c as it consists of all real numbers with a distance to c smaller than ε. In the sequel we will denote this ε-neighborhood by $U_\varepsilon(c)$.

In this chapter we will restrict ourselves to local extreme values at *interior* points. Such extreme values are also called *free extreme values*.

THEOREM 4.4 *Let f be a function on an interval I and let $c \in I$ such that $f(c)$ is a local extreme value of f. If c is an interior point of I and f is differentiable at c, then $f'(c) = 0$.*

PROOF Assume that c is an interior point of I and that f is differentiable at c. Also assume that $f(c)$ is a local maximum of f. Then an $\varepsilon > 0$ exists such that $f(x) \leq f(c)$ for every $x \in U_\varepsilon(c)$, in other words, $f(x) - f(c) \leq 0$.
Since f is differentiable at c, $\lim\limits_{x\to c} \dfrac{f(x)-f(c)}{x-c}$ equals $f'(c)$. Since for every $x \in (c-\varepsilon, c) \subset U_\varepsilon(c)$ we have

$$\frac{f(x)-f(c)}{x-c} \geq 0$$

we get $f'(c) \geq 0$. On the other hand, for every $x \in (c, c+\varepsilon) \subset U_\varepsilon(c)$ we have

$$\frac{f(x)-f(c)}{x-c} \leq 0$$

from which we infer that $f'(c) \leq 0$. Since both $f'(c) \geq 0$ and $f'(c) \leq 0$ we conclude that $f'(c) = 0$.
If $f(c)$ is a local minimum of f, then the proof proceeds in an analogous way. We leave this case to the reader. ◁

If therefore $f(c)$ is a local extreme value of f and c is an interior point of I, then according to Theorem 4.4 it must hold that $f'(c) = 0$. For this reason we call the condition '$f'(c) = 0$' a *necessary condition* for an extreme value at an interior point. The condition is certainly not sufficient for an extreme value: simple functions can be found for which $f'(c) = 0$, whereas $f(c)$ is not a local extreme value of f.

> **EXERCISE 4.10** Give an example of a differentiable function f on $\mathbb{R}$ and a point $c \in \mathbb{R}$, for which $f'(c) = 0$, whereas $f(c)$ is not a local extreme value of f.

Theorem 4.4 forms the core of the proof of Rolle's Theorem. Besides that, Theorem 4.4 is important in determining the extreme values of a function.

If we sketch the graph of a differentiable function, and this graph begins and ends at the same height, then it is intuitively clear that the graph must have a 'turning point' somewhere in between (see Figure 4.3).

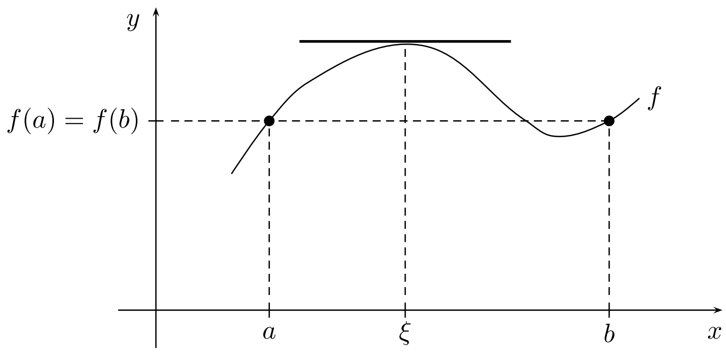

Figure 4.3: The theorem of Rolle.

The precise formulation of this fact is stated in the next theorem.

THEOREM 4.5 **(Rolle's Theorem)**
Let f be a continuous function on the interval $[a, b]$, which is differentiable on (a, b). If $f(a) = f(b)$, then at least one (intermediate) point $\xi \in (a, b)$ exists such that $f'(\xi) = 0$.

PROOF Assume that $f(a) = f(b)$. If f is constant (meaning that $f(x) = f(a)$ for every $x \in [a, b]$) then every interior point ξ satisfies the conclusion of the theorem.
We therefore assume that f is not constant. Then an $x \in (a, b)$ exists for which $f(x) \neq f(a)$. Assume that $f(x) > f(a)(= f(b))$. (The case $f(x) < f(a)$ will be discussed in an exercise.) According to the Theorem of Weierstrass f has a maximum on the interval $[a, b]$. Since $f(a) = f(b) < f(x)$, the maximal value is assumed at a point $\xi \in (a, b)$ and since f is differentiable on (a, b), we have $f'(\xi) = 0$ according to Theorem 4.4 on extreme values at interior points. $\triangleleft$

EXERCISE 4.11 Give the proof of Rolle's Theorem for the case that a point $x \in (a, b)$ exists for which $f(x) < f(a)$.

EXERCISE 4.12 Let g be a continuous function on the interval $[a, b]$, which is differentiable on (a, b) and for which $g'(x) \neq 0$ for every $x \in (a, b)$.
Prove that $g(x) \neq g(y)$ for all $x, y \in [a, b]$ with $x \neq y$.

EXERCISE 4.13 The function f on $\mathbb{R}$ is defined by

$$f(x) = x^5 + 2x^3 + x - 5.$$

a) Prove by means of the Intermediate Value Theorem that an $x^* \in \mathbb{R}$ exists for which $f(x^*) = 0$.

b) Prove that x^* is the only zero of f on $\mathbb{R}$.

An immediate consequence of Rolle's Theorem is the Mean Value Theorem.

THEOREM 4.6 **(Mean Value Theorem)**

Let f be a continuous function on $[a, b]$, which is differentiable on (a, b). Then at least one (intermediate) point $\xi \in (a, b)$ exists such that

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

PROOF Consider the auxiliary function g on $[a, b]$, defined by

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

(This auxiliary function is the difference between f and the linear function h on $[a, b]$, defined by $h(x) = \frac{f(b)-f(a)}{b-a} \cdot (x - a)$. The graph of h is a line segment, parallel to the chord connecting the points $\big(a, f(a)\big)$ and $\big(b, f(b)\big)$, see Figure 4.4.) Since g is continuous, differentiable on (a, b) and $g(a) = g(b)$, according to Rolle's Theorem, applied to the function g, a(n) (intermediate) point $\xi \in (a, b)$ exists such that $g'(\xi) = 0$:

$$f'(\xi) - \frac{f(b) - f(a)}{b - a} = 0 \quad \text{or} \quad \frac{f(b) - f(a)}{b - a} = f'(\xi). \qquad \triangleleft$$

The Mean Value Theorem can be illustrated graphically as follows. If we draw the chord between a pair of points P and Q on the graph of f, then there is at least one point R between P and Q on the graph, where the tangent line to the graph of f is parallel to that chord (see Figure 4.4).

So you can easily memorize the Mean Value Theorem (like Rolle's Theorem), by sketching a suitable graph.

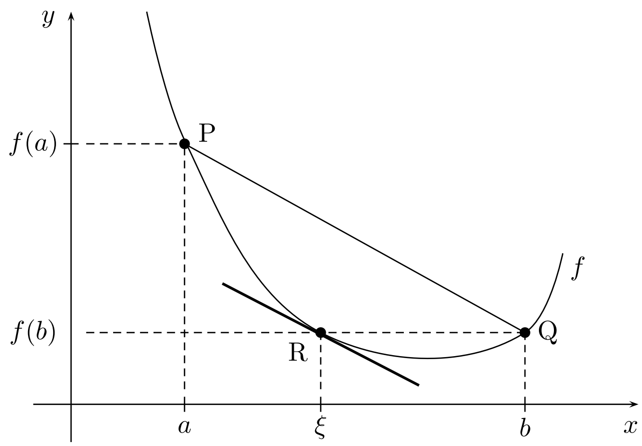

Figure 4.4: At R the tangent line is parallel to the chord between P and Q.

In a number of examples and exercises we determine the location of an intermediate point ξ, which is mentioned in the Mean Value Theorem.

EXAMPLE 4.7 Consider the function f on the interval $[a, b]$, defined by $f(x) = x^2$. The function f is continuous and differentiable on (a, b). According to the Mean Value Theorem a(n) (intermediate) point $\xi \in (a, b)$ exists for which $\frac{f(b)-f(a)}{b-a} = f'(\xi)$, for which it therefore holds that

$$\frac{b^2 - a^2}{b - a} = 2\xi \quad \text{or} \quad \xi = \frac{a+b}{2}.$$

So in the case of the quadratic function ξ lies precisely in the middle of the interval $[a, b]$. $\triangleleft$

EXERCISE 4.14 The function f on $\mathbb{R}$ is defined by $f(x) = x^3$.
Determine a point $\xi \in (-1, 3)$ such that $f'(\xi) = \frac{f(3)-f(-1)}{3-(-1)}$.

EXERCISE 4.15 The function g on $[-1, 1]$ is defined by $g(x) = x^3 - x$.
Determine a point $\eta \in (-1, 1)$ such that $g'(\eta) = \frac{g(1)-g(-1)}{1-(-1)}$.

EXAMPLE 4.8 Let f be a continuous function on the interval $[0, 1]$, which is differentiable on $(0, 1)$. Assume that $f(0) = 0$, $f(1) = 1$ and that $f'(x) \leq 1$ for every $x \in (0, 1)$. We will prove that $f(x) = x$ for every $x \in [0, 1]$.
Let $x \in (0, 1)$. According to the Mean Value Theorem, applied to f restricted to the interval

$[0, x]$, an intermediate point $\xi \in (0, x)$ exists for which

$$f'(\xi) = \frac{f(x) - f(0)}{x - 0} = \frac{f(x)}{x}.$$

Since $f'(x) \leq 1$ for every $x \in (0, 1)$, we have $f'(\xi) \leq 1$, so that $f(x) \leq x$.

According to the Mean Value Theorem, applied to f restricted to the interval $[x, 1]$, an intermediate point $\eta \in (x, 1)$ exists for which

$$f'(\eta) = \frac{f(1) - f(x)}{1 - x} = \frac{1 - f(x)}{1 - x}.$$

Since $f'(x) \leq 1$ for every $x \in (0, 1)$, it holds that $f'(\eta) \leq 1$, so that $\frac{1-f(x)}{1-x} \leq 1$, in other words, $f(x) \geq x$. So we can conclude: $f(x) = x$. $\triangleleft$

EXERCISE 4.16 Let f and g be differentiable functions on $\mathbb{R}$ with the property that $f'(x) > g'(x)$ for every $x \in \mathbb{R}$ and let $a \in \mathbb{R}$. Suppose that $f(a) = g(a)$.

Prove that $f(x) < g(x)$ for every $x < a$ and $f(x) > g(x)$ for every $x > a$.

4.5 Monotone and differentiable functions

In this paragraph we prove by means of the Mean Value Theorem a few theorems concerning monotone functions.

THEOREM 4.7 *Let f be a continuous function on $[a, b]$, which is differentiable on (a, b). If $f'(x) = 0$ for every $x \in (a, b)$ (in other words, $f' = 0$ on (a, b)), then f is constant.*

PROOF Assume that $f' = 0$ on (a, b). We prove that $f(x) = f(a)$ for every $x \in [a, b]$.

Let $x \in (a, b]$. According to the Mean Value Theorem, applied to f restricted to the interval $[a, x]$, a point $\xi \in (a, x)$ exists such that

$$f'(\xi) = \frac{f(x) - f(a)}{x - a} \quad \text{or} \quad f(x) - f(a) = f'(\xi)(x - a).$$

Since $f'(\xi) = 0$ according to the given, indeed we have $f(x) = f(a)$. $\triangleleft$

EXERCISE 4.17 Let f and g be continuous functions on $[a, b]$, which are differentiable on (a, b) and for which $f' = g'$ on (a, b).

Prove that a constant C exists such that $f(x) = g(x) + C$ for every $x \in [a, b]$.

