INTRODUCTORY COURSE IN ANALYSIS

Series of Lecture Notes and Workbooks for Teaching Undergraduate Mathematics

Algoritmuselmélet
Algoritmusok bonyolultsága
Analitikus módszerek a pénzügyekben
Bevezetés az analízisbe
Differential Geometry
Diszkrét optimalizálás
Diszkrét matematikai feladatok
Geometria
Igazságos elosztások
Interaktív analízis feladatgyűjtemény matematika BSc hallgatók számára
Introductory Course in Analysis
Matematikai pénzügy
Mathematical Analysis-Exercises 1-2
Mértékelmélet és dinamikus programozás
Numerikus funkcionálanalízis
Operációkutatás
Operációkutatási példatár
Optimális irányítások
Parciális differenciálegyenletek
Példatár az analízishez
Szimmetrikus kombinatorikai struktúrák
Többváltozós adatelemzés

István Faragó

Ágnes Havasi

István Mezei

Péter Simon

INTRODUCTORY COURSE IN ANALYSIS

Eötvös Loránd University
Faculty of Science

Typotex

2013

Reader: Dr. Bálint Nagy

ISBN 978 963 279 223 1
Prepared under the editorship of Typotex Publishing House (`http://www.typotex.hu`)
Responsible manager: Zsuzsa Votisky
Technical editor: József Gerner

Made within the framework of the project Nr. TÁMOP-4.1.2-08/2/A/KMR-2009-0045, entitled "Jegyzetek és példatárak a matematika egyetemi oktatásához" (Lecture Notes and Workbooks for Teaching Undergraduate Mathematics).

Nemzeti Fejlesztési Ügynökség
www.ujszechenyiterv.gov.hu
06 40 638 638

MAGYARORSZÁG MEGÚJUL

A projekt az Európai Unió támogatásával, az Európai Regionális Fejlesztési Alap társfinanszírozásával valósul meg.

KEY WORDS: Mathematical analysis, sets, real and complex numbers, sequences, series, differentiation, integration, analysis of functions, series of functions, vector analysis.

SUMMARY: This textbook has been written for the analysis education of non-mathematics students at the Eötvös Loránd University, Faculty of Science, but it can also be used as a supplementary material by students of mathematics. All subjects are presented at beginners' level, where mainly methods are taught. The book is strongly application-oriented. For example, vector calculus is included for students of geophysics, and contour and surface integrals are presented for physics student.

Contents

Chapter 1

Preface

These lecture notes are based on the series of lectures that were given by the authors at the Eötvös Loránd University for students in Physics, Geophysics, Meteorology and Geology. It is written firstly for these students, however, it can be also used by students in mathematics. People at the Department of Applied Analysis and Computational Mathematics have taught mathematics to science students for decades. The authors have taken part in this work for several years, they taught the topics dealt with in these lecture notes in many semesters. Their long term teaching and pedagogical experience is behind this work.

Concerning its contents the book is similar to other analysis textbooks, however, it is special because of several reasons. First of all, most of the textbooks are written for students in mathematics, or for non-mathematics students in some special field, for example students in engineering or economy. These lecture notes are customized to science students at the Eötvös Loránd University. According to our teaching experience the students do not acquire mathematical knowledge through the axiomatic set-up, instead they understand mathematical notions and methods gradually getting deeper and deeper synthesis. Hence the lecture notes follow an alternative way, all subjects are presented at the beginners level, when mainly methods are taught (for physics students this corresponds to the Calculus course). The book is strongly application oriented. For example, vector calculus is included for students in geophysics, complex functions, contour and surface integral is presented for physics student.

Acknowledgement

The authors are grateful to their colleagues at the Department of Applied Analysis and Computational Mathematics in the Institute of Mathematics at the Eötvös Loránd University, who supported the creation of the analysis course for science students at the University, and encouraged the efforts to make lecture notes electronically available to the students of these courses.

The authors appreciate the careful work of the referee, Bálint Nagy, who provided detailed comments, identified errors and typos, and helped to increase distinctness and consistency of the notes.

Chapter 2

Sets, relations, functions

We present the tools and frequently used notions of mathematics, and introduce some important conventions. We prepare a solid base for the further constructions. To abbreviate the words "every" or "arbitrary" we will often use the symbol $\forall$, while the notation $\exists$ will be employed for the expression "exists" or "there is". This chapter covers the following topics.

- Sets and operations on sets
- Relations
- Functions and their properties
- Composition and inverse functions

2.1 Sets, relations, functions

2.1.1 Sets and relations

A **set** is considered as given if we can decide about every well-defined object whether it belongs to the set or not. (A "clever thought", a "beautiful girl", a "sufficiently big number" or a "small positive number" cannot be considered as well-defined objects, so we will not ask if they belong to a set.)

Let A be a set, and x a well-defined object. If x belongs to the set, then we will denote this as $x \in A$. If x does not belong to the set, we will write $x \notin A$.

A set can be given by listing its elements, e.g., $A := \{a, b, c, d\}$, or by specifying a property, e.g., $B := \{x \mid x \text{ is a real number and } x^2 < 2\}$.

Definition 2.1. Let A and B be sets. We say that A is a subset of B if for all $x \in A$ $x \in B$ holds. Notation: $A \subset B$.

Definition 2.2. Let A and B be sets. Set A is equal to set B if both have the same elements. Notation: $A = B$.

It is easy to see that the following theorem holds.

Theorem 2.1. *Let A and B be sets. Then $A = B$ if and only if $A \subset B$ and $B \subset A$.*

We will show some procedures which yield further sets.

Definition 2.3. Let A and B be sets.
The union of A and B is the set $A \cup B := \{x \mid x \in A \text{ or } x \in B\}$.
The intersection of A and B is the set $A \cap B := \{x \mid x \in A \text{ and } x \in B\}$.
The difference of A and B is the set $A \setminus B := \{x \mid x \in A \text{ and } x \notin B\}$.

When taking the intersection or the difference of sets, it can happen that no object x possesses the required property. The set to which no well-defined object belongs is called **empty set**. Notation: $\emptyset$.

Let H be a set and $A \subset H$ a subset. The complement of A (with respect to set H) is defined as the set $\overline{A} := H \setminus A$. The following theorem is known as De Morgan's identities.

Theorem 2.2. *Let H be a set, $A, B \subset H$. Then we have*

$$\overline{A \cup B} = \overline{A} \cap \overline{B} \quad \text{and} \quad \overline{A \cap B} = \overline{A} \cup \overline{B}.$$

Let a and b be any objects. The set $\{a, b\}$ can obviously be written in several ways:

$$\{a, b\} = \{b, a\} = \{a, b, b, a\} = \{a, b, b, a, b, b\} = \text{etc.}$$

As opposed to this, we introduce the basic notion of the **ordered pair** (a, b), an essential property of which is

$$(a, b) = (c, d) \text{ if and only if } a = c \text{ and } b = d.$$

We define the product of sets with the aid of ordered pairs.

Definition 2.4. Let A, B be sets. The **Cartesian product** of A and B is the set of ordered pairs

$$A \times B := \{(a, b) \mid a \in A \text{ and } b \in B\}.$$

For example, let $A := \{2, 3, 5\}$ and $B := \{1, 3\}$, then

$$A \times B = \{(2, 1), (2, 3), (3, 1), (3, 3), (5, 1), (5, 3)\}.$$

Relations are based on the notion of ordered pair.

Definition 2.5. We say that a set r is a **relation** if each of its elements is an ordered pair.

A Hungarian-English dictionary is a relation because its elements are ordered pairs of a Hungarian word and the corresponding English word.

Definition 2.6. Let r be a relation. The **domain of definition** of r is

$$D(r) := \{x \mid \text{there exists an } y \text{ such that } (x, y) \in r\}.$$

The **range** of r is

$$R(r) := \{y \mid \text{there exists an } x \in D(r) \text{ such that } (x, y) \in r\}.$$

Obviously, $r \subset D(r) \times R(r)$.

For example, in the case of $r := \{(4, 2), (4, 3), (1, 2)\}$, $D(r) = \{4, 1\}$, $R(r) = \{2, 3\}$.

2.1.2 Functions

A function is a special relation.

Definition 2.7. Let f be a relation. We say that f is a **function** if for all $(x, y) \in f$ and $(x, z) \in f$ $y = z$.

For example, $r := \{(1, 2), (2, 3), (2, 4)\}$ is not a function since $(2, 3) \in r$ and $(2, 4) \in r$, but $3 \neq 4$; however, $f := \{(1, 2), (2, 3), (3, 3)\}$ is a function.

We introduce some conventions in connection with functions. If f is a function, then in case of $(x, y) \in f$ we call y the **value** of function f at x, and we say that f associates y to x or maps x to y. Notation: $y = f(x)$.

If f is a function, $A := D(f)$, and B is such a set that $R(f) \subset B$ (clearly, A is the domain of definition of the function, and B is (a) **range** of the function, then instead of the expression "$f \subset A \times B, f$ is a function" the notation $f : A \to B$ is employed ("the function f maps set A to set B").

If f is a function and $D(f) \subset A, R(f) \subset B$, then this is denoted by $f : A \rightarrowtail B$ ("f is a function that maps **from** set A to set B").

For example $f := \{(a, \alpha), (b, \beta), (g, \gamma), (d, \delta), (e, \varepsilon)\}$ is a function. One can see that β is the value of f at b: $\beta = f(b)$.

If L denotes the set of Latin letters and G the set of Greek letters, then $f : \{a, b, g, d, e\} \to G, f(a) = \alpha, f(b) = \beta, f(g) = \gamma, f(d) = \delta, f(e) = \varepsilon$. If we only want to refer to the type of the function, then it is sufficient to write $f \in L \rightarrowtail G$.

Obviously, any function has an inverse, however, it can happen that the inverse is not a function.

Definition 2.8. Let $f : A \to B$ be a function. We say that f is **one-to one** (injective) if it associates different elements of B to different elements $x_1, x_2 \in A$, that is, for all $x_1, x_2 \in A, x_1 \neq x_2$: $f(x_1) \neq f(x_2)$.

It is easy to see that the inverse of a one-to-one function is a function. In more detail:

Theorem 2.3. *Let f be a function, $A := D(f), B := R(f), f$ one-to-one. Then the inverse $f^{-1} : B \to A$ of f is such a function that maps any point $s \in B$ to $t \in A$ for which $f(t) = s$, (briefly: for any $s \in B$: $f(f^{-1}(s)) = s$).*

We can also prepare the composition of functions. Fortunately, it will always be a function.

Let $g : A \to B, f : B \to C$. Then by using the composition of relations one can show that

$$f \circ g : A \to C, \text{ and for all } x \in A : (f \circ g)(x) = f(g(x)).$$

For example, let the function g add 1 to the double of each number ($g\colon \mathbb{R} \to \mathbb{R}$, $g(x) := 2x + 1$); and the function f raise each number to the second power ($f : \mathbb{R} \to \mathbb{R}$, $f(x) := x^2$), then $f \circ g : \mathbb{R} \to \mathbb{R}$, $(f \circ g)(x) = (2x + 1)^2$ will be the composition of f and g.

Further useful notions

Let $f : A \to B$ and $C \subset A$. The **restriction** of a function f to C is the function $f_{|C} : C \to B$ for which $f_{|C}(x) := f(x)$ for all $x \in C$.

Let $f : A \to B,\ C \subset A$ and $D \subset B$. The set

$$f(C) := \{y \mid \text{ there exists } x \in C, \text{ such that } f(x) = y\}$$

is called the "image of set C under the function f". The set

$$f^{-1}(D) := \{x \mid f(x) \in D\}$$

is called the "preimage of set D under the function f". (Attention! The notation f^{-1} does not stand for the inverse function in this case.)

2.2 Exercises

1. Let $A := \{2, 4, 6, 3, 5, 9\}$, $B := \{4, 5, 6, 7\}$, $H := \{n \mid n \text{ is a whole number}, 1 \leq n \leq 20\}$. Prepare the sets $A \cup B, A \cap B, A \setminus B, B \setminus A$. What is the complement of A with respect to H?

2. Let $A := \{a, b\}$, $B := \{a, b, c\}$. $A \times B =?$ $B \times A =?$

3. Let $r := \{(x, y) \mid x, y \text{ real numbers}, y = x^2\}$. $r^{-1} =$? Is r a function? Is r^{-1} a function?

4. Let $f : \mathbb{R} \to \mathbb{R}, f(x) := \frac{x}{1+x^2}$. Prepare the functions $f \circ f$, $f \circ (f \circ f)$.

5. Think over how the inverse of a one-to-one function $f : A \to B$ can be illustrated.

6. Consider that the inverse of a function $f : A \to B$ can be obtained in the following steps:

   1) Write that $y = f(x)$.

   2) Swap the “variables” x and y: $x = f(y)$.

   3) From this equation express y with the aid of x: $y = g(x)$. This very g will be the inverse function f^{-1}.

   Example: $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 2x - 1$. (This is a one-to-one function.)

   1) $y = 2x - 1$

   2) $x = 2y - 1$

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

   So $f^{-1} : \mathbb{R} \to \mathbb{R}$, $f^{-1}(x) = \frac{1}{2}(x + 1)$.

   Draw the graphs of the functions f and f^{-1}.

7. Let $f : A \to B$, $C_1, C_2 \subset A, D_1, D_2 \subset B$. Show that

$$f(C_1 \cup C_2) = f(C_1) \cup f(C_2),$$
$$f(C_1 \cap C_2) \subset f(C_1) \cap f(C_2),$$
$$f^{-1}(D_1 \cup D_2) = f^{-1}(D_1) \cup f^{-1}(D_2),$$
$$f^{-1}(D_1 \cap D_2) = f^{-1}(D_1) \cap f^{-1}(D_2).$$

   Is it true that $C_1 \subset C_2$ implies $f(C_1) \subset f(C_2)$?

   Is it true that $D_1 \subset D_2$ implies $f^{-1}(D_1) \subset f^{-1}(D_2)$?

8. Let $f : A \to B$, $C \subset A, D \subset B$.

   Is it true that $f^{-1}(f(C)) = C$? Is it true that $f(f^{-1}(D)) = D$?

Chapter 3

Sets of numbers

We can calculate with real numbers since our childhood, we add, multiply and divide them, raise them to powers and take their absolute values. We re-arrange equations and inequalities. Now we lay down the relatively simple set of rules from which the learnt procedures can be derived. We will cover the following topics.

- The set of real numbers
- The set of natural numbers
- The sets of integers and rational numbers
- Upper bound, lower bound
- Interval and neighborhood
- Exponentiation and power law identities
- The set of complex numbers
- The trigonometric form of complex numbers, operations

3.1 Real numbers

3.1.1 The axiomatic system of real numbers

Let $\mathbb{R}$ be a nonempty set. Suppose there is a function $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ called addition and a function $\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ called multiplication satisfying the following properties:

a1. for all $a, b \in \mathbb{R}$, $a + b = b + a$ (commutativity);

a2. for all $a, b, c \in \mathbb{R}$, $a + (b + c) = (a + b) + c$ (associativity);

a3. there exists an element $0 \in \mathbb{R}$ such that for all $a \in \mathbb{R}$, $a + 0 = a$ (0 is a neutral element with respect to addition);

a4. for all $a \in \mathbb{R}$ there is an element $-a \in \mathbb{R}$ such that $a + (-a) = 0$;

m1. for all $a, b \in \mathbb{R}$, $a \cdot b = b \cdot a$;

m2. for all $a, b \in \mathbb{R}$, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$;

m3. there exists an element $1 \in \mathbb{R}$ such that for all $a \in \mathbb{R}$, $a \cdot 1 = a$ (1 is a neutral element with respect to the multiplication);

m4. for all $a \in \mathbb{R} \setminus \{0\}$ there exists a reciprocal element $\frac{1}{a} \in \mathbb{R}$ for which $a \cdot \frac{1}{a} = 1$;

d. for all $a, b, c \in \mathbb{R}$, $a \cdot (b + c) = ab + ac$ (multiplication is distributive with respect to addition).

It is easy to see that the fourth requirement of multiplication is essentially different from the laws of addition (otherwise the two operations would not differ from each other).

Axiom d also emphasizes the difference.

Assume that there exists an ordering relation $\leq$ (called less than or equal to) on $\mathbb{R}$, which has the following further properties:

r1. for all $a, b \in \mathbb{R}$ either $a \leq b$, or $b \leq a$ holds;

r2. in all cases where $a \leq b$ and $c \in \mathbb{R}$ are arbitrary numbers, $a + c \leq b + c$;

r3. in all cases where $0 \leq a$ and $0 \leq b$, $0 \leq ab$.

Let us fix that instead of $a \leq b$, $a \neq b$ the notation $a < b$ will be employed. (Unfortunately, $<$ is not an ordering relation, since it is not reflexive.)

On the basis of a1–a4, m1–m4, d, r1–r3 one can derive all "laws" related to equalities and inequalities. As a supplement, we mention three notions.

Definition 3.1. Let $a, b \in \mathbb{R}$, $b \neq 0$. Then $\frac{a}{b} := a \cdot \frac{1}{b}$.

So, division can be performed with real numbers.

Definition 3.2. Let $x \in \mathbb{R}$. The **absolute value** of x is

$$|x| := \begin{cases} x & \text{if } 0 \leq x \\ -x & \text{if } x \leq 0,\ x \neq 0. \end{cases}$$

Inequalities with absolute value are very useful.

1. For all $x \in \mathbb{R}$, $0 \leq |x|$.

2. Let $x \in \mathbb{R}$ and $\varepsilon \in \mathbb{R},\ 0 \leq \varepsilon$. Then $x \leq \varepsilon$ and $-x \leq \varepsilon \Longleftrightarrow |x| \leq \varepsilon$.
3. For all $a, b \in \mathbb{R}$, $|a+b| \leq |a| + |b|$ (triangle inequality).
4. For all $a, b \in \mathbb{R}$, $||a| - |b|| \leq |a-b|$.

These statements are simple to prove. Here we show the proof of A 4.

Consider the equality $a = a - b + b$. Then, by property 3,

$$|a| = |a-b+b| \leq |a-b| + |b|.$$

According to r2, by adding the number $-|b|$ to both sides, the inequality does not change.

$$|a| + (-|b|) = |a| - |b| \leq |a-b|. \tag{3.1}$$

Similarly,

$$\begin{aligned} b &= b - a + a, \\ |b| = |b-a+a| &\leq |b-a| + |a| \quad / - |a|, \\ |b| - |a| &\leq |b-a|, \\ -\left(|a| - |b|\right) &\leq |b-a| = |a-b|. \end{aligned} \tag{3.2}$$

The inequalities (3.1) and (3.2) according to property 2 (by the choice $x := |a| - |b|$; $\varepsilon := |a-b|$) exactly yield $||a| - |b|| \leq |a-b|$.

3.1.2 Natural, whole and rational numbers

Now we separate a famous subset of $\mathbb{R}$.

Let $\mathbb{N} \subset \mathbb{R}$ be such a subset for which

1^o $1 \in \mathbb{N}$,

2^o for all $n \in \mathbb{N}$, $n + 1 \in \mathbb{N}$,

3^o for all $n \in \mathbb{N}$, $n + 1 \neq 1$ (1 is the “first” element),

4^o the facts that a) $S \subset \mathbb{N}$,
b) $1 \in S$,
c) for all $n \in S$, $n + 1 \in S$

imply $S = \mathbb{N}$. (Complete induction.)

This subset $\mathbb{N}$ of $\mathbb{R}$ is called **the set of natural numbers**.

We supplement all this with the following definitions:

$\mathbb{Z} := \mathbb{N} \cup \{0\} \cup \{m \in \mathbb{R} \mid -m \in \mathbb{N}\}$ is **the set of integers**,

$\mathbb{Q} := \{x \in \mathbb{R} \mid$ there exists $p \in \mathbb{Z}, q \in \mathbb{N}$ such that $x = \frac{p}{q}\}$ is **the set of rational numbers**,

$\mathbb{Q}^* := \mathbb{R} \setminus \mathbb{Q}$ is **the set of irrational numbers**.

With the aid of $\mathbb{N}$, we impose a third requirement on $\mathbb{R}$ in addition to the laws of the operations and ordering.

Archimedes' axiom: For all $a, b \in \mathbb{R},\ 0 < a$ there exists $n \in \mathbb{N}$ such that $b < na$.

As a consequence of Archimedes' axiom, one can show that for all $K \in \mathbb{R}$ there exists a natural number $n \in \mathbb{N}$ for which $K < n$, since by the choice $a := 1$, $b := K$ the axiom provides such a natural number.

We also show that for all $\varepsilon \in \mathbb{R}, 0 < \varepsilon$ there exists a natural number $n \in \mathbb{N}$ such that $\frac{1}{n} < \varepsilon$, since let us choose $a := \varepsilon$ and $b := 1$. According to the axiom there is an $n \in \mathbb{N}$ such that $1 < n \cdot \varepsilon$. By applying the appropriate "law":

$$\begin{aligned}
&1 < n\varepsilon \qquad / + (-1),\\
&0 < n\varepsilon - 1 \qquad / \cdot \frac{1}{n},\\
&0 < \frac{1}{n}(n\varepsilon - 1) = \varepsilon - \frac{1}{n} \qquad / + \frac{1}{n},\\
&\frac{1}{n} < \varepsilon.
\end{aligned}$$

Even with the introduction of Archimedes' axiom $\mathbb{R}$ does not meet all demands. We need a final axiom, for which we make preparations by introducing some further notions.

3.1.3 Upper and lower bound

Definition 3.3. Let $A \subset \mathbb{R}, A \neq \emptyset$. We say that the set A is **bounded above** if there exists a $K \in \mathbb{R}$ such that for all $a \in A$, $a \leq K$. Such a number K is called an **upper bound** of set A.

Let $A \subset \mathbb{R}, A \neq \emptyset$ be bounded above. Consider

$$B := \{K \in \mathbb{R} \mid K \text{ is an upper bound of set } A\}.$$

Let $\alpha \in \mathbb{R}$ be the smallest element of set B, that is, a number for which

1^o $\alpha \in B$ (α is an upper bound of set A),

2^o for all upper bounds $K \in B$, $\alpha \leq K$.

The only question is whether there exists such an $\alpha \in \mathbb{R}$.

The least upper bound axiom: Every set $A \subset \mathbb{R}, A \neq \emptyset$ of real numbers having an upper bound must have a least upper bound.

Such a number $\alpha \in \mathbb{R}$ (which is not necessarily an element of A) is called **supremum** of A and denoted as

$$\alpha := \sup A.$$

Clearly, the following two properties of $\sup A$ hold:

1^o for all $a \in A$, $a \leq \sup A$,

2^o for all $0 < \varepsilon$ there exists $a' \in A$ such that $(\sup A) - \varepsilon < a'$.

The laws of the operations and ordering, Archimedes' axiom and the least upper bound axiom make the set of real numbers $\mathbb{R}$ complete. In this way we have laid down a solid base for the future calculations, too.

Some further conventions:

Definition 3.4. Let $A \subset \mathbb{R}, A \neq \emptyset$. We say that A is **bounded below** if there exists an $L \in \mathbb{R}$ such that for all $a \in A$, $L \leq a$. The number L is called (a) lower bound of set A.

Let A be a set of numbers that is bounded below. The greatest lower bound of A is called **infimum** of A. (The existence of this lower bound does not require any new axiom, it follows from the least upper bound axiom.) The infimum of A is denoted as

$$\inf A.$$

Obviously,

1^o for all $a \in A$, $\inf A \leq a$,

2^o for all $0 < \varepsilon$ there exists an $a' \in A$ such that $a' < (\inf A) + \varepsilon$.

3.1.4 Intervals and neighborhoods

Definition 3.5. Let $I \subset \mathbb{R}$. We say that I is an **interval** if for all $x_1, x_2 \in I$, $x_1 < x_2$: any $x \in \mathbb{R}$ for which $x_1 < x < x_2$ is in I.

Theorem 3.1. *Let $a, b \in \mathbb{R}, a < b$.*

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

$$[a, b) := \{x \in \mathbb{R} \mid a \leq x < b\},$$

$$(a, b] := \{x \in \mathbb{R} \mid a < x \leq b\},$$

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

$[a, +\infty) := \{x \in \mathbb{R} \mid a \leq x\}$,

$(a, +\infty) := \{x \in \mathbb{R} \mid a < x\}$; $(0, +\infty) =: \mathbb{R}^+$,

$(-\infty, a] := \{x \in \mathbb{R} \mid x \leq a\}$,

$(-\infty, a) := \{x \in \mathbb{R} \mid x < a\}$; $(-\infty, 0) =: \mathbb{R}^-$,

$(-\infty, +\infty) := \mathbb{R}$.

All these are intervals. We mention that $[a, a] = \{a\}$ and $(a, a) = \emptyset$ are degenerate intervals.

Definition 3.6. Let $a \in \mathbb{R}, r \in \mathbb{R}^+$. The neighborhood with radius r of point a is defined as the open interval

$$K_{\mathrm{r}}(a) := (a - r, a + r).$$

We say that $K(a)$ is a **neighborhood of point** a if there exists an $r \in \mathbb{R}^+$ such that $K(a) \subset K_{\mathrm{r}}(a)$.

3.1.5 The powers of real numbers

Definition 3.7. Let $a \in \mathbb{R}$. Then $a^1 := a, a^2 := a \cdot a, a^3 := a^2 \cdot a, \ldots, a^n := a^{n-1} \cdot a, \ldots$

Definition 3.8. Let $a \in \mathbb{R},\ 0 \leq a$. Denote by $\sqrt{a}$ the nonnegative number whose square is a, i.e., $0 \leq \sqrt{a},\ (\sqrt{a})^2 = a$.

Note that for all $a \in \mathbb{R}$, $\sqrt{a^2} = |a|$.

Definition 3.9. Let $a \in \mathbb{R}, k \in \mathbb{N}$. Denote by $\sqrt[2k+1]{a}$ the real number whose $(2k+1)$th power is a.

Note that if $0 < a$, then $\sqrt[2k+1]{a} > 0$, and if $a < 0$, then $\sqrt[2k+1]{a} < 0$.

Definition 3.10. Let $a \in \mathbb{R}, 0 \leq a, k \in \mathbb{N}$. Denote by $\sqrt[2k]{a}$ the nonnegative number whose $(2k)$th power is a.

Let us introduce the following notation: if $n \in \mathbb{N}$ and $a \in \mathbb{R}$ corresponds to the parity of n, then

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

Definition 3.11. Let $a \in \mathbb{R}^+, p, q \in \mathbb{N}$.

$$a^{\frac{p}{q}} := \sqrt[q]{a^p}.$$

Definition 3.12. Let $a \in \mathbb{R}^+, p, q \in \mathbb{N}$.

$$a^{-\frac{p}{q}} := \frac{1}{\sqrt[q]{a^p}}.$$

Definition 3.13. Let $a \in \mathbb{R} \setminus \{0\}$. Then $a^0 := 1$.

By this chain of definitions we have defined the number $a \in \mathbb{R}^+$ raised to any rational power $r \in \mathbb{Q}$. One can show that the numbers in the definitions uniquely exist, and the following identities are valid:

1^o for $a \in \mathbb{R}^+$, $r, s \in \mathbb{Q}$, $a^r \cdot a^s = a^{r+s}$,

2^o for $a \in \mathbb{R}^+$, $r \in \mathbb{Q}$, $a^r \cdot b^r = (ab)^r$,

3^o for $a \in \mathbb{R}^+$, $r, s \in \mathbb{Q}$, $(a^r)^s = a^{rs}$.

3.2 Exercises

1. Let $a, b \in \mathbb{R}$. Show that

$$\begin{aligned}(a+b)^2 : &= (a+b)(a+b) = a^2 + 2ab + b^2,\\ a^2 - b^2 &= (a-b)(a+b),\\ a^3 - b^3 &= (a-b)(a^2+ab+b^2),\\ a^3 + b^3 &= (a+b)(a^2-ab+b^2)\end{aligned}$$

2. Prove that for all $x \in \mathbb{R}$, $x \neq 1$ and $n \in \mathbb{N}$

$$\frac{x^{n+1}-1}{x-1} = 1 + x + x^2 + \cdots + x^n.$$

3. (Bernoulli's inequality)
   Let $h \in (-1, +\infty)$ and $n \in \mathbb{N}$. Show that

$$(1+h)^n \geq 1 + nh.$$

Solution: Let $S := \{n \in \mathbb{N} \mid (1+h)^n \geq 1 + nh\}$.

1^o $1 \in S$, since $(1+h)^1 = 1 + 1 \cdot h$.

2^o Let $k \in S$. Then $k+1 \in S$, since

$$\begin{aligned}(1+h)^{k+1} = (1+h)^k(1+h) &\geq (1+kh)(1+h) =\\ &= 1 + (k+1)h + kh^2 \geq 1 + (k+1)h.\end{aligned}$$

(In addition to the rules of ordering we have exploited the fact that $k \in S$, that is, $(1+h)^k \geq 1 + kh$.)

Keeping in mind requirement 4^o during the introduction of $\mathbb{N}$, this means that $S = \mathbb{N}$, so the inequality holds for all $n \in \mathbb{N}$. This method of proof is called **mathematical induction**.

4. Let $a, b \in \mathbb{R}^+$.

$$A_2 := \frac{a+b}{2}, \quad G_2 := \sqrt{ab}, \quad H_2 := \frac{2}{\frac{1}{a} + \frac{1}{b}}, \quad N_2 := \sqrt{\frac{a^2+b^2}{2}}.$$

Show that $H_2 \leq G_2 \leq A_2 \leq N_2$, and there is equality between the numbers if and only if $a = b$.

These equalities are also valid in a more general case.

Let $k \in \mathbb{N}$ $(k \geq 3)$ and $x_1, x_2, \ldots, x_k \in \mathbb{R}^+$.

$$A_k := \frac{x_1 + x_2 + \cdots + x_k}{k}, \quad G_k := \sqrt[k]{x_1 x_2 \cdots x_k},$$
$$H_k := \frac{k}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_k}}, \quad N_k := \sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_k^2}{k}}.$$

One can show that $H_k \leq G_k \leq A_k \leq N_k$, and there is equality between the numbers if and only if $x_1 = x_2 = \ldots = x_k$.

5. Let $h \in \mathbb{R}$ and $n \in \mathbb{N}$. Then

$$(1+h)^n = 1 + nh + \binom{n}{2} h^2 + \binom{n}{3} h^3 + \cdots + h^n,$$

where, exploiting the fact that $k! := 1 \cdot 2 \cdot \ldots \cdot k$,

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \quad k = 0, 1, 2, \ldots, n$$

(remember that $0! := 1$).

From this, one can prove the binomial theorem:

Let $a, b \in \mathbb{R},\ n \in \mathbb{N}$. Then

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

6. Let $A := \left\{\frac{n}{n+1} \mid n \in \mathbb{N}\right\}$. Show that A is bounded above. Find $\sup A$.

   Solution: Since for all $n \in \mathbb{N}$, $n < n+1$, therefore $\frac{n}{n+1} < 1$, so $K := 1$ is an upper bound. We show that $\sup A = 1$, since

   1^o For all $n \in \mathbb{N}$, $\frac{n}{n+1} < 1$.

   2^o Let $\varepsilon \in \mathbb{R}^+$. We seek such an index $n \in \mathbb{N}$ for which
$$\begin{aligned}\frac{n}{n+1} &> 1-\varepsilon,\\ n &> (1-\varepsilon)(n+1) = n - \varepsilon n + 1 - \varepsilon,\\ \varepsilon n &> 1-\varepsilon,\\ n &< \frac{1-\varepsilon}{\varepsilon}.\end{aligned}$$

   Since one can find a greater natural number than any real number, there is a greater natural number than $\frac{1-\varepsilon}{\varepsilon} \in \mathbb{R}$ as well, let it be $n' \in \mathbb{N}$, therefore for $\frac{n'}{n'+1} \in A$, $\frac{n'}{n'+1} > 1-\varepsilon$. So $\sup A = 1$.

7. * Let $E := \{(\frac{n+1}{n})^n \mid n \in \mathbb{N}\}$. Show that $E \subset \mathbb{R}$ is bounded above.

   Solution: We show that for all $n \in \mathbb{N}$
$$\left(\frac{n+1}{n}\right)^n \leq 4.$$

   Let $n \in \mathbb{N}$, and consider the number $\frac{1}{4}(\frac{n+1}{n})^n$. According to the equality between the algebraic (A_k) and geometric (G_k) means in Exercise 4:
$$\begin{aligned}\frac{1}{4}\left(\frac{n+1}{n}\right)^n &= \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{n+1}{n}\cdot\frac{n+1}{n}\cdots\frac{n+1}{n} \leq\\ &\leq \left(\frac{\frac{1}{2}+\frac{1}{2}+\frac{n+1}{n}+\frac{n+1}{n}\cdots\frac{n+1}{n}}{n+2}\right)^{n+2} = 1,\end{aligned}$$

   thus $(\frac{n+1}{n})^n \leq 4$, and so E is bounded above. According to the least upper bound axiom it **has** a supremum. Let $e := \sup E$.

   We remark that this supremum has never been and will never be conjectured (as opposed to Exercise 6...). It is approximately $e \approx 2.71$. The number e was introduced by Euler.

8. Let
$$P := \left\{\left(1-\frac{1}{2}\right)\cdot\left(1-\frac{1}{2^2}\right)\cdot\left(1-\frac{1}{2^3}\right)\cdots\left(1-\frac{1}{2^n}\right) \mid n \in \mathbb{N}\right\}.$$

   Is there an inf P? (When you have shown that inf P exists, do not get disappointed if you cannot find it. The problem is unsolved.)

3.3 Complex numbers

3.3.1 The concept of complex numbers, operations

We generalize the real numbers in such a way that the properties of the operations remain unchanged.

Let $\mathbb{C} := \mathbb{R} \times \mathbb{R}$ the set of real ordered pairs. Introduce the addition for any $(a, b), (c, d) \in \mathbb{C}$ as

$$(a, b) + (c, d) := (a + c, b + d);$$

and the multiplication as

$$(a, b) \cdot (c, d) := (ac - bd, ad + bc).$$

It is easy to check some properties of addition and multiplication.

a1. $\forall (a, b), (c, d) \in \mathbb{C}$, $(a, b) + (c, d) = (c, d) + (a, b)$ (commutativity).

a2. $\forall (a, b), (c, d), (e, f) \in \mathbb{C}$, $(a, b) + ((c, d) + (e, f)) = ((a, b) + (c, d)) + (e, f)$ (associativity).

a3. $\forall (a, b) \in \mathbb{C}$, $(a, b) + (0, 0) = (a, b)$.

a4. $\forall (a, b) \in \mathbb{C}$, $(-a, -b) \in \mathbb{C}$ is such that $(a, b) + (-a, -b) = (0, 0)$.

m1. $\forall (a, b), (c, d) \in \mathbb{C}$, $(a, b) \cdot (c, d) = (c, d) \cdot (a, b)$ (commutativity).

m2. $\forall (a, b), (c, d), (e, f) \in \mathbb{C}$, $(a, b) \cdot ((c, d) \cdot (e, f)) = ((a, b) \cdot (c, d)) \cdot (e, f)$ (associativity).

m3. $\forall (a, b) \in \mathbb{C}$, $(a, b) \cdot (1, 0) = (a, b)$.

m4. $\forall (a, b) \in \mathbb{C} \setminus \{(0, 0)\}$, $(\frac{a}{a^2+b^2}, -\frac{b}{a^2+b^2}) \in \mathbb{C}$ is such that

$$(a, b) \cdot \left(\frac{a}{a^2 + b^2}, -\frac{b}{a^2 + b^2} \right) = (1, 0).$$

d. $\forall (a, b), (c, d), (e, f) \in \mathbb{C}$

$$(a, b) \cdot [(c, d) + (e, f)] = (a, b) \cdot (c, d) + (a, b) \cdot (e, f)$$

(multiplication is distributive with respect to addition).

The properties a1–a4, m1–m4 and d ensure that operations and calculations performed with real numbers (containing only addition and multiplication and referring only to equalities) can be performed with complex numbers in the same way.

Let us identify the real number $a \in \mathbb{R}$ and the complex number $(a, 0) \in \mathbb{C}$. (Clearly, there is a one-to-one correspondence between $\mathbb{R}$ and the complex set $\mathbb{R} \times \{0\} \subset \mathbb{C}$.) We introduce the **imaginary unit** $i := (0, 1) \in \mathbb{C}$. Then for all complex number $(a, b) \in \mathbb{C}$

$$(a, b) = (a, 0) + (0, 1)(b, 0) = a + ib.$$

(The second equality is the consequence of the identification!)

Taking into account that $i^2 = (0, 1) \cdot (0, 1) = -1$, the addition becomes simple:

$$a + ib + c + id = a + c + i(b + d),$$

and so does the multiplication:

$$(a + ib) \cdot (c + id) = ac - bd + i(ad + bc).$$

Complex numbers can be illustrated as position vectors (Fig. 3.1).

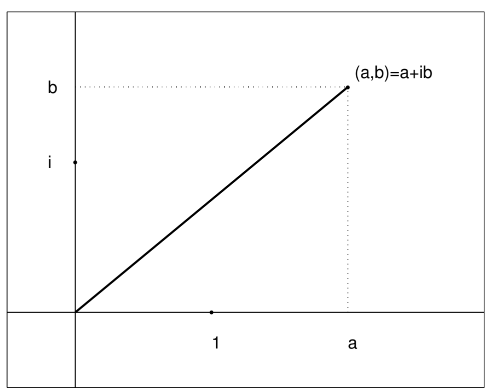


Figure 3.1

Addition corresponds to the addition of vectors in the plane by the "parallelogram rule" (Fig. 3.2).

3.3.2 The trigonometric form of complex numbers

To a complex number $a + ib \in \mathbb{C}$ we can assign its absolute value and its direction angle (Fig. 3.3).

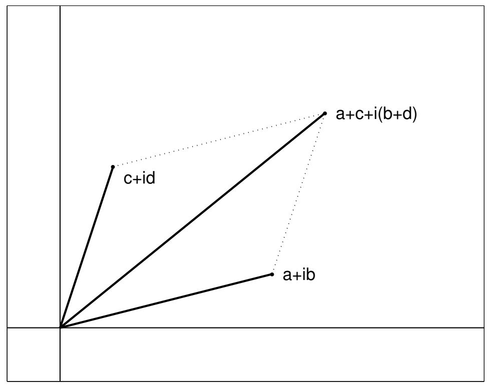


Figure 3.2

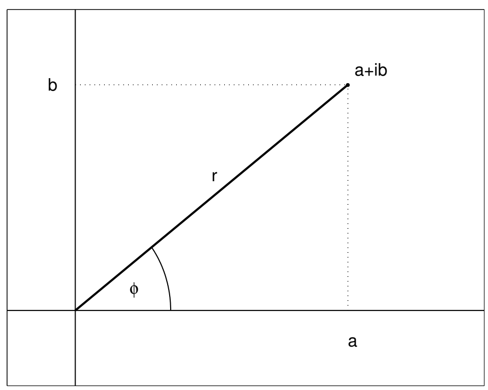


Figure 3.3

The absolute value: $r = \sqrt{a^2 + b^2}$.
The direction angle can be given in each quarter plane:

$$\phi = \begin{cases} \operatorname{arctg}\frac{b}{a} & \text{if } a > 0 \text{ and } b \geq 0, \\ \frac{\pi}{2} & \text{if } a = 0 \text{ and } b > 0, \\ \pi - \operatorname{arctg}|\frac{b}{a}| & \text{if } a < 0 \text{ and } b \geq 0, \\ \pi + \operatorname{arctg}|\frac{b}{a}| & \text{if } a < 0 \text{ and } b < 0, \\ \frac{3\pi}{2} & \text{if } a = 0 \text{ and } b < 0, \\ 2\pi - \operatorname{arctg}|\frac{b}{a}| & \text{if } a > 0 \text{ and } b < 0. \end{cases}$$

One can see that for the direction angle $\phi \in [0, 2\pi)$. We remark that for $a = 0, b = 0$: $r = 0$, and the direction angle is arbitrary.

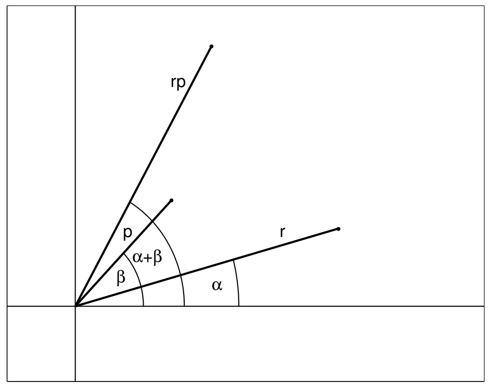


Figure 3.4

If a complex number $a+ib \in \mathbb{C}$ has absolute vale r and direction angle ϕ, then

$$a = r\cos\phi, \quad b = r\sin\phi,$$

therefore, $a+ib = r(\cos\phi + i\sin\phi)$. This is the **trigonometric form** of a complex number. With the aid of the trigonometric form the multiplication of complex numbers becomes geometrically meaningful.
Let $r(\cos\alpha + i\sin\alpha),\ p(\cos\beta + i\sin\beta) \in \mathbb{C}$, then

$$\begin{aligned} r(\cos\alpha + i\sin\alpha)\cdot\ & p(\cos\beta + i\sin\beta) = \\ &= rp(\cos\alpha\cos\beta - \sin\alpha\sin\beta + i(\sin\alpha\cos\beta + \cos\alpha\sin\beta)) = \\ &= rp(\cos(\alpha+\beta) + i\sin(\alpha+\beta)). \end{aligned}$$

So, by multiplication the absolute values are to be multiplied, and the direction angles to be added (Fig. 3.4).

Exponentiation also becomes fairly simple with the trigonometric form. If $z = a+ib = r(\cos\phi + i\sin\phi) \in \mathbb{C}$ and $n \in \mathbb{N}$, then

$$z^n = (a+ib)^n = [r(\cos\phi + i\sin\phi)]^n = r^n(\cos n\phi + i\sin n\phi),$$

so, when raising a complex number z to the nth power, the nth power of the absolute value and n times the direction angle are taken in the trigonometric form of z^n.

Chapter 4

Elementary functions

We present the major properties of functions defined on and mapping to the set of real numbers. We define the frequently used real functions called elementary functions. The following topics will be covered.

- Operations on real functions
- Bounded, monotone, periodic, odd and even functions
- Power functions
- Exponential and logarithmic functions
- Trigonometric functions and their inverses
- Hyperbolic function and their inverses
- Some peculiar functions

4.1 The basic properties of real functions

Definition 4.1. Let $f : \mathbb{R} \supset\!\!\to \mathbb{R},\ \lambda \in \mathbb{R}$. Then

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

Definition 4.2. Let $f, g : \mathbb{R} \supset\!\!\to \mathbb{R},\ D(f) \cap D(g) \neq \emptyset$. Then

$$\begin{aligned} f + g : D(f) \cap D(g) \to \mathbb{R}, &\quad (f+g)(x) := f(x) + g(x), \\ f \cdot g : D(f) \cap D(g) \to \mathbb{R}, &\quad (f \cdot g)(x) := f(x) \cdot g(x). \end{aligned}$$

Definition 4.3. Let $g : \mathbb{R} \supset\!\!\to \mathbb{R},\ H := D(g) \setminus \{x \in D(g) \mid g(x) = 0\} \neq \emptyset$. Then

$$1/g : H \to \mathbb{R},\ (1/g)(x) := \frac{1}{g(x)}.$$

Definition 4.4. Let $f, g : \mathbb{R} \supset\!\!\to \mathbb{R}$

$$\frac{f}{g} := f \cdot 1/g$$

Definition 4.5. Let $f : \mathbb{R} \supset\!\!\to \mathbb{R}$. We say that f is **bounded above** if the set $R(f) \subset \mathbb{R}$ is bounded above.

We say that f is **bounded below** if the set $R(f) \subset \mathbb{R}$ is bounded below.

We say that f is a **bounded** function if the set $R(f) \subset \mathbb{R}$ is bounded below and above.

Definition 4.6. Let $f : \mathbb{R} \supset\!\!\to \mathbb{R}$. We say that f is a **monotonically increasing** function if for all $x_1, x_2 \in D(f),\ x_1 < x_2$: $f(x_1) \le f(x_2)$.

The function f is **strictly monotonically increasing** if for all $x_1, x_2 \in D(f)$, $x_1 < x_2$: $f(x_1) < f(x_2)$.

We say that f is a **monotonically decreasing** function if for all $x_1, x_2 \in D(f)$, $x_1 < x_2$: $f(x_1) \ge f(x_2)$.

The function f is **strictly monotonically decreasing** if for all $x_1, x_2 \in D(f)$, $x_1 < x_2$: $f(x_1) > f(x_2)$.

Definition 4.7. Let $f : \mathbb{R} \supset\!\!\to \mathbb{R}$. We say that f is an **even** function if

1^o for all $x \in D(f)$, $-x \in D(f)$,

2^o for all $x \in D(f)$, $f(-x) = f(x)$.

Definition 4.8. Let $f : \mathbb{R} \supset\!\!\to \mathbb{R}$. We say that f is an **odd** function if

1^o for all $x \in D(f)$, $-x \in D(f)$,

2^o for all $x \in D(f)$, $f(-x) = -f(x)$.

Definition 4.9. Let $f : \mathbb{R} \supset\!\!\to \mathbb{R}$. We say that f is a **periodic** function if there exists a number $p \in \mathbb{R},\ 0 < p$ such that

1^o for all $x \in D(f)$, $x + p,\ x - p \in D(f)$,

2^o for all $x \in D(f)$, $f(x + p) = f(x - p) = f(x)$.

The number p is called a **period** of the function f.

4.2 Elementary functions

4.2.1 Power functions

Let $\mathrm{id} : \mathbb{R} \supset\!\!\to \mathbb{R},\ \mathrm{id}(x) := x$. As Fig. 4.1 shows, id is a strictly monotonically increasing function.

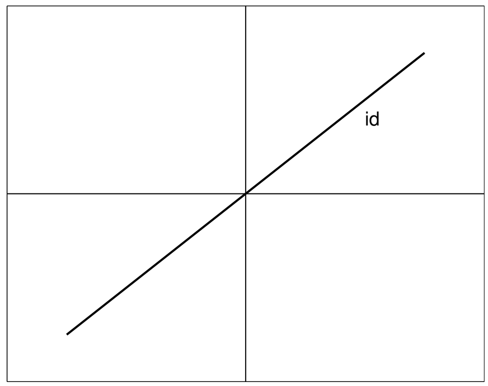


Figure 4.1

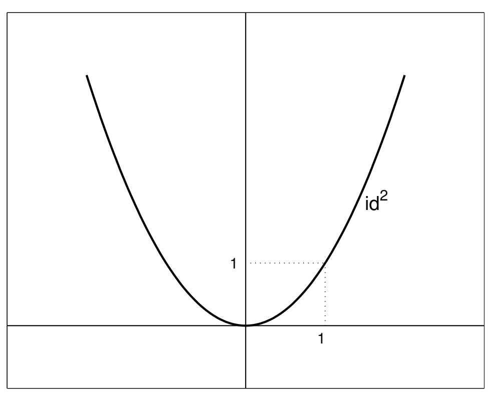


Figure 4.2

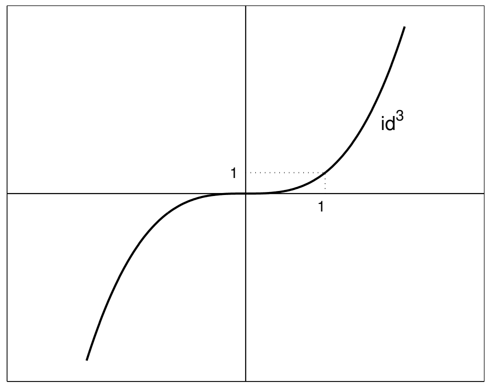


Figure 4.3

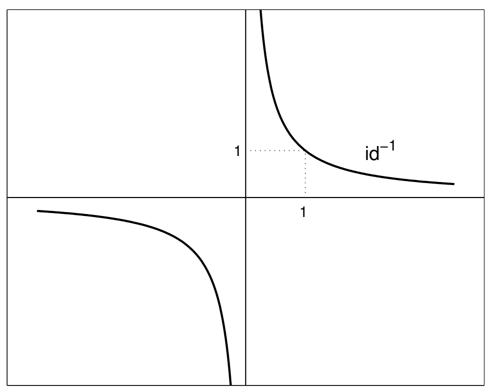


Figure 4.4

Let $\mathrm{id}^2 : \mathbb{R} \supset\!\to \mathbb{R},\ \mathrm{id}^2(x) := x^2$. Clearly, $\mathrm{id}^2{}_{|\mathbb{R}+}$ is a strictly monotonically increasing function, while $\mathrm{id}^2{}_{|\mathbb{R}-}$ is strictly monotonically decreasing. The function id^2 is even (Fig. 4.2).

Let $\mathrm{id}^3 : \mathbb{R} \supset\!\to \mathbb{R},\ \mathrm{id}^3(x) := x^3$. Function id^3 is strictly monotonically increasing and odd (Fig. 4.3). If $n \in \mathbb{N}$, then the function $\mathrm{id}^n : \mathbb{R} \to \mathbb{R}$, $\mathrm{id}^n(x) := x^n$ inherits the properties of id^2 for even n, and the properties of id^3 for odd n.

Let $\mathrm{id}^{-1} : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $\mathrm{id}^{-1}(x) := 1/x$. The functions $\mathrm{id}^{-1}{}_{|\mathbb{R}-}$ and $\mathrm{id}^{-1}{}_{|\mathbb{R}+}$ are strictly monotonically decreasing (however, id^{-1} is not monotone!). The function id^{-1} is odd (Fig. 4.4).

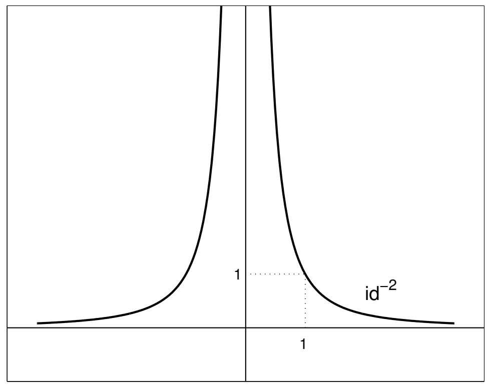


Figure 4.5

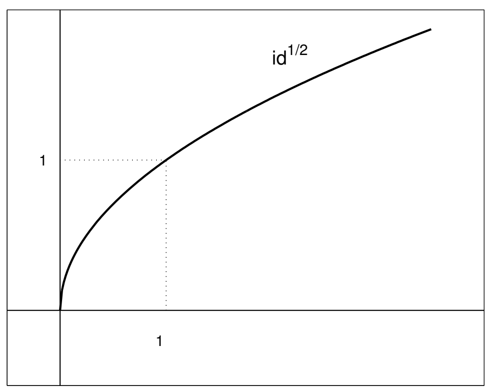


Figure 4.6

Let $\mathrm{id}^{-2} : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $\mathrm{id}^{-2}(x) := 1/x^2$. The function $\mathrm{id}^{-2}{}_{|_{\mathbb{R}^-}}$ is strictly monotonically increasing, while $\mathrm{id}^{-2}{}_{|_{\mathbb{R}^+}}$ is strictly monotonically decreasing. The function id^{-2} is even (Fig. 4.5).

Let $n \in \mathbb{N}$. The function $\mathrm{id}^{-n} : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $\mathrm{id}^{-n}(x) := 1/x^n$ inherits the properties of id^{-2} if n is even, and those of id^{-1} if n is odd.

Let $\mathrm{id}^{1/2} : [0, \infty) \to \mathbb{R}$, $\mathrm{id}^{1/2}(x) := \sqrt{x}$. The function $\mathrm{id}^{1/2}$ is strictly monotonically increasing (Fig. 4.6). We mention that $\mathrm{id}^{1/2}$ can also be defined as the inverse of the one-to-one function $\mathrm{id}^2{}_{|_{[0,\infty)}}$.

Let $r \in \mathbb{Q}$, and consider the function $\mathrm{id}^r : \mathbb{R}^+ \to \mathbb{R}$, $\mathrm{id}^r(x) := x^r$. For some values of r the functions id^r are plotted in Fig. 4.7.

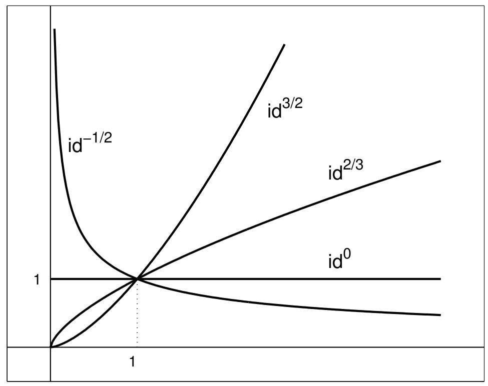


Figure 4.7

Finally, let $\mathrm{id}^0 : \mathbb{R} \to \mathbb{R}$, $\mathrm{id}^0(x) := 1$. The function id^0 is even, monotonically increasing, and at the same time monotonically decreasing. It is periodic by any number $p > 0$ (Fig. 4.7).

4.2.2 Exponential and logarithmic functions

Let $a \in \mathbb{R}^+$. The exponential function with base a is defined as

$$\exp_a : \mathbb{R} \to \mathbb{R}, \quad \exp_a(x) := a^x.$$

$\exp_a$ is strictly monotonically increasing if $a > 1$,

$\exp_a$ is strictly monotonically decreasing if $a < 1$,

$\exp_a = \mathrm{id}^0$ if $a = 1$ (monotonically increasing and decreasing at the same time) (Fig. 4.8).

If $a > 0$ and $a \neq 1$, then $R(\exp_a) = \mathbb{R}^+$, so $\exp_a$ only takes positive values (and it does take all positive values). For all $a > 0$ and by any $x_1, x_2 \in \mathbb{R}$:

$$\exp_a(x_1 + x_2) = \exp_a(x_1) \cdot \exp_a(x_2).$$

(This is the most important characteristic of the exponential functions.) A special role is played by the function $\exp_e =: \exp$ (Fig. 4.9) (where e is Euler's number introduced in Exercise 7* of the previous chapter).

Let $a > 0, a \neq 1$. Since $\exp_a$ is strictly monotone, therefore it is one-to-one, and so it has an inverse function:

$$\log_a := (\exp_a)^{-1}$$

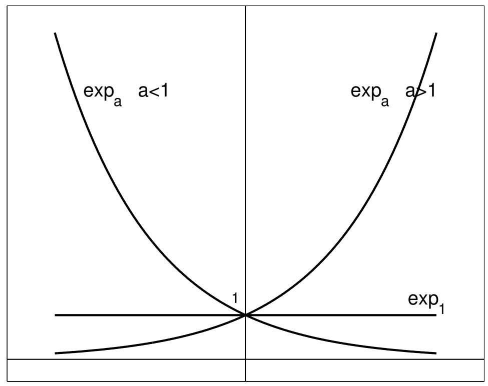


Figure 4.8

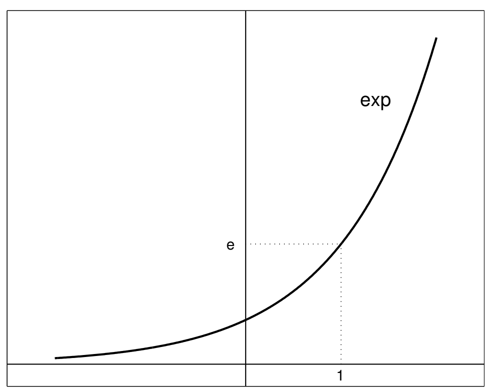


Figure 4.9

called logarithmic function with base a (Fig. 4.10). So

$$\log_a : \mathbb{R}^+ \to \mathbb{R}, \ \ \log_a(x) = y, \text{ for which } \exp_a(y) = x.$$

If $a > 1$, then $\log_a$ is strictly monotonically increasing, and if $a < 1$, then $\log_a$ is strictly monotonically decreasing. Logarithmic functions have the fundamental properties that

1^o for all $a > 0$, $a \neq 1$ and any $x_1, x_2 \in \mathbb{R}^+$

$$\log_a(x_1 x_2) = \log_a x_1 + \log_a x_2;$$

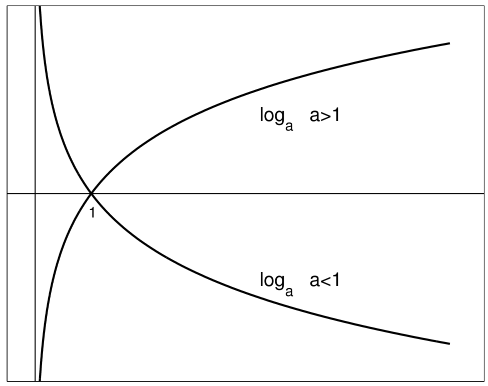


Figure 4.10

2^o for all $a > 0,\ a \neq 1$ and any $x \in \mathbb{R}^+$ and $k \in \mathbb{R}$

$$\log_a x^k = k \log_a x;$$

3^o for all $a, b > 0, a, b \neq 1$ and any $x \in \mathbb{R}^+$

$$\log_a x = \frac{\log_b x}{\log_b a}.$$

Property 3^o implies that all logarithmic functions can be obtained by multiplying any one logarithmic function by a real number. That is why the logarithmic function to the base e, called "natural logarithm" plays a special role:

$$\ln := \log_e$$

(Fig. 4.11).

4.2.3 Trigonometric functions and their inverses

Let $\sin : \mathbb{R} \to \mathbb{R},\ \sin x :=$ Do not expect a formula here! Draw a circle of radius 1. Then draw two straight lines perpendicular to each other through the center of the circle. One of them will be called axis (1), while the other axis (2). From the point where the (positive half) of axis (1) intersects the circle "measure the arc corresponding to the number $x \in \mathbb{R}$ to the circumference". [This operation requires considerable manual skills!...] The second coordinate of the end point P of the arc will be $\sin x$ (Fig. 4.12). The sine function is odd, and periodic with period $p = 2\pi$ (Fig. 4.13). $R(\sin) = [-1, 1]$.

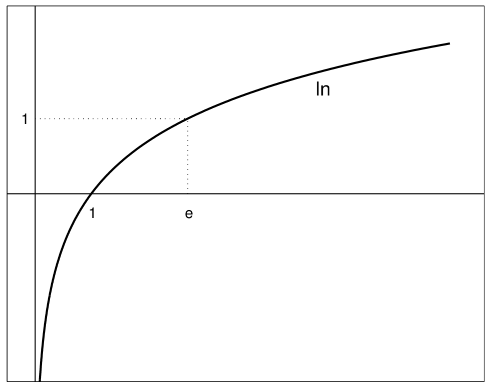


Figure 4.11

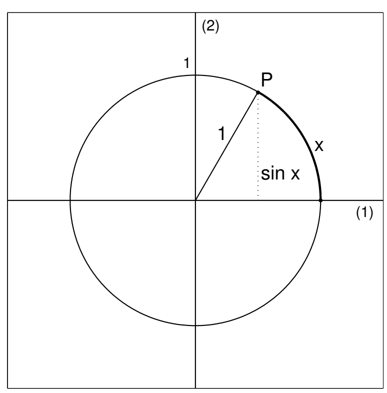


Figure 4.12

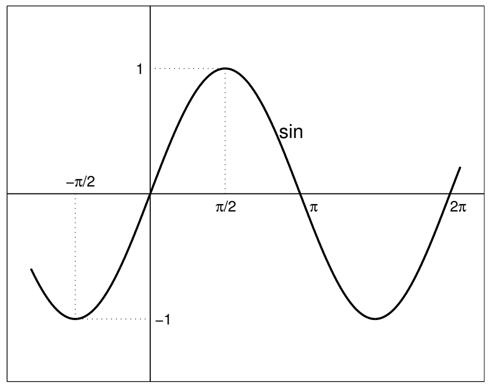


Figure 4.13

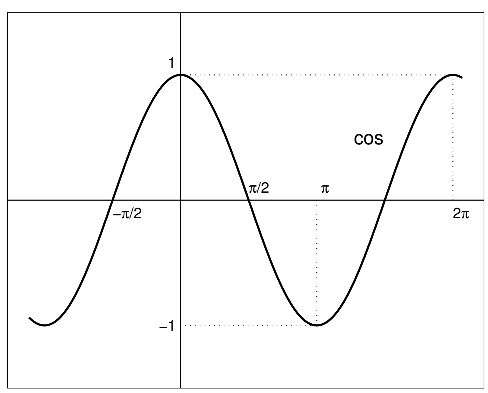


Figure 4.14

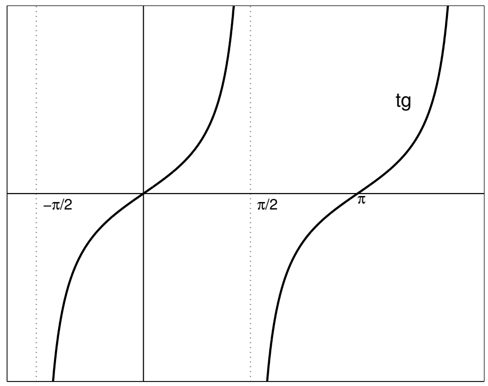


Figure 4.15

Let $\cos : \mathbb{R} \to \mathbb{R}$, $\cos x := \sin(x + \frac{\pi}{2})$. The cosine function is even, and periodic with period $p = 2\pi$ (Fig. 4.14). $R(\cos) = [-1, 1]$.

Fundamental relationships:

1^o For all $x \in \mathbb{R}$, $\cos^2 x + \sin^2 x = 1$.

2^o For all $x_1, x_2 \in \mathbb{R}$, $\sin(x_1 + x_2) = \sin x_1 \cos x_2 + \cos x_1 \sin x_2$,
$\cos(x_1 + x_2) = \cos x_1 \cos x_2 - \sin x_1 \sin x_2$.

Let $\mathrm{tg} := \frac{\sin}{\cos}$ and $\mathrm{ctg} := \frac{\cos}{\sin}$.

It follows from the definition that

$$D(\mathrm{tg}) = \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi \mid k \in \mathbb{Z} \right\}, \ D(\mathrm{ctg}) = \mathbb{R} \setminus \{ k\pi \mid k \in \mathbb{Z} \}.$$

The functions tg and ctg are odd, and periodic with period $p = \pi$ (Fig. 4.15 and Fig. 4.16).

Due to their periodicity, trigonometric functions are not one-to-one functions.

Consider the resriction $\sin_{|[-\frac{\pi}{2}, \frac{\pi}{2}]}$. This function is strictly monotonically increasing, therefore one-to-one, and so it has an inverse function:

$$\arcsin := (\sin_{|[-\frac{\pi}{2}, \frac{\pi}{2}]})^{-1}.$$

From the definition $\arcsin : [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$, $\arcsin x = \alpha$ for which $\sin \alpha = x$.

The arcsin function is strictly monotonically increasing and odd (Fig. 4.17).

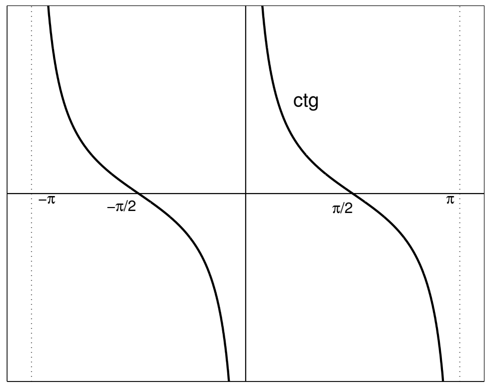


Figure 4.16

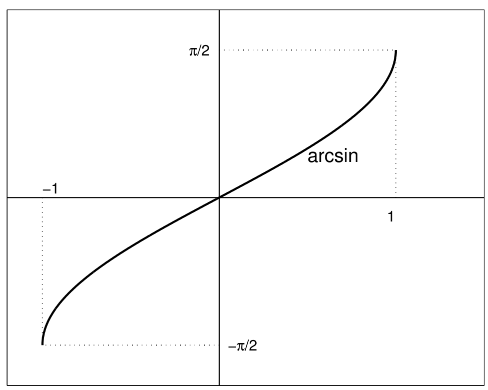


Figure 4.17

The restriction to the interval $[0, \pi]$ of the cosine function is strictly monotonically decreasing, therefore it has an inverse function:

$$\arccos := (\cos|_{[0,\pi]})^{-1}.$$

From the definition it follows that $\arccos : [-1, 1] \to [0, \pi]$, $\arccos x = \alpha$ for which $\cos \alpha = x$.

The arccos function is strictly monotonically decreasing (Fig. 4.18).

The restriction to the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$ of the tg function is strictly monotonically increasing, therefore it has an inverse function:

$$\mathrm{arctg} := (\sin|_{[-\frac{\pi}{2},\frac{\pi}{2}]})^{-1}.$$

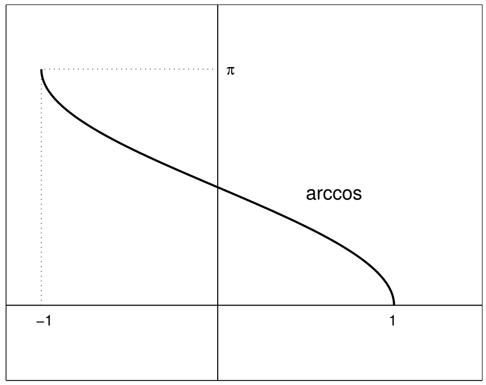


Figure 4.18

From the definition it follows that arctg: $\mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$, $\operatorname{arctg} x = \alpha$ for which $\operatorname{tg} \alpha = x$.

The arctg function is strictly monotonically increasing and odd (Fig. 4.19).

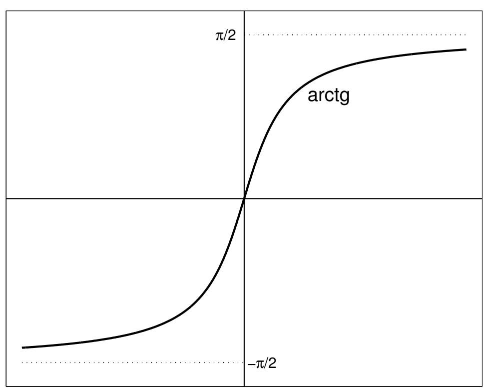


Figure 4.19

The restriction to the interval $(0, \pi)$ of the ctg function is strictly monotonically decreasing, therefore it has an inverse function:

$$\operatorname{arcctg} := (\operatorname{ctg}_{|_{[0,\pi]}})^{-1}.$$

From the definition it follows that $\operatorname{arcctg} : \mathbb{R} \to (0, \pi)$, $\operatorname{arcctg} x = \alpha$ for which $\operatorname{ctg} \alpha = x$.

The arcctg function is strictly monotonically decreasing (Fig. 4.20).

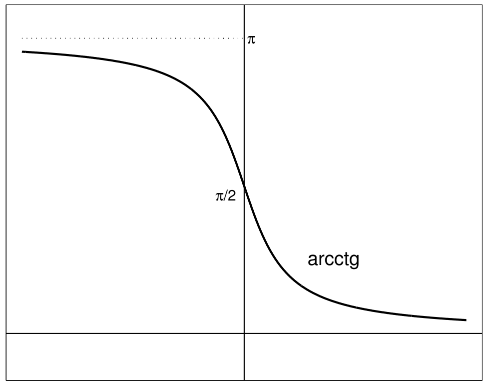


Figure 4.20

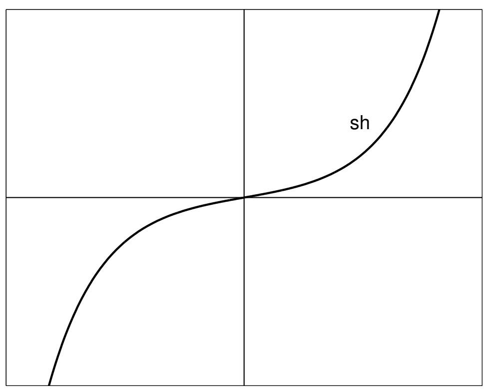


Figure 4.21

4.2.4 Hyperbolic functions and their inverses

Let $\mathrm{sh} : \mathbb{R} \to \mathbb{R}$, $\mathrm{sh}x := \frac{e^x - e^{-x}}{2}$. The sh function is strictly monotonically increasing and odd (Fig. 4.21).

Let $\mathrm{ch} : \mathbb{R} \to \mathbb{R}$, $\mathrm{ch}x := \frac{e^x + e^{-x}}{2}$. The function $\mathrm{ch}_{|_{\mathbb{R}-}}$ is strictly monotonically decreasing, while $\mathrm{ch}_{|_{\mathbb{R}+}}$ is strictly monotonically increasing. The ch function is even. $R(\mathrm{ch}) = [1, +\infty)$. This function is often called chain curve (Fig. 4.22).

Fundamental relationships:

1^o For all $x \in \mathbb{R}$, $\mathrm{ch}^2 x - \mathrm{sh}^2 x = 1$.

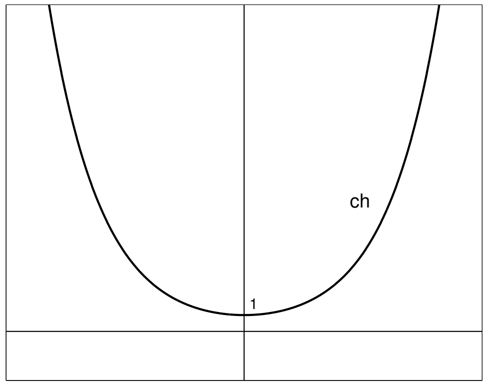


Figure 4.22

2^o For all $x_1, x_2 \in \mathbb{R}$

$$\mathrm{sh}(x_1 + x_2) = \mathrm{sh}x_1 \mathrm{ch}x_2 + \mathrm{ch}x_1 \mathrm{sh}x_2,$$

$$\mathrm{ch}(x_1 + x_2) = \mathrm{ch}x_1 \mathrm{ch}x_2 + \mathrm{sh}x_1 \mathrm{sh}x_2.$$

Let $\mathrm{th} := \frac{\mathrm{sh}}{\mathrm{ch}}$, $\mathrm{cth} := \frac{\mathrm{ch}}{\mathrm{sh}}$.

It follows from the definition that $\mathrm{th} : \mathbb{R} \to \mathbb{R}$, $\mathrm{th}\, x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$, $\mathrm{cth} : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $\mathrm{cth}\, x = \frac{e^x + e^{-x}}{e^x - e^{-x}}$. The th and cth functions are odd (Fig. 4.23).

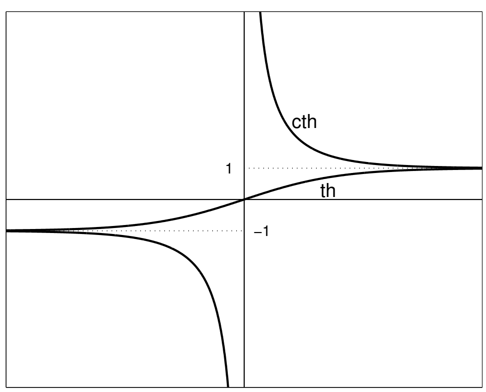


Figure 4.23

The th function is strictly monotonically increasing. $R(\mathrm{th}) = (-1, 1)$.

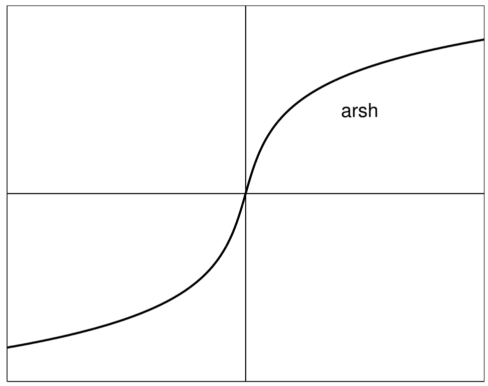


Figure 4.24

The function $\mathrm{cth}_{|_{\mathbb{R}^-}}$ is strictly monotonically decreasing, while $\mathrm{cth}_{|_{\mathbb{R}^+}}$ is strictly monotonically increasing. $R(\mathrm{cth}) = \mathbb{R} \setminus [-1, 1]$.

The sh function is strictly monotonically increasing, and so it has an inverse function:

$$\mathrm{arsh} := (\mathrm{sh})^{-1}.$$

It follows from the definition that $\mathrm{arsh} : \mathbb{R} \to \mathbb{R},\ \mathrm{arsh}\, x = \ln(x + \sqrt{x^2+1})$ (see Exercise 5). The arsh function is strictly monotonically increasing and odd (Fig. 4.24).

The restriction to the interval $[0, \infty)$ of the ch function is strictly monotonically increasing, therefore it has an inverse function:

$$\mathrm{arch} := (\mathrm{ch}_{|_{[0,\infty)}})^{-1}.$$

From the definition it follows that $\mathrm{arch} : [1, \infty) \to [0, \infty),\ \mathrm{arch}\, x = \ln(x + \sqrt{x^2-1})$. The arch function is strictly monotonically increasing (see Fig. 4.25).

The th function is strictly monotonically increasing, so it has an inverse function:

$$\mathrm{arth} := (\mathrm{th})^{-1}.$$

From the definition it follows that $\mathrm{arth} : (-1, 1) \to \mathbb{R},\ \mathrm{arth}\, x = \frac{1}{2} \ln \frac{1+x}{1-x}$. The arth function is strictly monotonically increasing and odd (Fig. 4.26).

The restriction to $\mathbb{R}^+$ of the cth function is strictly monotonically decreasing, therefore it has an inverse function:

$$\mathrm{arcth} := (\mathrm{cth}_{|_{\mathbb{R}^+}})^{-1}.$$

From the definition it follows that $\mathrm{arcth} : (1, +\infty) \to \mathbb{R}^+,\ \mathrm{arcth}\, x = \frac{1}{2} \ln \frac{x+1}{x-1}$. The arcth function is strictly monotonically decreasing (Fig. 4.27).