THEOREM 4.8 *Let f be a continuous function on $[a,b]$, which is differentiable on (a,b). The following holds:*
a) $f' \geq 0$ on (a,b) if and only if f is increasing on $[a,b]$
b) if $f' > 0$ on (a,b) then f is strictly increasing on $[a,b]$
c) $f' \leq 0$ on (a,b) if and only if f is decreasing on $[a,b]$
d) if $f' < 0$ on (a,b) then f is strictly decreasing on $[a,b]$.

PROOF We only give a proof of part a). The proof of the other parts is analogous to the proof of part a).
a) $(\Rightarrow)$ Assume that $f' \geq 0$ on (a,b). Let $x,\ y \in [a,b]$ with $x < y$. According to the Mean Value Theorem, applied to the function f restricted to the interval $[x,y]$, a point $\xi \in (x,y)$ exists such that

$$f(y) - f(x) = f'(\xi)(y - x).$$

Because $f'(\xi) \geq 0$ and $y - x > 0$, we have therefore $f(y) - f(x) \geq 0$, so that we also have $f(y) \geq f(x)$.
$(\Leftarrow)$ Assume that f is increasing on $[a,b]$. For an arbitrarily chosen point $c \in (a,b)$ it holds, both for every $x < c$ and for every $x > c$, that $\frac{f(x)-f(c)}{x-c} \geq 0$. Then we get

$$f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} \geq 0. \qquad \triangleleft$$

The reverse of the statement in part b) – if f is strictly increasing on $[a,b]$, then $f' > 0$ on (a,b) – is not true: strictly increasing functions exist of which the derivative at a point equals zero. (For example, take the function f on $\mathbb{R}$, defined by $f(x) = x^3$: f is strictly increasing, but $f'(0) = 0$.)

EXERCISE 4.18 For each of the functions below determine the intervals on which the function is (strictly) increasing and the intervals on which it is (strictly) decreasing:
a) the function f on $\mathbb{R}$, defined by $f(x) = \dfrac{1}{1+x^2}$
b) the function g on $\mathbb{R}\backslash\{0\}$, defined by $g(x) = x + \dfrac{4}{x}$.

EXERCISE 4.19 Let f be a differentiable function on the interval (a, b) and let $c \in (a, b)$ with $f'(c) = 0$. Assume that $f'(x) < 0$ for every $x \in (a, c)$ and that $f'(x) > 0$ for every $x \in (c, b)$.

Prove that $f(c)$ is the minimum of f.

4.6 Differentiability of the inverse function

In section 4.5 on monotone and differentiable functions we proved that a differentiable function with a fixed sign of its derivative is strictly monotone. According to Theorem 3.8 such a function is invertible. We will now prove that this inverse is a differentiable function and we will deduce a formula for its derivative. At some points this formula will enable us to calculate the derivative of the inverse, even when we are incapable of giving a function rule for the inverse. We will omit the proof.

THEOREM 4.9 *Let f be a differentiable function on an interval I such that either $f' > 0$ (in other words, $f'(x) > 0$ for every $x \in I$) or $f' < 0$ (in other words, $f'(x) < 0$ for every $x \in I$). The following holds:*

a) f is invertible on I

b) f^{-1} is a differentiable function on the interval $J = f(I)$

c) $\left(f^{-1}\right)'(y) = \dfrac{1}{f'(f^{-1}(y))}$, for every $y \in J$.

Once it is proved that a function has a differentiable inverse, the formula for the derivative of the inverse can easily be remembered by applying the chain rule for differentiation to the composite function $f \circ f^{-1}$. The composite function $f \circ f^{-1}$ has the property that for every $y \in J = f(I)$ it holds that $(f \circ f^{-1})(y) = f(f^{-1}(y)) = y$. According to the chain rule

$$(f \circ f^{-1})'(y) = f'(f^{-1}(y)) \cdot (f^{-1})'(y).$$

But because $(f \circ f^{-1})(y) = y$ for every $y \in J$, the derivative of $f \circ f^{-1}$ is also the constant function 1, so that for every $y \in J$

$$f'(f^{-1}(y)) \cdot (f^{-1})'(y) = 1 \quad \text{or} \quad (f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}.$$

In practice $f^{-1}(y)$ in the expression for the derivative of the inverse is sometimes replaced by x, so that the formula may also have the following appearance:

$$(f^{-1})'(y) = \frac{1}{f'(x)}, \quad \text{where} \ \ x = f^{-1}(y) \ \ \text{or} \ \ y = f(x).$$

In Theorem 4.9 the function f must satisfy the condition that either $f' > 0$ on I or $f' < 0$ on I. This is a necessary condition with regard to part b), the differentiability of the inverse function: if $f'(c) = 0$ for some $c \in I$, then the inverse function f^{-1} is not differentiable at $d = f(c)$. The function f on $\mathbb{R}$, defined by $f(x) = x^3$, is an example of this. The function is strictly monotone and therefore invertible. At 0 we have $f'(0) = 0$ and the inverse function $f^{-1}(y) = \sqrt[3]{y}$ is not differentiable at $f(0) = 0$.

EXAMPLE 4.9 In Example 3.8 we proved that the function f on $[0,3]$, defined by

$$f(x) = x^5 + x,$$

is strictly increasing and therefore invertible on $[0,3]$.
Since f is differentiable and $f'(x) = 5x^4 + 1 > 0$ for every $x \in [0,3]$, the inverse is a differentiable function on $f([0,3]) = [0,246]$, and for every $y \in [0,246]$

$$\left(f^{-1}\right)'(y) = \frac{1}{5x^4+1}, \quad \text{where} \ \ x = f^{-1}(y) \ \ \text{or} \ \ y = x^5 + x.$$

For example, at $y = 2$ (the image of $x = 1$) $\left(f^{-1}\right)'(2) = \frac{1}{6}$ and $\left(f^{-1}\right)'(34) = \frac{1}{81}$. Determining $(f^{-1})'(1)$ is by the way just as problematic as is determining $f^{-1}(1)$. $\triangleleft$

EXAMPLE 4.10 Let $n \in \mathbb{N}$ and let f be the function on $I = (0,\infty)$, defined by $f(x) = x^n$. The inverse function f^{-1} of f on $(0,\infty)$ is given by $f^{-1}(y) = y^{\frac{1}{n}}$ (refer to Exercise 3.14). For the function f it holds that $f'(x) = nx^{n-1} > 0$ for every $x \in (0,\infty)$. According to Theorem 4.9, f^{-1} is a differentiable function on $(0,\infty)$, and for every $y > 0$

$$\begin{aligned} \left(f^{-1}\right)'(y) &= \frac{1}{f'(f^{-1}(y))} = \frac{1}{n(f^{-1}(y))^{n-1}} \\ &= \frac{1}{n(y^{\frac{1}{n}})^{n-1}} = \frac{1}{ny^{\frac{n-1}{n}}} = \frac{1}{n}y^{\frac{1}{n}-1}. \triangleleft \end{aligned}$$

EXERCISE 4.20 Prove that for every $m,n \in \mathbb{N}$ the function h on $(0,\infty)$, defined by $h(x) = x^{\frac{m}{n}}$, is differentiable on $(0,\infty)$ and that $h'(x) = \frac{m}{n}x^{\frac{m}{n}-1}$ for every $x \in (0,\infty)$.

EXERCISE 4.21 Use the arithmetic rules for differentiable functions to determine the derivative of the function f on $(0,\infty)$, defined by $f(x) = 3\sqrt[3]{x}(\sqrt{x} + x^3)$.

4.7 Taylor's Theorem

If a function f on an interval I is differentiable and $c \in I$, then f is continuous in c,

$$\lim_{x \to c} f(x) = f(c) \quad \Longleftrightarrow \quad f(x) - f(c) \to 0 \text{ as } x \to c.$$

According to the Mean Value Theorem, applied to f, restricted to the interval with end points x and c, a point ξ_x exists between x and c such that

$$f(x) - f(c) = f'(\xi_x)(x - c).$$

If we interpret $f(x) - f(c)$ as the remainder for the constant approximation h of f at c (defined on $\mathbb{R}$ by $h(x) = f(c)$), then for a differentiable function we have thus obtained an expression for the remainder r corresponding with h, namely

$$r(x) = f(x) - h(x) = f'(\xi_x)(x - c).$$

If f is a twice differentiable function on I and if $c \in I$, then f is differentiable at c,

$$\begin{aligned} \lim_{x \to c} \frac{f(x) - f(c)}{x - c} = f'(c) \quad &\Longleftrightarrow \quad \frac{f(x) - f(c)}{x - c} - f'(c) \to 0 \quad \text{as } x \to c \\ &\Longleftrightarrow \quad \frac{1}{x - c} \left[f(x) - f(c) - f'(c)(x - c) \right] \to 0 \quad \text{as } x \to c. \end{aligned}$$

If we interpret $f(x) - f(c) - f'(c)(x - c)$ as the remainder $r(x) = f(x) - g(x)$ for the linear approximation g of f at c, (defined on $\mathbb{R}$ by $g(x) = f(c) + f'(c)(x-c)$), then for the remainder it is also true that $\frac{r(x)}{x-c} \to 0$ as $x \to c$ (and therefore also that $r(x) \to 0$ as $x \to c$).
Since it was assumed that f is twice differentiable, an explicit expression also exists for the remainder corresponding with the linear approximation. This result is known as Taylor's Theorem, which we provide without proof.

THEOREM 4.10 **(Taylor's Theorem)**
Let f be a twice differentiable function on an interval I and let $c \in I$. Then for every $x \in I \backslash \{c\}$ a point ξ_x between c and x exists such that

$$f(x) = f(c) + f'(c)(x - c) + \frac{f''(\xi_x)}{2} (x - c)^2.$$

The function p_1 on $\mathbb{R}$, defined by $p_1(x) = f(c) + f'(c)(x - c)$, is called the *first order Taylor polynomial* of f at c. (The first order Taylor polynomial coincides therefore with the linear approximation of f at c.) The term $r(x) = \frac{f''(\xi_x)}{2}\,(x - c)^2$ is called the *remainder corresponding with the first order Taylor polynomial* and the formula in Theorem 4.10,

$$f(x) = f(c) + f'(c)(x - c) + \frac{f''(\xi_x)}{2}\,(x - c)^2,$$

is known as the *Taylor formula* (for the first order polynomial).

In analogy to the above, the function p_0 on $\mathbb{R}$, defined by $p_0(x) = f(c)$, is called the *Taylor polynomial of order zero* of f at c. (So the Taylor polynomial of order zero is the constant approximation of f at c.) The term $f'(\xi_x)\,(x-c)$ is called the *remainder corresponding with the Taylor polynomial of order zero* and the formula $f(x) = f(c) + f'(\xi_x)\,(x - c)$ is known as the *Taylor formula* (*for the Taylor polynomial of order zero*).

EXAMPLE 4.11 The function f on $[0, \infty)$ is defined by

$$f(x) = \frac{3x}{1 + 2x}.$$

Since $f(1) = 1$ and $f'(1) = \frac{1}{3}$, the first order Taylor polynomial p_1 on $\mathbb{R}$ is given by

$$p_1(x) = 1 + \tfrac{1}{3}(x - 1).$$

Because $f''(x) = \frac{-12}{(1+2x)^3}$ for every $x > 0$, the remainder $r(x)$ corresponding with p_1 is given by

$$r(x) = \frac{f''(\xi_x)}{2}\,(x - 1)^2 = -\frac{6}{(1 + 2\xi_x)^3}\,(x - 1)^2$$

for every $x \in [0, \infty)$. $\triangleleft$

EXERCISE 4.22 Determine the first order Taylor polynomial p_1 of the function f at the point $c \in I$ and the remainder corresponding with p_1 if

a) $f(x) = \sqrt{x}$, $c = 4$ and $I = (1, 6)$

b) $f(x) = x + \frac{4}{x}$, $c = 2$ and $I = (1, 10)$.