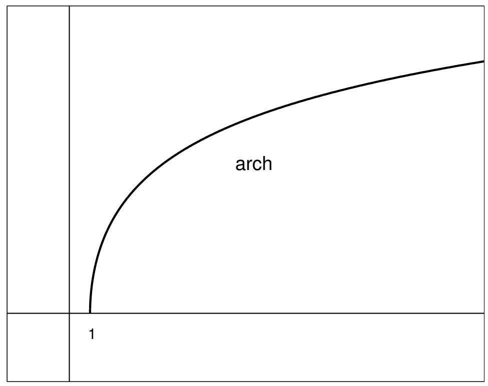


Figure 4.25

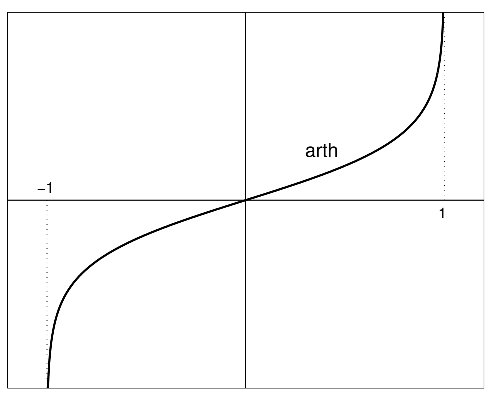


Figure 4.26

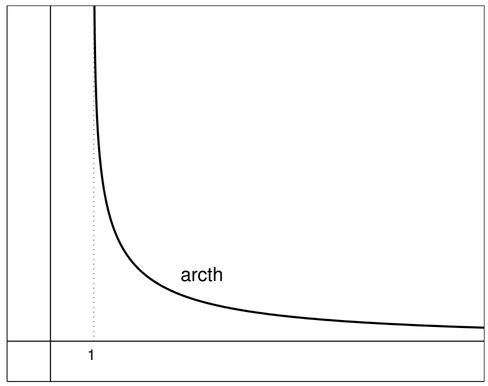


Figure 4.27

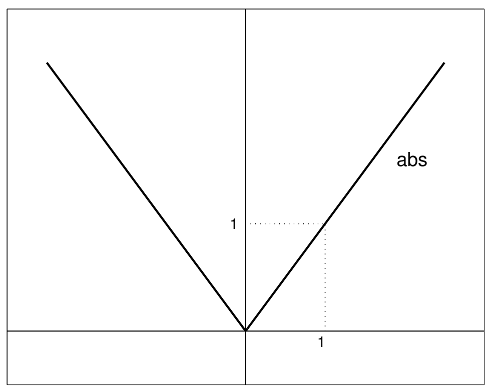


Figure 4.28

4.2.5 Some peculiar functions

1. Let $\mathrm{abs} : \mathbb{R} \to \mathbb{R},\ \mathrm{abs}(x) := |x|$, where (as we saw before)

$$|x| := \begin{cases} x, & \text{if } x \geq 0, \\ -x, & \text{if } x < 0 \end{cases} \qquad \text{(Fig. 4.28).}$$

2. Let $\mathrm{sgn} : \mathbb{R} \to \mathbb{R}, \mathrm{sgn}(x) := \begin{cases} 1, & \text{if } x > 0, \\ 0, & \text{if } x = 0, \\ -1, & \text{if } x < 0 \end{cases} \qquad \text{(Fig. 4.29).}$

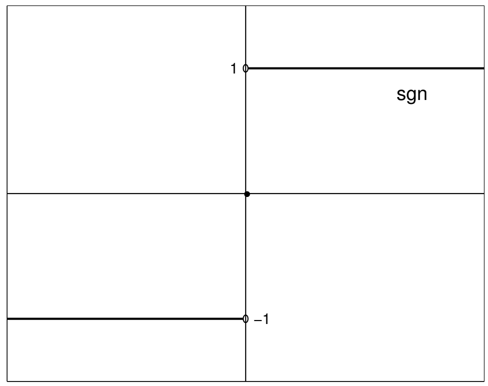


Figure 4.29

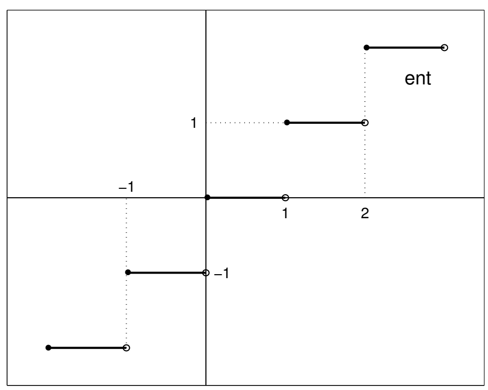


Figure 4.30

3. Let $\mathrm{ent} : \mathbb{R} \to \mathbb{R}, \mathrm{ent}(x) := [x]$, where

$$[x] := \max\{n \in \mathbb{Z} \mid n \leq x\}.$$

(The "integer part" of the number $x \in \mathbb{R}$ is the greatest integer that is less than or equal to x.) (Fig. 4.30.)

4. Let $d : \mathbb{R} \to \mathbb{R},\ d(x) := \begin{cases} 1 & \text{if } x \in \mathbb{Q}, \\ 0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$

This function is called Dirichlet's function, and we do not make an attempt to draw it.

5. Let $r : \mathbb{R} \to \mathbb{R}$
$$r(x) := \begin{cases} 0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \text{ or } x = 0, \\ \frac{1}{q} & \text{if } x \in \mathbb{Q},\ x = \frac{p}{q}, \end{cases}$$
where $p \in \mathbb{Z},\ q \in \mathbb{N}$, and p and q have no common divisor (different from 1). It is called Riemann's function, and again we do not try to plot it.

4.3 Exercises

1. Compute the following function values:

$$\mathrm{id}^0(7) = \qquad \mathrm{id}^3\left(\frac{1}{2}\right) = \qquad \mathrm{id}^{\frac{1}{2}}(4) = \qquad \mathrm{id}^{-6}(1) =$$
$$\mathrm{id}(6) = \qquad \mathrm{id}^3\left(-\frac{1}{2}\right) = \qquad \mathrm{id}^{\frac{3}{2}}(4) = \qquad \mathrm{id}^{-6}(2) =$$
$$\mathrm{id}^2(5) = \qquad \mathrm{id}^3(0) = \qquad \mathrm{id}^{-\frac{3}{2}}(4) = \qquad \mathrm{id}^{-6}\left(\frac{1}{2}\right) =$$

2. Arrange the following numbers in ascending order:

   a) $\sin 1,\ \sin 2,\ \sin 3,\ \sin 4$;

   b) $\ln 2,\ \exp_2 \frac{1}{2},\ \exp_{\frac{1}{2}} 2,\ \log_2 1$;

   c) $\mathrm{sh}\ 3,\ \mathrm{ch}\ (-2),\ \mathrm{arsh}\ 4,\ \mathrm{th}\ 1$;

   d) $\arcsin \frac{1}{2},\ \mathrm{arctg}\ 10,\ \mathrm{th}\ 10,\ \cos 1$.

3. Prove that $\mathrm{ch}^2 x - \mathrm{sh}^2 x = 1,\ \mathrm{ch}^2 x = \frac{\mathrm{ch}(2x)+1}{2}$ for all $x \in \mathbb{R}$.

4. Prove that for $x, y \in \mathbb{R}$

   a) $\sin 2x = 2 \sin x \cos x,\ \cos 2x = \cos^2 x - \sin^2 x,\ \cos^2 x = \frac{1+\cos 2x}{2}$, $\sin^2 x = \frac{1-\cos 2x}{2}$;

   b) $\sin x - \sin y = 2 \sin \frac{x-y}{2} \cos \frac{x+y}{2},\ \cos x - \cos y = 2 \sin \frac{y-x}{2} \sin \frac{x+y}{2}$.

5. Show that

   a) $\mathrm{arsh}\ x = \ln(x + \sqrt{x^2+1}) \quad (x \in \mathbb{R})$;

   b) $\mathrm{arch}\ x = \ln(x + \sqrt{x^2-1}) \quad (x \in [1, +\infty))$;

   c) $\mathrm{arth}\ x = \frac{1}{2} \ln \frac{1+x}{1-x} \quad (x \in (-1, 1))$.

   Solution: a)

   $1^o \quad y = \mathrm{sh}\ x = \frac{e^x - e^{-x}}{2}$;

2^o $\quad x = \frac{e^y - e^{-y}}{2};$
$2x = e^y - e^{-y} / \cdot e^y;$
$2xe^y = (e^y)^2 - 1;$
$(e^y)^2 - 2xe^y - 1 = 0;$
$(e^y)_{1,2} = \frac{2x \pm \sqrt{4x^2+4}}{2} = x \pm \sqrt{x^2+1}.$

Since the exp function only takes positive values, and for all $x \in \mathbb{R}$ $\sqrt{x^2+1} > \sqrt{x^2} = |x| \geq x$, therefore

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

From this

$$y = \ln(x + \sqrt{x^2+1}),$$

which means that

$$3^o \quad \text{arsh } x = \ln(x + \sqrt{x^2+1}).$$

6. Show that $\text{arctg} \neq \frac{\pi}{2}\text{th}$.

7. Sketch the following functions:

a) $f : \mathbb{R} \to \mathbb{R},\ f(x) := \begin{cases} \sin \frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$

b) $g : \mathbb{R} \to \mathbb{R},\ g(x) := \begin{cases} x^2 \sin \frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$

c) $h : \mathbb{R} \to \mathbb{R},\ h(x) := \begin{cases} x^2(\sin \frac{1}{x} + 2) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$

8. Let $f : \mathbb{R} \to \mathbb{R}$ be an arbitrary function. Show that for the functions $\phi, \psi : \mathbb{R} \to \mathbb{R}$

$$\phi(x) := \frac{f(x) + f(-x)}{2}, \quad \psi(x) := \frac{f(x) - f(-x)}{2}.$$

ϕ is even, ψ is odd, and $f = \phi + \psi$. If $f = \exp$, then what will be the functions ϕ and ψ?

9. Let $f, g : \mathbb{R} \to \mathbb{R}$. Assume that f is periodic with period $p > 0$, and g with period $q > 0$.

a) Show that if $\frac{p}{q} \in \mathbb{Q}$, then $f + g$ is periodic.

b) Give an example where $\frac{p}{q} \in \mathbb{R} \setminus \mathbb{Q}$, and $f + g$ is not periodic.

Solution: a) Let $\frac{p}{q} = \frac{k}{l}$, where $k, l \in \mathbb{N}$. Then $lp = kq$. Let $\omega := lp + kq > 0$. We show that $f + g$ is periodic with period ω.

1^o $D(f+g)=\mathbb{R}$.

2^o For all $x \in \mathbb{R}$

$$\begin{aligned}(f+g)(x+\omega) &= f(x+kq+lp)+g(x+lp+kq)\\ &= f(x+kq)+g(x+lp)=f(x+lp)+g(x+kq)\\ &= f(x)+g(x)=(f+g)(x).\end{aligned}$$

One can similarly prove that $(f+g)(x-\omega)=(f+g)(x)$.

Chapter 5

Sequences, series

Sequences are fairly simple functions. They are useful in studying the accuracy of approximations. They are important building stones of the later concepts. We will deal with the following topics.

- The concept of sequence, monotonicity, boundedness
- Limit and convergence
- Important limits
- The relationships between limits and operations
- The definition of the number e
- Cauchy's convergence criterion for sequences
- The convergence of series
- Convergence criteria for series

5.1 Sequences, series

5.1.1 The concept and properties of sequences

A sequence is a function defined on the set of natural numbers.

Let $H \neq \emptyset$ be a set. If $a : \mathbb{N} \to H$, then we have a sequence in H. For example, if H is the set of real numbers, then we have a number sequence; if H is a set of certain signals, then we have a signal sequence; if H is a set of intervals, then we have an interval sequence.

Let $a : \mathbb{N} \to \mathbb{R}$ be a number sequence. If $n \in \mathbb{N}$, then we denote the nth element of the sequence by a_n instead of $a(n)$. The number sequence $a : \mathbb{N} \to \mathbb{R}$ will also be denoted more briefly as (a_n), or we can emphasize by writing $(a_n) \subset \mathbb{R}$ that we have a number sequence.

For example, instead of $a : \mathbb{N} \to \mathbb{R}, a_n := \frac{1}{n}$ we can write $(\frac{1}{n})$.

Sometimes we use the lengthier notation $a_1, a_2, \dots, a_n, \dots$ instead of (a_n). For example, instead of (n^2) we can speak about the sequence $1, 4, 9, \dots, n^2, \dots$

Since sequences are functions, therefore the concepts of boundedness, monotonicity and the operations on sequences do not require new definitions. To remind the reader, we still reformulate some terms.

Definition 5.1. We say that the sequence (a_n) is bounded if there exists a $K \in \mathbb{R}$ such that for any $n \in \mathbb{N}$ $|a_n| \le K$.

Definition 5.2. We say that (a_n) is monotonically increasing if for all $n \in \mathbb{N}$ $a_n \le a_{n+1}$.

Definition 5.3. If (a_n) is a sequence, and $\lambda \in \mathbb{R}$, then

$$\lambda(a_n) := (\lambda a_n).$$

If $(a_n), (b_n)$ are two sequences, then

$$(a_n) + (b_n) := (a_n + b_n),$$
$$(a_n) \cdot (b_n) := (a_n \cdot b_n).$$

If moreover $b_n \neq 0$ $(n \in \mathbb{N})$, then

$$\frac{(a_n)}{(b_n)} := \left(\frac{a_n}{b_n}\right).$$

For example, the sequence $(\frac{n}{n+1})$ is bounded, since $n < n+1$ for all $n \in \mathbb{N}$, therefore

$$\left|\frac{n}{n+1}\right| = \frac{n}{n+1} < 1.$$

The sequence $(\frac{n}{n+1})$ is monotonically increasing, because for all $n \in \mathbb{N}$

$$a_n = \frac{n}{n+1} < \frac{n+1}{n+2} = a_{n+1},$$

since $n(n+2) < (n+1)^2$.

The sequence $(e_n) := ((\frac{n+1}{n})^n)$ is also monotonically increasing. To prove this, let $n \in \mathbb{N}$. Due to the inequality between the algebraic and geometric mean values:

$$e_n = \left(\frac{n+1}{n}\right)^n = 1 \cdot \frac{n+1}{n} \cdot \frac{n+1}{n} \cdots \frac{n+1}{n} \le \left(\frac{1 + n \cdot \frac{n+1}{n}}{n+1}\right)^{n+1}$$
$$= \left(\frac{n+2}{n+1}\right)^{n+1} = e_{n+1}.$$

The sequence (e_n) is bounded (proving this requires the same calculation as that shown in Exercise 7* of Chapter 3): for all $n \in \mathbb{N}$ $\left(\frac{n+1}{n}\right)^n \le 4$.

5.1.2 The limit of a sequence

Now we will learn about a completely new property of sequences. We call a sequence $a_1, a_2, \dots, a_n, \dots$ convergent if there is a value that the terms of the sequence "get close to eventually". More precisely:

Definition 5.4. We say that a number sequence (a_n) is **convergent**, if there exists an $A \in \mathbb{R}$ such that for any $\varepsilon > 0$ there exists an index $N \in \mathbb{N}$ for which $|a_n - A| < \varepsilon$ holds for all $n \in \mathbb{N},\ n > N$. If there is such a number A, then we call it **limit** of the sequence, and employ the notation $\lim a_n = A$ or $a_n \to A$.

For example, $\frac{1}{n} \to 0$, since for any $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ for which $N > \frac{1}{\varepsilon}$ (Archimedes' axiom). And if $n > N$, then $\frac{1}{n} < \frac{1}{N} < \varepsilon$, and so $|\frac{1}{n} - 0| < \varepsilon$.

As another example, consider a one meter long rod. If we cut it into half, then halve the half of the rod, then halve one of the pieces again, and so on, then we are led to the sequence

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{2^3}, \dots, \frac{1}{2^n}, \dots.$$

Clearly, the lengths of the remaining pieces of the rod form the sequence $(\frac{1}{2^n}) \to 0$, so the new pieces will be arbitrarily small.

One can see at once that whenever (a_n) is convergent, it is also bounded, since for $\varepsilon := 1$ there exists an index N_1 such that for all $n > N_1$

$$A - 1 < a_n < A + 1,$$

and the finitely many elements $a_1, a_2, \dots, a_N$ cannot spoil the boundedness of the sequence (a_n).

Convergent sequences nicely behave during operations.

Theorem 5.1. *If $a_n \to A$ and $\lambda \in \mathbb{R}$, then $\lambda a_n \to \lambda A$.*
If $a_n \to A$ and $b_n \to B$, then $a_n + b_n \to A + B$, $a_n b_n \to AB$.
If $b_n \to B$ and $B \neq 0$, then $\frac{1}{b_n} \to \frac{1}{B}$.
If $a_n \to A$ and $b_n \to B \neq 0$, then $\frac{a_n}{b_n} \to \frac{A}{B}$.

Let us look at some applications of these theorems.

$$\lim \frac{3n^2 - 2n + 1}{2n^2 + n} = \lim \frac{3 - 2 \cdot \frac{1}{n} + \frac{1}{n^2}}{2 + \frac{1}{n}} = \frac{3}{2},$$

since $\frac{1}{n} \to 0$, therefore $\frac{1}{n^2} = \frac{1}{n} \cdot \frac{1}{n} \to 0$. The denominator is $2 + \frac{1}{n} \to 2 + 0 \neq 0$, so the quotient sequence is convergent.

Further ways of determining whether a sequence is convergent:

Theorem 5.2 (The Sandwich Theorem for Sequences). *If (a_n), (x_n), (y_n) are such that*

1^o *for all* $n \in \mathbb{N}$, $x_n \leq a_n \leq y_n$,

2^o $\lim x_n = \lim y_n =: \alpha$,

then (a_n) *is convergent, and* $\lim a_n = \alpha$.

Theorem 5.3. *If* (a_n) *is monotone and bounded, then* (a_n) *is convergent.*

For example, we have seen that the sequence $(e_n) := \left(\left(\frac{n+1}{n}\right)^n\right)$ is monotonically increasing and bounded, and so it is convergent. Its limit is the number e, known from Exercise 7* of Chapter 3:

$$\lim \left(\frac{n+1}{n}\right)^n = e.$$

For further convergent sequences see the exercises.

The definition of the limit of a sequence involves a serious difficulty: we should conjecture the number $A \in \mathbb{R}$ to which the elements of the sequence get close eventually. This can be avoided with the aid of the following theorem:

Theorem 5.4 (Cauchy's convergence criterion). *The sequence* (a_n) *is convergent if any only if for all* $\varepsilon > 0$ *there exists an index* $N \in \mathbb{N}$ *such that for all* $m, n > N$ $|a_n - a_m| < \varepsilon$.

So, the property that the elements of a sequence get arbitrarily close to a number is equivalent to the property that the elements of the sequence get arbitrarily close to each other.

5.1.3 Series

Now consider the situation that somebody would like to glue the split pieces of the previous one meter long rod, i.e., they would like to prepare the „sum"

$$\frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \ldots + \frac{1}{2^n} + \ldots .$$

Then they would glue the $\frac{1}{2^2}$ meter long piece to the $\frac{1}{2}$ meter long piece, so it will be $\frac{1}{2} + \frac{1}{2^2}$ meter long; then they would glue the $\frac{1}{2^3}$ meter long piece to it, so it will be $\frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3}$ meter long, and so on. More generally: Let (a_n) be a number sequence. Prepare the new sequence

$$S_1 := a_1,\ S_2 := a_1 + a_2,\ S_3 := a_1 + a_2 + a_3, ..., S_n := a_1 + a_2 + ... + a_n, ...$$

The sequence of partial sums (S_n) prepared from the sums of the elements of (a_n) will be called **infinite series** and denoted as $\sum a_n$. Summing the infinite series is a convergent procedure, i.e., the infinite series $\sum a_n$ is **convergent** if the sequence (S_n) is convergent. If the sequence (S_n) is convergent, then the sum of the infinite series $\sum a_n$ is defined as the limit of the sequence (S_n), i.e., $\sum_{n=1}^{\infty} a_n := \lim S_n$.

For example, let $q \in \mathbb{R}$, $0 < q < 1$. Consider the sequence (q^n). Its nth partial sum is

$$S_n = q + q^2 + q^3 + \ldots + q^n = q\frac{q^n - 1}{q - 1}.$$

Since $q^n \to 0$ (see Exercise 3), therefore

$$\lim S_n = \lim q\frac{q^n - 1}{q - 1} = \frac{-q}{q - 1} = \frac{q}{1 - q}.$$

So, the infinite series $\sum q^n$ is convergent, and $\sum_{n=1}^{\infty} q^n = \frac{q}{1-q}$ is the sum of the infinite series.

If $\sum a_n$ is a convergent series, then (S_n) is convergent. Then, according to Cauchy's convergence criterion, for all $\varepsilon > 0$ there exists an index N such that for all $m > N$ and $n := m + 1 > N$

$$\varepsilon > |S_n - S_m| = |a_1 + a_2 + \ldots + a_m + a_{m+1} - (a_1 + a_2 + \ldots + a_m)| = |a_n|.$$

This exactly means that $a_n \to 0$. So the following theorem is valid:

Theorem 5.5. *If $\sum a_n$ is convergent, then $a_n \to 0$.*

The reverse statement is not true. Let $(a_n) := (\ln \frac{n+1}{n})$. Since $\frac{n+1}{n} = 1 + \frac{1}{n} \to 1$, therefore $\ln \frac{n+1}{n} \to \ln 1 = 0$. For all $n \in \mathbb{N}$

$$S_n = \ln\frac{2}{1} + \ln\frac{3}{2} + \ln\frac{4}{3} + \ldots + \ln\frac{n+1}{n} = \ln\frac{2}{1}\cdot\frac{3}{2}\cdot\frac{4}{3}\cdots\frac{n+1}{n} = \ln(n+1).$$

Let $K > 0$ be arbitrary. There exists an $n \in \mathbb{N}$ such that $n + 1 > e^K$. Then $S_n = \ln(n + 1) > \ln e^K = K$, so (S_n) is not bounded, but then it is not convergent, either, so $\sum a_n$ is not convergent.

The infinite series $\sum \frac{1}{n}$ behaves the same way: although $\frac{1}{n} \to 0$, the series $\sum \frac{1}{n}$ is not convergent.

It is possible to decide from the behavior of the summand whether the series is convergent or not.

Theorem 5.6 (Ratio test)**.** *Let (a_n) be a sequence for which there exists a number $0 < q < 1$ and an index N such that for all $n > N$ $|\frac{a_{n+1}}{a_n}| \le q$. Then $\sum a_n$ is convergent.*

Theorem 5.7 (Root test)**.** *Let (a_n) be a sequence for wich there exists a number $0 < q < 1$ and an index N such that for all $n > N$ $\sqrt[n]{|a_n|} \le q$. Then $\sum a_n$ is convergent.*

For example $\sum \frac{2^n}{n!}$ is a convergent series because

$$\frac{a_{n+1}}{a_n} = \frac{\frac{2^{n+1}}{(n+1)!}}{\frac{2^n}{n!}} = \frac{2}{n+1} < \frac{1}{2} \quad \text{if } n > 5.$$

There is an interesting theorem about alternating series.

Theorem 5.8 (Leibniz)**.** *Let (a_n) be a monotonically decreasing sequence of positive numbers for which $a_n \to 0$. Then the alternating series $\sum (-1)^{n+1} a_n$ is convergent.*

For example, the series $\sum (-1)^{n+1} \frac{1}{n}$ is convergent because $(\frac{1}{n})$ is monotonically decreasing, and $\frac{1}{n} \to 0$.

5.2 Exercises

1. Show that $a_n \to A$ if and only if $a_n - A \to 0$.

   Solution: Let $\varepsilon > 0$ be arbitrary. If $a_n \to A$, then there exists an N such that for $n > N$
   $$|a_n - A| < \varepsilon.$$
   Then $|a_n - A - 0| = |a_n - A| < \varepsilon$ also holds, and so $a_n - A \to 0$.

   One can similarly prove the converse statement.

2. Show that $a_n \to 0$ if and only if $|a_n| \to 0$.

   Solution: Let $\varepsilon > 0$. If $a_n \to 0$, then there exists an N such that for $n > N$, $|a_n - 0| = |a_n| < \varepsilon$, but then $||a_n| - 0| = |a_n| < \varepsilon$ is also true, from which $|a_n| \to 0$ follows.

   If $|a_n| \to 0$, then $-|a_n| \to 0$ also holds. Since $-|a_n| \le a_n \le |a_n|$ for all $n \in \mathbb{N}$, therefore due to the Sandwich Theorem $a_n \to 0$.

3. Let $q \in (-1, 1)$. Show that $q^n \to 0$.

   Is the sequence $(\frac{1}{3^n})$, $((\sin \frac{\pi}{4})^n)$, $(\frac{2^n}{3^n+10})$ convergent?

   Solution: If $q = 0$, then $0^n \to 0$. If $q \ne 0$, then $0 < |q| < 1$, therefore there exists an $h > 0$ such that $\frac{1}{|q|} = 1 + h$. Then due to Bernoulli's inequality for all $n \in \mathbb{N}$
   $$\left(\frac{1}{|q|}\right)^n = (1+h)^n \ge 1 + nh > nh,$$
   $$0 < |q|^n < \frac{1}{h} \cdot \frac{1}{n}.$$
   Since $0 \to 0$, $\frac{1}{h} \cdot \frac{1}{n} \to 0$, therefore the sandwiched sequence also tends to 0, that is $|q^n| = |q|^n \to 0$. According to Exercise 2, $q^n \to 0$ is also true.

4. Let $a > 1$. Show that $\frac{n}{a^n} \to 0$.

   Is the sequence $(\frac{n}{2^n})$, $(n \cdot 0.999^n)$ convergent?

   Solution: If $a > 1$, then there exists an $h > 0$ such that $a = 1 + h$. By Exercise 5 of Chapter 3 for all $n \in \mathbb{N},\ n > 1$

$$a^n = (1+h)^n > \binom{n}{2} h^2 = \frac{n(n-1)}{2} h^2.$$

   From this

$$0 < \frac{n}{a^n} < \frac{2}{h^2} \frac{1}{n-1}.$$

   Clearly, $\frac{1}{n-1} \to 0$, and so for the sandwiched sequence $\frac{n}{a^n} \to 0$.

5. Let $a > 1,\ k \in \mathbb{N}$. Show that $\frac{n^k}{a^n} \to 0$.

   Let $(a_n) := \left(\frac{n^{100}}{1{,}001^n}\right)$. Estimate the value of a_1, a_2, a_3, and give the limit $\lim a_n =$?

   Solution: For all $n \in \mathbb{N}$

$$\frac{n^k}{a^n} = \frac{n}{(\sqrt[k]{a})^n} \cdot \frac{n}{(\sqrt[k]{a})^n} \cdots \frac{n}{(\sqrt[k]{a})^n}$$

   Since $\sqrt[k]{a} > 1$, therefore $\frac{n}{(\sqrt[k]{a})^n} \to 0$ according to Exercise 4. Then the product of k sequences tending to 0 also tends to 0, consequently, $\frac{n}{(\sqrt[k]{a})^n} \to 0$.

6. Let $a > 0$. Show that $\frac{a^n}{n!} \to 0$.

   Solution: There exists a $k \in \mathbb{N}$ such that $a < k$. Let $n \in \mathbb{N},\ n > k$. Then

$$\frac{a^n}{n!} = \frac{a}{1} \cdot \frac{a}{2} \cdots \frac{a}{k} \cdot \frac{a}{k+1} \cdots \frac{a}{n-1} \cdot \frac{a}{n}.$$

   Let $\frac{a}{1} \cdot \frac{a}{2} \cdots \frac{a}{k} := L$; $\frac{a}{k+1} < 1, \ldots, \frac{a}{n-1} < 1$, then

$$0 < \frac{a^n}{n!} < L\frac{a}{n}.$$

   Since $\frac{La}{n} \to 0$, therefore for the sandwiched sequence $\frac{a^n}{n!} \to 0$.

7. Show that $\frac{n!}{n^n} \to 0$.

8. Let $a > 0$. Prove that $\sqrt[n]{a} \to 1$.

   Solution: First let $a > 1$. Let $p_n := \sqrt[n]{a} - 1 > 0$ $(n \in \mathbb{N})$. Due to Bernoulli's inequality, for all $n \in \mathbb{N}$

$$a = (1 + p_n)^n > np_n,$$

so

$$0 < p_n < \frac{a}{n}.$$

Since $\frac{a}{n} \to 0$, therefore for the sandwiched sequence $p_n \to 0$, but then, according to Exercise 1 $\sqrt[n]{a} \to 1$.

If $0 < a < 1$, then $\frac{1}{a} > 1$, therefore $\sqrt[n]{\frac{1}{a}} \to 1$, but then for the reciprocal sequence $\sqrt[n]{a} \to 1$ holds, too.

9. $\lim \sqrt[n]{5} = ?$ $\quad \lim \sqrt[n]{2^n + 1000} = ?$ $\quad \lim \sqrt[n]{2^n + 5^n} = ?$

10. Prove that $\sqrt[n]{n} \to 1$.

Is the sequence $(\sqrt[n]{n^2})$, $\left(\sqrt[n]{\frac{1}{n^2}}\right)$ convergent?

Solution: Let $p_n := \sqrt[n]{n} - 1 > 0$ $(n \in \mathbb{N})$. For all $n \in \mathbb{N}$, $n > 1$: $n = (1 + p_n)^n > \binom{n}{2} p_n^2$ according to Exercise 5 of Chapter 3. From this

$$n > \frac{n(n-1)}{2} p_n^2,$$
$$p_n^2 < \frac{2}{n-1},$$
$$0 < p_n < \frac{\sqrt{2}}{\sqrt{n-1}}.$$

It is easy to prove that $\frac{1}{\sqrt{n-1}} \to 0$, so the sandwiched sequence $p_n \to 0$, which is equivalent to the statement $\sqrt[n]{n} \to 1$ (see Exercise 1).

11. Prove that $\frac{1}{\sqrt[n]{n!}} \to 0$.

Solution: For all $n \in \mathbb{N}$ (temporarily assuming that n is even)

$$n! = 1 \cdot 2 \cdots \left(\frac{n}{2} - 1\right) \cdot \frac{n}{2} \cdot \left(\frac{n}{2} + 1\right) \cdots n > 1 \cdot 1 \cdots 1 \cdot \frac{n}{2} \cdot \frac{n}{2} \cdots \frac{n}{2},$$

so $n! > \left(\frac{n}{2}\right)^{\frac{n}{2}}$. From this

$$\sqrt[n]{n!} > \left(\frac{n}{2}\right)^{\frac{1}{2}},$$
$$0 < \frac{1}{\sqrt[n]{n!}} < \frac{\sqrt{2}}{\sqrt{n}}.$$

Since $\frac{1}{\sqrt{n}} \to 0$, therefore for the sandwiched sequence $\frac{1}{\sqrt[n]{n!}} \to 0$.

12. Show that $\sum \frac{1}{n(n+1)}$ is convergent.

Solution: Let $n \in \mathbb{N}$.

$$S_n = \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \ldots + \frac{1}{1\cdot n(n+1)}.$$

Since $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$, therefore

$$S_n = \frac{1}{1} - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \frac{1}{3} - \frac{1}{4} + \ldots + \frac{1}{n} - \frac{1}{n+1} = 1 - \frac{1}{n+1}.$$

$\lim S_n = \lim 1 - \frac{1}{n+1} = 1$, so $\sum \frac{1}{n(n+1)}$ is convergent, and $\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1$.

13. Show that $\sum \frac{1}{n^2}$ is convergent.

Solution: Let $n \in \mathbb{N},\ n > 1$. Then

$$S_n = \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \ldots + \frac{1}{n^2} < 1 + \frac{1}{1\cdot 2} + \frac{2}{2\cdot 3} + \ldots \frac{1}{(n-1)n}$$
$$= 1 + \frac{1}{1} - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \ldots + \frac{1}{n-1} - \frac{1}{n} = 1 + 1 - \frac{1}{n} < 2.$$

So the sequence (S_n) is bounded. On the other hand, for all $n \in \mathbb{N}$ $S_{n+1} = S_n + \frac{1}{(n+1)^2} > S_n$, therefore (S_n) is monotonically increasing. Therefore, S_n is convergent, that is, $\sum \frac{1}{n^2}$ is convergent.

14. Is the infinite series $\sum \frac{1}{n!}$, $\sum \frac{3^n}{n!}$ convergent?

For what values of $x \in \mathbb{R}$ is the infinite series $\sum \frac{x^n}{n!}$ convergent?

Solution: We show that for all $x \in \mathbb{R}$, $\sum \frac{x^n}{n!}$ convergent, since if $x \neq 0$, then

$$\left| \frac{\frac{x^{n+1}}{(n+1)!}}{\frac{x^n}{n!}} \right| = \frac{|x|}{n+1} \leq \frac{1}{2} \quad \text{if } n > [2|x| - 1],$$

therefore by the ratio test $\sum \frac{x^n}{n!}$ is convergent.

15. Is the infinite series $\sum \frac{3^n}{1+3^{2n}}$ convergent?

Solution: By the root test

$$\sqrt[n]{\frac{3^n}{1+3^{2n}}} < \sqrt[n]{\frac{3^n}{3^{2n}}} = \frac{1}{3} \quad (n \in \mathbb{N}),$$

therefore $\sum \frac{3^n}{1+3^{2n}}$ is convergent.

16. Can we decide by the ratio or root test whether $\sum \frac{1}{n^2}$ is convergent or not?

    Solution: No. The reason is that

$$\frac{\frac{1}{(n+1)^2}}{\frac{1}{n^2}} = \left(\frac{n}{n+1}\right)^2 < 1,$$

    but there exists no number $q < 1$ such that for some index N, $\left(\frac{n}{n+1}\right) \leq q$ for all $n > N$.

    According to the root test

$$\sqrt[n]{\frac{1}{n^2}} = \frac{1}{\sqrt[n]{n^2}} < 1$$

    holds, too, but since $\lim \frac{1}{\sqrt[n]{n^2}} = 1$, therefore there exists no number $q < 1$ such that from some index N $\frac{1}{\sqrt[n]{n^2}} < q$ for all $n > N$.

17. Is the infinite series $\sum \frac{\cos(n\pi)}{\sqrt{n}}$ convergent?

    Solution: $\cos(n\pi) = (-1)^n$, $\left(\frac{1}{\sqrt{n}}\right)$ tends to 0 monotonically decreasing, therefore by Leibniz's theorem $\sum \frac{\cos(n\pi)}{\sqrt{n}}$ is convergent.

18. * Prove the following statements:

    a) For all $\alpha, \beta, \gamma \in \mathbb{R}$

$$\lim \left(1 + \frac{\alpha}{n+\beta}\right)^{n+\gamma} = e^{\alpha}.$$

    b) $\sum_{n=0}^{\infty} \frac{1}{n!} = e$.

    c) For all $n \in \mathbb{N}$ there exists a $\vartheta \in (0,1)$ such that

$$e = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!} + \frac{\vartheta}{n!n}.$$

    d) $e \in \mathbb{R} \setminus \mathbb{Q}$.

Chapter 6

Continuity

Continuity is a local property of a function. It means that small changes of a point a result in small changes in the function value $f(a)$. The following topics will be discussed.

- The concept of continuous function
- The relationship between continuity and the operations
- The properties of continuous functions on an interval

6.1 Continuity

6.1.1 The concept and properties of a continuous function

Let $f_1 : \mathbb{R} \to \mathbb{R},\ f_1(x) := x,\ a := 2$. Consider another function $f_2 : \mathbb{R} \to \mathbb{R}$,

$$f_2(x) := \begin{cases} 1 \text{ if } x < 2, \\ 2 \text{ if } x = 2, \\ 3 \text{ if } x > 2 \end{cases} \qquad \text{(Fig. 6.1).}$$

Apparently, the function f_1 is such that when x is close to the point $a := 2$, then the values $f_1(x) = x$ are also close to the value $f_1(2) = 2$. The same does not hold for the function f_2. For any number x that is close to the point $a = 2$ ($x \neq 2$), the function values $f_2(x)$ are far from the value $f_2(2) = 2$ (clearly further than $\frac{1}{2}$). In view of the behavior of the function f_1 we formulate the concept of continuity.

Let $f : \mathbb{R} \rightarrowtail \mathbb{R},\ a \in D(f)$. We say that the function f is **continuous at some point** a, if for any $\varepsilon > 0$ there exists a $\delta > 0$, such that whenever

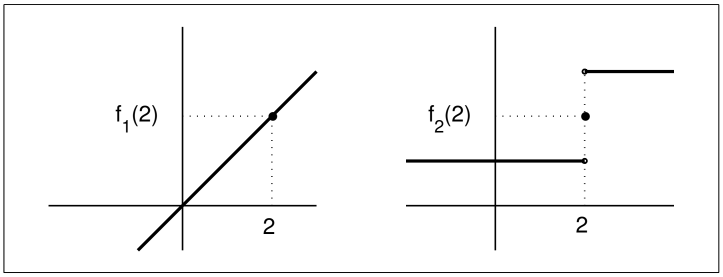


Figure 6.1

$x \in D(f)$ and $|x-a| < \delta$ (the point x is closer to the point a than δ), then $|f(x) - f(a)| < \varepsilon$ (the function value $f(x)$ is closer to $f(a)$ than ε). This property will be denoted as $f \in C[a]$.

Indeed, $f_1 \in C[2]$, since $\forall\varepsilon > 0$ the value $\delta := \varepsilon$ will do because $\forall x \in \mathbb{R}$, $|x-2| < \delta$: $|f_1(x) - f_1(2)| = |x-2| < \varepsilon$. However, $f_2 \notin C[2]$, since by $\varepsilon := \frac{1}{2}$ and $\forall\delta > 0$ there exists an $x \in \mathbb{R}$, for example $x := 2 + \frac{\delta}{2}$ for which $|x-2| = \frac{\delta}{2} < \varepsilon$, but $|f_2(x) - f_2(2)| = |3-2| > \varepsilon$, therefore the function f_2 is not continuous at the point $a := 2$.

A useful property of continuous functions is the **preservation of sign**. It means that if $f \in C[a]$ and $f(a) > 0$, then there exists a neighborhood $K(a) \subset D(f)$ such that $\forall x \in K(a)$ $f(x) > 0$, that is, the sign of $f(a)$ is inherited by function values taken in the neighborhood of a. To prove this property it is enough to consider the definition of continuity for the error bound $\varepsilon := \frac{f(a)}{2} > 0$, since $\exists\delta > 0$ such that $\forall x \in K_\delta(a)$: $f(a) - \varepsilon < f(x) < f(a) + \varepsilon$, that is,

$$0 < \frac{f(a)}{2} = f(a) - \varepsilon < f(x).$$

Continuity is also related to convergent sequences. If $f \in C[a]$ and $(x_n) \subset D(f)$ is an arbitrary sequence such that $x_n \to a$, then $f(x_n) \to f(a)$, that is, the function values taken at the points of the sequence (x_n) tend to $f(a)$. The converse is also true: if $\forall (x_n) \subset D(f),\ x_n \to a$: $f(x_n) \to f(a)$, then f is continuous at the point a. This characterization of continuity is symbolized by the equality $\lim f(x_n) = f(\lim x_n)$.

6.1.2 The relationship between continuity and the operations

Theorem 6.1. *If $f \in C[a]$ and $\lambda \in \mathbb{R}$, then $\lambda f \in C[a]$.*

Theorem 6.2. *If $f, g \in C[a]$, then $f + g \in C[a]$ and $f \cdot g \in C[a]$.*

Theorem 6.3. *If $f, g \in C[a]$ and $g(a) \neq 0$, then $\frac{f}{g} \in C[a]$.*

Theorem 6.4. *If $g \in C[a]$ and $f \in C[g(a)]$, then $f \circ g \in C[a]$.*

Note that the converse statements are not true. For instance, for $f := \operatorname{sgn}$ and $g := -\operatorname{sgn}$ the sum $f+g$ is the constant zero function, for which obviously $f + g = 0 \in C[0]$, however, $f \notin C[0]$ and $g \notin C[0]$.

The inverse function will only be continuous under rather special conditions.

Theorem 6.5. *Let $I \in \mathbb{R}$ be an interval, and $f : I \to \mathbb{R}$ a strictly monotone function. Assume that at the point $a \in I$ $f \in C[a]$. Moreover, let $b := f(a)$. Then $f^{-1} \in C[b]$.*

Let $[a, b] \subset D(f)$. The function f is **continuous on the closed interval** $[a, b]$ if $\forall \alpha \in [a, b]$, $f \in C[\alpha]$. This is denoted by $f \in C[a, b]$.

6.1.3 The properties of continuous functions on intervals

Functions defined on bounded, closed intervals have nice properties.

Theorem 6.6 (Bolzano)**.** *If $f \in C[a, b]$ and $f(a) < 0$, $f(b) > 0$, then $\exists c \in (a, b)$ for which $f(c) = 0$.*

This is a particular case of the following statement also known as Bolzano's theorem.

Theorem 6.7. *Let $f \in C[a, b]$, and d any number between $f(a)$ and $f(b)$. Then $\exists c \in [a, b]$ such that $d = f(c)$.*

This theorem states that if a continuous function defined on an interval takes two values, then it takes any value between these two values as well, that is, the continuous image of an interval is an interval.

Theorem 6.8 (Weierstrass)**.** *If $f \in C[a, b]$, then $\exists \alpha, \beta \in [a, b]$ such that $\forall x \in [a, b]$*

$$f(\alpha) \leq f(x) \leq f(\beta).$$

This theorem states that $f_{|_{[a,b]}}$ is bounded (since all values of the function are between $f(\alpha)$ and $f(\beta)$), moreover, the function $f_{|_{[a,b]}}$ has a **minimum** and a **maximum**.

As a consequence of Bolzano's and Weierstrass' theorems, the continuous image of a closed, bounded interval is a closed, bounded interval.

6.2 Exercises

1. Show that the function $f : [0, +\infty) \to \mathbb{R}$, $f(x) := \sqrt{x}$ is continuous at any point $a \geq 0$.

   Solution: First we show that if $a := 0$, then $f \in C[0]$.

   Let $\varepsilon > 0$ be arbitrary. Then due to $\sqrt{x} < \varepsilon \;\; \Leftrightarrow x < \varepsilon^2$ let $\delta := \varepsilon^2$. If $x \geq 0,\; x < \delta$, then $|f(x) - f(0)| = \sqrt{x} < \varepsilon$.

   Let now $a > 0$. Clearly, $\forall x \geq 0$

   $$|\sqrt{x} - \sqrt{a}| = \frac{|\sqrt{x} - \sqrt{a}| \cdot (\sqrt{x} + \sqrt{a})}{\sqrt{x} + \sqrt{a}} = \frac{|x - a|}{\sqrt{x} + \sqrt{a}} \leq \frac{|x - a|}{\sqrt{a}}.$$

   Let $\varepsilon > 0$ be arbitrary. In view of the previous inequality, $\delta := \varepsilon \cdot \sqrt{a}$. Then $\forall x \geq 0,\; |x - a| < \delta$

   $$|f(x) - f(a)| = |\sqrt{x} - \sqrt{a}| \leq \frac{|x - a|}{\sqrt{a}} < \frac{\delta}{\sqrt{a}} = \varepsilon,$$

   which means that $f \in C[a]$.

2. Show that the function $f : \mathbb{R} \to \mathbb{R},\; f(x) := x^2$ is continuous at any point $a \in \mathbb{R}$.

   Solution: Let $(x_n) \subset \mathbb{R},\; x_n \to a$ be an arbitrary sequence. Then $f(x_n) = (x_n)^2 = x_n \cdot x_n \to a \cdot a = f(a)$.

   Since for any sequence $(x_n) \subset \mathbb{R},\; x_n \to a$: $f(x_n) \to f(a)$, therefore, according to the definition of continuity in terms of limits of sequences $f \in C[a]$.

3. Show that the function $f : \mathbb{R} \to \mathbb{R},\; f(x) := \sin x$ is continuous at any point $a \in \mathbb{R}$.

   Solution: By using the definition of the sine function, one can see the inequality $|\sin x - \sin a| \leq |x - a| \;\forall a, x \in \mathbb{R}$ from Fig. 6.2.

   Let $\varepsilon > 0$ be arbitrary. If $\delta := \varepsilon$, then $\forall x \in \mathbb{R}, |x - a| < \delta$: $|f(x) - f(a)| = |\sin x - \sin a| \leq |x - a| < \varepsilon$, so $f \in C[a]$.

4. Let $f : \mathbb{R} \to \mathbb{R}$,

   $$f(x) := \begin{cases} \frac{\sin x}{x} & \text{if } x \neq 0, \\ 1 & \text{if } x = 0. \end{cases}$$

   Show that $f \in C[0]$.

5. Let $f : \mathbb{R} \to \mathbb{R}$. Suppose that there is a number $L > 0$ such that $\forall s, t \in D(f)$, $|f(s) - f(t)| \leq L|s - t|$. Show that $\forall a \in D(f)$, $f \in C[a]$.

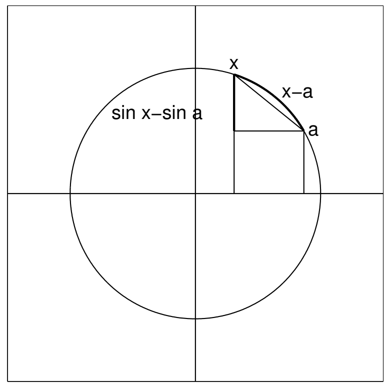


Figure 6.2

6. Show that the equation $x^5 + 4x - 3 = 0$ has a root on the interval $[0, 1]$.
7. Show that if $f \in C[a, b]$, f is a one-to-one function, then f is strictly monotone on the interval $[a, b]$.

Chapter 7

Limit of a function

The limit of a function at some point a is A if $f(x)$ is close to A whenever x is close to a. The following topics will be discussed.

- The concept of limit of a function
- Relationship between the limit and the operations
- Limit at infinity and infinite limit
- One-sided limit
- The limit of monotone functions

7.1 Limit of a function

7.1.1 Finite limit at a finite point

Examine three functions, which look rather similar. Let

$$\begin{array}{ll} f_1 : \mathbb{R} \to \mathbb{R} & f_1(x) := x + 2, \\ f_2 : \mathbb{R} \setminus \{2\} \to \mathbb{R} & f_2(x) := \frac{x^2-4}{x-2} = \frac{(x-2)(x+2)}{x-2} = x + 2, \\ f_3 : \mathbb{R} \to \mathbb{R} & f_3(x) := \begin{cases} x + 2 & \text{if } x \neq 2, \\ 1 & \text{if } x = 2 \end{cases} \qquad \text{Fig. (7.1).} \end{array}$$

We are interested in the behavior of these functions around the point $a := 2$. The function f_1 is continuous at this point, which means that whenever x is close to 2, the values $f_1(x) = x + 2$ are close to 4, which is none other than $f_1(2)$.

Although the function f_2 is not defined in 2, when x is close to 2, the values $f_2(x) = x + 2$ change just a little around 4.

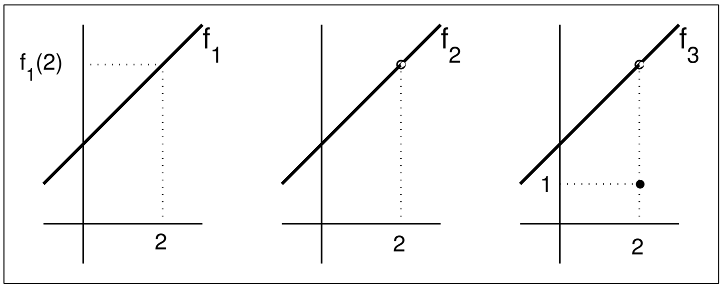


Figure 7.1

The function f_3 is defined in 2. When x is close to 2, (but $x \neq 2$), then the values $f_3(x) = x + 2$ (similarly to the functions f_1 and f_2) change just a little around 4 (independently of the fact that $f(2) = 1$).

We formulate the concept of limit of a function in view of the phenomena experienced in the above examples.

We study functions $f : \mathbb{R} \supset\!\!\to \mathbb{R}$ whose domain of definition $D(f)$ contains points other than a arbitrarily close to a (possibly $a \notin D(f)$).

Definition 7.1. We say that the function f has a limit at the point a if there exists a number $A \in \mathbb{R}$ such that for any error bound $\varepsilon > 0$ there is a distance $\delta > 0$ such that for all points $x \in D(f)$ closer to a than δ ($|x - a| < \delta$), but $x \neq a$, the function values $f(x)$ are closer to A than the error bound ε ($|f(x) - A| < \varepsilon$).

This property of the function f is designated by any of the notations

$$\lim_a f = A;$$
$$\lim_{x \to a} f(x) = A;$$
$$\text{if } x \to a, \text{ then } f(x) \to A.$$

By comparing the limit of a function f with the definition of continuity, one can see that $\lim_a f = A$ exactly means that considering the function

$$\tilde{f} : D(f) \cup \{a\} \to \mathbb{R},\ \tilde{f}(x) := \begin{cases} f(x) & \text{if } x \neq a, \\ A & \text{if } x = a \end{cases}$$

instead of the function f, $\tilde{f}$ will be continuous at the point a. In other words, the function f has a limit at the point a if it can be made continuous at the point a. Therefore, whenever $a \in D(f)$ and there exists $\lim_a f$, the function f is continuous at the point a if and only if $\lim_a f = f(a)$.

From this observation it follows that operations performed with limits originate in operations performed with continuous functions.

Theorem 7.1. *If* $\lim\limits_{a} f = A$ *and* $\lambda \in \mathbb{R}$*, then* $\lim\limits_{a} \lambda f = \lambda A$.

Theorem 7.2. *If* $\lim\limits_{a} f = A$ *and* $\lim\limits_{a} g = B$*, then* $\lim\limits_{a}(f + g) = A + B$.

Theorem 7.3. *If* $\lim\limits_{a} f = A$ *and* $\lim\limits_{a} g = B$*, then* $\lim\limits_{a} f \cdot g = AB$.

Theorem 7.4. *If* $\lim\limits_{a} g = B$ *and* $B \neq 0$*, then* $\lim\limits_{a} \frac{1}{g} = \frac{1}{B}$.

Theorem 7.5. *If* $\lim\limits_{a} f = A$ *and* $\lim\limits_{a} g = B$*,* $B \neq 0$*, then* $\lim\limits_{a} \frac{f}{g} = \frac{A}{B}$.

Theorem 7.6. *If* $\lim\limits_{a} g = B$ *and* $f \in C[b]$*, then* $\lim\limits_{a} f \circ g = f(b)$.

7.1.2 Limit at infinity, infinite limit

One can notice that the concept of limit is based on the tendency of how the function values change. We can extend the so-called "finite limit at a finite point" (that we have considered so far). Let us look at what possibilities we have:

Let $f : \mathbb{R} \supset\!\!\to \mathbb{R}$.

1^o If $D(f)$ is a set that is not bounded above, and there exists an $A \in \mathbb{R}$ such that for any error bound $\varepsilon > 0$ there is a number $\omega \in \mathbb{R}$ such that for all points $x > \omega$, $x \in D(f)$: $|f(x) - A| < \varepsilon$, then we say that the limit of the function f at $(+\infty)$ is A.

Notation: $\lim\limits_{+\infty} f = A$ or $\lim\limits_{x\to+\infty} f(x) = A$ or $f(x) \to A$ when $x \to +\infty$.

For example, $\lim\limits_{x\to+\infty} \frac{1}{x} = 0$.

2^o If $D(f)$ is a set that is not bounded below, and there exists an $A \in \mathbb{R}$ such that for any error bound $\varepsilon > 0$ there is $\omega \in \mathbb{R}$ such that for all points $x < \omega$, $x \in D(f)$: $|f(x) - A| < \varepsilon$, we say that the limit of the function f at $(-\infty)$ is A.

Notation: $\lim\limits_{-\infty} f = A$ or $\lim\limits_{x\to-\infty} f(x) = A$ or $f(x) \to A$ as $x \to -\infty$.

For example, $\lim\limits_{x\to-\infty} \frac{1}{x} = 0$.

3^o If $a \in \mathbb{R}$, moreover, the domain of definition $D(f)$ contains points other than a arbitrarily close to a, and for all $K \in \mathbb{R}$ there exists a $\delta > 0$ such that for all $x \in D(f)$, $x \neq a$, $|x - a| < \delta$: $f(x) > K$, then we say that the limit of f at a is $+\infty$.

Notation: $\lim\limits_a f = +\infty$ or $\lim\limits_{x\to a} f(x) = +\infty$ or $f(x) \to +\infty$ as $x \to a$.

For example, $\lim\limits_{x\to 0} \frac{1}{x^2} = +\infty$.

4^o If $a \in \mathbb{R}$, moreover, the domain of definition $D(f)$ contains points other than a arbitrarily close to a, and for all $K \in \mathbb{R}$ there exists a $\delta > 0$ such that for all $x \in D(f),\ x \neq a,\ |x-a| < \delta$: $f(x) < K$, then we say that the limit of f at a is $-\infty$.

Notation: $\lim\limits_a f = -\infty$ or $\lim\limits_{x\to a} f(x) = -\infty$ or $f(x) \to -\infty$ as $x \to a$.

For example, $\lim\limits_{x\to 0}(-\frac{1}{x^2}) = -\infty$.

5^o If $D(f)$ is not bounded above, and for any number $K \in \mathbb{R}$ there exists an $\omega \in \mathbb{R}$ such that for any point $x > \omega,\ x \in D(f)$: $f(x) > K$, then we say that the limit of the function f at $(+\infty)$ is $+\infty$.

Notation: $\lim\limits_{+\infty} f = +\infty$ or $\lim\limits_{x\to+\infty} f(x) = +\infty$ or $f(x) \to +\infty$ as $x \to +\infty$.

For example, $\lim\limits_{x\to+\infty} x^2 = +\infty$.

6^o If $D(f)$ is not bounded below, and for any number $K \in \mathbb{R}$ there exists an $\omega \in \mathbb{R}$ such that for all points $x < \omega,\ x \in D(f)$: $f(x) > K$, then we say that the limit of the function f at $(-\infty)$ is $+\infty$.

Notation: $\lim\limits_{-\infty} f = +\infty$ or $\lim\limits_{x\to-\infty} f(x) = +\infty$ or $f(x) \to +\infty$ as $x \to -\infty$.

For example, $\lim\limits_{x\to-\infty} x^2 = +\infty$.

7^o If $D(f)$ is not bounded above, and for any number $K \in \mathbb{R}$ there exists an $\omega \in \mathbb{R}$ such that for all points $x > \omega,\ x \in D(f)$: $f(x) < K$, then we say that the limit of f at $(+\infty)$ is $-\infty$.

Notation: $\lim\limits_{+\infty} f = -\infty$ or $\lim\limits_{x\to+\infty} f(x) = -\infty$ or $f(x) \to -\infty$ as $x \to +\infty$.

For example, $\lim\limits_{x\to+\infty}(-x^2) = -\infty$.

8^o If $D(f)$ is not bounded below, and for any number $K \in \mathbb{R}$ there exists an $\omega \in \mathbb{R}$ such that for all points $x < \omega,\ x \in D(f)$: $f(x) < K$, then we say that the limit of f at $(-\infty)$ is $-\infty$.

Notation: $\lim\limits_{-\infty} f = -\infty$ or $\lim\limits_{x\to-\infty} f(x) = -\infty$ or $f(x) \to -\infty$ as $x \to -\infty$.

For example, $\lim\limits_{x\to-\infty}(-x^2) = -\infty$.

Sequences are functions whose domain of definition is $\mathbb{N}$. The set $\mathbb{N}$ is not bounded above, therefore a function $a : \mathbb{N} \to \mathbb{N}$, that is, a sequence (a_n) may have a limit at $(+\infty)$. Comparing the definitions of $a_n \to A$, $a_n \to +\infty$ or $a_n \to -\infty$ with the definition of the limit of a function at $(+\infty)$, we obtain that

$$\begin{aligned}
\lim a_n = A &\iff \lim_{+\infty} a = A, \\
\lim a_n = +\infty &\iff \lim_{+\infty} a = +\infty, \\
\lim a_n = -\infty &\iff \lim_{-\infty} a = -\infty.
\end{aligned}$$

7.1.3 One-sided limit

It can happen that there are points in the domain of definition arbitrarily close to the point a both on the left and on the right side of a, however, the function f has no limit at a. Sometimes we can say something about the behavior of the function even in such a case.

9^o If the point $a \in \mathbb{R}$ is such that there are points $x \in D(f)$, $x > a$ arbitrarily close to a, and there exists a number $A \in \mathbb{R}$ such that for any error bound $\varepsilon > 0$ there exists $\delta > 0$ such that for all points $x \in D(f)$, $a < x < a+\delta$: $|f(x)-A| < \varepsilon$, then we say that the right-side limit of f at a is A.

Notation: $\lim\limits_{a+0} f = A$ or $\lim\limits_{x\to a+0} f(x) = A$. Sometimes $f(a+0)$ denotes the right-side limit of the function f at a. [Traditionally, in the case $a = 0$, instead of "$0 + 0$" we simply write "$0+$" everywhere.

For example the function

$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) := \begin{cases} 1 & \text{if } x \geq 0, \\ -1 & \text{if } x < 0 \end{cases}$$

has no limit at 0, however, $\lim_{x\to 0+} f(x) = 1$ or $f(0+) = 1$.]

10^o If $a \in \mathbb{R}$ is such that there are points $x \in D(f)$, $x > a$ arbitrarily close to a, and for any number $K \in \mathbb{R}$ there exists a $\delta > 0$ such that for all points $x \in D(f)$, $a < x < a + \delta$: $f(x) > K$, then we say that the right-side limit of f at a is $+\infty$.

Notation: $\lim\limits_{a+0} f = +\infty$ or $\lim\limits_{x\to a+0} f(x) = +\infty$.

For example, the limit $\lim\limits_{x\to 0} \frac{1}{x}$ does not exist, but $\lim_{x\to 0+} \frac{1}{x} = +\infty$.

11^o If $a \in \mathbb{R}$ is such that there are points $x \in D(f)$, $x > a$ arbitrarily close to a, and for any number $K \in \mathbb{R}$ there exists a $\delta > 0$ such that for all points $x \in D(f),\ a < x < a + \delta$: $f(x) < K$, then we say that the right-side limit of f at a is $-\infty$.

Notation: $\lim\limits_{a+0} f = -\infty$ or $\lim\limits_{x \to a+0} f(x) = -\infty$.

For example, the limit $\lim\limits_{x \to 0}(-\frac{1}{x})$ does not exist, but $\lim_{x \to 0+}(-\frac{1}{x}) = -\infty$.

12^o If $a \in \mathbb{R}$ is such that there are points $x \in D(f),\ x < a$ arbitrarily close to a, and there exists an $A \in \mathbb{R}$ such that for any error bound $\varepsilon > 0$ there exists a $\delta > 0$ such that for all points $x \in D(f),\ a - \delta < x < a$: $|f(x) - A| < \varepsilon$, then we say that the left-side limit of f at a is A.

Notation: $\lim\limits_{a-0} f = A$ or $\lim\limits_{x \to a-0} f(x) = A$. Sometimes $f(a - 0)$ denotes the left-side limit of f at a. [Traditionally, in case $a = 0$, instead of "$0 - 0$" we write "$0-$" everywhere.

For example, in the example after definition 9^o $\lim\limits_{x \to 0-} f(x) = -1$ or $f(0-) = -1$.]

13^o If $a \in \mathbb{R}$ is such that there are points $x \in D(f),\ x < a$ arbitrarily close to a, and for any number $K \in \mathbb{R}$ there exists a $\delta > 0$ such that for all points $x \in D(f),\ a - \delta < x < a$: $f(x) > K$, then we say that the left-sided limit of f at a is $+\infty$.

Notation: $\lim\limits_{a-0} f = +\infty$ or $\lim\limits_{x \to a-0} f(x) = +\infty$.

For example $\lim\limits_{x \to 0-}(-\frac{1}{x}) = +\infty$.

14^o If $a \in \mathbb{R}$ is such that there are points $x \in D(f),\ x < a$ arbitrarily close to a, and for any number $K \in \mathbb{R}$ there exists a $\delta > 0$ such that for all points $x \in D(f),\ a - \delta < x < a$: $f(x) < K$, then we say that the left-sided limit of f at a is $-\infty$.

Notation: $\lim\limits_{a-0} f = -\infty$ or $\lim\limits_{x \to a-0} f(x) = -\infty$.

For example $\lim_{x \to 0-} \frac{1}{x} = -\infty$.

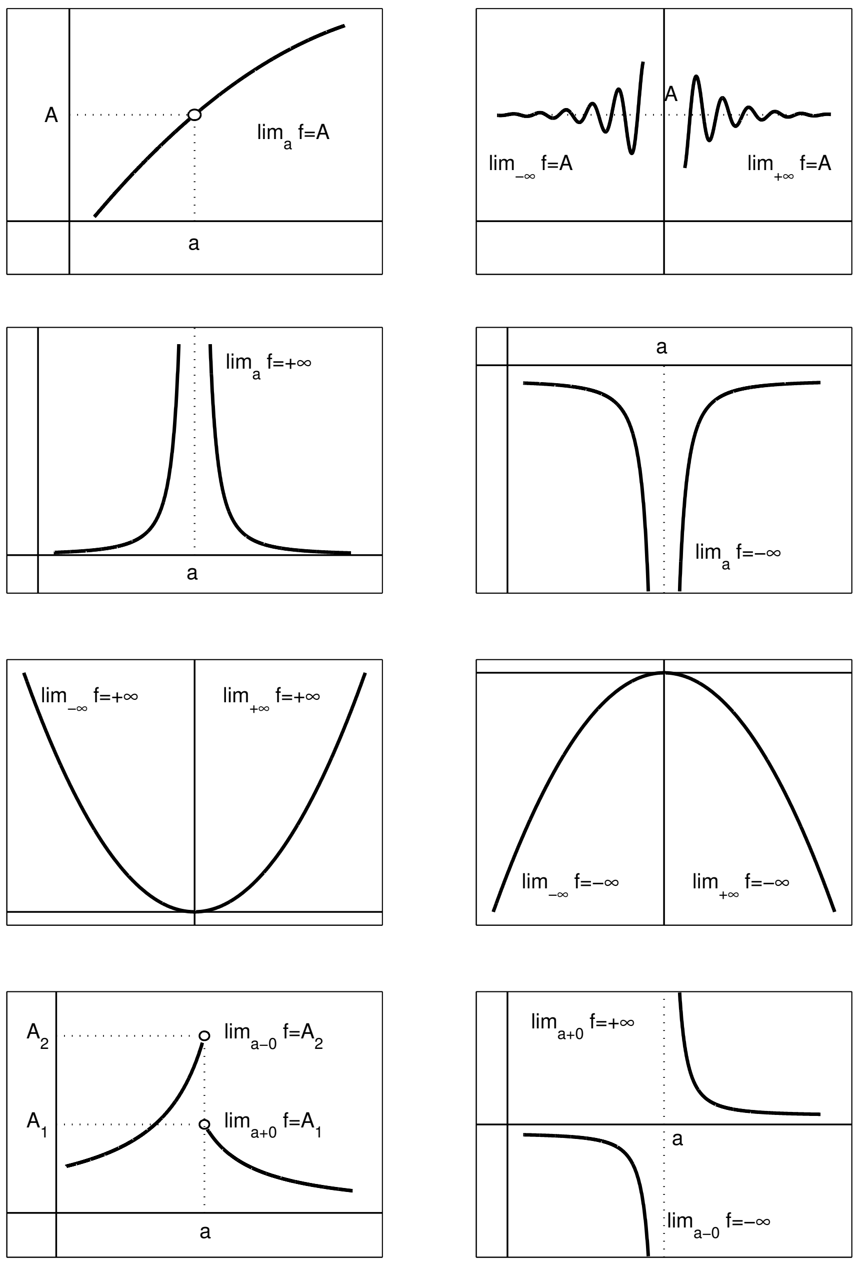
The icons in Fig. 7.2 summarize the different cases of limits.

We can also formulate the relationship between the one-sided limits and the limit:

If there exists $\lim\limits_{a-0} f$ as well as $\lim\limits_{a+0} f$, and $\lim\limits_{a-0} f = \lim\limits_{a+0} f$, then the function f has a limit at a, and

$$\lim_{a} f = \lim_{a-0} f = \lim_{a+0} f.$$

Note that whenever $a \in \mathbb{R}$ is a point where only the right-sided or only the left-sided limit may exist at the point a and it does exist, then the limit of f at a will be this very one-sided limit.



A
lim_a f=A
a
A
lim_−∞ f=A
lim_+∞ f=A
lim_a f=+∞
a
a
lim_a f=−∞
lim_−∞ f=+∞
lim_+∞ f=+∞
lim_−∞ f=−∞
lim_+∞ f=−∞
A_2
lim_a−0 f=A_2
A_1
lim_a+0 f=A_1
a
lim_a+0 f=+∞
a
lim_a−0 f=−∞

Figure 7.2

7.2 Exercises

1. $\lim_{x\to 2} \frac{2x^2-x-6}{x^2-x-2} = ?$ $\qquad \lim_{x\to\infty} \frac{2x^2-x-6}{x^2-x-2} = ?$
2. $\lim_{x\to 1} \frac{x^4-2x^2-3}{x^2-3x+2} = ?$ $\qquad \lim_{x\to 2-0} \frac{x^4-2x^2-3}{x^2-3x+2} = ?$ $\qquad \lim_{x\to 2+0} \frac{x^4-2x^2-3}{x^2-3x+2} = ?$
3. $\lim_{x\to 1} \left(\frac{3}{1-x^3} - \frac{2}{1-x^2}\right) = ?$
4. $\lim_{x\to 0} \frac{\sin 3x}{x} = ?$ $\qquad \lim_{x\to 0} \frac{\sin 3x}{\sin 5x} = ?$ $\qquad \lim_{x\to 0} \frac{\mathrm{tg} 2x}{x} = ?$
5. $\lim_{x\to 0} \frac{1-\cos x}{x^2} = ?$ $\qquad \lim_{x\to 0} \frac{\mathrm{tg} x - \sin x}{x^3} = ?$
6. $\lim_{x\to 0} \frac{e^{2x}-1}{x} = ?$ $\qquad \lim_{x\to 0} \frac{2^x-1}{x} = ?$
7. $\lim_{x\to 0} \frac{\mathrm{sh}(x+2)}{\mathrm{sh}(x-2)} = ?$
8. $\lim_{x\to+\infty} \sqrt{x^2+2} - \sqrt{x^2+2x-3} = ?$
   $\lim_{x\to-\infty} \sqrt{x^2+2} - \sqrt{x^2+2x-3} = ?$
9. Is there a number $k \in \mathbb{R}$ for which the limit
   $$\lim_{x\to 3} \frac{x^3 - 9x^2 + kx - 27}{x^2 - 6x + 9}$$
   exists and is real?

Chapter 8

Differentiability

Differentiability means the smoothness of a function. A differentiable function is continuous, and its graph cannot contain any breaks or spires. The following topics will be discussed.

- The concept of derivative and its geometric meaning
- Derivatives of elementary functions
- Differentiation rules
- Monotonicity and extreme values
- Convexity and inflection
- Function analysis
- Taylor polynomial
- L'Hospital's rule

8.1 Differentiability

8.1.1 The concept of derivative and its geometric meaning

Let us examine two simple functions: $f_1 : \mathbb{R} \to \mathbb{R}$, $f_1(t) := t^2$ and $f_2 : \mathbb{R} \to \mathbb{R}$, $f_2(t) := |t|$. Let us fix the point $x := 0$. As one can easily check, f_1 and f_2 are even; bounded below and not bounded above; increasing on the set of positive numbers, decreasing on the set of negative numbers; attain their minima at the point $x = 0$ with minimum value 0; and both are continuous at the point $x = 0$.

In spite of the many similarities, it is apparent that the function f_1 is smooth at the point $x = 0$, while the function f_2 breaks at this point.

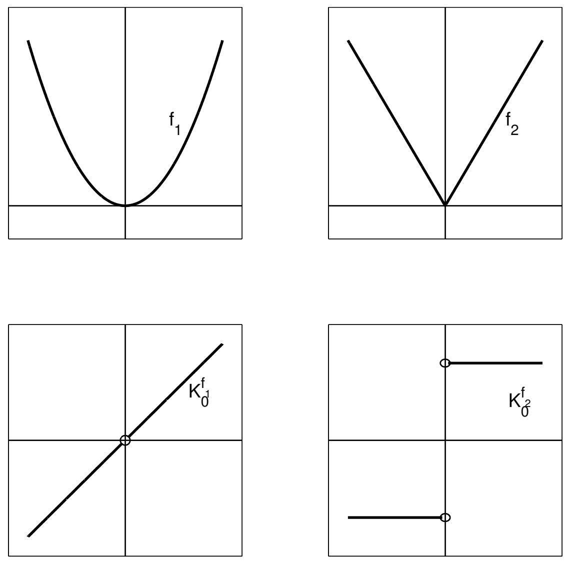


Figure 8.1

Is there a "device" which would be able to detect whether a function is smooth at a point, while another one is not smooth?

Let $f : \mathbb{R} \supset\!\!\to \mathbb{R}$ be an arbitrary function, $x \in D(f)$ a given point. We introduce the difference quotient of f belonging to x as the function

$$K_x^f : D(f) \setminus \{x\} \to \mathbb{R}, \quad K_x^f(t) := \frac{f(t) - f(x)}{t - x}.$$

Apply this "device" to the functions f_1 and f_2 at the point $x := 0$ (Fig. 8.1).

As one can see, in the case of the smooth function f_1, the difference quotient function $K_0^{f_1}$ has a limit (it can be made continuous) at 0, while for the function f_2 with the break the difference quotient function $K_0^{f_2}$ has no limit at 0.

This investigation motivates that a function whose difference quotient has a limit at the point it belongs to be called **differentiable** at that point. We will denote this property as $f \in D[x]$.

If $f \in D[x]$, then the limit of the difference quotient function of f is called **derivative** of the function f at x:

$$\lim_{t \to x} \frac{f(t) - f(x)}{t - x} =: f'(x).$$

It is easy to show that $t - x \to 0$ in case $t \to x$, however, $\frac{f(t)-f(x)}{t-x}$ does not go to infinity, which can only be the case if $f(t) - f(x) \to 0$, so if f is differentiable at x, then it must be continuous at x, too.

How did we create our "device", which is able to detect the smoothness of a function? First we will show a geometric approach. Let $f \in D[x]$. Draw a straight line (a so-called secant) through two different points, $(x, f(x))$ and $(t, f(t))$ of the coordinate system. The slope of this line is

$$\frac{f(t) - f(x)}{t - x}.$$

[That is what we denoted as $K_x^f(t)$.]

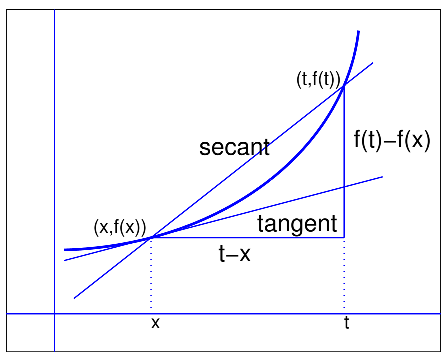


Figure 8.2

If t tends to x, then the secants tend to a limit position called **tangent**, so the slopes of the secants tend to the slope of the tangent (Fig. 8.2). [It is this limit value that we called derivative.]

The other one is a physical interpretation. Assume that the movement of a point is described by the position-time function $t \mapsto s(t)$. The average velocity during the time interval $[t_0, t]$ is the quotient of the displacement $s(t) - s(t_0)$ and the total time $t - t_0$, that is

$$\frac{s(t) - s(t_0)}{t - t_0}.$$

[This quotient is often denoted by $\frac{\Delta s}{\Delta t}$.] If we shorten the time interval "beyond all bounds", the average velocity will be arbitrarily close to a number (by the assumption that the position-time function is smooth), and this number is called instantaneous velocity:

$$\lim_{t\to t_0}\frac{s(t)-s(t_0)}{t-t_0} =: v(t_0) \quad \text{or} \quad \lim_{\Delta t\to 0}\frac{\Delta s}{\Delta t} = v.$$

[One can see that the instantaneous velocity is the limit of the average velocity and the derivative of the position-time function: $s'(t_0) = v(t_0)$.]

The function $f : \mathbb{R} \to \mathbb{R},\ f(t) := t^2$ does not only look smooth at the point $x := 0$. Let $x \in \mathbb{R}$ be an arbitrary real number. Let us examine whether the difference quotient of the function f has a limit at x.

$$\lim_{t\to x}\frac{f(t)-f(x)}{t-x} = \lim_{t\to x}\frac{t^2-x^2}{t-x} = \lim_{t\to x}\frac{(t-x)(t+x)}{t-x} = \lim_{t\to x}(t+x) = 2x.$$

So, $f \in D[x]$ and $f'(x) = 2x$.

The function which gives the difference quotient of f at each x (where the function is differentiable) is called derivative of the function f, and is denoted by f'. In our example $f' : \mathbb{R} \to \mathbb{R},\ f'(x) = 2x$.

The function $f : \mathbb{R} \to \mathbb{R},\ f(t) := t^2$ is often mentioned as the x^2 function, and its derivative is denoted by $(x^2)'$. With this convention in mind we can write

$$(x^2)' = 2x.$$

8.1.2 Derivatives of the elementary functions, differentiation rules

Let us look at some further examples. Let $f : \mathbb{R} \to \mathbb{R}$, $f(t) := t^3$, $x \in \mathbb{R}$.

$$\begin{aligned}\lim_{t\to x}\frac{f(t)-f(x)}{t-x} &= \lim_{t\to x}\frac{t^3-x^3}{t-x}\\ &= \lim_{t\to x}\frac{(t-x)(t^2+tx+x^2)}{t-x} = \lim_{t\to x}(t^2+tx+x^2) = 3x^2,\end{aligned}$$

thus $f \in D[x]$ and $f'(x) = 3x^2$, or briefly $(x^3)' = 3x^2$.

Let $f : \mathbb{R} \to \mathbb{R},\ f(t) := \sin t,\ x \in \mathbb{R}$.

$$\begin{aligned}\lim_{t\to x}\frac{f(t)-f(x)}{t-x} &= \lim_{t\to x}\frac{\sin t-\sin x}{t-x} = \lim_{t\to x}\frac{2\sin\frac{t-x}{2}\cos\frac{t+x}{2}}{t-x}\\ &= \lim_{t\to x}\left(\frac{\sin\frac{t-x}{2}}{\frac{t-x}{2}}\cos\frac{t+x}{2}\right) = 1\cdot\cos x = \cos x.\end{aligned}$$

(During the transformation we used a consequence of the trigonometric addition formulas. Since $\lim_{u\to 0}\frac{\sin u}{u} = 1$, therefore in case $t \to x$, $u := \frac{t-x}{2} \to 0$, so $\lim_{t\to x} \frac{\sin\frac{t-x}{2}}{\frac{t-x}{2}} = 1$.) So, $f \in D[x]$, that is, the sine function is differentiable at all $x \in \mathbb{R}$, and $f'(x) = \cos x$, or briefly $(\sin x)' = \cos x$.