Now we will present two applications of Taylor's Theorem. The first application concerns extreme values of a function and the second concerns the concepts of convexity and concavity of a function, which we must still introduce.

We proved in Theorem 4.4 that $f'(c) = 0$ is a *necessary condition* for an extreme value at an interior point of an interval I (under the assumption that f is differentiable). And it was made clear through simple examples that the condition is not sufficient for an extreme value at an interior point. Examples that do present *sufficient conditions* for an extreme value of f at an interior point $c \in I$ can be formed for instance by demanding that both $f'(c) = 0$ and the first order derivative of f switches sign at c (see Exercise 4.19).

It is also possible to formulate sufficient conditions for an extreme value at an interior point in terms of the second order derivative of the function.

THEOREM 4.11 **(second order condition for extreme value)**

Let f be a twice differentiable function on an interval I and let c be an interior point of I such that $f'(c) = 0$ and f'' is continuous in c. The following holds:

a) if $f''(c) > 0$, then $f(c)$ is a local minimum of f

b) if $f''(c) < 0$, then $f(c)$ is a local maximum of f.

PROOF a) Assume that $f''(c) > 0$. Since f'' is continuous in c, an $\varepsilon > 0$ exists, such that $f''(x) > 0$ for every $x \in (c - \varepsilon, c + \varepsilon)$.

Let $x \in (c - \varepsilon, c + \varepsilon)\backslash\{c\}$. According to Taylor's Theorem for x a point ξ_x between c and x exists, such that

$$f(x) = f(c) + f'(c)(x - c) + \frac{f''(\xi_x)}{2}(x - c)^2 = f(c) + \frac{f''(\xi_x)}{2}(x - c)^2.$$

Since $x \in (c - \varepsilon, c + \varepsilon)$, the intermediate point ξ_x belongs to $(c - \varepsilon, c + \varepsilon)$ so that $f''(\xi_x) > 0$. Since $(x - c)^2 \geq 0$, we have therefore

$$f(x) = f(c) + \frac{f''(\xi_x)}{2}(x - c)^2 \geq f(c) + 0 = f(c).$$

So $f(c)$ is a local minimum of f.

b) Assume that $f'(c) = 0$ and $f''(c) < 0$. Then $(-f)'(c) = 0$ and $(-f)''(c) > 0$. Then according to part a) $(-f)(c)$ is a local minimum of $-f$, so that $f(c)$ is a local maximum of f. $\triangleleft$

EXERCISE 4.23 Determine the extreme value of the function f on $(0, \infty)$, defined by $f(x) = x + \frac{4}{x}$.

For the second application of Taylor's Theorem we consider a twice differentiable function, whose second order derivative has a fixed sign. For such a function a simple geometric relationship exists between the graph of the function and the tangent lines to this graph.

THEOREM 4.12 *Let f be a twice differentiable function on an interval I. It holds that:*
a) if $f'' \geq 0$, then $f \geq g_c$ for every $c \in I$
b) if $f'' \leq 0$, then $f \leq g_c$ for every $c \in I$,
where g_c is the linear approximation of f at c.

In other words: if $f'' \geq 0$, then the graph of f lies above or on each tangent line to the graph of f and if $f'' \leq 0$, then the graph of f lies below or on each tangent line to the graph of f. In case a) the function is called *convex*, in case b) *concave*.

EXERCISE 4.24 Give a proof of Theorem 4.12.

EXERCISE 4.25

a) Determine the interval on which the function f on $\mathbb{R}$, defined by $f(x) = \frac{1}{1+x^2}$, is convex and the interval on which f is concave.

b) Show that the function g on $(0, \infty)$, defined by $g(x) = x + \frac{4}{x}$ is convex.

If a function is differentiable even more times, then we can determine higher order Taylor polynomials, and as we did in Theorem 4.10 we can deduce a formula for the remainder: if f is a three times differentiable function on an interval I and c an interior point of I then for every $x \in I \backslash \{c\}$ a point ξ_x between c and x exists such that

$$f(x) = f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \frac{f'''(\xi_x)}{3!}(x-c)^3.$$

The quadratic function p_2 on $\mathbb{R}$, defined by $p_2(x) = f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2$, is called the *second order Taylor polynomial* of f at c.

An approximation of f at c by means of a second order Taylor polynomial is a better approximation than an approximation by means of a first order Taylor polynomial, which in turn is a better approximation than an approximation by means of a Taylor polynomial of order zero. Of course this process of determining ever better approximations of f at c by means of higher order Taylor polynomials does not stop at the second order.

EXERCISE 4.26 Determine the second order Taylor polynomial p_2 of the function f at the point $c \in I$ and the remainder with p_2 if

a) $f(x) = \sqrt{x}$, $c = 1$ and $I = (0, \infty)$

b) $f(x) = \frac{3x}{1+2x}$, $c = 1$ and $I = (0, \infty)$.

4.8 Rule of de l'Hôpital

In Chapter 2 we proved the quotient rule for limits of functions:

$$\text{if } \lim_{x \to c} f(x) = l \text{ and } \lim_{x \to c} g(x) = m \neq 0, \text{ then } \lim_{x \to c} \frac{f(x)}{g(x)} = \frac{l}{m}.$$

The quotient rule gives no definite answer for the case $m = 0$. If $m = 0$, then it can easily be proved that the quotient $\frac{f(x)}{g(x)}$ can only have a limit for $x \to c$ if $l = 0$ as well.

If f and g are functions with the property that $\lim_{x \to c} f(x) = 0$ and $\lim_{x \to c} g(x) = 0$, then we call the quotient $\frac{f(x)}{g(x)}$ an *indefinite form at c of the type $\frac{0}{0}$*. There are indefinite forms of the type $\frac{0}{0}$ which have a limit at c ($\frac{\alpha x}{x} \to \alpha$ for $x \to 0$) and there are indefinite forms of the type $\frac{0}{0}$ which have no limit at c.

EXERCISE 4.27 Find two functions f and g such that $\frac{f(x)}{g(x)}$ is an indefinite form at 0 of the type $\frac{0}{0}$, which has no limit as $x \to 0$.

The rule of de l'Hôpital concerns the limit of an indefinite form. We formulate a simple version of the rule of de l'Hôpital.

THEOREM 4.13 **(Rule of de l'Hôpital)**
Let f and g be differentiable functions on an interval I and let $c \in I$. Assume that $g'(x) \neq 0$ for every $x \in I$. The following holds:

$$\textit{if } f(c) = g(c) = 0, \textit{then } \lim_{x \to c} \frac{f(x)}{g(x)} = \frac{f'(c)}{g'(c)}.$$

PROOF First we will show that it is meaningful to speak of the quotient $\frac{f(x)}{g(x)}$ on the interval I. By means of a proof by contradiction we prove that the conditions formulated in the theorem imply that $g(x) \neq 0$ for every $x \in I \backslash \{c\}$.

Assume that a $d \in I, d \neq c$, exists for which $g(d) = 0$. According to Rolle's Theorem, applied to the function g and the interval with end points c and d, a point ξ between c and d exists such that $g'(\xi) = 0$. This contradicts the given fact that $g'(x) \neq 0$ for every $x \in I$.
For every $x \in I, x \neq c$, the quotient $\frac{f(x)}{g(x)}$ can be rewritten by means of the following subsequent steps:

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(c)}{g(x) - g(c)} = \frac{\frac{f(x)-f(c)}{x-c}}{\frac{g(x)-g(c)}{x-c}}.$$

Because the limits of numerator and denominator at c exist and $g'(c) \neq 0$, according to the quotient rule for limits

$$\lim_{x \to c} \frac{f(x)}{g(x)} = \frac{\lim\limits_{x \to c} \frac{f(x)-f(c)}{x-c}}{\lim\limits_{x \to c} \frac{g(x)-g(c)}{x-c}} = \frac{f'(c)}{g'(c)}. \qquad \triangleleft$$

EXERCISE 4.28 Determine, if possible, the limits below:
a) $\lim\limits_{x \to 1} \dfrac{2x^2 - x - 1}{x^3 - 1}$ b) $\lim\limits_{x \to 0} \dfrac{x}{2x}$ c) $\lim\limits_{x \to 0} \dfrac{2x}{x + 8}$.

The following theorem provides a version of the rule of de l'Hôpital for functions f and g which are not necessarily differentiable at c. We will omit the proof.

THEOREM 4.14 *Let f and g be continuous functions on an interval I and let $c \in I$. Assume that f and g are differentiable on $I \setminus \{c\}$ and that $g'(x) \neq 0$ for every $x \in I \setminus \{c\}$. The following holds:*

$$\textit{if } f(c) = g(c) = 0 \ \textit{ and } \ \lim_{x \to c} \frac{f'(x)}{g'(x)} = L, \textit{then } \lim_{x \to c} \frac{f(x)}{g(x)} = L.$$

EXAMPLE 4.12 In order to show that

$$\lim_{x \to 1} \frac{x^3 - 1}{\sqrt[3]{(x-1)^2}} = 0,$$

we introduce the functions f and g on $[1, \infty)$ defined by $f(x) = x^3 - 1$ and $g(x) = \sqrt[3]{(x-1)^2}$. The functions f and g are continuous on $[1, \infty)$, they are differentiable on $(1, \infty)$ and $g'(x) = \frac{2}{3\sqrt[3]{(x-1)^2}} \neq 0$ for all $x > 1$.
Because $f(1) = g(1) = 0$ and

$$\frac{f'(x)}{g'(x)} = \tfrac{9}{2} x^2 \sqrt[3]{x - 1} \to 0 \ \text{ as } \ x \to 1,$$

Theorem 4.14 implies that $\lim_{x \to 1} \frac{x^3-1}{\sqrt[3]{(x-1)^2}} = 0$. $\triangleleft$

Apart from Theorem 4.14, other versions exist of the rule of de l'Hôpital where the quotient of the derivatives diverges to $+\infty$ or to $-\infty$. In that case the quotient of the functions themselves also diverges to $+\infty$ or $-\infty$ respectively. And in other versions of the rule of de l'Hôpital we have so-called indefinite forms of the type $\frac{\infty}{\infty}$. The treatment of each of these versions falls outside the scope of this course. You will have to consult the existing literature for them.

4.9 Mixed exercises

EXERCISE 4.29 Let f be a function on $\mathbb{R}$ which is continuous in 0 and for which $\lim_{x \to 0} \dfrac{f(x)}{x}$ exists.

a) Prove that $f(0) = 0$.

b) Prove that f is differentiable in 0.

EXERCISE 4.30 Show by means of the definition of derivative of a function that the function f on $(0, \infty)$, defined by $f(x) = \frac{1}{x}$, is differentiable and determine the derivative function.

EXERCISE 4.31 Let f be a differentiable function on an interval I.

Prove that for every $n \in \mathbb{N}$ the function f^n $(= f \cdot f \cdot \cdots \cdot f$, ($n$ times)) is differentiable and that

$$(f^n)'(x) = n\,(f(x))^{n-1}\,f'(x) \quad \text{for every} \quad x \in I.$$

EXERCISE 4.32 Determine the linear approximation in 1 of the function f on $\mathbb{R}$, defined by $f(x) = x^3$, and the equation of the tangent to the graph of f in the point $(1, 1)$.

EXERCISE 4.33 Define the function g on $\mathbb{R}$ by $g(x) = x^2$. Let f be a differentiable function on $\mathbb{R}$ with the property that for every $x \in \mathbb{R}$ it holds that

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

Prove that $f(1) = 1$ or $f'(1) = 0$.