One can show the differentiability of several other functions in a similar manner, and as a result of the calculations we obtain the derivatives as well.

In the following summary we have collected the derivatives of some important functions:

$$
\begin{aligned}
(x^\alpha)' &= \alpha x^{\alpha-1}, \quad \alpha \in \mathbb{R}, & (\ln x)' &= \frac{1}{x},\\
(\sin x)' &= \cos x, & (\log_a x)' &= \frac{1}{x\ln a} \quad (a>0,\ a\neq 1),\\
(\cos x)' &= -\sin x, & (\arcsin\, x)' &= \frac{1}{\sqrt{1-x^2}},\\
(e^x)' &= e^x, & (\arccos\, x)' &= -\frac{1}{\sqrt{1-x^2}},\\
(a^x)' &= a^x \ln a \quad (a>0), & (\mathrm{arctg}\ x)' &= \frac{1}{1+x^2},\\
(\mathrm{tg}\ x)' &= \frac{1}{\cos^2 x}, & (\mathrm{arsh}\ x)' &= \frac{1}{\sqrt{x^2+1}},\\
(\mathrm{ctg}\ x)' &= -\frac{1}{\sin^2 x}, & (\mathrm{arch}\ x)' &= \frac{1}{\sqrt{x^2-1}} \quad (x>1),\\
(\mathrm{sh}\ x)' &= \mathrm{ch}\ x, & (\mathrm{arth}\ x)' &= \frac{1}{1-x^2} \quad (-1<x<1),\\
(\mathrm{ch}\ x)' &= \mathrm{sh}\ x, & (\mathrm{arcth}\ x)' &= \frac{1}{1-x^2} \quad (|x|>1),\\
(\mathrm{th}\ x)' &= \frac{1}{\mathrm{ch}^2 x}, & &\\
(\mathrm{cth}\ x)' &= -\frac{1}{\mathrm{sh}^2 x}. & &
\end{aligned}
$$

During operations with differentiable functions we often obtain a differentiable function. For example, if $f, g \in D[x]$, then

$$
\begin{aligned}
\lim_{t\to x} \frac{(f+g)(t)-(f+g)(x)}{t-x} &= \lim_{t\to x} \frac{f(t)-f(x)+g(t)-g(x)}{t-x}\\
&= \lim_{t\to x} \frac{f(t)-f(x)}{t-x} + \lim_{t\to x} \frac{g(t)-g(x)}{t-x} = f'(x)+g'(x).
\end{aligned}
$$

Consequently, the function $f+g$ is also differentiable at the point x, and $(f+g)'(x) = f'(x) + g'(x)$.

One can verify the following theorems in a similar manner:

Theorem 8.1. *If $f \in D[x]$ and $\lambda \in \mathbb{R}$, then $\lambda f \in D[x]$ and $(\lambda f)'(x) = \lambda f'(x)$.*

Theorem 8.2. *If $f, g \in D[x]$, then $f + g \in D[x]$ and $(f+g)'(x) = f'(x) + g'(x)$, moreover, $f \cdot g \in D[x]$ and $(f \cdot g)'(x) = f'(x)g(x) + f(x)g'(x)$.*

Theorem 8.3. *If $g \in D[x]$ and $g(x) \neq 0$, then $\frac{1}{g} \in D[x]$ and $\left(\frac{1}{g}\right)'(x) = -\frac{g'(x)}{g^2(x)}$.*

Theorem 8.4. *If $f, g \in D[x]$ and $g(x) \neq 0$, then $\frac{f}{g} \in D[x]$ and*

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

Theorem 8.5. *If $g \in D[x]$ and $f \in D[g(x)]$, then $f \circ g \in D[x]$ and $(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$.*

Theorem 8.6. *If $f \in D[x]$, $f'(x) \neq 0$, and the inverse function f^{-1} exists, then by using the notation $u := f(x)$ we have $f^{-1} \in D[u]$, and $(f^{-1})'(u) = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(u))}$.*

8.1.3 The relationship between the derivative and the properties of the function

How do we benefit form the fact that a function is differentiable (smooth) and we know its derivative?

a) Let $f \in D[x]$. This means that whenever t is close to x, $\frac{f(t)-f(x)}{t-x}$ is close to $f'(x)$. From this follows a further expressive and useful meaning of the differentiability. Namely, if $t \approx x$, then

$$\frac{f(t)-f(x)}{t-x} \approx f'(x), \text{ which implies } f(t) - f(x) \approx f'(x)(t-x) \text{ or}$$

$$f(t) \approx f(x) + f'(x)(t-x).$$

This means that at points t that are close to x the function values can be approximated by values of a first degree polynomial (straight line). The function $e_x(t) := f(x) + f'(x)(t-x)$ $(t \in \mathbb{R})$ is the tangent of f at the point $(x, f(x))$.

b) The sign of the derivative informs us about the increase of the function.

Let $f \in D[x]$ and $f'(x) > 0$. Then

$$\frac{f(t)-f(x)}{t-x} \approx f'(x) \text{ if } t \approx x.$$

Since $f'(x) > 0$, therefore $\frac{f(t)-f(x)}{t-x} > 0$ if $t \approx x$. This means that if $t_1 < x$, then $f(t_1) < f(x)$, and if $t_2 > x$, then $f(t_2) > f(x)$. So, for any points t_1, t_2 for which t_1 and t_2 are both close to x and $t_1 < x < t_2$, $f(t_1) < f(x) < f(t_2)$. One can also show that if $f'(x) > 0$ at every point x of an interval I, then the function f is strictly monotonically increasing on the interval I, that is, for all $x_1, x_2 \in I,\ x_1 < x_2$: $f(x_1) < f(x_2)$.

c) The derivative can also be applied for seeking local extreme values. A function f has a **local minimum** at the point $a \in D(f)$ if there exists an interval around a such that for any $x \in D(f)$, chosen from this interval, $f(x) \geq f(a)$.

   If $f \in D[a]$, and the function f has a minimum at the point a, then $f'(a) = 0$. Should $f'(a) \neq 0$, for example $f'(a) > 0$ hold, then there should be two points $t_1 < a < t_2$ close to a for which $f(t_1) < f(a) < f(t_2)$, which contradicts the fact that f has a local minimum at a.

   Consequently, a function that is differentiable at all points of an open interval may have a local extreme value only at a point where its derivative is zero.

   Attention! If $f \in D[a]$ and $f'(a) = 0$, then it is possible that there is no extreme value at a. For example, in case $f : \mathbb{R} \to \mathbb{R},\ f(t) := t^3$: $(t^3)' = 3t^2$, therefore $f'(0) = 3 \cdot 0^2 = 0$, but the function f has no local extreme value at 0.

d) The derivative also tells us about the shape of a function. The function f is called **convex** on the interval I if for all $x_1, x_2 \in I, x_1 < x_2$ the graph of the function is below the chord connecting the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$ on the interval $[x_1, x_2]$.

   It can be verified that a differentiable function f is convex on the interval I if and only if its derivative f' is monotonically increasing on this interval.

   We can decide if f' is monotonically increasing by examining the sign of its derivative. If f' is differentiable, then by introducing the second derivative $f'' := (f')'$, we are led to the theorem: if $f''(x) > 0$ $(x \in I)$, then f is convex on I.

   (Similarly, we can define the concept of a **concave** function, and we can obtain a similar sufficient condition for such a function as well.)

   A point $a \in D(f)$ is called **inflection point** of the function f if the shape of the function is different on intervals before and after this point (that is, the function changes from convex to concave or from concave to convex at this point). For example, the function $f(x) = x^3$ has an inflection at 0.

   One can show that if $a \in D(f)$ is an inflection point of the twice differentiable function f, then $f''(a) = 0$.

   Attention! If f is twice differentiable at the point a, and $f''(a) = 0$, then it may not have an inflection at a. For example, in case of the function

$f : \mathbb{R} \to \mathbb{R},\ f(t) := t^4$ the second derivative is $f''(t) = 12t^2$, and so $f''(0) = 0$, but the function f is convex on the whole interval $\mathbb{R}$ (and not concave on any sub-interval).

e) How can we use our previous results for sketching the curve of a function? It is advisable to follow the steps on the example of Exercise 3.

1. Prepare the derivative function f'.
2. Find the zeros of f' (or the points where f' may change sign).
3. Calculate the second derivative f''.
4. Find the zeros of f'' (or the points where f'' may change sign).
5. The domain of definition of the function is slashed into open intervals by the zeros (or the points of possible sign changes) of f' and f''. We examine the signs of the derivatives on these intervals, from which we can determine the monotonicity and shape properties of the function. (It is helpful to prepare a table for the function analysis.)
6. We calculate some important characteristics: the values of the local minima and maxima, if they exist, the limits (possibly one-sided ones, too) of the function at each point where the function is not defined, but there may be a limit.
7. We sketch the curve of the function.

8.1.4 Multiple derivatives and the Taylor polynomial

We have seen the roles of the first and second derivatives of a function. As a generalization, let us introduce the higher-order derivatives.

If f' is differentiable, then $f'' := (f')'$.

If f'' is differentiable, then $f''' := (f'')'$.

$\vdots$

If $f^{(k)}$ is differentiable, then $f^{(k+1)} := (f^{(k)})',\ k = 1, 2, \ldots$.

We remark that the apostrophes are only used for denoting the first three derivatives, so $f^{(1)} := f'$, $f^{(2)} := f''$, $f^{(3)} := f'''$. Sometimes it is useful to follow the convention that $f^{(0)} := f$.

"Sufficiently smooth" functions can be well approximated with polynomials. We have already seen that if $f \in D[a]$, then for the tangent function

$$e_a(t) := f(a) + f'(a)(t - a) \quad (t \in \mathbb{R})$$

we have

$$e_a(a) = f(a);$$

$e_a'(t) = f'(a)$, therefore $e_a'(a) = f'(a)$, so the derivative of e_a and that of f are equal at a.

One can also see that

$$\lim_{t\to a}\frac{f(t)-e_a(t)}{t-a}=\lim_{t\to a}\frac{f(t)-(f(a)+f'(a)(t-a))}{t-a}$$
$$=\lim_{t\to a}\frac{f(t)-f(a)}{t-a}-f'(a)=0,$$

which expresses the fact that the tangent function e_a is such an approximation to the function f that even if the difference $f(t)-e_a(t)$ is magnified by dividing it by $(t-a)$, the obtained quotient will be close to 0 whenever t is close to a.

The tangent function e_a is just a first degree polynomial. What should that higher-order polynomial look like which would provide an even better approximation?

Let $P(x):=3-2x+4x^2-5x^3$. Then $P(0)=3$.

$$\begin{aligned} P'(x)&=-2+8x-15x^2, & P'(0)&=-2;\\ P''(x)&=8-30x, & P''(0)&=8;\\ P'''(x)&=-30, & P'''(0)&=-30. \end{aligned}$$

One can easily check that for all $x\in\mathbb{R}$

$$P(x)=P(0)+P'(0)x+\frac{P''(0)}{2!}x^2+\frac{P'''(0)}{3!}x^3,$$

that is, a polynomial was fairly well (in this case exactly) approximated with the aid of a polynomial, the coefficients of which are the higher-order derivatives of the function at a point (now this point was 0), divided by the factorial of the order of the derivative in each term.

The above two observations lead us to the so-called Taylor formula. Assume that f is so "smooth" that even the derivative function $f^{(n+1)}$ is continuous in a neighborhood $K(a)\subset D(f)$ of $a\in D(f)$. Let $T_n:\mathbb{R}\to\mathbb{R}$,

$$T_n(x):=f(a)+f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2+\ldots+\frac{f^{(n)}(a)}{n!}(x-a)^n$$

the so-called nth Taylor polynomial. (One can see that $T_1=e_a$.) It is easy to check that $T_n(a)=f(a)$, $T_n'(a)=f'(a)$, $T_n''(a)=f''(a),\ldots,T_n^{(n)}(a)=f^{(n)}(a)$ (that is, the Taylor polynomial T_n possesses the same property as the tangent function e_a.) It can be verified that under such a condition for all $x\in K(a)$ there exists a point c between x and a such that

$$f(x)=T_n(x)+\frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1},$$

which means that the function f is so well approximated by the Taylor polynomial that

$$\frac{f(x)-T_n(x)}{(x-a)^n}=\frac{f^{(n+1)}(c)}{(n+1)!}(x-a)\approx 0 \quad \text{if } x\approx a.$$

So, the Taylor polynomial approximates the function f fairly well (in the order n); that is, the values of f at points close to a can be approximated very accurately by function values of a polynomial.

8.1.5 L'Hospital's rule

With the aid of the derivatives we can calculate some seemingly complicated function limits. One of L'Hospital's rules says that if f and g are differentiable on an open interval I and at the point a (which can either be an element or one of the end points of the interval, and can even be $+\infty$ or $-\infty$), and

$$\lim_{x\to a} f(x)=\lim_{x\to a} g(x)=0,$$

however, the quotient of the derivatives has a limit

$$\lim_{x\to a}\frac{f'(x)}{g'(x)}=: L,$$

then the quotient of f and g has a limit as well, and

$$\lim_{x\to a}\frac{f(x)}{g(x)}=L.$$

The same holds when at the point a f and g tends to $+\infty$ or $-\infty$ instead of 0 [the two infinite limits may have opposite signs, too].

Apply L'Hospital's rule to calculate the limit $\lim\limits_{x\to 0}\dfrac{\cos x-\cos 3x}{x^2}$.

Both the numerator and the denominator are 0 at 0, therefore it is enough to calculate the limit of the quotient of the derivatives.

$$\begin{aligned}\lim_{x\to 0}\frac{(\cos x-\cos 3x)'}{(x^2)'}&=\lim_{x\to 0}\frac{-\sin x+3\sin 3x}{2x}=-\frac{1}{2}\lim_{x\to 0}\frac{\sin x}{x}+\frac{3}{2}\lim_{x\to 0}\frac{\sin 3x}{x}\\&=-\frac{1}{2}\cdot 1+\frac{9}{2}\lim_{x\to 0}\frac{\sin 3x}{3x}=-\frac{1}{2}+\frac{9}{2}=4.\end{aligned}$$

In this manner,

$$\lim_{x\to 0}\frac{\cos x-\cos 3x}{x^2}=4.$$

[The limit of the quotient of the derivatives could have also been calculated by using L'Hospital's rule:

$$\lim_{x\to 0}\frac{-\sin x+3\sin 3x}{2x}=\lim_{x\to 0}\frac{-\cos x+9\cos 3x}{2}=\frac{-1+9}{2}=4.]$$

Unfortunately, not even L'Hospital's rule is able to give a simple answer to every "critical" limit value problem. For example,

$$\lim_{x\to\infty}\operatorname{sh}(x+2)=\lim_{x\to\infty}\operatorname{sh}(x-2)=+\infty.$$

If we consider the derivatives, then

$$\lim_{x\to\infty}\operatorname{ch}(x+2)=\lim_{x\to\infty}\operatorname{ch}(x-2)=+\infty,$$

then the derivatives of these will read as

$$\lim_{x\to\infty}\operatorname{sh}(x+2)=\lim_{x\to\infty}\operatorname{sh}(x-2)=+\infty,$$

and so on and so on. We will never obtain the limit $\lim\limits_{x\to\infty}\dfrac{\operatorname{sh}(x+2)}{\operatorname{sh}(x-2)}$ by applying L'Hospital's rule. [We remark that

$$\lim_{x\to\infty}\frac{\operatorname{sh}(x+2)}{\operatorname{sh}(x-2)}=\lim_{x\to\infty}\frac{e^{x+2}-e^{-(x+2)}}{e^{x-2}-e^{-(x-2)}}=\lim_{x\to\infty}\frac{e^2-\frac{e^{-2}}{e^{2x}}}{e^{-2}-\frac{e^2}{e^{2x}}}=e^4.]$$

8.2 Exercises

1. Differentiate the function $f(x):=3x^5+\sqrt{x}+\ln\sin^2(\frac{1}{x})$.

   Solution: $f'(x)=3\cdot 5x^4+\frac{1}{2}x^{-\frac{1}{2}}+\frac{1}{\sin^2(\frac{1}{x})}\cdot 2\sin(\frac{1}{x})\cdot\cos(\frac{1}{x})\cdot(-\frac{1}{x^2})$.

2. Differentiate the functions

$$\begin{aligned}&g(x):=4x^3-2x^2+5x-3;\\&h(x):=(x-2)^3\sin(4x);\\&l(x):=x^a+a^x+ax+\tfrac{x}{a}+\tfrac{a}{x}\quad\text{(a>0)};\\&k(x):=(\sin x)^{\cos x};\\&m(x):=\operatorname{arctg}\tfrac{\operatorname{tg}x+1}{1-\operatorname{tg}x}.\end{aligned}$$

3. Sketch the curve of the function $f:\mathbb{R}\to\mathbb{R}$, $f(x):=\frac{2x-1}{\sqrt{x^2+1}}$.

   Solution:

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

b) $\dfrac{x+2}{(x^2+1)^{\frac{3}{2}}} = 0,\ x = -2$ (the fraction does not change sign at any other point since the denominator is positive).

c) $$f''(x) = \frac{(x^2+1)^{\frac{3}{2}} - (x+2)\frac{3}{2}(x^2+1)^{\frac{1}{2}}2x}{(x^2+1)^3}$$
$$= \frac{x^2+1-(3x^2+6x)}{(x^2+1)^{\frac{5}{2}}} = \frac{-2x^2-6x+1}{(x^2+1)^{\frac{5}{2}}}.$$

d) $$\frac{-2x^2-6x+1}{(x^2+1)^{\frac{5}{2}}} = 0,\quad -2x^2-6x+1=0 \quad \begin{cases} x_1 = \frac{6+\sqrt{44}}{-4} \approx -3.16, \\ x_2 = \frac{6-\sqrt{44}}{-4} \approx 0.16. \end{cases}$$

e)

	-3.16 $\quad$ -2 $\quad$ 0.16
f'	- | + + + + + + + + + + +
f mon.	decreasing $\quad$ min $\quad$ increasing
f''	- - - - - - - | + + + + + + + + + + + + + | - - - - - -
f shape	concave | inflection | convex | inflection | concave

f) $f(-2) = -\frac{5}{\sqrt{5}} = -\sqrt{5} \approx -2.23,$

$\lim_{x\to-\infty} \frac{2x-1}{\sqrt{x^2+1}} = \lim_{x\to\infty} \frac{2-\frac{1}{x}}{-\sqrt{1+\frac{1}{x^2}}} = -2,$

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

g)

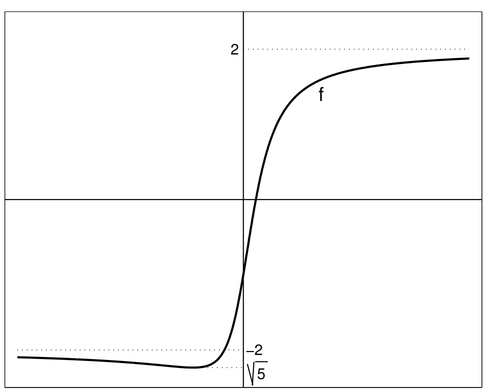


Figure 8.3

4. Analyze the curves of the following functions:

$$g : \mathbb{R} \to \mathbb{R},\ g(x) := e^{-x^2},$$

$h : \mathbb{R} \setminus \{-2, 8\},\ h(x) := \frac{x}{x^2-6x-16},$
$l : \mathbb{R}^+ \to \mathbb{R},\ l(x) := x \ln x,$
$k : \mathbb{R} \to \mathbb{R},\ k(x) := \frac{e^x}{1+e^x}.$

5. Prepare the Taylor polynomial of the order $n := 11$ for the function $f(x) := \sin x$ around the point $a := 0$.

   Solution:

$$\begin{aligned}
f(x) &= \sin x, & f(0) &= 0,\\
f'(x) &= \cos x, & f'(0) &= 1,\\
f''(x) &= -\sin x, & f''(0) &= 0,\\
f'''(x) &= -\cos x, & f'''(0) &= -1,\\
f^{(4)}(x) &= \sin x, & f^{(4)}(0) &= 0,\\
f^{(5)}(x) &= \cos x, & f^{(5)}(0) &= 1,\\
&\vdots & &\vdots\\
f^{(11)}(x) &= -\cos x & f^{(11)}(0) &= -1,\\
f^{(12)}(x) &= \sin x & f^{(12)}(0) &= 0.
\end{aligned}$$

   [One can see that $f = f^{(4)} = f^{(8)} = \ldots = f^{(4k)} = \ldots = \sin$.]

   Thus,

$$T_{11}(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \frac{1}{9!}x^9 - \frac{1}{11!}x^{11}$$

   Remark: If the sine function is approximated by its Taylor polynomial T_{11}, then, e.g., at $x := 0,1$ we have

$$\begin{aligned}
|\sin 0,1 - T_{11}(0,1)| &= \frac{|\sin c|}{12!}0,1^{12} \le \frac{0,1^{12}}{12!}\\
&= \frac{10^{-12}}{479001600} < 2 \cdot 10^{-9} \cdot 10^{-12} = 2 \cdot 10^{-21}.
\end{aligned}$$

   Moreover, if $0 \le x \le \frac{\pi}{2}$, then (by exploiting the fact that $x \le \frac{\pi}{2} < 2$)

$$\begin{aligned}
|\sin x - T_{11}(x)| &= \frac{|\sin c|}{12!}x^{12} \le \frac{1}{12!}\left(\frac{\pi}{2}\right)^{12} < \frac{2^{12}}{12!}\\
&\le 2 \cdot 10^{-9} \cdot 2^{12} = 8192 \cdot 10^{-9} < 10^{-5}.
\end{aligned}$$

   Clearly, the values of T_{11} approximate the values of the sine function fairly well (at least with an accuracy of four decimal digits) on the interval $[0, \frac{\pi}{2}]$.

6. Prepare the Taylor polynomials for the following functions:

$$\begin{aligned} g(x) &:= e^x, & a &:= 0, & n &:= 10, \\ h(x) &= \cos x, & a &:= 0, & n &:= 12, \\ l(x) &= \sqrt{1+x}, & a &:= 0, & n &:= 2, \\ k(x) &= \frac{1}{\sqrt{1+x^2}}, & a &:= 0, & n &:= 2. \end{aligned}$$

7. Calculate the limit $\lim_{x\to 0} x^2 \ln x$.

   Solution:
$$\lim_{x\to 0} x^2 \ln x = \lim_{x\to 0} \frac{\ln x}{x^{-2}}.$$

   Since
$$\lim_{x\to 0} \frac{(\ln x)'}{(x^{-2})'} = \lim_{x\to 0} \frac{x^{-1}}{-2x^{-3}} = \lim_{x\to 0} -\frac{1}{2}x^2 = 0,$$

   therefore
$$\lim_{x\to 0} \frac{\ln x}{x^{-2}} = 0, \text{ and so } \lim_{x\to 0} x^2 \ln x = 0.$$

8. Calculate the following limits:

   a) $\lim_{x\to 0} \frac{x\mathrm{tg}x-1}{x^2}$.
   b) $\lim_{x\to 0} \frac{\sqrt{\cos x}-1}{\sin^2 2x}$.
   c) $\lim_{x\to 0} \frac{1-\sqrt{\cos x}}{1-\cos\sqrt{x}}$.
   d) $\lim_{x\to \infty} \frac{x^{\ln x}}{(\ln x)^x}$.

Chapter 9

Integrability, integration

Integrability of a function means that the "domain under the graph" of the function has an area. We will elaborate a method for determining this area. We will trace several problems to the calculation of the area under the graph. The following topics will be discussed.

- The concept and geometric meaning of the Riemann integral
- Integration rules
- The Newton–Leibniz formula
- Primitive functions
- Primitive functions of the elementary functions
- Some geometric and physical applications of the integral
- Fourier series
- Improper integral

9.1 Integration

9.1.1 The concept and geometric meaning of the Riemann integral

It is well known that the area of a rectangle of length $u > 0$ and width $v > 0$ is uv. In case of $u > 0$ and $v < 0$ let us call uv the "signed area". The areas of curvilinear shapes were already investigated in the early ages of mathematics. Let us now examine the area of the domain "under the parabola"

$$H := \{(x, y) \mid x \in [0, 1],\ y \in [0, x^2]\}.$$

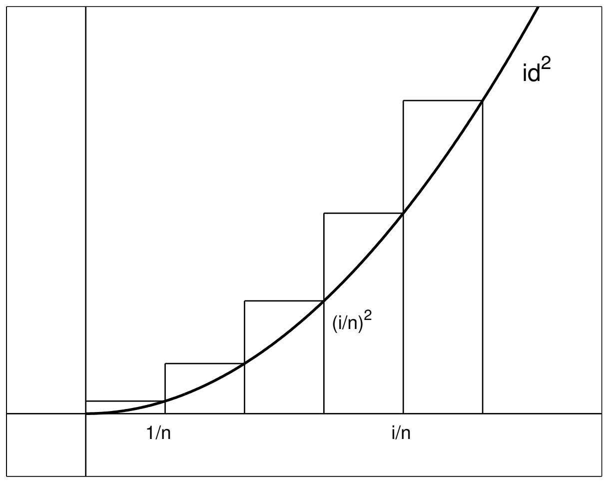


Figure 9.1

Divide the interval $[0,1]$ into n equal parts. The mesh points are then $x_0 = 0$, $x_1 = \frac{1}{n}$, $x_2 = \frac{2}{n}, \ldots, x_n = \frac{n}{n}$. Let $S_n := \frac{1}{n} \cdot (\frac{1}{n})^2 + \frac{1}{n} \cdot (\frac{2}{n})^2 + \ldots + \frac{1}{n} \cdot (\frac{n}{n})^n$, that is, the sum of the areas of those rectangles whose base is $\frac{1}{n}$, and whose height is the function value of id^2 at the mesh points (Fig. 9.1).

S_n is the area of a "staircase shape". If we increase the number n of the mesh points, then the staircase shapes will fit better and better to the set H, and so we can expect that the limit of the sequence (S_n) will exactly be the area of the set H. By using the fact that for all $k \in \mathbb{N}$, $1^2 + 2^2 + \ldots + k^2 = \frac{k(k+1)(2k+1)}{6}$,

$$\begin{aligned} \lim S_n &= \lim \frac{1}{n^3}(1^2 + 2^2 + \ldots + n^2) = \lim \frac{1}{n^3} \frac{n(n+1)(2n+1)}{6} \\ &= \lim \frac{2n^2 + 3n + 1}{6n^2} = \lim \frac{2 + \frac{3}{n} + \frac{1}{n^2}}{6} = \frac{1}{3}. \end{aligned}$$

So, let the area of the set H be $\frac{1}{3}$.

This train of thought will be generalized.

Let $f : [a,b] \to \mathbb{R}$ be a function.

Let

$$\tau := \{x_0, x_1, x_2, \ldots, x_{i-1}, x_i, \ldots, x_n\} \subset [a,b],$$

with

$$a = x_0 < x_1 < x_2 < \ldots < x_{i-1} < x_i < \ldots < x_n = b$$

be a **partition** of the interval $[a,b]$.

Let us set a point ξ_i within all sub-intervals $[x_{i-1}, x_i]$ ($i = 1, 2, \ldots, n$.) Prepare the sum approximation of the function f for the partition τ:

$$\begin{aligned}\sigma(\tau) :&= f(\xi_1)(x_1 - x_0) + f(\xi_2)(x_2 - x_1) + \ldots + f(\xi_n)(x_n - x_{n-1}) \\ &= \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}).\end{aligned}$$

(This $\sigma(\tau)$ corresponds to the area S_n of the staircase shape, where the point ξ_i was taken at the right end of each sub-interval.)

We call a function integrable if the sum approximations $\sigma(\tau)$ get infinitely close to a number as the partition is "refined". More precisely:

Definition 9.1. We say that the function $f : [a, b] \to \mathbb{R}$ is **integrable on the interval [a,b]**, if there is a number $I \in \mathbb{R}$ such that for any error bound $\varepsilon > 0$ there is a $\delta > 0$ such that for all partitions τ of the interval $[a, b]$ for which

$$\max\{x_i - x_{i-1} \mid i = 1, 2, \ldots, n\} < \delta$$

and for any points $\xi_i \in [x_{i-1}, x_i]$ taken from the sub-intervals $[x_{i-1}, x_i]$ of the partition τ the sum approximation $\sigma(\tau) = \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1})$ satisfies

$$|\sigma(\tau) - I| < \varepsilon.$$

If f is integrable on the interval $[a, b]$, then this will be denoted as $f \in R[a, b]$ (in honour of Riemann, who introduced the integral in this manner), and let

$$\int_a^b f := I.$$

("integral from a to b"). Moreover, in this case we say that the set

$$H := \{(x, y) \mid x \in [a, b],\ y \in [0, f(x)] \text{ if } f(x) \geq 0,\ \text{ or } y \in [f(x), 0] \text{ if } f(x) < 0\}$$

("domain under the graph") has a signed area, and this area is the number $I \in \mathbb{R}$.

It is common to refer to this concept by introducing the notation $\Delta x_i := x_i - x_{i-1}$ and writing

$$\lim_{\Delta x_i \to 0} \sum f(\xi_i)\Delta x_i = I$$

or

$$\lim_{\Delta x \to 0} \sum f(\xi)\Delta x = \int_a^b f(x)\mathrm{d}x.$$

(It is worth having a look at the metamorphose of the symbols!)

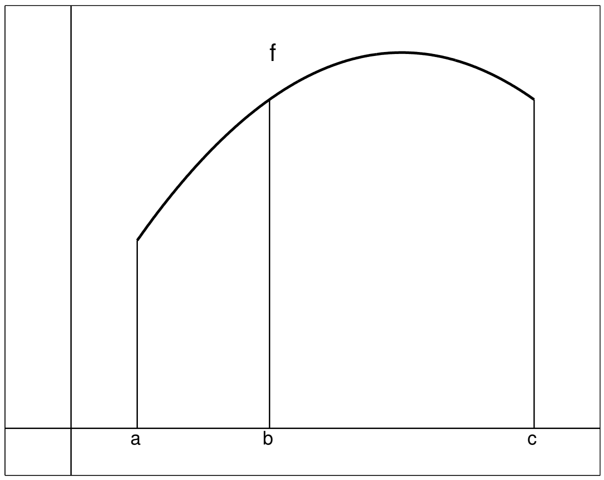


Figure 9.2

It is easy to see that if $f : [a, b] \to \mathbb{R}$, $f(x) = c$ is a constant function, then

$$\lim_{\Delta x_i \to 0} \sum f(\xi_i)\Delta x_i = \lim_{\Delta x_i \to 0} \sum_{i=1}^{n} c(x_i - x_{i-1}) = c(b - a),$$

as we have expected from the geometry, so $f \in R[a, b]$ and $\int_a^b f = c(b - a)$.

9.1.2 Relationship between the Riemann integral and the operations

It can be proved that if $f \in C[a, b]$, then $f \in R[a, b]$. It follows also from the geometric approach that if $f \in R[a, b]$ and $f \in R[b, c]$, then $f \in R[a, c]$, what is more,

$$\int_a^b f + \int_b^c f = \int_a^c f \qquad \text{(Fig. 9.2)}.$$

It is not so obvious, but one can verify the following

Theorem 9.1. *If $f \in R[a, b]$ and $\lambda \in \mathbb{R}$, then $\lambda f \in R[a, b]$, and*

$$\int_a^b \lambda f = \lambda \int_a^b f.$$

Theorem 9.2. *If $f, g \in R[a, b]$, then $f + g \in R[a, b]$, and*

$$\int_a^b (f + g) = \int_a^b f + \int_a^b g.$$

Theorem 9.3. *If $f, g \in R[a,b]$, and $f(x) \geq g(x)$ for all $x \in [a,b]$, then*

$$\int_a^b f \geq \int_a^b g.$$

An important consequence of the latter theorem is the statement that if $f \in R[a,b]$, then $|f| \in R[a,b]$, and due to

$$-|f| \leq f \leq |f|$$

it follows that

$$-\int_a^b |f| \leq \int_a^b f \leq \int_a^b |f|,$$

thus

$$\left|\int_a^b f\right| \leq \int_a^b |f|.$$

9.1.3 Newton–Leibniz formula

It is apparent from the geometry that the following theorem holds.

Theorem 9.4. *If $f \in C[a,b]$, then there exists a $c \in [a,b]$ such that*

$$\int_a^b f = f(c) \cdot (b-a) \qquad \textit{(Fig. 9.3).}$$

The number $\frac{\int_a^b f}{b-a}$ is called **mean value** of the function f. This is a generalization of the mean value of finitely many numbers. (The theorem says that the mean value is a function value.)

The statements that we have seen so far are expressive, however, we still owe the reader a comfortable way of calculating the integral.

For simplicity, let $f \in C[a,b]$. We introduce the **area function** as $T : [a,b] \to \mathbb{R}$, $T(x) := \int_a^x f$ (Fig. 9.4).

Let $\alpha \in (a,b)$ be an arbitrary point, and let us examine the difference quotient

$$\frac{T(x) - T(\alpha)}{x - \alpha}$$

in case $x \in (a,b),\ x \neq \alpha$.

$$\frac{T(x) - T(\alpha)}{x - \alpha} = \frac{\int_a^x f - \int_x^\alpha f}{x - \alpha} = \frac{1}{x-\alpha}\int_\alpha^x f = \frac{1}{x-\alpha} f(c)(x-\alpha) = f(c),$$

where $c \in [\alpha, x]$ (Fig. 9.5). From this, by exploiting the fact that $f \in C[\alpha]$,

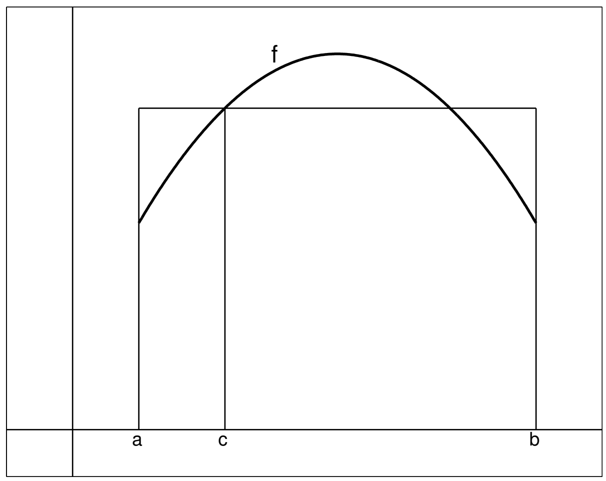


Figure 9.3

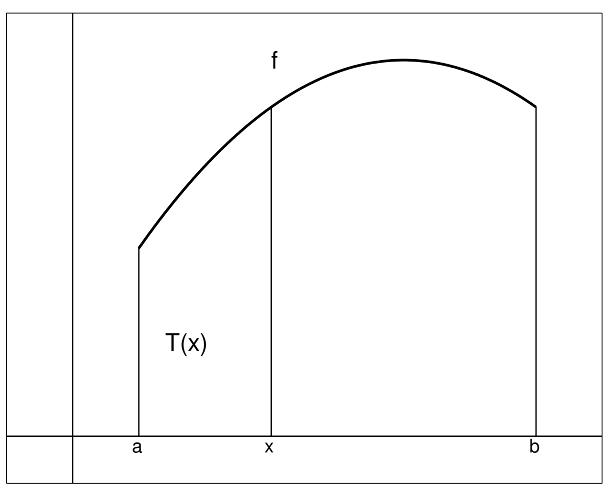


Figure 9.4

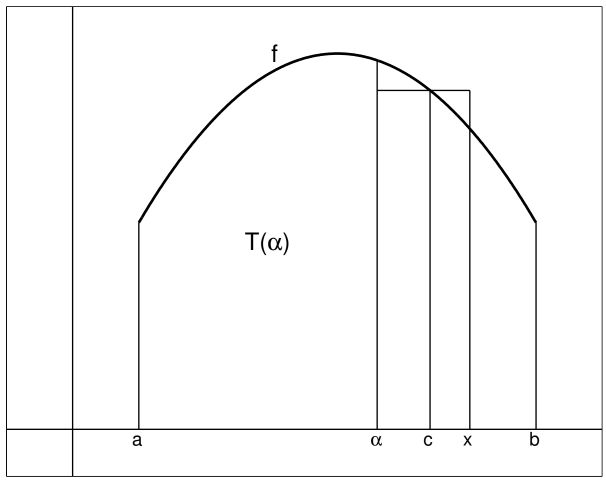


Figure 9.5

$$\lim_{x \to \alpha} \frac{T(x) - T(\alpha)}{x - \alpha} = \lim_{x \to \alpha} f(c) = f(\alpha),$$

on the other hand,

$$\lim_{x \to \alpha} \frac{T(x) - T(\alpha)}{x - \alpha} = T'(\alpha).$$

So, the area function T is such that its derivative is f. Since $T(a) = 0$ and $T(b) = \int_a^b f$, therefore

$$\int_a^b f = T(b) - T(a).$$

We have arrived at a famous result (in a rather heuristic way).

Theorem 9.5 (Newton–Leibniz formula)**.** *If $f \in C[a, b]$, and T is such a function that $T' = f$, then*

$$\int_a^b f = T(b) - T(a).$$

For example, if $f : [0, 1] \to \mathbb{R}$, $f(x) = x^2$, then the function $T : [0, 1] \to \mathbb{R}$, $T(x) := \frac{x^3}{3}$ will be appropriate (since $(\frac{x^3}{3})' = x^2$), thus

$$\int_0^1 f = T(1) - T(0) = \frac{1^3}{3} - \frac{0^3}{3} = \frac{1}{3},$$

which is in accordance with the result obtained in the introducing example.