EXERCISE 4.34 Prove the following statement or disprove it by means of a counter-example:
if f is a function on $\mathbb{R}$, which for every $x \in \mathbb{R}\backslash\{1\}$ is differentiable in x and $f'(x) = 0$, then f is a constant function.

EXERCISE 4.35 The function f on $[0, 2]$ is defined by $f(x) = x^5 - 16x$.
Determine a point $\xi \in (0, 2)$ such that $f'(\xi) = \frac{f(2)-f(0)}{2}$.

EXERCISE 4.36 The function p on $\mathbb{R}$ is defined by $p(x) = x^3 + ax + b$, where $a > 0$ and $b \in \mathbb{R}$.
Prove that p has precisely one zero.

EXERCISE 4.37 Determine for each of the functions below the intervals, on which the function is (strictly) increasing, and the intervals on which the function is (strictly) decreasing:
a) the function f on $(0, \infty)$, defined by $f(x) = \sqrt{x} - 2\sqrt{x+2}$
b) the function g on $\mathbb{R}$, defined by $g(x) = \frac{x}{2+x^2}$
c) the function h on $[0, 2]$, defined by $h(x) = 1 - \sqrt[3]{(x-1)^2}$.

EXERCISE 4.38 Determine the first order Taylor polynomial p_1 of the function f at the point $c \in I$ and the remainder corresponding with p_1 if
a) $f(x) = x^2 + \sqrt{x} \quad c = 1$ and $I = (\frac{1}{4}, 4)$
b) $f(x) = x^2 \quad c = 2$ and $I = (0, 28)$
c) $f(x) = \frac{1}{2+x} \quad c = 0$ and $I = (-1, 1)$.

EXERCISE 4.39
a) Let f be a continuous function on $[a, b]$, which is differentiable on (a, b). Prove the statement:
if $\lim\limits_{x \to a} f'(x) = l$, then f is differentiable at a and $f'(a) = l$.
b) Give an example of a function f which is continuous on $[0, 1]$ and differentiable on $(0, 1)$, but whose derivative at 0 does not exist.

EXERCISE 4.40 Determine, if possible, the following limits:
a) $\lim\limits_{x \to 0} \frac{1}{x}(\sqrt{x+1} - 1 + \frac{1}{2}x)$ b) $\lim\limits_{x \to 2} \dfrac{x^3 - 7x^2 + 16x - 12}{x^2 - 4x + 4}$.

EXERCISE 4.41 The quadratic function f on $I = (-\frac{1}{2}, \frac{1}{2})$ is defined by $f(x) = x^2$.

a) Determine the first order Taylor polynomial p_1 of f at 0.

b) Show that for the remainder $r(x)$ corresponding with p_1 on I it holds that $|r(x)| \leq \frac{1}{4}$ for every $x \in I$.

5 POWER SERIES

In this chapter we consider a special class of differentiable functions, the power series. In section 5.1 we consider an introductory example and in section 5.2 we consider general power series. We will introduce the concept of 'radius of convergence', which specifies the domain of the power series and we provide a theorem about differentiability of a power series.

5.1 An introductory example

The family of polynomial functions is wellknown. These are functions like $x^2 + x + 2$, $7x^3 + 18x - 12$, $x^{2018} + 3x^3 - 17$ etcetera. In general, a polynomial function of *degree* n is specified by $n+1$ constants $a_0, a_1, a_2, \ldots, a_n$ ($a_n \neq 0$) and defined by

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = \sum_{k=0}^{n} a_k \cdot x^k$$

for every $x \in \mathbb{R}$. It is obvious that $f(x)$ can be computed for every $x \in \mathbb{R}$, in other words $D_f = \mathbb{R}$.

Power series are polynomials of infinite degree, i.e. they are specified by an infinite sequence of constants $a_0, a_1, a_2, a_3, a_4, \ldots$. So, a power series is a function that is defined in the following way:

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + \cdots = \sum_{k=0}^{\infty} a_k \cdot x^k. \tag{1}$$

Two questions arise when considering an expression like (**??**). The first question is: for which values of x does the infinite summation (**??**) provide a meaningful numerical outcome? In other words, what is the domain D_f of the function f? The second question is: can we determine the outcome of the infinite summation for every $x \in D_f$? In the sequel we will see that the answer to the first question is not that difficult, whereas the answer to the second one in many cases is impossible to give. Let us first consider a relatively simple example.

EXAMPLE 5.1 Consider the power series specified by the sequence of constants $a_0, a_1, a_2, a_3, a_4, \ldots$ such that $a_k = 1$ for every $k = 0, 1, 2, \ldots$. So consider the function

$$f(x) = 1 + x + x^2 + x^3 + x^4 + \cdots = \sum_{k=0}^{\infty} x^k.$$

What is the domain of f? Let us try some values of x. First try $x = 0$. Then

$$f(0) = 1 + 0 + 0 + 0 + 0 + \cdots \quad .$$

It is not surprising that the outcome of this infinite summation is 1, so $0 \in D_f$ and $f(0) = 1$. Now take $x = \frac{1}{2}$. Then we get

$$f(\frac{1}{2}) = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \frac{1}{64} + \cdots \quad .$$

Can we compute the outcome of this infinite summation? Well, let us take our calculator and add some terms. Taking the first term we get 1, adding the second one we get $1\frac{1}{2}$, adding the third one we get $1\frac{3}{4}$, adding the fourth one we get $1\frac{7}{8}$, adding the fifth one we get $1\frac{15}{16}$, etcetera. So the display of our calculator will show a number that is slightly smaller than 2, but gets arbitrarily close to 2 if we add a sufficient number of terms. The limit of the numbers shown in the display is 2. In other words, the outcome of the infinite summation is 2, so $\frac{1}{2} \in D_f$ and $f(\frac{1}{2}) = 2$.

Now take $x = 2$. We have

$$f(2) = 1 + 2 + 4 + 8 + 16 + \cdots \quad .$$

Again taking our calculator for computing the sums of a finite collection of terms we consecutively get $1, 3, 7, 15, 31, 63, 127, \ldots$ in the display of our calculator. In other words, this sequence of numbers tends to infinity, indicating that the infinite summation does not yield a numerical outcome. In other words, $2 \notin D_f$ so $f(2)$ is not well-defined.

Can we also take negative values for x? Why not, take for example $x = -\frac{1}{3}$? We have

$$f(-\frac{1}{3}) = 1 + (-\frac{1}{3}) + (-\frac{1}{3})^2 + (-\frac{1}{3})^3 + (-\frac{1}{3})^4 + \cdots \quad .$$

The display of our calculator shows $1, \frac{2}{3}, \frac{7}{9}, \frac{20}{27}, \frac{61}{81}, \ldots$ and these values seem to get closer and closer to $\frac{3}{4}$. Is $-\frac{1}{3} \in D_f$ and $f(-\frac{1}{3}) = \frac{3}{4}$?

Let us consider the problem in a more systematic way. Take an arbitrary $x \in \mathbb{R}$. We would like to check whether $x \in D_f$ and, if so, compute $f(x)$. The display of our calculator shows 1, $1+x$, $1+x+x^2$, $1+x+x^2+x^3, \ldots$ and, after n steps, $1+x+x^2+\cdots+x^{n-1} = \sum_{k=0}^{n-1} x^k$. Can we write this number in a different way? The answer is yes. In Exercise 1.10 b) we have seen that, in case $x \neq 1$,

$$1 + x + x^2 + \cdots + x^{n-1} = \frac{1 - x^n}{1 - x}.$$

If $-1 < x < 1$ then x^n converges to 0 if $n \to \infty$ so the number in the display of the calculator converges to $\frac{1}{1-x}$ if we add more and more terms. If $x < -1$ or $x > 1$ then x^n becomes, in absolute value, larger and larger and the numbers in the display of our calculator do not converge. If $x = -1$ then we are considering

$$f(-1) = 1 + (-1) + 1 + (-1) + 1 + (-1) + \cdots$$

and the numbers displayed are $1, 0, 1, 0, 1, 0, \ldots$. This sequence does not converge. Lastly, if $x = 1$ then

$$f(1) = 1 + 1 + 1 + 1 + 1 + 1 + \cdots$$

and the sequence displayed is $1, 2, 3, 4, 5, 6, \ldots$. This sequence does not converge as well. Summarizing, we find that $D_f = (-1, 1)$ and

$$f(x) = 1 + x + x^2 + x^3 + x^4 + \cdots = \frac{1}{1 - x}$$

for every $x \in (-1, 1)$. Note that indeed $f(\frac{1}{2}) = \frac{1}{1-\frac{1}{2}} = 2$ and $f(-\frac{1}{3}) = \frac{1}{1-(-\frac{1}{3})} = \frac{3}{4}$. ◁

5.2 Power series in general

In this section we consider general power series. So let $a_0, a_1, a_2, a_3, \ldots$ or $(a_k)_{k=0}^{\infty}$ be an infinite sequence of constants and consider

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + \cdots \quad .$$

The following theorem provides an answer to the question about the domain of f. We will not provide the proof of this theorem.

THEOREM 5.1 **(radius of convergence)**

Let $(a_k)_{k=0}^{\infty}$ be a sequence of real numbers with $a_k \neq 0$ for every $k \in \mathbb{N}$. Consider

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + \cdots \quad .$$

a) If $\lim_{k\to\infty} |\frac{a_k}{a_{k+1}}| = R$ and $R \neq 0$, then for every $x \in \mathbb{R}$ we have:
if $|x| < R$, then $x \in D_f$.
if $|x| > R$, then $x \notin D_f$.
So either $D_f = (-R, R)$ or $D_f = (-R, R]$ or $D_f = [-R, R)$ or $D_f = [-R, R]$.
(We call R the radius of convergence *of the power series $\sum_{n=1}^{\infty} a_n x^n$.)*
b) If $(|\frac{a_k}{a_{k+1}}|)_{k=1}^{\infty}$ is divergent to ∞, then $D_f = \mathbb{R}$.
(The radius of convergence of the power series $\sum_{n=1}^{\infty} a_n x^n$ is infinite.)
c) If $\lim_{k\to\infty} |\frac{a_k}{a_{k+1}}| = 0$, then $D_f = \{0\}$.
(The radius of convergence of the power series $\sum_{k=1}^{\infty} a_k x^k$ is zero.)

EXAMPLE 5.2 We determine the radius of convergence of the power series

$$f(x) = \sum_{k=1}^{\infty} \frac{x^k}{k \cdot 2^k}.$$

Let $a_0 = 0$ and for every $k \in \mathbb{N}$ we define $a_k = \frac{1}{k \cdot 2^k}$. We have $|\frac{a_k}{a_{k+1}}| = |\frac{(k+1)\cdot 2^{k+1}}{k \cdot 2^k}| = |\frac{2(k+1)}{k}| = \frac{2(k+1)}{k} = 2 + \frac{2}{k}$ so $\lim_{k\to\infty} |\frac{a_k}{a_{k+1}}| = 2$.
So, the radius of convergence is 2 and $D_f = (-2, 2)$ or $D_f = (-2, 2]$ or $D_f = [-2, 2)$ or $D_f = [-2, 2]$. $\triangleleft$

EXERCISE 5.1 Determine the radius of convergence of each of the following power series:

a) $\sum_{k=1}^{\infty} (-1)^k k! x^k$
b) $\sum_{k=1}^{\infty} x^k$
c) $\sum_{k=1}^{\infty} \frac{x^k}{k}$.

EXAMPLE 5.3 We will show that the power series

$$f(x) = 1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \cdots$$

has $\mathbb{R}$ as its domain. Define for every $k = 0, 1, 2, \ldots$, $a_k = \frac{1}{k!}$. Then $|\frac{a_k}{a_{k+1}}| = |\frac{(k+1)!}{k!}| = k + 1$. So, the sequence $|\frac{a_k}{a_{k+1}}|$ is divergent to ∞. Hence, from Theorem 5.1 it follows that $D_f = \mathbb{R}$.

This power sequence is famous, by the way. One can show that $f(x) = e^x$ for every $x \in \mathbb{R}$. $\triangleleft$

The following theorem shows that you can differentiate a power series termwise. Again we state the result without proof.

THEOREM 5.2 *Let $a_0, a_1, a_2, \ldots$ be a sequence of real numbers and R the radius of convergence of the power series $\sum_{k=0}^{\infty} a_k x^k$. We have: the function f on $(-R, R)$, defined by*

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

is differentiable (and hence continuous) and

$$f'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1} \qquad \text{for every } x \in (-R, R).$$

EXERCISE 5.2

a) Consider the power series $f(x)$ in Example 5.3. Show that $f'(x) = f(x)$ for every $x \in \mathbb{R}$ (most likely you remember this property of the function e^x from high school).

b) Determine the radius of convergence R of the power series $\sum_{k=1}^{\infty} kx^k$.

c) Determine for every $x \in \mathbb{R}$ with $|x| < R$ the power series $\sum_{k=1}^{\infty} kx^k$.

5.3 Mixed exercises

EXERCISE 5.3 Determine the radius of convergence of each of the following power series:

a) $\sum_{n=0}^{\infty} \frac{n^n}{n!} x^n$. Here you can use that $\lim_{n\to\infty} (1 + \frac{x}{n})^n = e^x$.

b) $\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^k$

c) $\sum_{k=0}^{\infty} \frac{x^{k+2}}{(k+2)k!}$.

EXERCISE 5.4 Consider the power series $f(x) = \sum_{k=0}^{\infty} k^2 x^k$.

a) Determine the radius of convergence R of the power series.

b) Determine for every $x \in \mathbb{R}$ with $|x| < R$ the value $\sum_{k=0}^{\infty} k^2 x^k$.

EXERCISE 5.5

a) Determine an explicit expression for $\sum\limits_{k=1}^{\infty} x^k = x + x^2 + x^3 + \ldots$ for $x \in (-1, 1)$, see also exercise 5.1b.

b) Determine $\sum\limits_{k=1}^{\infty} kx^{k-1}$ for $x \in (-1, 1)$.

EXERCISE 5.6

Consider the power series

$$\sum_{k=0}^{\infty}(3^k + 4^k)x^k.$$

a) Determine the radius of convergence R of this power series.

b) Determine an explicit expression for $\sum\limits_{k=0}^{\infty}(3^k + 4^k)x^k$ for every $x \in (-R, R)$.

6 INTEGRATION

Most likely you remember from high school that integration has something to do with the determination of 'the area below the graph of a function' and that integration is the 'reverse' of differentiation.

In section 6.1 we provide the formal definition of the integral of a continuous function over a compact interval. In section 6.2 we develop a technique for the calculation of integrals. Since this technique strongly relies upon the *antiderivative* of a function, an operation that is 'reverse' to differentiation, we come up in section 6.3 with a list of antiderivatives of well-known functions. In section 6.4 we discuss two important integration techniques: integration by parts and integration by substitution. Indefinite integrals show up in section 6.5.

6.1 The integral

In this section we will give the formal definition of the integral of a continuous function over a compact interval $[a, b]$. In order to be able to give this definition we need the concepts of 'partition' and 'under sum'.

DEFINITION A *partition* P of the compact interval $[a, b]$ is a finite sequence of points $(x_0, x_1, \ldots, x_n)$ such that

$$a = x_0 < x_1 < \ldots < x_n = b.$$

If $P = (x_0, x_1, \ldots, x_n)$ is a partition of $[a, b]$, then the interval $[a, b]$ can be considered as the union of the subintervals $[x_0, x_1], [x_1, x_2], \ldots, [x_{n-1}, x_n]$.

EXAMPLE 6.1 Examples of partitions of the interval $[0, 1]$ are

$$P_1 = (0,\ \frac{1}{10},\ \frac{2}{10},\ \frac{3}{10},\ \frac{4}{10},\ \frac{5}{10},\ \frac{6}{10},\ \frac{7}{10},\ \frac{8}{10},\ \frac{9}{10},\ 1),$$

$$P_2 = (0,\ \frac{1}{6},\ \frac{1}{4},\ \frac{3}{7},\ \frac{1}{2},\ \frac{3}{4},\ \frac{7}{8},\ 1),$$

$$P_3 = (0,\ 0.12,\ 0.15,\ 0.32,\ 0.43,\ 0.55,\ 0.65,\ 0.83,\ 1).$$

The partition P_1 has the property that each subinterval $[x_{i-1}, x_i]$ has the same length (namely $\frac{1}{10}$). Therefore, P_1 is also called an *equidistant* (= same distance) *partition* of $[0, 1]$. $\triangleleft$

EXERCISE 6.1 Determine a partition $P = (x_0, x_1, \dots, x_6)$ of $[0, 1]$ such, that the length of each subinterval $[x_{i-1}, x_i]$ is at most $\frac{1}{4}$.

Let us assume that f is a continuous function on $[a, b]$, which is non-negative (that is, $f(x) \geq 0$ for every $x \in [a, b]$) and that $P = (x_0, x_1, \dots, x_n)$ is a partition of $[a, b]$. For each subinterval $[x_{i-1}, x_i]$ we can consider the rectangle with height equal to the minimum of the function on this subinterval (as drawn in Figure 6.1). Since the area of this rectangle equals the length $(x_i - x_{i-1})$ times he height $(\min\{f(x)\ :\ x \in [x_{i-1}, x_i]\})$, the number

$$\sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \min\{f(x)\ :\ x \in [x_{i-1}, x_i]\}$$

represents the total area of the histogram.

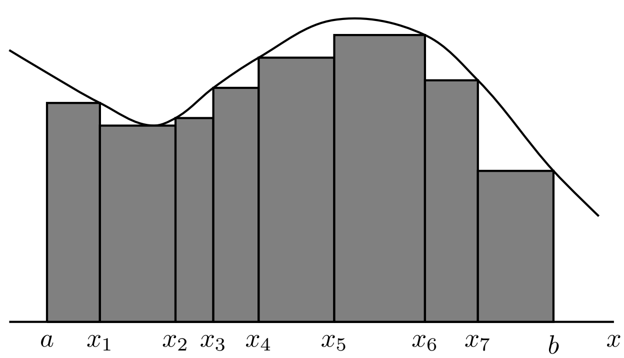

Figure 6.1: The area of the histogram.

Since the existence of the minimum of f on each of the subintervals $[x_{i-1}, x_i]$ is a consequence of the continuity of f (according to the Theorem of Weierstrass, applied to f and the compact interval $[x_{i-1}, x_i]$), we can define the sum above without problems for an arbitrary continuous function, so also for a continuous function with possibly negative values. This leads to the definition of under sum of a function for some partition.

DEFINITION Let f be a continuous function on the interval $[a, b]$ and let $P = (x_0, x_1, \ldots, x_n)$ be a partition of $[a, b]$. The number

$$\underline{S}_P = \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \min\{f(x) \ : \ x \in [x_{i-1}, x_i]\}$$

is called the *under sum* of f, corresponding to the partition P.

EXERCISE 6.2 The function f is defined on $[1, 4]$ by $f(x) = (x - 3)^2 - 1$.
Determine the under sum of f corresponding to the partition $P = (1, \frac{3}{2}, 2, \frac{5}{2}, 3, 4)$.

EXERCISE 6.3 The function f is defined on $[0, 1]$ by $f(x) = x^2$.
Determine a partition P of $[0, 1]$ such, that for the under sum $\underline{S}_P$ of f, corresponding to the partition P, we have

$$\underline{S}_P > \frac{1}{8}.$$

Again, if the function f is non-negative, then the under sum $\underline{S}_P$ of f, corresponding to the partition P, represents the area of a histogram. Since this histogram lies completely under the graph of f, this under sum is not more than the area of the region under the graph of f (see Figure 6.1). We get a better approximation of the area of the region under the graph of f if we add more points to the points of the partition P of $[a, b]$ and determine subsequently the under sum of f for this 'finer' partition. The finer the partition of $[a, b]$ is, the better is the approximation of the area of the region under the graph by the corresponding under sum. Intuitively it is clear that the supremum of the set of under sums equals the area of the region under the graph.

This observation is the basis for the definition of the integral of an arbitrary continuous function, so also of a continuous function with possibly negative values. Of course, we first have to show that the supremum of the set of under sums exist, in other words, that the set of under sums is bounded from above.

THEOREM 6.1 *Let f be a continuous function on $[a, b]$ and $M = \max\{f(x) \ : \ x \in [a, b]\}$. For every partition P of $[a, b]$ we have:*

$$\underline{S}_P \leq M(b - a).$$

PROOF Let $P = (x_0, x_1, \ldots, x_n)$ be a partition of $[a, b]$. We have:

$$\begin{aligned}
\underline{S}_P &= \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \min\{f(x) \ : \ x \in [x_{i-1}, x_i]\} \\
&\leq \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot M \\
&= M \cdot \sum_{i=1}^{n} (x_i - x_{i-1}) = M(b - a).
\end{aligned}$$

◁

Hence, the set of under sums of a function f on the interval $[a, b]$ is a nonempty set of numbers that is bounded from above. According to the Axiom of Completeness this set has a supremum. We define the integral of f over $[a, b]$ as the supremum of this set.

DEFINITION Let f be a continuous function on the interval $[a, b]$.
The number

$$\sup\{\underline{S}_P \ : \ P \text{ is a partition of } [a, b]\}$$

is called the *integral of f over* $[a, b]$ and denoted by $\int_a^b f(x)\, dx$.

The integral of f over $[a, b]$ is also called the *integral of f from a to b.* The integral of f over $[a, b]$ is a number and not a function. The x in $\int_a^b f(x)\, dx$ has to be considered as a dummy variable that may be replaced with any other variable without changing the integral.

EXERCISE 6.4 Let f be a continuous function on $[a, b]$.

a) Show, by using the definition of the integral, that $\int_a^a f(x)\, dx = 0$.

b) Show the property: if f is nonnegative, then $\int_a^b f(t)\, dt \geq 0$.

EXERCISE 6.5 The function f on $\mathbb{R}$ is defined by $f(x) = 10 - 2x$.

a) Determine $\underline{S}_P$ if

(i) $P = (0, 2\frac{1}{2}, 5)$,

(ii) $P = (0, 1, 2, 3, 4, 5)$,

(iii) $P = (0, \frac{1}{2}, 1, 1\frac{1}{2}, 2, 2\frac{1}{2}, 3, 3\frac{1}{2}, 4, 4\frac{1}{2}, 5)$.

b) Determine (as area of the region under the graph of f)

$$\int_0^5 f(x)\, dx.$$

Exercise 6.5 illustrates that the calculation of an integral by definition is impossible. We need a technique for the calculation of integrals. This is provided by the Principal Theorem of Integral Calculus, Theorem 6.3.