9.1.4 Primitive functions

In some sense, a primitive function represents the opposite of derivation (also called antiderivative).

Definition 9.2. Let $I \subset \mathbb{R}$ be an open interval, and $f : I \to \mathbb{R}$. The differentiable function $F : I \to \mathbb{R}$ is a **primitive function** of f if $F' = f$.

If F and G are primitive functions of f, then $F' = f$ and $G' = f$, and so $(F - G)' = 0$, but on an interval the derivative of a function can only be zero if the function itself is constant. So, there exists a $c \in \mathbb{R}$ such that $F - G = c$, that is, the primitive functions of a function can only differ from each other by a constant. (Also, the area function T can only differ from any other primitive function by a constant.)

The calculation of the integral has become extremely simple, since we only have to find a primitive function of f. If we denote it by F, then, according to the traditional way of writing

$$\int_a^b f(x)\mathrm{d}x = [F(x)]_a^b,$$

where $[F(x)]_a^b := F(b) - F(a)$.

For example, for the calculation of $\int_0^\pi \sin x \mathrm{d}x$ $F(x) := -\cos x$ is a suitable primitive function ($(-\cos x)' = \sin x$), thus

$$\int_0^\pi \sin x \mathrm{d}x = [-\cos x]_0^\pi = 1 - (-1) = 2.$$

Let us agree that a primitive function of f will be denoted by $\int f$ instead of F, and $\int f(x)\mathrm{d}x$ instead of $F(x)$. Searching for a primitive function is the "inverse" of derivation. Some simple methods for finding primitive functions (which you can easily check by derivation!) are as follows.

$\int \lambda f = \lambda \int f, \quad \int (f + g) = \int f + \int g.$

$\int f'g = fg - \int fg' \quad$ (integration by parts).

$\int \phi^\alpha \cdot \phi' = \frac{\phi^{\alpha+1}}{\alpha+1} \quad$ if $\alpha \neq -1$.

$\int \frac{\phi'}{\phi} = \ln \circ \phi$ if $\phi(x) > 0$ $(x \in I)$.

If $\int f(x)\mathrm{d}x = F(x)$, then $\int f(ax + b) = \frac{1}{a} F(ax + b)$.

$(\int f) \circ \phi = \int (f \circ \phi \cdot \phi') \quad$ (integration by substitution).

The inversion of the differentiation "rules" yields the following table of integrals.

$\int x^\alpha \mathrm{d}x = \frac{x^{\alpha+1}}{\alpha+1} \quad (\alpha \neq -1)$,

$\int \frac{1}{x}\mathrm{d}x = \ln x$ if $x > 0$, and $\int \frac{1}{x}\mathrm{d}x = \ln(-x)$ if $x < 0$

$\int e^x \mathrm{d}x = e^x$, $\qquad \int a^x \mathrm{d}x = \frac{a^x}{\ln a}$,

$\int \sin x \mathrm{d}x = -\cos x$, $\qquad \int \cos x \mathrm{d}x = \sin x$,

$\int \frac{1}{\cos^2 x} = \mathrm{tg}x$, $\qquad \int \frac{1}{\sin^2 x} = -\mathrm{ctg}x$,

$\int \mathrm{sh}x \mathrm{d}x = \mathrm{ch}x$, $\qquad \int \mathrm{ch}x \mathrm{d}x = \mathrm{sh}x$,

$\int \frac{1}{1+x^2}\mathrm{d}x = \mathrm{arctg}x$, $\qquad \int \frac{1}{\sqrt{1-x^2}}\mathrm{d}x = \mathrm{arcsin}x,\ x \in (-1,1)$,

$\int \frac{1}{\sqrt{x^2+1}}\mathrm{d}x = \mathrm{arsh}x$, $\qquad \int \frac{1}{\sqrt{x^2-1}}\mathrm{d}x = \mathrm{arch}x,\ x \in (1,+\infty)$.

9.1.5 The applications of integrals

1. If $f, g \in R[a,b]$, and for all $x \in [a,b]$, $f(x) \geq g(x)$, then the **area** of the set
$$H := \{(x,y) \mid x \in [a,b] \text{ és } g(x) \leq y \leq f(x)\}$$
is given by the formula
$$T = \int_a^b f - \int_a^b g = \int_a^b (f-g) \left(= \int_a^b f(x) - g(x)\mathrm{d}x \right)$$
(Fig. 9.6). (Note that the functions f and g are not necessarily nonnegative!)

2. It is known from geometry that the volume of a brick of length a, width b and height m is $V = ab \cdot m$. By generalization, the volume of a prism is the product of the area of the base and the height, that is $V = T \cdot m$.

   Consider now a three-dimensional shape H (for example a potato). Prepare all plane sections of H perpendicular to the x axis of the coordinate system. (In the example of the potato a knitting needle, stuck through the potato could play the role of the x axis. We split at right angles to this needle.) Assume that the plane section $S(x)$ obtained at x has an area, and the function
$$S : [a,b] \to \mathbb{R}, \quad x \mapsto S(x),$$
is continuous on the interval $[a,b]$ (if x' and x'' are close to each other, then $S(x')$ and $S(x'')$ are close, too) (Fig. 9.7).

   Divide the interval $[a,b]$:
$$\tau : a = x_0 < x_1 < x_2 < \ldots < x_{i-1} < x_i < \ldots < x_n = b,$$

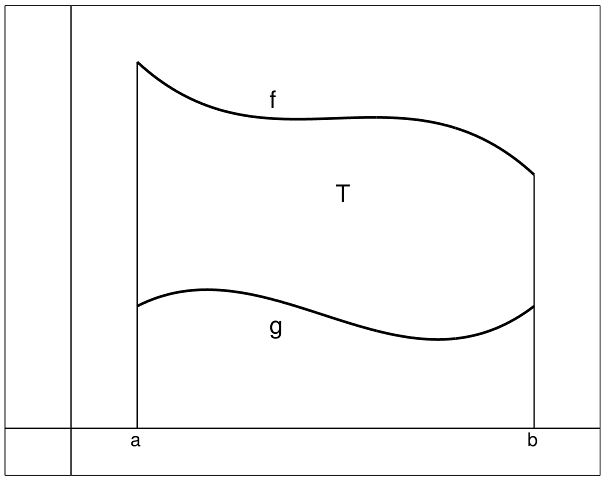


Figure 9.6

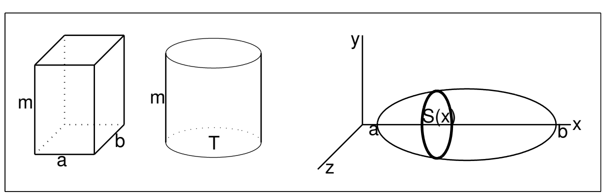


Figure 9.7

and take arbitrary points $\xi_i \in [x_{i-1}, x_i]$ $(i = 1, 2, \dots, n)$ (Fig. 9.8).

$S(\xi_i)(x_i - x_{i-1})$ is the volume of a prism whose "area of base" is $S(\xi_i)$, and whose height is $(x_i - x_{i-1})$. Summing these areas we get a sum approximation:

$$\sum_{i=1}^{n} S(\xi_i)(x_i - x_{i-1}).$$

By refining the partition of the interval $[a, b]$, the sum approximations have a limit ($S \in C[a, b]$, therefore $S \in R[a, b]$), which will be the volume of the shape:

$$V = \lim_{x_i - x_{i-1} \to 0} \sum S(\xi_i)(x_i - x_{i-1}) = \int_a^b S(x)\mathrm{d}x.$$

Calculating the volume becomes especially simple if H is a "solid of revolution", obtained by "rotating a function $f : [a, b] \to \mathbb{R}$, $f \in C[a, b]$, $f(x) \geq 0$ $(x \in [a, b])$ around the x axis" (Fig. 9.9). Then the area of the plane section $S(x)$ is the area of a circle:

$$S(x) = \pi f^2(x),$$

thus

$$V = \int_a^b \pi f^2(x)\mathrm{d}x.$$

It is easy to see **Cavalieri's principle**, too, which says that if two shapes have pairwise equal plane sections, parallel with a plane (that is, $S_1(x) = S_2(x)$ for all x, where the x axis is a straight line perpendicular to the plane), and the functions S_1 and S_2, obtained in this manner, are continuous, then the two shapes have the same volume, since due to $S_1 = S_2$

$$\int_a^b S_1(x)\mathrm{d}x = \int_a^b S_2(x)\mathrm{d}x.$$

3. Let $f : [a, b] \to \mathbb{R}$ a continuously differentiable function. The set

$$H := \{(x, f(x) \mid x \in [a, b])\}$$

is often called the **graph** of f. Our aim is to calculate its arc length. Divide the interval $[a, b]$ again:

$$\tau : a = x_0 < x_1 < x_2 < \ldots < x_{i-1} < x_i < \ldots < x_n = b.$$

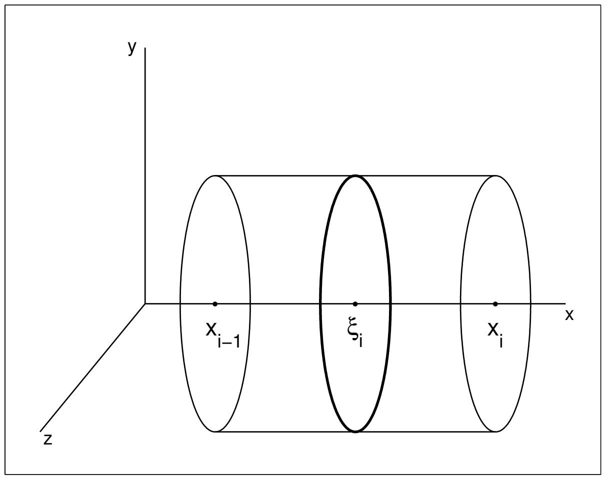


Figure 9.8

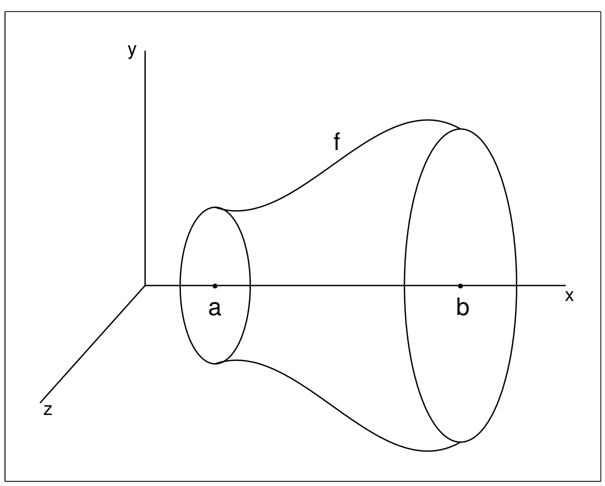


Figure 9.9

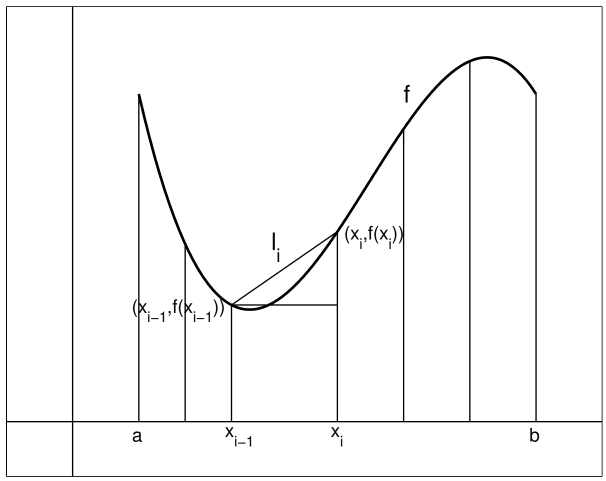


Figure 9.10

The length of the line segment connecting the points $(x_{i-1}, f(x_{i-1}))$ and $(x_i, f(x_i))$ (Fig. 9.10) is

$$\begin{aligned} l_i :&= \sqrt{(x_i - x_{i-1})^2 + (f(x_i) - f(x_{i-1}))^2} \\ &= (x_i - x_{i-1})\sqrt{1 + \left[\frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}\right]^2}. \end{aligned}$$

By Lagrange's mean value theorem there exists a $\xi_i \in (x_{i-1}, x_i)$ for which $f(x_i) - f(x_{i-1}) = f'(\xi_i) \cdot (x_i - x_{i-1})$, thus

$$l_i = (x_i - x_{i-1})\sqrt{1 + [f'(\xi_i)]^2}.$$

One can see that the length of the broken line, approximating the graph of f is

$$\sum_{i=1}^{n} l_i = \sum_{i=1}^{n} \sqrt{1 + [f'(\xi_i)]^2}(x_i - x_{i-1}),$$

which is a sum approximation of the integral for the function $\phi : [a, b] \to \mathbb{R}$, $\phi(x) := \sqrt{1 + [f'(x)]^2}$. So, the **arc length of the graph of f** is

$$I(f) = \lim_{x_i - x_{i-1} \to 0} \sum \sqrt{1 + [f'(\xi_i)]^2}(x_i - x_{i-1}) = \int_a^b \sqrt{1 + [f'(x)]^2}\mathrm{d}x.$$

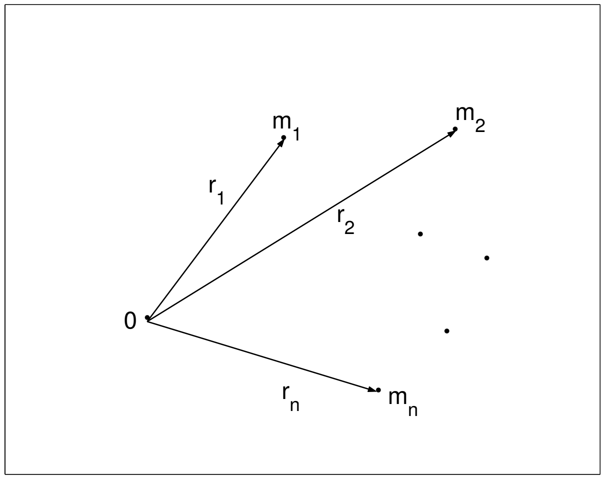


Figure 9.11

4. If $f : [a,b] \to \mathbb{R}$, $f(x) \geq 0$ $(x \in [a,b])$ is a continuously differentiable function, then, in a similar manner, the side of the solid of revolution obtained by rotating the graph of f around the x axis has the **area**
$$P(f) = \int_a^b 2\pi f(x)\sqrt{1+[f'(x)]^2}\mathrm{d}x.$$
5. It is known that the vector pointing to the center of a mass-point system is given by
$$\underline{r}_s = \frac{m_1\underline{r}_1 + m_2\underline{r}_2 + \ldots + m_n\underline{r}_n}{m_1 + m_2 + \ldots + m_n},$$
where m_i is the mass of the ith mass point, and $\underline{r}_i$ is the position vector pointing from a fixed point to the ith mass point (Fig. 9.11). Let $f \in R[a,b]$, $f(x) \geq 0$, $x \in [a,b]$, and
$$H := \{(x,y) \mid x \in [a,b],\ y \in [0, f(x)]\}$$
a homogeneous disc of density ρ (Fig. 9.12).

To find the center of mass of the disc let us divide the interval $[a,b]$.
$$\tau : a = x_0 < x_1 < x_2 < \ldots < x_{i-1} < x_i < \ldots < x_n = b.$$
By choosing the points $\xi_i := \frac{x_{i-1}+x_i}{2}$ $(i = 1,2,\ldots,n)$, the position vector directed to the center of mass of the rectangle $[x_{i-1}, x_i] \times [0, f(\xi_i)]$ is
$$\underline{r}_i\left(\xi_i, \frac{f(\xi_i)}{2}\right),$$

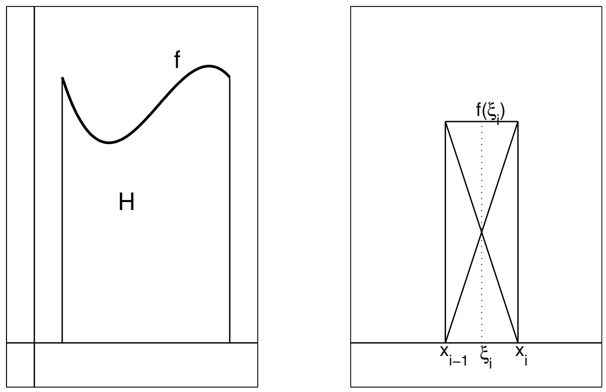


Figure 9.12

and the mass of the point, substituting the rectangle is

$$m_i = \rho f(\xi_i)(x_i - x_{i-1}).$$

The approximate value of the first coordinate of the center of mass is

$$\frac{m_1\xi_1 + m_2\xi_2 + \ldots + m_n\xi_n}{m_1 + m_2 + \ldots + m_n} = \frac{\sum_{i=1}^{n} \rho f(\xi_i)\cdot\xi_i(x_i - x_{i-1})}{\sum_{i=1}^{n} \rho f(\xi_i)(x_i - x_{i-1})},$$

while that of the second coordinate is

$$\frac{m_1\frac{f(\xi_1)}{2} + m_2\frac{f(\xi_2)}{2} + \ldots + m_n\frac{f(\xi_n)}{2}}{m_1 + m_2 + \ldots + m_n} = \frac{\frac{1}{2}\sum_{i=1}^{n} \rho f^2(\xi_i)(x_i - x_{i-1})}{\sum_{i=1}^{n} \rho f(\xi_i)(x_i - x_{i-1})}.$$

Apparently, both expressions contain sum approximations of integrals, therefore no surprise that the vector $\underline{r}_s = (x_s, y_s)$ pointing to the **center of mass of the disc** will be as follows:

$$x_s = \frac{\int_a^b x f(x)\mathrm{d}x}{\int_a^b f(x)\mathrm{d}x}; \quad y_s = \frac{\frac{1}{2}\int_a^b f^2(x)\mathrm{d}x}{\int_a^b f(x)\mathrm{d}x}.$$

6. The moment of inertia of a mass point of mass m rotating around the point O is $\Theta = ml^2$, where l is the distance of the mass point from O (Fig. 9.13).

   If a rod of length L and mass M rotates around an axis, fixed to the end of the rod and perpendicular to it, then we can calculate the moment of inertia of the rod with respect to the axis. Divide the interval $[0, L]$ as

$$\tau : 0 = x_0 < x_1 < x_2 < \ldots < x_{i-1} < x_i < \ldots < x_n = L.$$

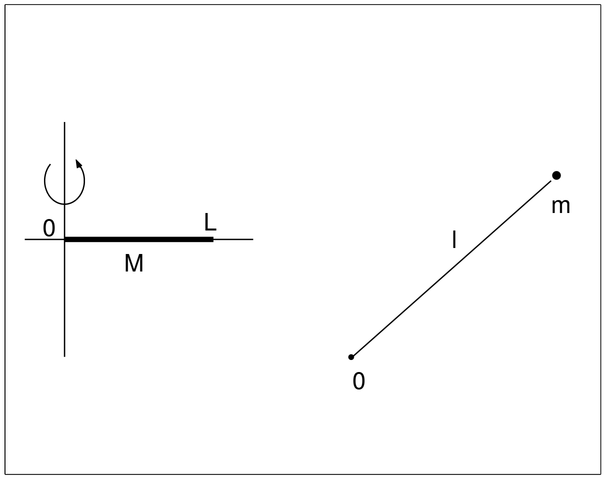


Figure 9.13

The mass of the rod segment $[x_{i-1}, x_i]$ is

$$\frac{M}{L} \cdot (x_i - x_{i-1}),$$

while the distance to the rotation axis can be chosen as

$$\xi_i := x_i,$$

so the moment of inertia of this segment can be given as

$$\frac{M}{L} \cdot (x_i - x_{i-1})\xi_i^2.$$

The sum of the moments of inertia of the segments approximates the moment of inertia of the whole rod:

$$\sum_{i=1}^{n} \frac{M}{L} \xi_i^2 (x_i - x_{i-1}).$$

From this, the moment of inertia of the rod is

$$\Theta = \lim_{x_i - x_{i-1} \to 0} \sum_{i=1}^{n} \frac{M}{L} \xi_i^2 (x_i - x_{i-1}) = \int_0^L \frac{M}{L} x^2 \mathrm{d}x,$$

which, by the Newton–Leibniz formula can be calculated as

$$\int_0^L \frac{M}{L} x^2 \mathrm{d}x = \left[\frac{M}{L} \frac{x^3}{3} \right]_0^L = \frac{M}{L} \frac{L^3}{3} = \frac{1}{3} M L^2,$$

thus $\Theta = \frac{1}{3}ML^2$.

These few examples have shown what a wide variety of problems can be traced to integrals.

We will sketch a further important application.

9.1.6 Fourier series

Let $f : \mathbb{R} \to \mathbb{R}$ be a periodic function with period 2π. (If the period of f is $p > 0$, then by a simple transformation ($x := \frac{2\pi}{p}t$) we can make it periodic with period 2π.) Our aim is to compose f as a "sum" of the well known functions $\cos \circ n\mathrm{id}$ ($n = 0, 1, 2, \ldots$) and $\sin \circ n\mathrm{id}$ ($n = 0, 1, 2, \ldots$), that is, to construct such sequences $a_0, a_1, a_2, \ldots, a_n, \ldots$ and $b_1, b_2, \ldots, b_n, \ldots$ that for all $x \in \mathbb{R}$

$$\begin{aligned} f(x) = {} & \frac{a_0}{2} + a_1 \cos x \\ & + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \ldots + a_n \cos nx + b_n \sin nx + \ldots . \end{aligned} \tag{9.1}$$

It is not clear at this point for what functions f it is possible to do this. Moreover, if we do obtain some sequences (a_n) and (b_n), then how do we know that the sum on the right-hand side returns f? Now, by formal reasoning, let us start from the assumption that $f \in C(\mathbb{R})$ and for all $x \in \mathbb{R}$ it is equal to the sum of the infinite series

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \sin nx.$$

1. Integrate the equality (9.1) from $(-\pi)$ to π by the assumption that the sum can be integrated termwise:

$$\int_{-\pi}^{\pi} f(x)\mathrm{d}x = \int_{-\pi}^{\pi} \frac{a_0}{2}\mathrm{d}x + \sum_{n=1}^{\infty} a_n \int_{-\pi}^{\pi} \cos nx\mathrm{d}x + b_n \int_{-\pi}^{\pi} \sin nx\mathrm{d}x.$$

Since

$$\int_{-\pi}^{\pi} \frac{a_0}{2}\mathrm{d}x = \frac{a_0}{2}2\pi,$$

$$\int_{-\pi}^{\pi} \cos nx\mathrm{d}x = \left[\frac{\sin nx}{n}\right]_{-\pi}^{\pi} = 0, \qquad \int_{-\pi}^{\pi} \sin nx\mathrm{d}x = \left[\frac{-\cos nx}{n}\right]_{-\pi}^{\pi} = 0,$$

therefore

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\mathrm{d}x.$$

2. Let $k \in \mathbb{N}$ be a fixed index. Multiply the equality (9.1) by $(\cos kx)$, and integrate from $(-\pi)$ to π:
$$\int_{-\pi}^{\pi} f(x)\cos kx\mathrm{d}x = \frac{a_0}{2}\int_{-\pi}^{\pi}\cos kx\mathrm{d}x$$
$$+\sum_{n=1}^{\infty} a_n\int_{-\pi}^{\pi}\cos nx\cos kx\mathrm{d}x + b_n\int_{-\pi}^{\pi}\sin nx\cos kx\mathrm{d}x.$$
By trigonometric formulas, for $n \neq k$
$$\int_{-\pi}^{\pi}\cos nx\cos kx\mathrm{d}x = \int_{-\pi}^{\pi}\frac{1}{2}(\cos(n+k)x + \cos(n-k)x)\mathrm{d}x$$
$$= \frac{1}{2}\left(\left[\frac{\sin(n+k)x}{n+k}\right]_{-\pi}^{\pi} + \left[\frac{\sin(n-k)x}{n-k}\right]_{-\pi}^{\pi}\right) = 0,$$
and for $n = k$
$$\int_{-\pi}^{\pi}\cos^2 kx\mathrm{d}x = \int_{-\pi}^{\pi}\frac{1+\cos 2kx}{2}\mathrm{d}x = \left[\frac{1}{2}x + \frac{\sin 2kx}{4k}\right]_{-\pi}^{\pi} = \pi.$$
By a similar calculation
$$\int_{-\pi}^{\pi}\sin nx\cos kx\mathrm{d}x = 0.$$
So, the terms of the infinite series are all zero with the exception of one, hence
$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kx\mathrm{d}x.$$
If the equality (9.1) is multiplied by $(\sin kx)$ and integrated from $(-\pi)$ to π, then the same way of calculation yields
$$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin kx\mathrm{d}x.$$

3. When the function f is continuous, then the obtained numbers
$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\mathrm{d}x,\ \ a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kx\mathrm{d}x,\ \ b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin kx\mathrm{d}x,$$
$k = 1, 2, \ldots$ are called **Fourier coefficients** of f. It can be verified that (apart from some very exceptional functions, unimaginable in practise) f is given as the sum of the **Fourier series**, made up with the above coefficients:
$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k\cos kx + b_k\sin kx \quad (x \in \mathbb{R}).$$
This method can be used for the investigation of vibrations and waves.

9.1.7 Improper integrals

So far we have investigated integrability and calculated integrals over closed and bounded intervals $[a, b]$ only. Now we shall extend our concepts.

Definition 9.3. Let $f : [a, +\infty) \to \mathbb{R}$ be a function for which $\forall \omega \in \mathbb{R}$, $\omega > a$: $f \in R[a, \omega]$. We say that the **improper integral of f is convergent** on the interval $[a, +\infty)$ if

$$\exists \lim_{\omega \to \infty} \int_a^\omega f \in \mathbb{R}.$$

Let us denote this as $f \in R[a, +\infty)$. If $f \in R[a, +\infty)$, then

$$\int_a^{+\infty} f := \lim_{\omega \to \infty} \int_a^\omega f.$$

If the limit $\lim_{\omega \to \infty} \int_a^\omega f$ does not exist, or if it exists, but not finite, then we say that the improper integral of f is divergent.

For example,

$$\int_1^{+\infty} \frac{1}{x^2} \mathrm{d}x = \lim_{\omega \to \infty} \int_1^\omega \frac{1}{x^2} \mathrm{d}x = \lim_{\omega \to \infty} \left[-\frac{1}{x} \right]_1^\omega = \lim_{\omega \to \infty} \left(-\frac{1}{\omega} + 1 \right) = 1,$$

thus $\mathrm{id}^{-2} \in R[1, +\infty)$.

Since

$$\lim_{\omega \to \infty} \int_1^\omega \frac{1}{x} \mathrm{d}x = \lim_{\omega \to \infty} [\ln x]_1^\infty = \lim_{\omega \to \infty} \ln \omega = +\infty,$$

therefore $\mathrm{id}^{-1} \notin R[1, +\infty)$, that is, the improper integral of id^{-1} is divergent.

There are other extensions as well.

Definition 9.4. Let $f : (a, b] \to \mathbb{R}$ be a function for which $\forall \mu \in (a, b)$, $f \in R[\mu, b]$. We say that $f \in R[a, b]$ if

$$\exists \lim_{\mu \to a} \int_\mu^b f \in \mathbb{R}.$$

Then

$$\int_a^b f := \lim_{\mu \to a} \int_\mu^b f.$$

If the limit $\lim_{\mu \to a} \int_\mu^b f$ does not exist or it is infinite, then we say that the integral of f is divergent on the interval $[a, b]$. (This can even happen to bounded functions, but in most cases the victims are unbounded functions.)

For example,

$$\int_0^1 \frac{1}{\sqrt{x}}\mathrm{d}x = \lim_{\mu\to 0}\int_\mu^1 \frac{1}{\sqrt{x}}\mathrm{d}x = \lim_{\mu\to 0}[2\sqrt{x}]_\mu^1 = \lim_{\mu\to 0} 2 - 2\sqrt{\mu} = 2,$$

so $\mathrm{id}^{-\frac{1}{2}} \in R[0,1]$.

$$\int_0^1 \frac{1}{x}\mathrm{d}x = \lim_{\mu\to 0}\int_\mu^1 \frac{1}{x}\mathrm{d}x = \lim_{\mu\to 0}[\ln x]_\mu^1 = \lim_{\mu\to 0}(\ln 1 - \ln \mu) = +\infty,$$

hence the integral of the function id^{-1} is divergent on the interval $[0,1]$ as well.

One of the main results obtained for improper integrals is that

$$\int_0^\infty e^{-x^2}\mathrm{d}x = \frac{\sqrt{\pi}}{2}.$$

This implies (by further extension of the concepts) that if $m \in \mathbb{R}$ and $\delta > 0$, then

$$\int_{-\infty}^{+\infty} e^{-\frac{(x-m)^2}{2\sigma^2}}\mathrm{d}x = \sqrt{2\pi}\sigma,$$

which plays an important role in probability theory.

9.2 Exercises

1. Check the methods for finding primitive functions.

   Solution: If $\alpha \neq -1$, then

   a) $(\frac{\phi^{\alpha+1}}{\alpha+1})' = \frac{1}{\alpha+1}\cdot(\alpha+1)\phi^\alpha \cdot \phi'$, therefore $\int \phi^\alpha \cdot \phi' = \frac{\phi^{\alpha+1}}{\alpha+1}$.

   b) $(f\cdot g - \int f\cdot g')' = f'\cdot g + f\cdot g' - f\cdot g' = f'\cdot g$, therefore $\int f'\cdot g = f\cdot g - \int f\cdot g'$ (integration by parts).

   c) $(\int f\circ\phi\cdot\phi')' = f\circ\phi\cdot\phi'$, on the other hand, $((\int f)\circ\phi)' = f\circ\phi\cdot\phi'$. Since the derivatives of the two functions are equal on an interval, therefore the functions may only differ by a constant, thus $(\int f)\circ\phi = \int(f\circ\phi\cdot\phi')$ (integration by substitution).

2. Find the following primitive functions:

$\int \sin^3 x \cos x \mathrm{d}x = ?$	$\int \mathrm{tg}x \cos^5 x \mathrm{d}x = ?$	$\int \frac{2x+3}{(x^2+3)^4}\mathrm{d}x = ?$
$\int \frac{x}{1+x^2}\mathrm{d}x = ?$	$\int \mathrm{tg}x \mathrm{d}x = ?$	$\int \frac{2x+3}{x^2+3x+10}\mathrm{d}x = ?$
$\int \cos(5x-1)\mathrm{d}x = ?$	$\int \frac{1}{x^2+2}\mathrm{d}x = ?$	$\int \frac{1}{x^2+3x+10}\mathrm{d}x = ?$

3. Find the following primitive functions by integration by parts:

$$\int xe^{2x}\mathrm{d}x = ? \qquad \int x^2e^{2x}\mathrm{d}x = ? \qquad \int e^{2x}\sin 3x\mathrm{d}x = ?$$

$$\int e^{ax}\cos bx\mathrm{d}x = ? \qquad \int \ln x\mathrm{d}x = ? \qquad \int \sqrt{1-x^2}\mathrm{d}x = ?$$

Solution:

$$\begin{aligned}\int \sqrt{1-x^2}\mathrm{d}x = \int 1\cdot\sqrt{1-x^2}\mathrm{d}x &= x\sqrt{1-x^2} - \int x\frac{-x}{\sqrt{1-x^2}}\mathrm{d}x \\ &= x\sqrt{1-x^2} - \int \frac{1-x^2-1}{\sqrt{1-x^2}}\mathrm{d}x \\ &= x\sqrt{1-x^2} - \int \sqrt{1-x^2} - \frac{1}{\sqrt{1-x^2}}\mathrm{d}x \\ &= x\sqrt{1-x^2} - \int \sqrt{1-x^2}\mathrm{d}x + \mathrm{arcsin}x.\end{aligned}$$

From this $\quad 2\int \sqrt{1-x^2}\mathrm{d}x = x\sqrt{1-x^2} + \mathrm{arcsin}x$, and so

$$\int \sqrt{1-x^2}\mathrm{d}x = \frac{1}{2}(x\sqrt{1-x^2} + \mathrm{arcsin}x).$$

4. The graph of the function $f : [0, r] \to \mathbb{R}$, $f(x) := \sqrt{r^2 - x^2}$ is a quarter circle of radius r centered at the origin. Calculate the area and the circumference of the circle, and the volume and the surface of a sphere of radius r.

5. Let $a > 0$. Calculate the area under the graph and the arc length of the function $\mathrm{ch}_{|_{[0,a]}}$.

6. Where is the center of mass of the homogeneous semicircle of radius r?

7. Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) := x^2$ if $x \in [-\pi, \pi]$, and for all $x \in \mathbb{R}$ $f(x + 2\pi) = f(x - 2\pi) =: f(x)$ (Fig. 9.14).

   a) Find the Fourier coefficients of f.

   b) Expand f into Fourier series.

   c) What does this series give for $x := 0,\ x := \pi$?

8. Calculate the following improper integrals:

$$\int_1^\infty \frac{1}{x^\alpha}\mathrm{d}x = ? \ \ (\alpha > 1) \qquad \int_0^\infty e^{-at}\mathrm{d}t = ? \ \ (a > 0)$$

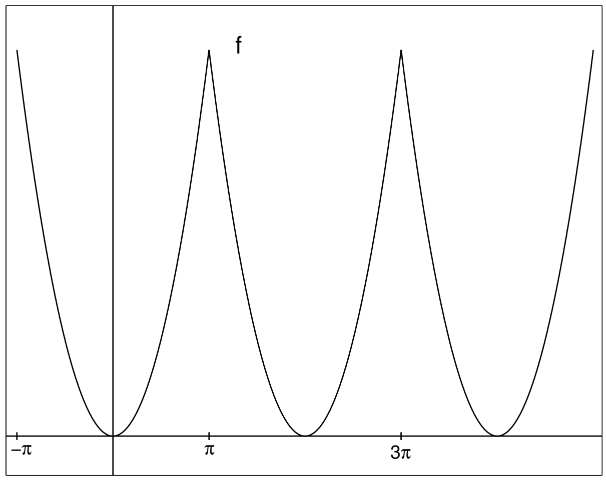


Figure 9.14

9. The gamma function.

Let $\Gamma : [0, \infty) \to \mathbb{R}$, $\Gamma(\alpha) := \int_0^\infty t^\alpha e^{-t}\mathrm{d}t$. Show that $\Gamma(0) = 1$, $\Gamma(1) = 1$, and for any $\alpha > 0$ $\Gamma(\alpha + 1) = (\alpha + 1)\Gamma(\alpha)$.

Solution: $\Gamma(0) = \int_0^\infty e^{-t}\mathrm{d}t = [e^{-t}]_0^\infty = 1$ (Here we have used an abbreviation: instead of $\lim_{\omega\to\infty}[e^{-t}]_0^\omega$ we wrote $[e^{-t}]_0^\infty$.)

$$\begin{aligned}\Gamma(\alpha+1) = \int_0^\infty t^{\alpha+1}e^{-t}\mathrm{d}t &= [-e^{-t}t^{\alpha+1}]_0^\infty - \int_0^\infty -e^{-t}(\alpha+1)t^\alpha\mathrm{d}t \\ &= 0 + (\alpha+1)\int_0^\infty t^\alpha e^{-t}\mathrm{d}t = (\alpha+1)\Gamma(\alpha).\end{aligned}$$

Therefore,

$$\begin{aligned}&\Gamma(1) = (0+1)\Gamma(0) = 1\\ &\Gamma(2) = (1+1)\Gamma(1) = 2\cdot 1 = 2!\\ &\Gamma(3) = (2+1)\Gamma(2) = 3\cdot 2! = 3!\\ &\vdots\\ &\Gamma(n) = n! \quad (n \in \mathbb{N}).\end{aligned}$$

The values $\Gamma(\alpha)$, $\alpha \notin \mathbb{N}$ can also be calculated (approximately).

10. Calculate the mean value of $\sin^2_{|[0,2\pi]}$.

Solution:

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi}\sin^2 x\mathrm{d}x &= \frac{1}{2\pi}\int_0^{2\pi}\frac{1-\cos 2x}{2}\mathrm{d}x\\ &= \frac{1}{2\pi}\left[\frac{1}{2}x-\frac{\sin 2x}{4}\right]_0^{2\pi} = \frac{1}{2\pi}\frac{1}{2}2\pi = \frac{1}{2}.\end{aligned}$$

The mean value of $\sin^2_{|[0,2\pi]}$ is $\frac{1}{2}$.

Chapter 10

Sequences and series of functions

This is a supplementary chapter. Some problems arising in practice (such as the approximation of functions or the approximate solution of ordinary and partial differential equations) necessitate the forthcoming concepts and results. The following topics will be discussed.

- The domain of convergence of a sequence of functions
- Pointwise and uniform convergence
- Continuity, differentiability and integrability of the limit function
- Convergence of a series of functions
- Weierstrass' majority criterion
- Continuity, differentiability and integrability of the sum of a series of functions
- Power series
- The Cauchy–Hadamard theorem
- Differentiability of the sum of a series, Abel's theorem

10.1 Sequences and series of functions

10.1.1 Sequences of functions

Let $H \subset \mathbb{R}$, $H \neq \emptyset$ be a set, and $\phi_1, \phi_2, \ldots, \phi_n, \ldots$ functions for which $\phi_n : H \to \mathbb{R}$ $(n \in \mathbb{N})$. Then we say that the sequence of functions (ϕ_n) is "defined on the set H".