We have defined the integral of a continuous function f on an interval $[a, b]$ as the supremum of all under sums of f over the interval $[a, b]$, where the under sum of f for a partition $P = (x_0, x_1, \ldots, x_n)$ of $[a, b]$ is defined by means of the minimum of f over each of the subintervals $[x_{i-1}, x_i]$. We could have used a similar procedure by taking the maximum of f over each of the subintervals $[x_{i-1}, x_i]$ of $[a, b]$ instead of the minimum of f. This procedure leads to the definition of the *upper sum* $\overline{S}_P$ of a function f for a partition P of $[a, b]$,

$$\overline{S}_P = \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \max\{f(x) \ : \ x \in [x_{i-1}, x_i]\}.$$

We can show that for every partition P of $[a, b]$ we have

$$m(b - a) \leq \overline{S}_P,$$

where $m = \min\{f(x) \ : \ x \in [a, b]\}$. Hence the set of upper sums of f is a nonempty set of numbers that is bounded from below and admits an infimum. We could also define the integral of f from a to b as the infimum of this set of upper sums,

$$\int_a^b f(x)\, dx = \inf\{\overline{S}_P \ : \ P \text{ is a partition of } [a, b]\}.$$

Since for a continuous function f it can be shown that

$$\begin{gathered}\inf\{\overline{S}_P\ :\ P \text{ is a partition of } [a,b]\}\\ =\\ \sup\{\underline{S}_P\ :\ P \text{ is a partition of } [a,b]\},\end{gathered}$$

the choice for the definition with upper sums or the definition with lower sums does not matter: the outcome is the same.

6.2 The Principal Theorem of Integral Calculus

In section 6.1 we have defined the concept 'integral' for a continuous function. We already observed that using this definition is definitely not the way for calculating integrals. It is difficult to compute for all partitions of the interval $[a,b]$ the corresponding under sum and subsequently the supremum of the set of under sums (of course there are infinitely many partitions of $[a,b]$).

However, there is a technique to find the integral of a function on an interval. This technique strongly relies upon the method of finding an antiderivative of the function, an operation that is reverse to differentiation. In order to prove that this technique is correct we need two auxiliary lemmas. Since the proofs of these lemmas are quite technical, we do not present them here.

LEMMA 6.1 *Let f be a continuous function on the interval $[a,b]$.*
a) For every upper bound u of f (that is, $f(x) \leq u$ for every $x \in [a,b]$) we have:

$$\int_a^b f(t)\ dt \leq u(b-a).$$

b) For every lower bound l of f (that is, $l \leq f(x)$ for every $x \in [a,b]$) we have:

$$l(b-a) \leq \int_a^b f(t)\ dt.$$

LEMMA 6.2 *Let f be a continuous function on $[a,b]$ and $a \leq c \leq b$. We have:*

$$\int_a^b f(t)\ dt = \int_a^c f(t)\ dt + \int_c^b f(t)\ dt.$$

In the following theorem you will see that there is a relation between integration and differentiation. ◁

THEOREM 6.2 *Let f be a continuous function on $[a, b]$. The function G on $[a, b]$, defined by*

$$G(x) = \int_a^x f(t)\, dt,$$

is differentiable and we have

$$G'(x) = f(x) \text{ for every } x \in [a, b].$$

PROOF Let $c \in [a, b]$. The function G is differentiable in c with derivative $f(c)$ if

$$\lim_{x \to c} \frac{G(x) - G(c)}{x - c} = f(c).$$

Let $\varepsilon > 0$.

Since f is continuous, there exists a $\delta > 0$ such, that for every $t \in [a, b]$ with $|t - c| < \delta$ we have $|f(t) - f(c)| < \varepsilon$, or,

$$f(c) - \varepsilon < f(t) < f(c) + \varepsilon.$$

Let $\delta_\varepsilon = \delta$ and let subsequently $x \in [a, b]$ be such, that $0 < |x - c| < \delta_\varepsilon$. Now we distinguish between two cases.

Case 1: $c < x < c + \delta_\varepsilon$. According to Lemma 6.2 we have

$$G(x) - G(c) = \int_a^x f(t)\, dt - \int_a^c f(t)\, dt = \int_c^x f(t)\, dt.$$

Since $f(c) - \varepsilon < f(t) < f(c) + \varepsilon$ for every $t \in [c, x]$ it follows from Lemma 6.1 that

$$(f(c) - \varepsilon)(x - c) \le \int_c^x f(t)\, dt \le (f(c) + \varepsilon)(x - c).$$

This implies that

$$f(c) - \varepsilon \le \frac{G(x) - G(c)}{x - c} \le < f(c) + \varepsilon,$$

or,

$$|\frac{G(x) - G(c)}{x - c} - f(c)| \le \varepsilon.$$

Case 2: $c - \delta_\varepsilon < x < c$. According to Lemma 6.2 we have

$$G(x) - G(c) = \int_a^x f(t)\, dt - \int_a^c f(t)\, dt = -\int_x^c f(t)\, dt.$$

Since $f(c) - \varepsilon < f(t) < f(c) + \varepsilon$ for every $t \in [x, c]$, it follows from Lemma 6.1 that

$$(f(c) - \varepsilon)(c - x) \leq \int_x^c f(t)\, dt \leq (f(c) + \varepsilon)(c - x),$$

so that

$$(f(c) - \varepsilon)(x - c) \geq -\int_x^c f(t)\, dt \geq (f(c) + \varepsilon)(x - c).$$

This implies that ($x - c < 0$!)

$$f(c) - \varepsilon \leq \frac{G(x) - G(c)}{x - c} \leq f(c) + \varepsilon,$$

or,

$$\left|\frac{G(x) - G(c)}{x - c} - f(c)\right| \leq \varepsilon. \qquad \triangleleft$$

EXERCISE 6.6 Determine the derivative of the function G, defined by

a) $G(x) = \int_0^x e^{-t^2}\, dt$

b) $G(x) = x \int_1^x e^{-u^2}\, du$.

The function G, defined in Theorem 6.2, has the property that $G'(x) = f(x)$ for every $x \in [a, b]$. Such a function is called an antiderivative of f.

DEFINITION Let f be a function on $[a, b]$.

A differentiable function F on $[a, b]$ is called an *antiderivative* of f if

$$F'(x) = f(x) \text{ for every } x \in [a, b].$$

According to Theorem 6.2 every continuous function on a compact interval has an antiderivative. It is not correct to mention *the* antiderivative of f, since an antiderivative is not unique. For example, the function f on $\mathbb{R}$, defined by $f(x) = x^2$, has as antiderivatives:

$$\frac{1}{3}x^3, \;\; \frac{1}{3}x^3 + 1, \;\; \frac{1}{3}x^3 - 5, \;\; \ldots .$$

The result of Theorem 6.2 is the basis for a technique for the computation of the integral of a function.

THEOREM 6.3 **(Principal Theorem for Integral Calculus)**
Let f be a continuous function on $[a, b]$.
If F is an antiderivative of f, then we have:

$$\int_a^b f(t)\, dt = F(b) - F(a).$$

PROOF Consider the function G on $[a, b]$, defined by

$$G(x) = \int_a^x f(t)\, dt.$$

According to Theorem 6.2 we have $G'(x) = f(x)$ for every $x \in [a, b]$. Since we also have that $F'(x) = f(x)$ for every $x \in [a, b]$, we have $G' = F'$, or, $(G - F)' = 0$. According to Theorem 4.7 $G - F$ is a constant function, in other words, there exists a constant C such that $G(x) = F(x) + C$ for every $x \in [a, b]$. Since $G(a) = 0$ and hence $F(a) + C = 0$, we find that $C = -F(a)$. This yields:

$$\int_a^b f(t)\, dt = G(b) = F(b) + C = F(b) - F(a). \qquad \triangleleft$$

The expression $F(b) - F(a)$ is sometimes denoted as $[F(x)]_{x=a}^{x=b}$, so that

$$\int_a^b f(x)\, dx = [F(x)]_{x=a}^{x=b}.$$

With the Principal Theorem of Integral Calculus the problem of computing the integral of a function f is reduced to the determination of an antiderivative of f.

EXAMPLE 6.2

$$\int_3^6 x^2\, dx = \left[\frac{1}{3}x^3\right]_{x=3}^{x=6} = (\frac{1}{3}6^3) - (\frac{1}{3}3^3) = 72 - 9 = 63. \qquad \triangleleft$$

EXERCISE 6.7 Determine $\int_0^3 x^3\, dx$.

The main conclusion of the Principal Theorem of Integral Calculus is that the integral of a function, of which we know an antiderivative, can be computed 'directly' (see section 6.3 for a list of functions and their antiderivatives).
There are also integrals of functions which can not be computed directly, since an antiderivative can not be found easily. For these integrals we may use a detour: via arithmetic rules

for integrals we may rewrite the integral of a function as an integral of another function of which we do know an antiderivative (see section 6.4 for the arithmetic rules).

Finally we mention the fact that there are functions, which do not have an antiderivative in the form of 'elementary' functions. An example is the (continuous) function f with $f(x) = e^{-x^2}$. From Theorem 6.2 we know that f has an antiderivative, but an antiderivative of f in terms of elementary functions does not exist. An antiderivative of this function plays an important role in probability theory and is called the *error function* $\mathrm{erf}(x)$. (The problem, that a function can not be expressed in terms of 'elementary' functions, also occurred when we tried to determine the inverse of an invertible function: although we know that a function is invertible, it is not possible to determine the inverse in terms of 'elementary' functions (see also the comment in section 3.4).)

In case a function has no antiderivative in terms of 'elementary' functions, we have to compute (by approximation) the integral of the function by so-called 'numerical methods'.

In the end we mention a number of properties of the integral, like the the linearity property, which is a consequence of the corresponding property for differentiation.

THEOREM 6.4 *Let f and g be continuous functions on an interval $[a, b]$ and $\alpha, \beta \in \mathbb{R}$. We have:*

$$\int_a^b (\alpha f + \beta g)(x)dx = \alpha \int_a^b f(x)dx + \beta \int_a^b g(x)dx.$$

PROOF Let F be an antiderivative of f and G an antiderivative of g. Since $(\alpha F + \beta G)' = \alpha F' + \beta G' = \alpha f + \beta g$, $\alpha F + \beta G$ is an antiderivative of $\alpha f + \beta g$. Hence

$$\begin{aligned}
\int_a^b (\alpha f + \beta g)(x)\,dx &= [(\alpha F + \beta G)(x)]_{x=a}^{x=b} \\
&= \alpha F(b) + \beta G(b) - (\alpha F(a) + \beta G(a)) \\
&= \alpha\,(F(b) - F(a)) + \beta\,(G(b) - G(a)) \\
&= \alpha\,[F(x)]_{x=a}^{x=b} + \beta\,[G(x)]_{x=a}^{x=b} \\
&= \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx.
\end{aligned}$$

◁

6.3 A list with antiderivatives

In this section we provide a list of some well-known functions and their antiderivatives. By differentiating the function F in the second column of the list below the function f in the first column is obtained.

$f(x)$	$F(x)$
$x^m (m \in \mathbb{N})$	$\frac{1}{m+1} x^{m+1}$
$x^q (q \in \mathbb{Q}, q \neq -1)$	$\frac{1}{q+1} x^{q+1}$
$\frac{1}{x}$	$\ln \lvert x \rvert$
e^x	e^x
$\cos x$	$\sin x$
$\sin x$	$-\cos x$
$\frac{1}{x^2+1}$	$\arctan x$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x$
$\frac{-1}{\sqrt{1-x^2}}$	$\arccos x$

In the table above F is an antiderivative of f on every interval that is contained in the domain of f. (For a more extensive list of functions and their antiderivatives we refer to the existing literature, such as Gradshteyn, I.S., Ryzhik, I.M., *Table of Integrals, Series and Products*, New York: Academic Press, 1990.)

EXERCISE 6.8 Determine an antiderivative of the function f, defined by

a) $f(x) = \sqrt[3]{x^2}, \quad (x \in \mathbb{R})$

b) $f(x) = \frac{1}{\sqrt[5]{x^3}}, \quad (x > 0)$

c) $f(x) = \frac{1}{x}, \quad (x > 0)$.

EXERCISE 6.9 Compute

a) $\int_0^5 \left(x^5 + \sqrt{x}\right)\ dx$

b) $\int_1^4 (2t\sqrt{t} - \sqrt[3]{t^2})\ dt$.

6.4 Arithmetic rules for integration

In this section we provide two arithmetic rules for integration, which are based upon arithmetic rules for differentiable functions. We start with integration by parts, a consequence of the product rule for differentiation.

THEOREM 6.5 **(Integration by parts)**
Let f and g be continuously differentiable functions on an interval $[a, b]$. We have:

$$\int_a^b f'(x)g(x)dx = [f(x)g(x)]_{x=a}^{x=b} - \int_a^b f(x)g'(x)dx.$$

PROOF Since $(f \cdot g)' = f' \cdot g + f \cdot g'$, $f \cdot g$ is an antiderivative of $f' \cdot g + f \cdot g'$. We have

$$\begin{aligned}
\int_a^b f'(x)g(x)\ dx + \int_a^b f(x)g'(x)\ dx &= \int_a^b (f' \cdot g + f \cdot g')(x)\ dx \\
&= [(f \cdot g)(x)]_{x=a}^{x=b} \\
&= [(f(x)g(x)]_{x=a}^{x=b},
\end{aligned}$$

or,

$$\int_a^b f'(x)g(x)dx = [f(x)g(x)]_{x=a}^{x=b} - \int_a^b f(x)g'(x)dx. \qquad \triangleleft$$

We can use integration by parts successfully for the computation of $\int_a^b h(x)\ dx$ if

a) h can be written as $h = f' \cdot g$ and

b) the integral of $f \cdot g'$ can be computed more easily (for example directly) than the integral of $f' \cdot g$.

So, with integration by parts you move the problem of finding an antiderivative of $f' \cdot g$ to the problem of finding an antiderivative of $f \cdot g'$. We give some examples how integration by parts can be used.

EXAMPLE 6.3

a) The antiderivative of $f \cdot g'$ can be determined directly, for example by checking the list with antiderivatives.

(i)

$$\begin{aligned}\int_a^b x\cos x\ dx &= [x\sin x]_{x=a}^{x=b} - \int_a^b \sin x\ dx \\ &= [x\sin x]_{x=a}^{x=b} - [-\cos x]_{x=a}^{x=b} \\ &= b\sin b + \cos b - a\sin a - \cos a.\end{aligned}$$

Here we have made the choice that $f(x) = \sin x$ (with $f'(x) = \cos x$) and $g(x) = x$.

(ii)

$$\begin{aligned}\int_1^e \ln x\ dx &= [x\ln x]_{x=1}^{x=e} - \int_1^e x\frac{1}{x}\ dx \\ &= e - (e-1) = 1.\end{aligned}$$

Here we have made the choice that $f(x) = x$ (with $f'(x) = 1$) and $g(x) = \ln x$.

b) We use repeated integration by parts in order to end up with a function of which an antiderivative can be found directly.

$$\begin{aligned}\int_0^1 x^2e^x\ dx &= \left[x^2e^x\right]_{x=0}^{x=1} - \int_0^1 2xe^x\ dx \\ &= e - 2([xe^x]_{x=0}^{x=1} - \int_0^1 e^x\ dx) \\ &= e - 2e + 2(e-1) = e - 2.\end{aligned}$$

c) Integration by parts leads to a recurrent relation for the integral.
For $n = 0, 1, 2, \ldots$ the number I_n is defined as $I_n = \int_0^1 x^n e^{-x}\ dx$. Let $n \geq 1$. We have subsequently:

$$\begin{aligned} I_n &= \int_0^1 x^n e^{-x}\, dx \\ &= \left[-x^n e^{-x}\right]_{x=0}^{x=1} + n\int_0^1 x^{n-1}e^{-x}\, dx \\ &= -e^{-1} + nI_{n-1}. \end{aligned}$$

We have derived a *recurrent relation* for I_n. $\triangleleft$

EXERCISE 6.10 Determine $I_0 = \int_0^1 x^0 e^{-x}\, dx$ and compute subsequently $I_1,\ I_2$ and I_3 if

$$I_n = -e^{-1} + nI_{n-1}, \quad (n \in \mathbb{N}).$$

EXERCISE 6.11 Compute each of the integrals below via integration by parts:

a) $\int_1^{e^2} x \ln x\, dx$

b) $\int_0^{\sqrt{3}} x \arctan x\, dx$

c) $\int_0^1 \frac{x^3}{\sqrt{4+x^2}}\, dx$

d) $\int_0^{\pi} e^t \sin t\, dt$.

EXERCISE 6.12 For $m = 0, 1, 2, \ldots$ the number I_m is defined by $I_m = \int_0^{\pi} \sin^m x\, dx$.

a) Show by integration by parts that

$$I_m = \frac{m-1}{m} I_{m-2} \text{ for every natural number } m \geq 2.$$

b) Determine I_4 and I_5.

The chain rule for differentiation leads to integration by substitution.

THEOREM 6.6 **(Integration by substitution)**

Let g be a continuously differentiable function on the interval $[a,b]$ and f a continuous function on the interval $g([a,b])$. We have:

$$\int_a^b f(g(x))g'(x)\, dx = \int_{g(a)}^{g(b)} f(y)\, dy.$$

PROOF Let F be an antiderivative of f. Since $(F \circ g)'(x) = F'(g(x))g'(x) = f(g(x))g'(x) = ((f \circ g) \cdot g')(x)$, $F \circ g$ is an antiderivative of $(f \circ g) \cdot g'$. We have subsequently:

$$\begin{aligned}\int_a^b f(g(x))g'(x)\,dx &= [(F \circ g)(x)]_{x=a}^{x=b} = [(F(g(x))]_{x=a}^{x=b} = [F(y)]_{y=g(a)}^{y=g(b)} \\ &= \int_{g(a)}^{g(b)} f(y)\,dy.\end{aligned}$$

◁

We say that the integral $\int_a^b f(g(x))g'(x)\,dx$ is transformed by the substitution $y = g(x)$ into the integral $\int_{g(a)}^{g(b)} f(y)\,dy$. One could say that by the substitution $y = g(x)$ the factor $f(g(x))$ is transformed into $f(y)$, the expression $g'(x)\,dx$ into dy and the boundaries a and b into $g(a)$ and $g(b)$. Why the expression $g'(x)\,dx$ has been transformed into dy can be explained intuitively, in case g is a strictly increasing in the following way.

If $P = (x_0, \ldots, x_n)$ is a partition of $[a, b]$ and $g(x_i) = y_i$ for $i = 0, \ldots, n$, then $P' = (y_0, \ldots, y_n)$ is a partition of $[g(a), g(b)]$. For the factor $y_i - y_{i-1}$ in the under sum of f corresponding to the partition P' we have, according to the Mean Value Theorem applied to g and the interval $[x_{i-1}, x_i]$,

$$y_i - y_{i-1} = g(x_i) - g(x_{i-1}) = g'(\xi_i)(x_i - x_{i-1}),\ \text{where } \xi_i \in (x_i, x_{i-1}).$$

In other words, the length of the subinterval $[y_{i-1}, y_i]$ is $g'(\xi_i)$ times the length of the corresponding subinterval $[x_{i-1}, x_i]$. This is in line with the observation that, if $g'(\xi_i) > 1$, the interval $[x_{i-1}, x_i]$ is 'stretched' to the interval $[y_{i-1}, y_i]$ and, if $g'(\xi) < 1$, the interval has 'shrunk' (see Figure 6.2). The factor $g'(\xi)$ has to be regarded as a correction, that moreover depends upon x.

If g is a strictly decreasing function, then we are, after the substitution $y = g(x)$, confronted with an integral of which the lower boundary $g(a)$ is larger than the upper boundary $g(b)$. In general, we define the integral of f from a to b if $a > b$ as

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

EXERCISE 6.13 Let f be a continuous function on $[a, b]$ and let the function H on $[a, b]$ be defined by $H(x) = \int_x^b f(t)\,dt$. Show that:

$$H'(x) = -f(x),\ \text{for every } x \in [a, b].$$

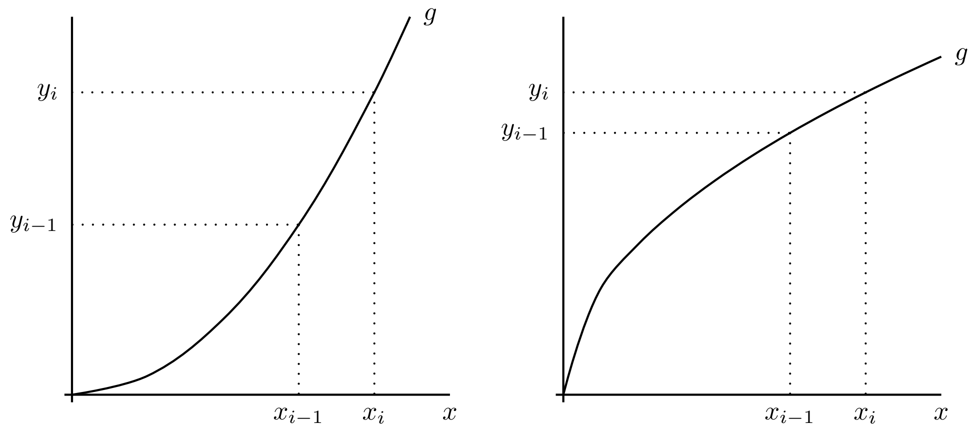

Figure 6.2: The image $[y_{i-1}, y_i] = [g(x_{i-1}), g(x_i)]$ of the subinterval $[x_{i-1}, x_i]$ if $g'(x) > 1$ and if $g'(x) < 1$.

Integration by substitution can successfully be used for the computation of $\int_a^b h(x)\, dx$ if

a) $h(x)$ can be written as $h(x) = f(g(x))g'(x)$ and

b) the integral of f can be computed.

EXAMPLE 6.4 Via integration by substitution we compute the integral

$$\int_0^1 \frac{2x}{1+x^2}\, dx.$$

Define the function g on $[0,1]$ by $g(x) = 1 + x^2$ and the function f on $g([0,1])$ by $f(y) = \frac{1}{y}$. Then $f(g(x)) = \frac{1}{1+x^2}$ and, since $g'(x) = 2x$, is $f(g(x))g'(x) = \frac{2x}{1+x^2}$ (in other words, consider the numerator as the derivative of the denominator). Subsequently we have:

$$\begin{aligned} \int_0^1 \frac{2x}{1+x^2} dx &= \int_0^1 \frac{1}{1+x^2}\, 2x\, dx \\ &=^{*)} \int_1^2 \frac{1}{y}\, dy \\ &= [\ln|y|]_{y=1}^{y=2} = \ln 2. \end{aligned}$$

$(^{*)}$ by the substitution $y = 1 + x^2$ the lower boundary becomes $g(0) = 1$ and the upper boundary $g(1) = 2$.) ◁

Integration by substitution can also be written as

$$\int_a^b f(g(x))g'(x)\ dx = \int_a^b f(g(x))\ d\left(g(x)\right) = \int_{g(a)}^{g(b)} f(y)\ dy.$$

Using this notation the computation of the integral in Example 6.4 looks as follows:

$$\int_0^1 \frac{2x}{1+x^2}\ dx = \int_0^1 \frac{1}{1+x^2}\ d(1+x^2) = \int_1^2 \frac{1}{y}\ dy = [\ln|y|]_{y=1}^{y=2} = \ln 2.$$

EXAMPLE 6.5 We compute two other integrals via integration by substitution.

a) Let $a \neq 0$. Then

$$\begin{aligned}
\int_0^1 \frac{dx}{x^2+a^2} &= \frac{1}{a}\int_0^1 \frac{1}{(x/a)^2+1}\ d(x/a) \\
&= \frac{1}{a}\int_0^{\frac{1}{a}} \frac{1}{y^2+1}\ dy \\
&= \frac{1}{a}\left[\arctan y\right]_{y=0}^{y=\frac{1}{a}} = \frac{1}{a}\arctan\frac{1}{a}.
\end{aligned}$$

b)

$$\begin{aligned}
\int_0^4 \tfrac{x}{1+\sqrt{x}}\ dx &=^{*)} 2\int_0^4 \tfrac{x\sqrt{x}}{1+\sqrt{x}}\tfrac{1}{2\sqrt{x}}\ dx \\
&= 2\int_0^4 \tfrac{x\sqrt{x}}{1+\sqrt{x}}\ d(1+\sqrt{x}) \\
&=^{**)} 2\int_1^3 \tfrac{(y-1)^3}{y}\ dy \\
&= 2\int_1^3 \left(y^2-3y+3-\tfrac{1}{y}\right)\ dy \\
&= 2\left[\tfrac{1}{3}y^3-\tfrac{3}{2}y^2+3y-\ln y\right]_{y=1}^{y=3} = 5\tfrac{1}{3} - 2\ln 3.
\end{aligned}$$

(*) we carry out the substitution $y = 1+\sqrt{x}$, so that $g(x) = 1+\sqrt{x}$ and $g'(x) = \frac{1}{2\sqrt{x}}$;

**) if $y = 1+\sqrt{x}$, then $\sqrt{x} = y-1$ and $x = (y-1)^2$.) ◁

EXERCISE 6.14 Compute each of the integrals below by integration by substitution:

a) $\int_0^1 \cos(3x)\ dx$

b) $\int_0^1 \frac{e^x}{1+e^x}\ dx$

c) $\int_1^e \frac{\ln^2 x}{x}\ dx$

d) $\int_{-\frac{\pi}{4}}^0 \tan x\ dx$.