For example,

1. (id^n) [here $D(\mathrm{id}^n) = \mathbb{R}$, $n \in \mathbb{N}$];

2. if $n \in \mathbb{N}$, then $\phi_n : [0,1] \to \mathbb{R}$, $\phi_n(x) := \begin{cases} 0 & \text{if } x = 0, \\ 1 & \text{if } 0 < x < \frac{1}{n}, \\ 0 & \text{if } \frac{1}{n} \le x \le 1; \end{cases}$

3. if $n \in \mathbb{N}$, then see Fig. 10.1;

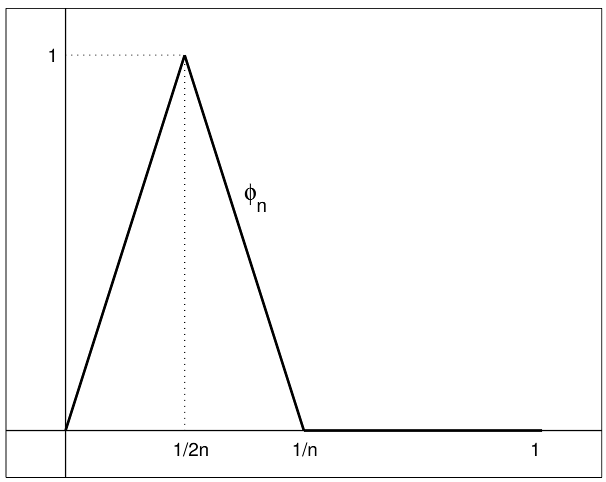


Figure 10.1

4. if $n \in \mathbb{N}$, then see Fig. 10.2;

5. * $(\sum_{k=1}^{n} \sin \circ (k \cdot \mathrm{id}))$ [here again $\mathbb{R}$ is the domain of definition of each function].

It is interesting to raise the question of whether the members of the sequence of functions(ϕ_n) get closer to some function if n is increased.

Definition 10.1. The **convergence domain** of the sequence of functions (ϕ_n) $[D(\phi_n) = H$, $n \in \mathbb{N}]$ is

$$KH(\phi_n) := \{x \in H \mid \text{the number sequence } (\phi_n(x)) \text{ is convergent}\}.$$

In Example 1, $KH(\mathrm{id}^n) = (-1, 1]$, because if $-1 < x < 1$, then $\mathrm{id}^n(x) = x^n \to 0$; if $x := 1$, then $\mathrm{id}^n(1) = 1 \to 1$, but if $x > 1$ or $x \le -1$, then $(\mathrm{id}^n(x))$ does not converge.

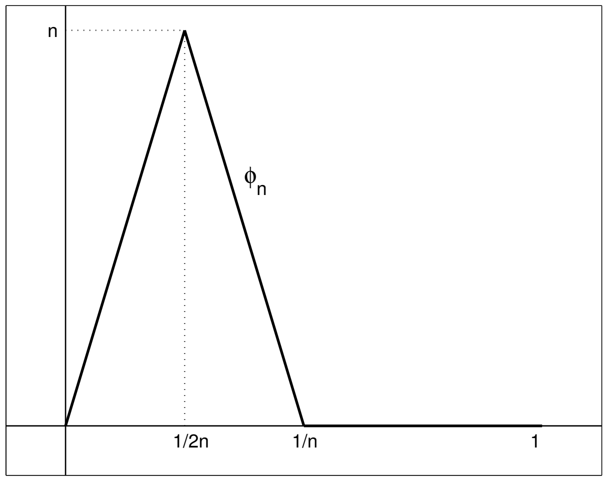


Figure 10.2

In Example 2, $KH(\phi_n) = [0,1]$, since $\phi_n(0) = 0 \to 0$. If $0 < x < 1$, then there exists an N such that $\frac{1}{N} < x$, and then the sequence $(\phi_n(x))$ is $1, 1, \ldots, 1, 0, 0, 0, \ldots, 0, \ldots$, which tends to 0 (if $n \le N$, then $\phi_n(x) = 1$, if $n > N$, then $\phi_n(x) = 0$).

The same holds for Examples 3 and 4.

Example 5* is a bit more difficult. If $x = l\pi$ ($l \in \mathbb{Z}$), then the number sequence is $(\sum_{k=1}^{n} \sin(kl\pi)) = (0)$, which tends to 0. So

$$KH\left(\sum_{k=1}^{n} \sin\circ(k \cdot \mathrm{id})\right) \supset \{l\pi \mid l \in \mathbb{Z}\}.$$

Should there be further $x \in \mathbb{R}$, $x \neq l\pi$ ($l \in \mathbb{Z}$) in the convergence domain of the function sequence, the sequence $(\sin kx)$ should converge to 0. Assume that $\sin kx \to 0$. Then $\sin(k+1)x \to 0$ would also hold, that is,

$$\sin(kx + x) = \sin kx \cos x + \cos kx \sin x \to 0.$$

Since $\sin x \neq 0$, $\sin kx \to 0$, therefore $\cos kx \to 0$ is also true. In this manner, $\sin^2 kx + \cos^2 kx \to 0$ would also hold, but it is impossible, since $\sin^2 kx + \cos^2 kx = 1$. Thus,

$$KH\left(\sum_{k=1}^{n} \sin\circ(k \cdot \mathrm{id})\right) = \{l\pi \mid l \in \mathbb{Z}\}.$$

This is an important example because the topic of Fourier series gives rise to several difficulties of the same kind.

Definition 10.2. Let (ϕ_n) be a sequence of functions defined on a set H. Assume that $KH(\phi_n) \neq \emptyset$. The **limit function** of the sequence of functions(ϕ_n) is the function $f : KH(\phi_n) \to \mathbb{R}$ for which

$$f(x) := \lim \phi_n(x)$$

for all $x \in KH(\phi_n)$

The notation $f := \lim \phi_n$ is often used.

In Example 1

$$\lim \mathrm{id}^n : (-1,1] \to \mathbb{R},\ (\lim \mathrm{id}^n)(x) := \begin{cases} 0 & \text{if } -1 < x < 1, \\ 1 & \text{if } x = 1. \end{cases}$$

In Examples 2, 3 and 4

$$\lim \phi_n : [0,1] \to \mathbb{R},\ (\lim \phi_n)(x) := 0.$$

In Example 5*

$$\lim \sum_{k=1}^{n} \sin \circ (k \cdot \mathrm{id}) = \sum_{k=1}^{\infty} \sin \circ (k \cdot \mathrm{id}) : \{l\pi \mid k \in \mathbb{Z}\} \to \mathbb{R},$$

$$\left(\lim \sum_{k=1}^{n} \sin \circ (k \cdot \mathrm{id}) \right)(x) = \sum_{k=1}^{\infty} \sin kx := 0.$$

When we compare the properties of the members of a sequence of functions with the properties of the limit function in the examples, we can find some interesting differences. In Example 1, id^n is continuous and differentiable on $\mathbb{R}$, while $\lim \mathrm{id}^n$ is not continuous, and so of course not differentiable, either. In Example 2 none of the functions ϕ_n is continuous, however, $\lim \phi_n$ is not only continuous, but also differentiable. In Examples 3 and 4 both the members ϕ_n and $\lim \phi_n$ are continuous, ϕ_n is not differentiable, and $\lim \phi_n$ is smooth. But here we should notice another exciting difference:

In Example 3

$$\int_0^1 \phi_n = \frac{1}{2}\frac{1}{n} \cdot 1 = \frac{1}{2n} \to 0, \quad \text{and} \quad \int_0^1 \lim \phi_n = \int_0^1 0 = 0,$$

while in Example 4

$$\int_0^1 \phi_n = \frac{1}{2}\frac{1}{n} \cdot n = \frac{1}{2} \to \frac{1}{2}, \quad \text{but} \quad \int_0^1 \lim \phi_n = \int_0^1 0 = 0,$$

so, the limit of the sequence consisting of the integrals of the members in the sequence of functions is not the integral of the limit function.

In Example 5* the members of the sequence of functions are nice trigonometric, smooth and periodic functions on the whole $\mathbb{R}$, while the limit function has rather poor properties, it is only periodic...

The above examples show that the concept of “pointwise convergence” is not sufficient for the nice properties of the members of the sequence of functions to be inherited by the limit function. We will try to do something about this.

Definition 10.3. Let (ϕ_n) be a sequence of functions defined on the set $H \neq \emptyset$. Assume that $KH(\phi_n) \neq \emptyset$, and let $E \subset KH(\phi_n)$, $E \neq \emptyset$. We say that the sequence of functions (ϕ_n) is **uniformly convergent** on the set E if for any error bound $\varepsilon > 0$ there exists an index $N \in \mathbb{N}$ such that $\forall n > N$ and for all $x \in E$

$$|\phi_n(x) - (\lim \phi_n)(x)| < \varepsilon.$$

Let us denote this fact by $\phi_n \hookrightarrow_E \lim \phi_n$.

This means that the index N is independent of x, that is why this kind of convergence is called uniform.

In Example 1 the convergence of (id^n) is not uniform on the set $(-1, 1]$. But not even on $(-1, 1)$! However, if $\delta > 0$, then on the interval $E := [-1 + \delta, 1 - \delta]$

$$\mathrm{id}^n \hookrightarrow_E 0.$$

In Examples 2, 3 and 4 the convergence is not uniform on the interval $[0, 1]$, but for $\delta > 0$

$$\phi_n \hookrightarrow_{[\delta,1]} 0$$

already holds.

Although in Example 5* the convergence is uniform on the set $E := KH(\sum_{k=1}^n \sin \circ (k \cdot \mathrm{id}))$, it is of little avail...

What are the consequences of the uniform convergence of a sequence of functions? We can put an end to the confusion seen in the examples.

Theorem 10.1. *Let $\phi_n \in C[a, b]$ $(n \in \mathbb{N})$. Assume that $\phi_n \hookrightarrow_{[a,b]} f$. Then $f \in C[a, b]$.*

Theorem 10.2. *Let $I \subset \mathbb{R}$ be an open interval, $\phi_n : I \to \mathbb{R}$ $(n \in \mathbb{N})$. Assume that there is an $x_0 \in I$ such that $(\phi_n(x_0))$ is convergent. Suppose ϕ_n is continuously differentiable on the interval I ($\phi_n \in C_1(I)$, $n \in \mathbb{N}$) and $\phi_n' \hookrightarrow_I g$. Then the sequence of functions (ϕ_n) is also uniformly convergent to a function $f : I \to \mathbb{R}$ on the interval I ($\phi_n \hookrightarrow_I f$), and $f \in D(I)$, moreover, $f'(x) = g(x) = \lim \phi_n'(x)$ for all $x \in I$.*

Briefly, the theorem states that

$$(\lim \phi_n)' = \lim \phi_n',$$

so the limit and the differentiation are interchangeable.

Theorem 10.3. *Let $\phi_n \in R[a,b]$, $(n \in \mathbb{N})$. Assume that $\phi_n \hookrightarrow_{[a,b]} f$. Then $f \in R[a,b]$, and $\lim \int_a^b \phi_n = \int_a^b f$.*

To put it more briefly, in case of uniform convergence

$$\lim \int_a^b \phi_n = \int_a^b \lim \phi_n,$$

that is, the limit and the integration are interchangeable.

10.1.2 Series of functions

The concepts defined for sequences of functions can be extended to series of functions with some obvious modifications.

Definition 10.4. Let (ϕ_n) be a sequence of functions defined on the set $H \neq \emptyset$. Let $S_n := \phi_1 + \phi_2 + \ldots + \phi_n$ $(n \in \mathbb{N})$ be the nth partial sum. The **series of functions** contributing to the sequence of functions (ϕ_n) is defined as the sequence of functions (S_n), that is, $\sum \phi_n := (S_n)$.

Definition 10.5. $KH \sum \phi_n := KH(S_n)$.

One can see that $KH \sum \phi_n = \{x \in H \mid \sum \phi_n(x) \text{ is convergent}\}$.

Definition 10.6. Assume that $KH \sum \phi_n \neq \emptyset$. We call

$$\sum_{n=1}^{\infty} \phi_n : KH \sum \phi_n \to \mathbb{R}, \ \left(\sum_{n=1}^{\infty} \phi_n\right)(x) := \sum_{n=1}^{\infty} \phi_n(x)$$

the **sum function** of the series of functions.

Clearly, $\left(\sum_{n=1}^{\infty} \phi_n\right)(x) = \lim S_n(x)$ for all $x \in KH(S_n)$.
For example, in case $\phi_n := \mathrm{id}^n$ $(n \in \mathbb{N})$, at any point $x \in \mathbb{R}$

$$S_n(x) := x + x^2 + \ldots + x^n = \begin{cases} x\frac{x^n-1}{x-1} & \text{if } x \neq 1, \\ n & \text{if } x = 1. \end{cases}$$

Therefore, $\lim S_n(x) = \frac{x}{1-x}$ if $x \in (-1,1)$; if $x \in \mathbb{R} \setminus (-1,1)$, then $(S_n(x))$ is divergent. Thus, $KH \sum \mathrm{id}^n = (-1,1)$, and $(\sum_{n=1}^{\infty} \mathrm{id}^n)(x) = \frac{x}{1-x}$.

Definition 10.7. Let (ϕ_n) be such a sequence of functions for which

$$KH\sum \phi_n \neq \emptyset.$$

Let $E \subset KH\sum \phi_n$. We say that $\sum \phi_n$ is **uniformly convergent** on the set E if the sequence of partial sums (S_n) is uniformly convergent on the set E.

Notation: $\sum \phi_n \hookrightarrow_E \sum_{n=1}^{\infty} \phi_n$.

A useful sufficient condition for the uniform convergence of a series of functions:

Theorem 10.4 (Weierstrass' majorant criterion)**.** *Let (ϕ_n) be a sequence of functions defined on the set $H \neq \emptyset$ for which there is a sequence $(a_n) \subset \mathbb{R}^+$ of positive numbers such that $|\phi_n(x)| \leq a_n$ $(n \in \mathbb{N})$ for all $x \in H$, and also $\sum a_n$ is convergent. Then*

$$\sum \phi_n \hookrightarrow_H \sum_{n=1}^{\infty} \phi_n.$$

Due to the conditions for the uniform convergence of the sequence of partial sums of the series of functions, the following theorems (sketchily given here) are valid:

- If $\phi_n \in C[a,b]$, and $\sum \phi_n \hookrightarrow_{[a,b]} \sum_{n=1}^{\infty} \phi_n$, then $\sum_{n=1}^{\infty} \phi_n \in C[a,b]$.
- If $\phi_n \in C_1(I)$, and $\sum \phi_n' \hookrightarrow_I g$, then $\sum_{n=1}^{\infty} \phi_n \in C_1(I)$, and

  $$\left(\sum_{n=1}^{\infty} \phi_n\right)' = \sum_{n=1}^{\infty} \phi_n' = g.$$

  (The summation is interchangeable with the differentiation.)
- If $\phi_n \in R[a,b]$, and $\sum \phi_n \hookrightarrow_{[a,b]} \sum_{n=1}^{\infty} \phi_n$, then $\sum_{n=1}^{\infty} \phi_n \in R[a,b]$, and

  $$\int_a^b \sum_{n=1}^{\infty} \phi_n = \sum_{n=1}^{\infty} \int_a^b \phi_n.$$

  (A series of functions can be integrated termwise.)

10.1.3 Power series

Power series are special series of functions.

Definition 10.8. Let $c_0, c_1, c_2, \dots, c_n, \dots$ be a number sequence, and $a \in \mathbb{R}$ a number. The series of functions

$$\sum c_n(\mathrm{id} - a)^n$$

is called **power series** with sequence of coefficients (c_n) and "center" $a \in \mathbb{R}$.

Let us consider $a := 0$ in the sequel for the sake of simpler formulation.

Theorem 10.5 (The Cauchy–Hadamard theorem)**.** *Let $\sum c_n id^n$ be a power series.*

1^o *If $(\sqrt[n]{|c_n|})$ is bounded, and* $\limsup \sqrt[n]{|c_n|} \neq 0$*, then let*

$$R := \frac{1}{\limsup \sqrt[n]{|c_n|}}$$

(R is called **convergence radius** *of the power series).*

Then

$$(-R, R) \subset KH \sum c_n id^n \subset [-R, R].$$

2^o *If $(\sqrt[n]{|c_n|})$ is not bounded above, then*

$$KH \sum c_n id^n = \{0\}.$$

3^o *If* $\limsup \sqrt[n]{|c_n|} = 0$*, then*

$$KH \sum c_n id^n = \mathbb{R}.$$

One can see that the convergence domain of a power series is always an interval (in case 2^o it is a degenerate interval), but in case 1^o this interval can either be $(-R, R)$, $(-R, R]$, $[-R, R)$ or $[-R, R]$. Compare the convergence domain of a power series with the convergence domain of the series of functions $\sum \sin \circ (k \cdot \mathrm{id})$ of Example 5*, the difference is salient.

Theorem 10.6. *Let $\sum c_n id^n$ be a power series for which $(\sqrt[n]{|c_n|})$ is bounded above. Then the sum function*

$$f : KH \sum c_n id^n \to \mathbb{R}, \quad f(x) := \sum_{n=0}^{\infty} c_n x^n$$

is not only continuous, but also differentiable on the open interval

$$\mathrm{int}\ KH \sum c_n id^n;$$

what is more, for all $x \in \operatorname{int} KH \sum c_n \mathit{id}^n$

$$f'(x) = \sum_{n=0}^{\infty} n c_n x^{n-1}.$$

This theorem implies that $(\sqrt[n]{|nc_n|}) = (\sqrt[n]{n}\sqrt[n]{|c_n|})$ is bounded above if $(\sqrt[n]{|c_n|})$ is bounded above, moreover, the convergence domain of the power series $\sum nc_n \mathrm{id}^{n-1}$ remains the same as the convergence domain of $\sum c_n \mathrm{id}^n$. Therefore, the sum function of this power series is also differentiable, and

$$f''(x) = \sum_{n=2}^{\infty} n(n-1) c_n x^{n-2}$$

for all $x \in \operatorname{int} KH \sum c_n \mathrm{id}^n$.

This train of thought can be continued:

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\dots(n-k+1) c_n x^{n-k}, \quad x \in \operatorname{int} KH \sum c_n \mathrm{id}^n.$$

Note that $f(0) = c_0,\ f'(0) = c_1\ \dots,\ f^{(k)}(0) = k! c_k, \dots$

Theorem 10.7 (Abel)**.** *Let* $\sum c_n \mathit{id}^n$ *be a power series with convergence radius* $R > 0$*. Assume that* $\sum c_n R^n$ *is convergent. Then* $f \in C[R]$*, that is,* $\lim_{x\to R} f(x) = \sum_{n=0}^{\infty} c_n R^n$.

10.2 Exercises

1. Find the domains of convergence for the following power series:

$$\sum \mathrm{id}^n, \quad \sum \frac{1}{n}\mathrm{id}^n, \quad \sum \frac{(-1)^n}{n}\mathrm{id}^n, \quad \sum \frac{1}{n^2}\mathrm{id}^n, \quad \sum \frac{1}{5^n n}\mathrm{id}^n,$$

$$\sum \frac{1}{n!}\mathrm{id}^n, \quad \sum n^n \mathrm{id}^n.$$

Solution: $\limsup \sqrt[n]{1} = 1$, $\limsup \sqrt[n]{\frac{1}{n^2}} = 1$, $\limsup \sqrt[n]{\frac{1}{n!}} = 0$,

$\limsup \sqrt[n]{\frac{1}{n}} = 1$, $\limsup \sqrt[n]{\frac{1}{5^n n}} = \frac{1}{5}$, $(\sqrt[n]{n^n}) = (n)$ is not bounded above.

$$KH \sum \mathrm{id}^n = (-1, 1),$$

$$KH \sum \frac{1}{n}\mathrm{id}^n = [-1, 1) \quad \text{(Leibniz's theorem)},$$

$$KH\sum\frac{(-1)^n}{n}\mathrm{id}^n = (-1,1],$$
$$KH\sum\frac{1}{n^2}\mathrm{id}^n = [-1,1],$$
$$KH\sum\frac{1}{5^n n}\mathrm{id}^n = [-5,5),$$
$$KH\sum\frac{1}{n!}\mathrm{id}^n = \mathbb{R},$$
$$KH\sum n^n\mathrm{id}^n = \{0\}.$$

2.
$$1 + x + x^2 + x^3 + \ldots + x^n + \ldots = \frac{1}{1-x} \text{ if } |x| < 1.$$

Is it true that

$$1 + 2x + 3x^2 + \ldots + nx^{n-1} + \ldots = \left(\frac{1}{1-x}\right)' = \frac{1}{(1-x)^2} \text{ if } |x| < 1?$$

Is it true that

$$x + 2x^2 + 3x^3 + \ldots + nx^n + \ldots = \frac{x}{(1-x)^2} \text{ if } |x| < 1?$$

Calculate the sum

$$1 + 2^2x + 3^2x^2 + \ldots + n^2x^{n-1} + \ldots \text{ if } |x| < 1.$$

3.
$$1 + x + x^2 + \ldots + x^n + \ldots = \frac{1}{1-x} \text{ if } |x| < 1.$$

Let $x := -t$. Then

$$1 - t + t^2 - \ldots + (-1)^n t^n + \ldots = \frac{1}{1+t} \text{ if } |t| < 1.$$

Is it true that

$$t - \frac{t^2}{2} + \frac{t^3}{3} - \ldots + (-1)^n\frac{t^{n+1}}{n+1} + \ldots = \ln(1+t) \text{ if } |t| < 1?$$

Is it true that

$$1 - \frac{1}{2} + \frac{1}{3} - \ldots + (-1)^n\frac{1}{n+1} + \ldots = \ln 2?$$

4.
$$1 + x + x^2 + \ldots + x^n + \ldots = \frac{1}{1-x} \text{ if } |x| < 1.$$

Let $x := -t^2$. Then

$$1 - t^2 + t^4 - \ldots + (-1)^n t^{2n} + \ldots = \frac{1}{1+t^2} \text{ if } |t| < 1.$$

Is it true that

$$t - \frac{t^3}{3} + \frac{t^5}{5} - \ldots + (-1)^n \frac{t^{2n+1}}{2n+1} + \ldots = \operatorname{arctg} t \text{ if } |t| < 1?$$

Is it true that

$$1 - \frac{1}{3} + \frac{1}{5} - \ldots + (-1)^n \frac{1}{2n+1} + \ldots = \frac{\pi}{4}?$$

Chapter 11

Multivariable functions

There are numerous phenomena that cannot be described by real functions. Therefore we generalize our concepts, introduced so far for real functions. The following topics will be discussed.

- Operations on vectors and matrices
- The concept and illustration of multivariable functions
- The limit of a vector sequence
- The limit and continuity of a multivariable function

11.1 Multivariable functions

11.1.1 The n-dimensional space

In the Linear Algebra course we learnt about the vector space $\mathbb{R}^n$. If $x \in \mathbb{R}^n$ is a vector, then $x = (x_1, x_2, \ldots, x_n)$, where $x_i \in \mathbb{R}$ is the ith coordinate of x. The **norm** (length) of a vector x is

$$\|x\| := \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} \in \mathbb{R}.$$

The norm of a vector satisfies the following properties:

1^o $\|x\| \geq 0$, and $\|x\| = 0 \Leftrightarrow x = 0 \in \mathbb{R}^n$,

2^o $\lambda \in \mathbb{R} \quad \|\lambda x\| = |\lambda| \|x\|$,

3^o $\|x + y\| \leq \|x\| + \|y\|$.

Let $e_i := (0, \ldots, 1^{i)}, \ldots 0) \in \mathbb{R}^n$ be the ith unit vector ($\|e_i\| = 1$), $i = 1, 2, \ldots, n$. Then $x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$.

The **inner product** of the vectors $a, b \in \mathbb{R}^n$ is defined as the number

$$\langle a, b\rangle := a_1 b_1 + a_2 b_2 + \ldots + a_n b_n \in \mathbb{R}$$

The properties of the inner product:

1^o $\langle a, b\rangle = \langle b, a\rangle$,

2^o $\langle a + b, c\rangle = \langle a, c\rangle + \langle b, c\rangle$,

3^o if $\lambda \in \mathbb{R}$, $\langle \lambda a, b\rangle = \langle a, \lambda b\rangle = \lambda\langle a, b\rangle$,

4^o $\langle a, a\rangle = \|a\|^2 \geq 0$,

5^o $|\langle a, b\rangle| \leq \|a\| \cdot \|b\|$ (Cauchy–Bunyakovsky–Schwarz inequality).

The vectors a and b are **orthogonal** (perpendicular) if $\langle a, b\rangle = 0$.

We are familiar with matrices, too. If A is a matrix of m rows and n columns, then $A \in \mathbb{R}^{m\times n}$, where the jth element of the ith row is a_{ij}.

Let $A, B \in \mathbb{R}^{m\times n}$ and $\lambda \in \mathbb{R}$. Then $C := A + B \in \mathbb{R}^{m\times n}$, where $c_{ij} = a_{ij} + b_{ij}$, and $D := \lambda A \in \mathbb{R}^{m\times n}$, where $d_{ij} = \lambda a_{ij}$.

If $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times p}$, then $S := A \cdot B \in \mathbb{R}^{m\times p}$, where $s_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$.

Let us agree that we identify the vector space $\mathbb{R}^n$ with the space of column matrices $\mathbb{R}^{n\times 1}$, which implies that the vector $x \in \mathbb{R}^n$, $x = (x_1, x_2, \ldots, x_n)$ is identified with the column matrix

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n\times 1}.$$

(We do not even make any distinction in their notation, what is more, we say vector, but write a column matrix.) For example, if $a, b \in \mathbb{R}^n$, then their scalar product can be considered in the form

$$\langle a, b\rangle = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n = [a_1\ a_2\ \ldots\ a_n] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

too, that is, as the matrix product of a row matrix of $\mathbb{R}^{1\times n}$ and a column matrix of $\mathbb{R}^{n\times 1}$.

11.1.2 Multivariable functions

Let $f : \mathbb{R}^n \supset\!\!\!\to \mathbb{R}^k$ be a function of n variables with vector values of dimension k. If $x \in D(f)$ and $x = (x_1, x_2, \ldots, x_n)$, then $f(x) \in \mathbb{R}^k$, and $f(x) = (f_1(x), f_2(x), \ldots, f_k(x))$, where $f_i : \mathbb{R}^n \rightarrowtail \mathbb{R}$ is the ith coordinate function of f $(i = 1, 2, \ldots, k)$. Such a function f can be given in the form

$$f = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_k \end{bmatrix}.$$

For example, let

$$f : \mathbb{R}^2 \to \mathbb{R}^3, \quad f(x_1, x_2) := \begin{bmatrix} \sin(x_1 x_2) \\ x_1 + x_2 \\ x_2 \end{bmatrix}.$$

Here $f_1 : \mathbb{R}^2 \to \mathbb{R},\ f_1(x_1, x_2) := \sin(x_1 x_2)$ is the first coordinate function, while $f_2 : \mathbb{R}^2 \to \mathbb{R},\ f_2(x_1, x_2) := x_1 + x_2$ and $f_3 : \mathbb{R}^2 \to \mathbb{R},\ f_3(x_1, x_2) = x_2$ are the second and third coordinate functions, respectively.

Let us look at some special cases.

1^o $n = 1,\ k = 1,\ f \in \mathbb{R} \rightarrowtail \mathbb{R}$ is a real function that we have discussed so far.

2^o $n > 1,\ k = 1,\ f \in \mathbb{R}^n \to \mathbb{R}$ is an n-variable, real-valued or scalar-valued function. It can be illustrated for $n = 2$ as follows. On a line perpendicular to the plane at the point $(x_1, x_2) \in D(f)$ we measure the number $f(x_1, x_2) \in \mathbb{R}$. The points obtained in this manner form a **surface** (Fig. 11.1). Another way of illustration for the case $n = 2$ is as follows. Let $c \in \mathbb{R}$ and

$$N_c := \{(x_1, x_2) \in D(f) \mid f(x_1, x_2) = c\}.$$

N_c is the **level curve** of the function f corresponding to the value c. Plotting the contour lines for a few values $c_1 < c_2 < \ldots < c_s$ tells a lot about the function f. In cartography this is called contour map (Fig. 11.2).

3^o $n := 1,\ k > 1,\ r \in \mathbb{R} \rightarrowtail \mathbb{R}^k$ is a vector-valued function in k dimensions of a single variable.

Such a function can be illustrated for $k = 3$ as follows. Let $D(r) := [\alpha, \beta]$. To a parameter value $t \in [\alpha, \beta]$ we assign a point in $\mathbb{R}^3$ given by

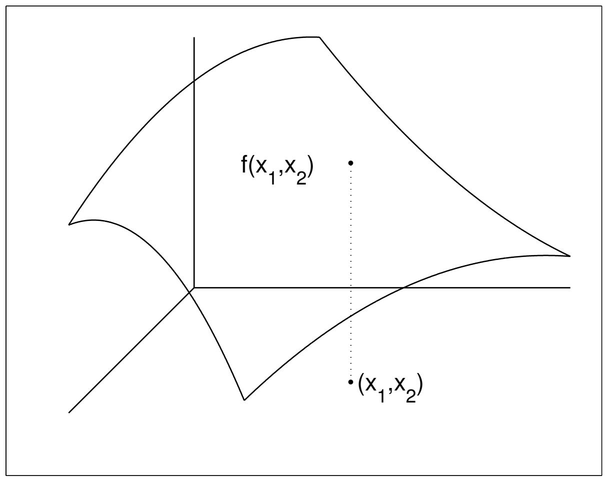


Figure 11.1

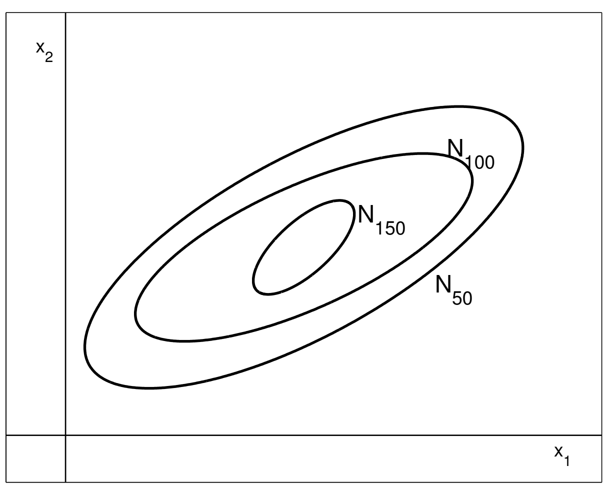


Figure 11.2

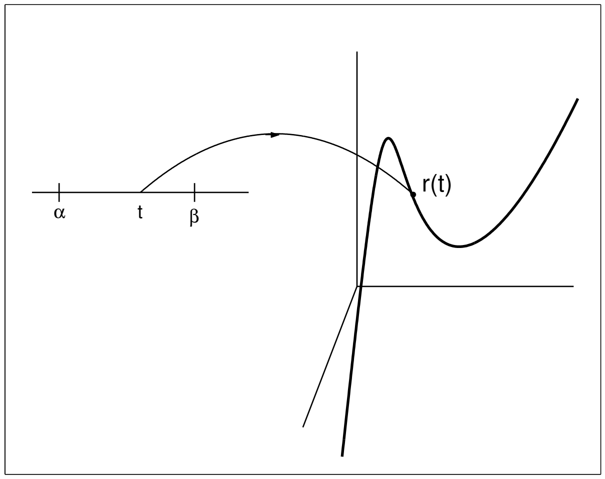


Figure 11.3

the coordinates $r(t) := (x(t), y(t), z(t))$. The points obtained in this way form a **space curve** (Fig. 11.3). (Note that the space curve is the range of the function r!)

For example,

$$r : [0, 6\pi] \to \mathbb{R}^3,\ r(t) := \begin{bmatrix} 2\cos t \\ 2\sin t \\ 0.5t \end{bmatrix}$$

is the section of a helix of height 3π that spirals up around a cylinder of radius 2 (Fig. 11.4).

4^o $n > 1,\ k > 1,\ f \in \mathbb{R}^n \rightarrowtail \mathbb{R}^k$ is a multivariable function of vector values.

When we assign the wind velocity vector to each point of the atmosphere:

$$v \in \mathbb{R}^3 \rightarrowtail \mathbb{R}^3,$$

then we obtain the so-called velocity function (or velocity field). The gravitational field of a mass (for example, of a star) can also be defined by assigning a vector, namely, the gravitational force, acting at the given point, to each point of the space, and so a function $g \in \mathbb{R}^3 \rightarrowtail \mathbb{R}^3$ describes the gravitational field of a mass (or a star).

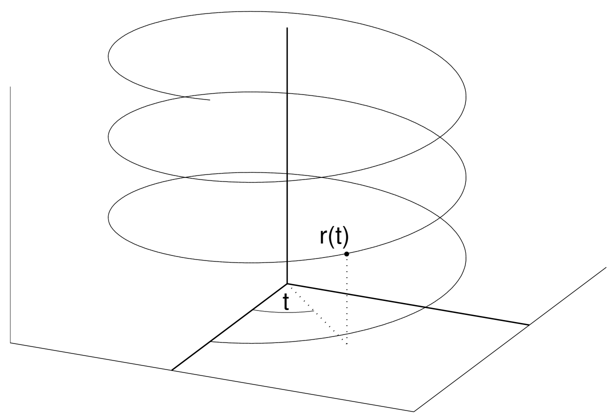


Figure 11.4

11.1.3 Limit and continuity

Let us examine what properties can be extended to such functions.

Let $a := (a_1, a_2, \dots, a_m) : \mathbb{N} \to \mathbb{R}^m$ be a vector sequence. The vector sequence (a_n) is convergent if it gets arbitrarily close to some point, more precisely:

Definition 11.1. We say that the vector sequence (a_n) is **convergent** if there exists a vector $A \in \mathbb{R}^m,\ A = (A_1, A_2, \dots, A_m)$ such that for any error bound $\varepsilon > 0$ there exists an index N such that for all $n > N$

$$\|a_n - A\| < \varepsilon.$$

This will also be denoted by $\lim a_n = A$ or $a_n \to A$. It is easy to see that

$$\|a_n - A\| < \varepsilon \Leftrightarrow |a_{in} - A_i| < \frac{\varepsilon}{\sqrt{m}}, \quad i = 1, 2, \dots, m,$$

so, a vector sequence is convergent if and only if each of its coordinate sequences (number sequences) are convergent. For example, the vector sequence $\left(\left(\frac{1}{n}, \frac{n}{n+1}\right)\right)$ is convergent because $\frac{1}{n} \to 0$, $\frac{n}{n+1} \to 1$, therefore $\lim\left(\frac{1}{n}, \frac{n}{n+1}\right) = (0, 1)$. The vector sequence $((\frac{1}{n}, (-1)^n))$ is not convergent (divergent) because $((-1)^n)$ is not convergent.

Let $f \in \mathbb{R}^n \rightarrowtail \mathbb{R}^k$ and $a \in D(f)$. The function f is called continuous at the point a if at points close to a the function values are close to $f(a)$, more precisely:

Definition 11.2. We say that f is continuous at the point a if for all error bounds $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $x \in D(f)$, $\|x - a\| < \delta$: $\|f(x) - f(a)\| < \varepsilon$.

This will also be denoted by $f \in C[a]$.

One can see that for $f = (f_1, f_2, \ldots, f_k)$ we have

$$\|f(x) - f(a)\| < \varepsilon \Leftrightarrow |f_i(x) - f_i(a)| < \frac{\varepsilon}{\sqrt{k}}, \quad i = 1, 2, \ldots, k,$$

and so f is continuous at a if and only if each of its coordinate functions is continuous at a. This can also be defined in terms of sequences:

Theorem 11.1. *Let $f \in \mathbb{R}^n \rightarrowtail \mathbb{R}^k$, $a \in D(f)$. Then $f \in C[a] \iff$ for all sequences $(x_n) \subset D(f)$, $x_n \to a : f(x_n) \to f(a)$.*

For example,

$$f : \mathbb{R}^2 \to \mathbb{R}^3, \; f(x_1, x_2) := \begin{bmatrix} x_1 x_2 \\ x_1^2 \\ x_2 \end{bmatrix}$$

is continuous at $a := (1, 3)$ because for any sequence $(x_{1n}, x_{2n}) \to (1, 3)$, $x_{1n} \cdot x_{2n} \to 1 \cdot 3$, $(x_{1n})^2 \to 1^2$ and $x_{2n} \to 3$, therefore $f(x_{1n}, x_{2n}) \to f(1, 3)$. So, $f \in C[(1, 3)]$.

Sometimes we need matrix-valued functions. If $f \in \mathbb{R}^n \rightarrowtail \mathbb{R}^{k \times p}$, then $f_{ij} \in \mathbb{R}^n \rightarrowtail \mathbb{R}$ is the (i, j)th component of f. Let us call f continuous at the point $a \in D(f)$ if each of its components f_{ij} is continuous at a. (It is sufficient to consider that $\mathbb{R}^{k \times p}$ is identified with the vector space $\mathbb{R}^{kp}$.)