EXERCISE 6.15

a) Let $a > 0$ and let f be a continuous function on $[-a, a]$. Prove the following property.

If $f(-x) = -f(x)$ for every $x \in [-a, a]$, then $\int_{-a}^a f(x)\ dx = 0$.

b) Compute each of the following integrals:

(i) $\int_{-1}^1 x^3 \cos^2 x\ dx$ (ii) $\int_1^3 \frac{(x-2)^5}{\sqrt{2-(x-2)^2}}\ dx$.

EXERCISE 6.16 Compute each of the following integrals:

a) $\int_0^2 \frac{xe^x}{(1+x)^2}\ dx$

b) $\int_3^4 \sqrt{x}e^{\sqrt{x}}\ dx$

c) $\int_0^1 \frac{1}{8+\sqrt{4+x}}\ dx$

d) $\int_0^{\ln 2} \frac{e^{2x}-e^x}{e^x+2}\ dx$.

6.5 Indefinite integral

The concept of integral has been defined in section 6.1 for a continuous function on an interval of the form $[a, b]$, so on a closed and bounded interval. Hence this definition of section 6.1 can not be applied to a function like $f(x) = \frac{1}{x^2}$, defined on $[1, \infty)$ or a function like $f(x) = \frac{1}{\sqrt{x}}$ on $(0, 1]$. Nevertheless both functions admit a region under their graph that possibly has a finite area, although the region is unbounded. In order to deal with these situations we extend the concept of integral to so-called indefinite integrals.

DEFINITION Let f be a continuous function on an interval $[a, \infty)$.
If $\lim_{b\to\infty} \int_a^b f(x)\ dx$ exists, then the limit is called the *indefinite integral of the first kind* of f from a to ∞ and denoted by $\int_a^\infty f(t)\ dt$, so

$$\int_a^\infty f(t)\ dt = \lim_{b\to\infty} \int_a^b f(t)\ dt.$$

In an analogous way we can define the indefinite integral of a continuous function on an interval of the form $(-\infty, b]$.

EXAMPLE 6.6 We check whether the indefinite integral (of the first kind) $\int_1^\infty \frac{1}{x^2}\ dx$ exists. For $b > 1$ we have

$$\int_1^b \frac{1}{x^2}\ dx = \left[-\frac{1}{x}\right]_{x=1}^{x=b} = 1 - \frac{1}{b}.$$

Since $\lim_{b\to\infty} \int_1^b \frac{1}{x^2}\ dx = \lim_{b\to\infty} \left(1 - \frac{1}{b}\right) = 1$, the indefinite integral $\int_1^\infty \frac{1}{x^2}\ dx$ exists and

$$\int_1^\infty \frac{1}{x^2}\ dx = 1. \qquad \triangleleft$$

EXERCISE 6.17 Check which of the following indefinite integrals exist and determine their value.

a) $\int_0^\infty \frac{1}{(x+1)^2}\ dx$

b) $\int_0^\infty \sin t\ dt$

c) $\int_{-\infty}^0 e^x\ dx$

d) $\int_2^\infty \frac{1}{x(x-1)}\ dx$

DEFINITION Let f be a continuous function on the interval $(a, b]$.
If $\lim_{a'\to a} \int_{a'}^b f(x)\ dx$ exists, then the limit is called the *indefinite integral of the second kind* of f from a to b and denoted by $\int_a^b f(x)\ dx$, so

$$\int_a^b f(x)\ dx = \lim_{a'\to a} \int_{a'}^b f(x)\ dx.$$

In an analogous way we can define the indefinite integral of a continuous function on an interval of the form $[a, b)$,

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

EXAMPLE 6.7 We show that the indefinite integral (of the second kind) $\int_0^1 \frac{1}{\sqrt{x}}\ dx$ exists. For $0 < a' < 1$ we have

$$\int_{a'}^1 \frac{1}{\sqrt{x}}\ dx = \left[2\sqrt{x}\right]_{x=a'}^{x=1} = 2 - 2\sqrt{a'}.$$

Since $\lim_{a' \to 0} \int_{a'}^1 \frac{1}{\sqrt{x}}\ dx = \lim_{a' \to 0}(2 - 2\sqrt{a'}) = 2$, the indefinite integral $\int_0^1 \frac{1}{\sqrt{x}}\ dx$ exists and

$$\int_0^1 \frac{1}{\sqrt{x}}\ dx = 2. \qquad \triangleleft$$

EXERCISE 6.18 Check which of the following indefinite integrals exist and determine their value.

a) $\int_0^1 \frac{1}{x}\ dx$

b) $\int_0^1 \frac{1}{\sqrt{1-x}}\ dx$

c) $\int_0^1 \ln x\ dx$, here you are allowed to use the standard limit $\lim\limits_{a \downarrow 0}(a \ln a) = 0$

d) $\int_0^1 (1-x)^{-\frac{3}{2}}\ dx$

It also may happen that an integral is indefinite to both sides, that is that f is a continuous function, defined on (a, b) or (a, ∞) or $(-\infty, b)$ or even $(-\infty, \infty)$. In that case we choose an arbitrary point c in such an interval and consider the indefinite integrals $\int_a^c f(x)\ dx$ and $\int_c^b f(x)\ dx$ separately (here we possibly have that '$a = -\infty$' and \ or '$b = \infty$'). Only if both indefinite integrals exists, we state that the indefinite integral $\int_a^b f(x)\ dx$ exists and we have

$$\int_a^b f(x)\ dx = \int_a^c f(x)\ dx + \int_c^b f(x)\ dx.$$

Moreover, it turns out that the outcome does not depend upon the choice of the auxiliary point c.

EXAMPLE 6.8 We show that the indefinite integral $\int_{-\infty}^{\infty} \frac{1}{1+x^2}\ dx$ exists. Let $c \in \mathbb{R}$. First we determine the indefinite integral $\int_{-\infty}^{c} \frac{1}{1+x^2}\ dx$. Let $a < c$. We have

$$\int_a^c \frac{1}{1+x^2}\ dx = [\arctan x]_{x=a}^{x=c} = \arctan c - \arctan a.$$

Since $\lim_{a\to-\infty}\int_a^c \frac{1}{1+x^2}\,dx = \lim_{a\to-\infty}(\arctan c - \arctan a) = \arctan c + \frac{\pi}{2}$, the indefinite integral $\int_{-\infty}^c \frac{1}{1+x^2}\,dx$ exists and

$$\int_{-\infty}^c \frac{1}{1+x^2}\,dx = \arctan c + \frac{\pi}{2}.$$

Subsequently we determine the indefinite integral $\int_c^\infty \frac{1}{1+x^2}\,dx$. Let $b > c$. We have

$$\int_c^b \frac{1}{1+x^2}\,dx = [\arctan x]_{x=c}^{x=b} = \arctan b - \arctan c.$$

Since $\lim_{b\to\infty}\int_c^b \frac{1}{1+x^2}\,dx = \lim_{b\to\infty}(\arctan b - \arctan c) = \frac{\pi}{2} - \arctan c$, the indefinite integral $\int_c^\infty \frac{1}{1+x^2}\,dx$ exists and

$$\int_c^\infty \frac{1}{1+x^2}\,dx = \frac{\pi}{2} - \arctan c.$$

Therefore, the indefinite integral $\int_{-\infty}^\infty \frac{1}{1+x^2}\,dx$ exists and we have

$$\begin{aligned}\int_{-\infty}^\infty \frac{1}{1+x^2}\,dx &= \arctan c + \frac{\pi}{2} + \frac{\pi}{2} - \arctan c\\ &= \pi.\end{aligned}$$

◁

EXERCISE 6.19 Check whether the indefinite integral

$$\int_0^1 \frac{1}{\sqrt{x(1-x)}}\,dx$$

exists and, if yes, determine its value.

When introducing the concept of integral in section 6.1 we explicitly assumed that the function is continuous. For, in order to determine the under sum for a partition we assumed that the function has a minimum on each of the (compact) subintervals of the partition P (Theorem of Weierstrass). So, if we would like to introduce the concept of integral for a function which is not continuous, we have to follow another way. For this we refer to the existing literature on integration theory.

6.6 Mixed exercises

EXERCISE 6.20 The function f on $[0,4]$ is defined by $f(x) = 1-(x-1)^2$.

Determine for the partition $P = (0, \frac{1}{2}, 1, 2, \frac{5}{2}, 3, \frac{7}{2}, 4)$ the corresponding under sum of f.

EXERCISE 6.21 Determine the derivative of the function G, defined by

a) $G(x) = \int_x^2 e^{-t^2}\ dt$

b) $G(x) = \int_0^{x^2} \frac{1}{\sqrt{1+t^2}}\ dt$.

EXERCISE 6.22 Compute (if possible) each of the following integrals:

a) $\int_1^{16} \frac{2}{x^{\frac{3}{4}}(x^{\frac{1}{4}}+3)}\ dx$

b) $\int_1^2 \frac{1}{\sqrt{x}+\sqrt[3]{x}}\ dx$

(Hint: substitute $y = \sqrt[6]{x}$.)

c) $\int_1^\infty \frac{1}{\sqrt{x}}\ dx$

d) $\int_0^\infty xe^{-x^2}\ dx$

e) $\int_0^4 \sqrt{4-\sqrt{x}}\ dx$

f) $\int_0^{\frac{1}{2}\pi} x^2 \sin x\ dx$

g) $\int_0^1 x^2 \arctan x\ dx$

h) $\int_0^{\frac{\pi}{4}} \arctan x\ dx$

i) $\int_0^{\frac{\pi}{4}} \frac{\ln(\cos x)^2}{(\cos x)^2}\ dx$.

(Hint: if $f(x) = \tan x$, then $f'(x) = \frac{1}{(\cos x)^2}$.)

EXERCISE 6.23 For $n = 0, 1, 2, \ldots$ the number I_n is defined by $I_n = \int_0^{\frac{\pi}{2}} x^n \sin x\ dx$.

a) Show that for $n \geq 2$ we have $I_n = n(\frac{\pi}{2})^{n-1} - n(n-1)I_{n-2}$.

b) Compute I_3 by using the recursive formula of item a).