A matrix-valued function is

$$f : \mathbb{R}^2 \to \mathbb{R}^{2 \times 2}, \quad f(x_1, x_2) := \begin{bmatrix} x_1 x_2 & e^{x_1} \\ 0 & x_1 + x_2 \end{bmatrix}.$$

Moreover, this function f is continuous at every point $(a_1, a_2) \in \mathbb{R}^2$.

Let $a \in \mathbb{R}^n$ and $r > 0$. The neighborhood of radius r of the point a is defined as

$$K_r(a) := \{x \in \mathbb{R}^n \mid \|x - a\| < r\}.$$

Let $H \subset \mathbb{R}^n$ and $a \in \mathbb{R}^n$. The point a is called **accumulation point** of the set H if in any neighborhood $K(a)$ of a there are infinitely many points of H. Let us denote this by $a \in \dot{H}$.

Let $f \in \mathbb{R}^n \rightarrowtail \mathbb{R}^k$ and $a \in (D(f))^{\cdot}$.

Definition 11.3. We say that the function f has a **limit** at the point a if there exists an $A \in \mathbb{R}^k$ such that for all error bounds $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $x \in D(f)$, $\|x - a\| < \delta$, $x \neq a$:

$$\|f(x) - A\| < \varepsilon.$$

This will be denoted by $\lim_a f = A$ or $\lim_{x\to a} f(x) = A$ or $f(x) \to A$ whenever $x \to a$.

It is easy to see that the function $f \in \mathbb{R}^n \rightarrowtail \mathbb{R}^k$ has a limit at the point $a \in (D(f))'$ if and only if each of its coordinate functions $f_i \in \mathbb{R}^n \rightarrowtail \mathbb{R}$ has a limit at a.

The following theorem is again valid:

Theorem 11.2. $\lim_a f = A \iff$ *for all* $(x_n) \subset D(f),\ x_n \to a,\ x_n \neq a$: $f(x_n) \to A$.

11.2 Exercises

1. Verify the Cauchy–Bunyakovsky–Schwarz inequality: for all vectors $a, b \in \mathbb{R}^n$, $a = (a_1, a_2, \ldots, a_n)$ and $b = (b_1, b_2, \ldots, b_n)$

$$|\langle a, b\rangle| \leq \|a\| \cdot \|b\|,$$

or

$$|a_1b_1 + a_2b_2 + \ldots + a_nb_n| \leq \sqrt{a_1^2 + a_2^2 + \ldots + a_n^2} \cdot \sqrt{b_1^2 + b_2^2 + \ldots + b_n^2}.$$

Solution:

1^o If $b = (0, 0, \ldots, 0)$, then the statement is obviously true.

2^o If $b \neq 0$, then for all $\lambda \in \mathbb{R}$

$$\begin{aligned} 0 \leq \langle a + \lambda b, a + \lambda b\rangle &= \langle a, a\rangle + 2\langle a, b\rangle\lambda + \langle b, b\rangle\lambda^2 \\ &= \|b\|^2\lambda^2 + 2\langle a, b\rangle\lambda + \|a\|^2. \end{aligned}$$

Due to the fact that $\|b\| \neq 0$, this is such a second degree polynomial which is nonnegative for all $\lambda \in \mathbb{R}$. Therefore its discriminant $D \leq 0$. Thus,

$$\begin{aligned} 4\langle a, b\rangle^2 - 4\|b\|^2\|a\|^2 &\leq 0, \\ \langle a, b\rangle^2 &\leq \|a\|^2\|b\|^2, \\ |\langle a, b\rangle| &\leq \|a\| \cdot \|b\|. \end{aligned}$$

2. Think over that the function $f : \mathbb{R}^2 \to \mathbb{R},\ f(x_1, x_2) := x_1^2 + x_2^2$ can be illustrated as a surface of revolution. How does the surface $g : \mathbb{R}^2 \to \mathbb{R}$, $g(x_1, x_2) := x_1^2 + x_2^2 - 2x_1 + 4x_2 + 1$ look like?

3. Let $F : \mathbb{R}^3 \setminus \{0\} \to \mathbb{R}^3,\ F(r) := -\frac{Mr}{\|r\|^3}$, where $r = (x_1, x_2, x_3),\ M > 0$. Find the coordinate functions of $F =: (P, Q, R)$.

4. Let

$$g : \mathbb{R}^2 \to \mathbb{R}^3, \quad g(x,y) := \begin{bmatrix} xy \\ x+y \\ x-y \end{bmatrix},$$

$$f : \mathbb{R}^3 \to \mathbb{R}, \quad f(u,v,w) := u^2v + w^3.$$

Find the composition $f \circ g$.

5. Let

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) := \begin{cases} \frac{xy}{x^2+y^2} & \text{if } x^2+y^2 \neq 0, \\ 0 & \text{if } x^2+y^2 = 0. \end{cases}$$

Does the function f have a limit at the point $(0,0) \in \mathbb{R}^2$?

Solution: First let $(x_n, y_n) := (\frac{1}{n}, 0)$ $(n \in \mathbb{N})$. $(x_n, y_n) \to (0,0)$, but $(x_n, y_n) \neq (0,0)$.

$$f(x_n, y_n) = \frac{\frac{1}{n} \cdot 0}{\frac{1}{n^2} + 0} = 0 \to 0.$$

If, however, $(x_n, y_n) := (\frac{1}{n}, \frac{1}{n})$ $(n \in \mathbb{N})$, then though $(x_n, y_n) \to (0,0)$ and $(x_n, y_n) \neq (0,0)$,

$$f(x_n, y_n) = \frac{\frac{1}{n} \cdot \frac{1}{n}}{\frac{1}{n^2} + \frac{1}{n^2}} = \frac{1}{2} \to \frac{1}{2}.$$

Since for two appropriate sequences, converging to $(0,0)$, the sequences of the function values converge to different limits, f itself does not have a limit at $(0,0)$.

6. Let

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) := \begin{cases} \frac{x^2y^2}{x^2+y^2} & \text{if } x^2+y^2 \neq 0, \\ 0 & \text{if } x^2+y^2 = 0. \end{cases}$$

Show that $f \in C[(0,0)]$.

Chapter 12

Differentiation of multivariable functions

We introduce the concepts of partial derivatives, the differentiability of multivariable functions and the directional derivative. We will compute extreme values with the aid of derivatives. The following topics will be discussed.

- Partial derivatives
- Differentiation of multivariable functions, the derivative matrix
- The relation between partial derivatives and the derivative matrix
- The tangent plane to a surface
- The tangent of a space curve
- Extreme values and their necessary condition
- Young's theorem
- Second derivative, Taylor's formula
- The sufficient condition of an extreme value

12.1 Multivariable differentiation

12.1.1 Partial derivatives

Let $f : \mathbb{R}^2 \supset\!\!\to \mathbb{R}$ be a function. Consider an internal point $a = (x, y) \in \operatorname{int} D(f)$ of the domain of definition. Draw a straight line through the point a parallel with the x axis, a point of which will be

$$(x + t, y),\ t \in \mathbb{R},$$

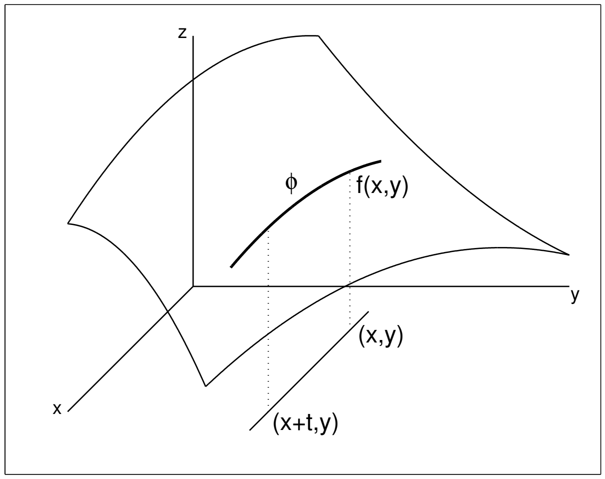


Figure 12.1

then take the values of the function at these points: $f(x+t, y)$. In this manner we have defined a real function $\phi : \mathbb{R} \supset\!\to \mathbb{R}$, $\phi(t) := f(x+t, y)$, the graph of which is a curve lying on the surface (Fig. 12.1).

Definition 12.1. We say that the function f is **partially differentiable with respect to the first variable** at the point (x, y) if ϕ is differentiable at $t = 0$.

If $\phi \in D[0]$, then the **partial derivative** of f with respect to the first variable at (x, y) is defined as $\phi'(0)$, that is,

$$\partial_1 f(x, y) := \phi'(0).$$

Keeping in mind the differentiability of real functions, the partial derivative is none other than

$$\partial_1 f(x, y) = \lim_{t \to 0} \frac{f(x+t, y) - f(x, y)}{t}.$$

One can see that the partial differentiability with respect to the first variable only means the smoothness of the curve on the surface at the point $t = 0$, and $\partial_1 f(x, y)$ gives the slope of this curve. Apparently,

$$\frac{f(x+t, y) - f(x, y)}{t} \approx \partial_1 f(x, y) \text{ if } t \approx 0,$$

which tells us that moving from (x, y) in the direction of the first axis

$$f(x+t, y) \approx f(x, y) + \partial_1 f(x, y) \cdot t \text{ if } t \approx 0.$$

Similarly, if we draw a straight line through the point (x, y) parallel with the y axis, then we get another surface curve $\psi : \mathbb{R} \supset\!\!\to \mathbb{R}$, $\psi(t) := f(x, y+t)$. If $\psi \in D[0]$, then f is partially differentiable with respect to the second variable at the point (x, y), and

$$\partial_2 f(x,y) := \psi'(0) = \lim_{t\to 0} \frac{f(x, y+t) - f(x,y)}{t}$$

will be the partial derivative of f with respect to the second variable at the point (x, y). The meaning of $\partial_2 f(x, y)$ can be given similarly.

Instead of $\partial_1 f(x, y)$ the notations $\frac{\partial f}{\partial x}(x, y)$, $f'_x(x, y)$ and $D_1 f(x, y)$ are also frequently used for the partial derivative. Corresponding notations are employed for $\partial_2 f(x, y)$, too.

Observe that when differentiating f partially with respect to the first variable, the second coordinate y does not change, it remains constant. That is why if we want to calculate the first partial derivative of, say, the function

$$f(x,y) := x^2y^3 + 2x + y \quad ((x,y) \in \mathbb{R}^2)$$

at an arbitrary point (x, y), then the variable y should be considered as constant during the differentiation, so

$$\partial_1 f(x,y) = 2xy^3 + 2 + 0 \quad ((x,y) \in \mathbb{R}^2).$$

In a similar manner, during partial differentiation with respect to the second variable, x should be considered as constant, that is

$$\partial_2 f(x,y) = x^2 3y^2 + 1 \quad ((x,y) \in \mathbb{R}^2).$$

Unfortunately, the partial differentiability of f, even with respect to both variables, does not even guarantee the continuity of f at that point. For example, for the function

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) := \begin{cases} 1 & \text{if } xy = 0, \\ 0 & \text{if } xy \neq 0. \end{cases}$$

$\partial_1 f(0, 0) = 0$ and $\partial_2 f(0, 0) = 0$, but $f \notin C[(0, 0)]$.

12.1.2 The derivative matrix

Now our aim is to create such a differentiability concept which will be an actual generalization of the differentiability of real functions.

Let $f : \mathbb{R}^2 \supset\!\!\to \mathbb{R}$, $(x, y) \in \operatorname{int} D(f)$.

Definition 12.2. We say that f is **differentiable** at the point (x, y) if there exist $A_1, A_2 \in \mathbb{R}$ and a function $\alpha : \mathbb{R}^2 \supset\!\!\to \mathbb{R}$ such that for any vector $h = (h_1, h_2) \in \mathbb{R}^2$ for which $(x + h_1, y + h_2) \in D(f)$, the equality

$$f(x + h_1, y + h_2) - f(x, y) = A_1 h_1 + A_2 h_2 + \alpha(h_1, h_2)$$

is satisfied, where

$$\lim_{h\to 0} \frac{\alpha(h)}{\|h\|} = 0.$$

The limit $\lim_{h\to 0} \frac{\alpha(h)}{\|h\|} = 0$ ensures that the remainder $\alpha(h)$ be "small". Clearly, $\lim_{h\to 0} \alpha(h) = 0$ holds too, but if we divide the values of $\alpha(h)$ by the small number $\|h\| \approx 0$, then we "magnify" the values of $\alpha(h)$, so, if even this quotient tends to zero, then $\alpha(h)$ will be really "small".

When h is of the form $h := (h_1, 0)$, then after rearranging and taking the limit

$$\lim_{h\to 0} \frac{f(x + h_1, y) - f(x, y)}{h_1} = \lim_{h_1\to 0} \left(A_1 + \frac{\alpha(h_1, 0)}{|h_1|} \right) = A_1,$$

which means that if f is differentiable at the point (x, y), then A_1 can only be $\partial_1 f(x, y)$.

Similarly, for vectors of the form $h := (0, h_2)$ we would obtain that A_2 can only be $\partial_2 f(x, y)$.

So, if f is differentiable at the point (x, y), then the change of the function $f(x + h_1, y + h_2) - f(x, y)$ can be well approximated by the "linear function"

$$\partial_1 f(x, y) h_1 + \partial_2 f(x, y) h_2,$$

what is more, the error of the approximation, $\alpha(h_1, h_2)$ is negligibly small: even the magnified quotient $\frac{\alpha(h)}{\|h\|}$ is close to 0 when $\|h\|$ is small.

By using matrices, the differentiability of f means that there exists a function $\alpha : \mathbb{R}^2 \supset\!\!\to \mathbb{R}$ such that

$$f(x + h_1, y + h_2) - f(x, y) = [\partial_1 f(x, y)\ \partial_2 f(x, y)] \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} + \alpha(h_1, h_2)$$

and

$$\lim_{h\to 0} \frac{\alpha(h)}{\|h\|} = 0.$$

The differentiability of the function f at the point $(x, y) \in \operatorname{int} D(f)$ will be denoted as $f \in D[(x, y)]$, and the derivative of f at this point as

$$f'(x, y) := [\partial_1 f(x, y)\ \partial_2 f(x, y)] \in \mathbb{R}^{1\times 2}.$$

If $f \in D[(x,y)]$, then $f \in C[(x,y)]$, too, since

$$\begin{aligned}&\lim_{h_1,h_2\to 0} f(x+h_1,y+h_2)-f(x,y)\\&=\lim_{h_1,h_2\to 0} \partial_1 f(x,y)h_1+\partial_2 f(x,y)h_2+\alpha(h_1,h_2)=0.\end{aligned}$$

In the applications of mathematics the notations $h_1 =: \Delta x$, $h_2 =: \Delta y$ are often used; and the change of the function is denoted by

$$\Delta f := f(x+h_1,y+h_2)-f(x,y).$$

Then

$$\Delta f \approx \frac{\partial f}{\partial x}\Delta x+\frac{\partial f}{\partial y}\Delta y$$

refers to the fact that the change of the function Δf is well approximated by the linear function prepared from the partial derivatives. This has a version using "infinitesimal quantitites", which hardly makes any sense:

$$\mathrm{d}f=\frac{\partial f}{\partial x}\mathrm{d}x+\frac{\partial f}{\partial y}\mathrm{d}y.$$

Here $\mathrm{d}f$ is called "differential of the function f".

The concepts that we introduced for functions of two variables can be generalized to functions of more than two variables without any difficulty.

Let $f:\mathbb{R}^n \supset\!\!\!\to \mathbb{R},\ x=(x_1,x_2,\dots,x_i,\dots,x_n)\in \operatorname{int} D(f)$. Then

$$\partial_i f(x) := \lim_{t\to 0}\frac{f(x_1,x_2,\dots,x_i+t,\dots,x_n)-f(x_1,x_2,\dots,x_i,\dots,x_n)}{t}$$

is the partial derivative of f with respect to the i-th variable.

We call the function $f:\mathbb{R}^n \supset\!\!\!\to \mathbb{R}$ differentiable at the point $x\in \operatorname{int} D(f)$ if there exist

$$A:=[A_1\ A_2\ \dots\ A_n]\in\mathbb{R}^{1\times n},\ \text{and a function } \alpha:\mathbb{R}^n \supset\!\!\!\to \mathbb{R}$$

such that for any vector $h\in\mathbb{R}^n$

$$f(x+h)-f(x)=Ah+\alpha(h),\ \text{where } \lim_{h\to 0}\frac{\alpha(h)}{\|h\|}=0.$$

It is again true that $A_i=\partial_i f(x),\ i=1,2,\dots,n$. If $f\in D[x]$, then

$$f'(x)=[\partial_1 f(x)\ \partial_2 f(x)\ \dots\ \partial_n f(x)].$$

Finally, let $f:\mathbb{R}^n \supset\!\!\!\to \mathbb{R}^k,\ x\in \operatorname{int} D(f)$. The function f is differentiable at the point x if there exists an $A\in\mathbb{R}^{k\times n}$ and $\alpha:\mathbb{R}^n \supset\!\!\!\to \mathbb{R}^k$ such that for all $h\in\mathbb{R}^n$

$$f(x+h)-f(x)=Ah+\alpha(h),\ \text{where } \lim_{h\to 0}\frac{\alpha(h)}{\|h\|}=0.$$

Now $A_{ij} = \partial_j f_i(x)$, and so

$$f'(x) = \begin{bmatrix} \partial_1 f_1(x) & \partial_2 f_1(x) & \dots & \partial_n f_1(x) \\ \partial_1 f_2(x) & \partial_2 f_2(x) & \dots & \partial_n f_2(x) \\ \vdots & & & \\ \partial_1 f_k(x) & \partial_2 f_k(x) & \dots & \partial_n f_k(x) \end{bmatrix} \in \mathbb{R}^{k\times n}$$

is the derivative of f at the point x. It is called Jacobian matrix.

For example,

1. if $f(x,y,z) := xyz$, then $f'(x,y,z) = [yz\ xz\ xy]$,
2. if $r(t) := \begin{bmatrix} \cos t \\ \sin t \\ t \end{bmatrix}$, then $r'(t) := \begin{bmatrix} -\sin t \\ \cos t \\ 1 \end{bmatrix}$,
3. if $F(x,y,z) = \begin{bmatrix} P(x,y,z) \\ Q(x,y,z) \\ R(x,y,z) \end{bmatrix}$, then $F'(x,y,z) = \begin{bmatrix} \partial_x P & \partial_y P & \partial_z P \\ \partial_x Q & \partial_y Q & \partial_z Q \\ \partial_x R & \partial_y R & \partial_z R \end{bmatrix}$.

12.1.3 Tangent

Tangent plane to a surface

Let $f : \mathbb{R}^2 \supset\!\!\to \mathbb{R}$, $(x_0, y_0) \in \text{int } D(f)$, and assume that $f \in D[(x_0, y_0)]$. This means that

$$f(x,y) - f(x_0,y_0) \approx \partial_1 f(x_0,y_0)(x - x_0) + \partial_2 f(x_0,y_0)(y - y_0)$$

if $(x,y) \approx (x_0,y_0)$, and this approximation is "sufficiently good". Let $z_0 := f(x_0,y_0)$, then by

$$z := \partial_1 f(x_0,y_0)(x - x_0) + \partial_2 f(x_0,y_0)(y - y_0) + z_0$$

$f(x,y) \approx z$ if $(x,y) \approx (x_0,y_0)$. Note that if

$$n := (\partial_1 f(x_0,y_0), \partial_2 f(x_0,y_0), -1),$$

$$r_0 := (x_0, y_0, z_0),$$

$$r := (x, y, z),$$

then we have shown that the plane described by the equation $\langle n, r - r_0\rangle = 0$ "sufficiently well" approximates the surface given by the function f.

The plane $\langle n, r - r_0\rangle = 0$ is called **tangent plane** to the surface f at the point $(x_0, y_0, f(x_0, y_0))$.

Tangent line to a space curve

Let $r : \mathbb{R} \supset\!\!\rightarrow \mathbb{R}^3,\ t_0 \in \operatorname{int} D(r)$, and assume that $r \in D[t_0]$. If

$$r(t) = \begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix},$$

then

$$r(t) - r(t_0) = \begin{bmatrix} x(t) - x(t_0) \\ y(t) - y(t_0) \\ z(t) - z(t_0) \end{bmatrix} \approx \dot{r}(t_0) \cdot (t - t_0) = \begin{bmatrix} \dot{x}(t_0) \\ \dot{y}(t_0) \\ \dot{z}(t_0) \end{bmatrix} (t - t_0),$$

and the approximation is "sufficiently good". This means that for any point $\underline{r} := \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ of the straight line with direction vector $\underline{v} := \begin{bmatrix} \dot{x}(t_0) \\ \dot{y}(t_0) \\ \dot{z}(t_0) \end{bmatrix}$ and passing through the point $\underline{r}_0 := \begin{bmatrix} x(t_0) \\ y(t_0) \\ z(t_0) \end{bmatrix}$ we have

$$\begin{aligned} x &= x(t_0) + \dot{x}(t_0) \cdot (t - t_0), \\ y &= y(t_0) + \dot{y}(t_0) \cdot (t - t_0), \\ z &= z(t_0) + \dot{z}(t_0) \cdot (t - t_0), \end{aligned}$$

and this line runs close to the curve, that is, $r(t) \approx \underline{r}$ if $t \approx t_0$.

The line $\underline{r} = \underline{r}_0 + \underline{v}(t - t_0)$ is called **tangent line** to the space curve r at the parameter value t_0, whose direction vector is the **tangent vector** $\dot{r}(t_0)$.

(Traditionally, the derivative of a space curve is denoted by a point instead of a comma.)

12.1.4 Extreme values

Let $f : \mathbb{R}^2 \supset\!\!\rightarrow \mathbb{R},\ a = (a_1, a_2) \in D(f)$.

Definition 12.3. We say that the function f has a **local minimum** at the point a if there is a neighborhood $K(a)$ of a that for all $x = (x_1, x_2) \in K(a) \cap D(f)$

$$f(x_1, x_2) \geq f(a_1, a_2) \text{ or } f(x) \geq f(a).$$

The **local maximum** can be defined in a similar way.

Theorem 12.1 (The necessary condition of a local extreme value)**.** *Let $f : \mathbb{R}^2 \rightarrowtail \mathbb{R},\ a = (a_1, a_2) \in \operatorname{int} D(f)$ and $f \in D[a]$. If f has a local extreme value (either minimum or maximum) at a, then $f'(a) = 0$.*

[$f'(a) = 0 \iff \partial_1 f(a_1, a_2) = 0$ and $\partial_2 f(a_1, a_2) = 0$.]

To prove this, it is enough to consider that if f has a local minimum at the point (a_1, a_2), then the function

$$\phi : \mathbb{R} \supset\!\!\to \mathbb{R}, \quad \phi(t) := f(t, a_2)$$

will have a local minimum at the point $t = a_1$.

Since $f \in D[(a_1, a_2)]$, therefore $\phi \in D[a_1]$, and so $\phi'(a_1) = 0$, which exactly means that $\partial_1 f(a_1, a_2) = 0$.

The same holds for the function

$$\psi : \mathbb{R} \supset\!\!\to \mathbb{R}, \quad \psi(t) := f(a_1, t).$$

So, $\psi'(a_2) = 0$, that is, $\partial_2 f(a_1, a_2) = 0$.

This method can be applied for finding the extreme value points of a differentiable function.

Our results can be extended to functions $f : \mathbb{R}^n \supset\!\!\to \mathbb{R}$ with slight modifications.

Theorem 12.2. *Let $f : \mathbb{R}^n \supset\!\!\to \mathbb{R}$, $a \in \operatorname{int} D(f)$ and $f \in D[a]$. If f has a local extreme value at a, then $f'(a) = 0$.*

$[f'(a) = 0 \Longleftrightarrow \partial_1 f(a) = 0,\ \partial_2 f(a) = 0, \dots, \partial_n f(a) = 0.]$

If $f : \mathbb{R}^n \supset\!\!\to \mathbb{R}$ and $f \in D[a]$, then instead of the row matrix $f'(a) \in \mathbb{R}^{1\times n}$ the vector $\operatorname{grad} f(a) := (f'(a))^T$ is used.

So,

$$\operatorname{grad} f(a) = \begin{bmatrix} \partial_1 f(a) \\ \partial_2 f(a) \\ \vdots \\ \partial_n f(a) \end{bmatrix}$$ (a column matrix that can be identified with a vector).

The meaning of $\operatorname{grad} f(a)$ will be illustrated in Exercise 4.

12.2 Exercises

1. Imagine the surfaces defined by the functions $f : \mathbb{R}^2 \to \mathbb{R}$,

   $$f(x, y) := x^2 + y^2;$$
   $$f(x, y) := x^2 + y^2 + 4x - 2y + 10;$$
   $$f(x, y) := 2x^2 + 5y^2.$$

   How could the surface of the functions

   $$h : \{(x, y) \mid x^2 + y^2 < 100\} \to \mathbb{R}, \quad h(x, y) := -x^2 - y^2 + 100$$

   look like?

2. Find the tangent plane to the function $f : \mathbb{R}^2 \to \mathbb{R},\ f(x, y) := x^2y^3$ at the point $(x_0, y_0) := (1, 2)$.

3. Find the tangent vector to the space curve $r : [0, 4\pi] \to \mathbb{R}^3,\ r(t) := (2\cos t, 2\sin t, t)$ at any point $t_0 \in (0, 4\pi)$. Calculate the scalar product $\langle \dot{r}(t_0), e_3\rangle$ ($e_3 := (0, 0, 1)$). Explain the result.

4. Let $f : \mathbb{R}^2 \to \mathbb{R},\ a \in \text{int } D(f)$ and $e \in \mathbb{R}^2$, for which $\|e\| = 1$. The **directional derivative** of f along a vector e at a point a is defined as the limit
$$\partial_e f(a) := \lim_{t\to 0} \frac{1}{t}(f(a + te) - f(a))$$
it this limit exists.

   If $f \in D[a]$, then one can show that
$$\partial_e f(a) = \langle \text{grad} f(a), e\rangle.$$

   Verify that at the point a the surface is steepest in the direction which is parallel with the vector $\text{grad} f(a)$.

   Solution:

   The direction that we look for is the vector $\hat{e} \in \mathbb{R}^2,\ \|\hat{e}\| = 1$ for which $\partial_e f(a) \le \partial_{\hat{e}} f(a)$ for any $e \in \mathbb{R}^2, \|e\| = 1$. We know from Linear Algebra that for planar vectors
$$\langle \text{grad} f(a), e\rangle = \|\text{grad} f(a)\| \cdot \|e\| \cos\alpha,$$
   where α is the angle of the two vectors. Since $\|\text{grad} f(a)\|$ does not change ($a \in \text{int } D(f)$ is fixed), and $\|e\| = 1$, therefore, the product is maximal if $\cos\alpha = 1$, that is, if e is parallel with the vector $\text{grad} f(a)$.

   This has the consequence that mountain streams and glaciers always move along the direction of the gradient at each point.

5. **The method of least squares**

   Assume that we make measurements in order to verify some relation. Denote by y_i the measurement corresponding to the value x_i. Our conjecture is that the points $(x_i, y_i),\ i = 1, 2, \ldots, n$ should be located along a line. Let us find the straight line $y = Ax + B$ that best fits the measurement points.

   Solution:

   The sum $\sum_{i=1}^{n}(Ax_i + B - y_i)^2$ is the square sum of the differences between the measurements and the straight line. We want this sum to be as small as possible.

Let $e(A,B) := \sum_{i=1}^n (Ax_i+B-y_i)^2$. The function e can be minimal where $e'(A,B)=0$, that is

$$\begin{aligned}\partial_A e(A,B) &= \sum 2(Ax_i+B-y_i)x_i = 0,\\ \partial_B e(A,B) &= \sum 2(Ax_i+B-y_i) = 0.\end{aligned}$$

In more detail,

$$\begin{aligned}A\sum x_i^2 + B\sum x_i &= \sum x_i y_i,\\ A\sum x_i + Bn &= \sum y_i.\end{aligned}$$

This is a system of linear equations in two unknowns (A and B), whose solution (which always exists if the points x_i are all different) is

$$A = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}, \quad B = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{n\sum x_i^2 - (\sum x_i)^2}.$$

(The summation indices run from 1 to n everywhere.) It can be shown that for such values of A and B the line $y = Ax+B$ truly runs closest to the points.

6. Let $f:\mathbb{R}^2\to\mathbb{R},\ f(x,y) := e^{xy}\cos(x^2y^3)$.

   Calculate the partial derivatives $\partial_x f(x,y),\ \partial_y f(x,y), \partial_y(\partial_x f)(x,y)$ and $\partial_x(\partial_y f)(x,y)$. What do you observe?

7. Let $f:\mathbb{R}^2\to\mathbb{R}$,

$$f(x,y) := \begin{cases} xy\frac{x^2-y^2}{x^2+y^2} & \text{if } x^2+y^2\neq 0,\\ 0 & \text{if } x^2+y^2 = 0.\end{cases}$$

   Show that

$$\partial_y(\partial_x f)(0,0) \neq \partial_x(\partial_y f)(0,0).$$

8. Find the local extreme values of the function $f:\mathbb{R}^2\to\mathbb{R},\ f(x,y) := x^4+y^4-2x+3y+1$.

9. The equality $2x^5y^3+x^3y^5-3x^4y^2+5xy^3 = 6x^2-1$ is satisfied for $x=1$ and $y=1$. Does this equation have any other solution?

   Solution:

   Let $f:\mathbb{R}^2\to\mathbb{R}$, $f(x,y) := 2x^5y^3+x^3y^5-3x^4y^2+5xy^3-6x^2+1$. Clearly, $f\in C^1$ and $f(1,1)=0$.

   $\partial_2 f(x,y) = 6x^5y^2+5x^3y^4-6x^4y+15xy^2$, therefore $\partial_2 f(1,1) = 20 \neq 0$.

Due to the implicit function theorem, there exist neighborhoods $K_\mu(1)$ and $K_\rho(1)$ and a differentiable function $\phi : K_\mu(1) \to K_\rho(1)$ such that for all $x \in (1-\mu, 1+\mu)$, $f(x, \phi(x)) = 0$, so the equation has infinitely many solutions. (Of course this does not mean that besides the solution $(1,1)$ it has any other solution that is made up of two integers!) Since

$$\partial_1 f(x,y) = 10x^4y^3 + 3x^2y^5 - 12x^3y^5 + 5y^3 - 12x, \ \partial_1 f(1,1) = -6,$$

therefore

$$\phi'(1) = -\frac{\partial_1 f(1,1)}{\partial_2 f(1,1)} = \frac{3}{10}.$$

Exploiting this, we can get an approximation of the function ϕ:

$$\phi(x) \approx \phi(1) + \phi'(1)(x-1) \text{ if } x \approx 1,$$

that is,

$$\phi(x) \approx 1 + \frac{3}{10}(x-1) \text{ if } x \approx 1.$$

10. Assume that there is a differentiable function $y : \mathbb{R} \supset\!\!\to \mathbb{R}$ that is defined by the equation $xy + e^{x+y} - y^2 + 5 = 0$. Calculate its derivative!

Solution:

Let $f : \mathbb{R}^2 \supset\!\!\to \mathbb{R}, f(x,y) := xy + e^{x+y} - y^2 + 5$. Due to the assumption, for all $x \in D(y)$

$$h(x) := f(x, y(x)) = 0,$$

therefore, the derivative of the function h is 0, too, that is, for all $x \in D(y)$

$$\begin{aligned} h'(x) &= (xy(x) + e^{x+y(x)} - y^2(x) + 5)' \\ &= y(x) + xy'(x) + e^{x+y(x)} \cdot (1 + y'(x)) - 2y(x)y'(x) = 0. \end{aligned}$$

From this we can express $y'(x)$ as

$$y'(x) = -\frac{y(x) + e^{x+y(x)}}{x + e^{x+y(x)} - 2y(x)} \quad (x \in D(y)).$$

This result is often written in the superficial form

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

We remark that in this case the derivation rule of the implicitly defined function is applied without checking the conditions.

11. The state of a gas is given by the equation of state $F(p, V, T) = 0$. (For an ideal gas it has the form $pV - nRT = 0$.) This equation defines three implicit functions:

$$\begin{aligned} p &= p(V, T), \\ V &= V(T, p), \\ T &= T(p, V). \end{aligned}$$

Show that

$$\partial_V p(V, T) \cdot \partial_T V(T, p) \cdot \partial_p T(p, V) = -1.$$

Solution:

Assuming that the assumptions of the implicit function theorem are satisfied and substituting the implicit functions we obtain that

$$\begin{aligned} (V, T) &\mapsto F(p(V, T), V, T) = 0, \\ (T, p) &\mapsto F(p, V(T, p), T) = 0, \\ (p, V) &\mapsto F(p, V, T(p, V)) = 0. \end{aligned}$$

The partial derivative of the constant 0 function is 0, therefore

$$\begin{aligned} \partial_V F(p(V, T), V, T) &= \partial_1 F \cdot \partial_V p + \partial_2 F \cdot \partial_V V + \partial_3 F \cdot \partial_V T \\ &= 0 \Rightarrow \partial_V p = -\frac{\partial_2 F}{\partial_1 F}, \\ \partial_T F(p, V(T, p), T) &= \partial_1 F \cdot \partial_T p + \partial_2 F \cdot \partial_T V + \partial_3 F \cdot \partial_T T \\ &= 0 \Rightarrow \partial_T V = -\frac{\partial_3 F}{\partial_2 F}, \\ \partial_p F(p, V, T(p, V)) &= \partial_1 F \cdot \partial_p p + \partial_2 F \cdot \partial_p V + \partial_3 F \cdot \partial_p T \\ &= 0 \Rightarrow \partial_p T = -\frac{\partial_1 F}{\partial_3 F}. \end{aligned}$$

From this we have

$$\partial_V p \cdot \partial_T V \cdot \partial_p T = \left(-\frac{\partial_2 F}{\partial_1 F}\right)\left(-\frac{\partial_3 F}{\partial_2 F}\right)\left(-\frac{\partial_1 F}{\partial_3 F}\right) = -1.$$

We remark that those thinking superficially may be surprised that by using the traditional notations and treating the partial derivatives as quotients the expected result would be

$$\frac{\partial p}{\partial V} \cdot \frac{\partial V}{\partial T} \cdot \frac{\partial T}{\partial p} = 1 \ldots.$$

Check by calculation in the case $pV - nRT = 0$ that the product is really (-1).

12. Let $\Phi : \mathbb{R}^2 \to \mathbb{R}^3, \Phi(u,v) = \begin{bmatrix} x(u,v) \\ y(u,v) \\ z(u,v) \end{bmatrix} := \begin{bmatrix} u+v \\ u^2+v^2 \\ u^3+v^3 \end{bmatrix}$ be a "surface parameterized with two parameters". Find a normal vector to the tangent plane of this surface at the point $(u_0, v_0) := (1,2)$.

Solution:

The inverse of the function

$$g : (u,v) \mapsto \begin{bmatrix} x(u,v) \\ y(u,v) \end{bmatrix} = \begin{bmatrix} u+v \\ u^2+v^2 \end{bmatrix}$$

would be

$$g^{-1} : (x,y) \mapsto \begin{bmatrix} u(x,y) \\ v(x,y) \end{bmatrix}.$$

By substituting this into the function z, the surface Φ would be given as the two-variable real-valued function

$$z \circ g^{-1} : (x,y) \mapsto z(u(x,y), v(x,y)).$$

A normal vector of its tangent plane is

$$\underline{n} = (\partial_x z(u(x_0,y_0), v(x_0,y_0)), \partial_y z(u(x_0,y_0), v(x_0,y_0)), -1).$$

One can see that $(z \circ g^{-1})'(x_0,y_0)$ is the derivative we just need. Using the inverse function theorem we obtain that

$$\begin{aligned}
(z \circ g^{-1})'(x_0,y_0) &= z'(g^{-1}(x_0,y_0)) \cdot (g^{-1})'(x_0,y_0) \\
&= \begin{bmatrix} \partial_u z(u_0,v_0) & \partial_v z(u_0,v_0) \end{bmatrix} \cdot (g'(u_0,v_0))^{-1} \\
&= \begin{bmatrix} 3u_0^2 & 3v_0^2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 2u_0 & 2v_0 \end{bmatrix}^{-1} \\
&= \begin{bmatrix} 3 & 12 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}^{-1} \\
&= \begin{bmatrix} 3 & 12 \end{bmatrix} \cdot \begin{bmatrix} 2 & -\frac{1}{2} \\ -1 & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} -6 & \frac{9}{2} \end{bmatrix}.
\end{aligned}$$

So, $\partial_x z(u(1,2), v(1,2)) = -6$ and $\partial_y z(u(1,2), v(1,2)) = \frac{9}{2}$, thus, a normal vector of the tangent plane is $\underline{n}(-6, \frac{9}{2}, -1)$.

13. Find the local extreme value points of the function $f : \mathbb{R}^2 \to \mathbb{R}, f(x,y) := x^2 + y^2 - 2x + 4y - 1$ on the closed set

$$Q := \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}.$$

Solution:

The function f may have a local extreme value on the open circle

$$\mathrm{int}Q = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 < 1\}$$

only where

$$\begin{aligned} \partial_1 f(x,y) &= 2x-2 = 0, \\ \partial_2 f(x,y) &= 2y+4 = 0. \end{aligned}$$

From this we obtain $(x_0, y_0) = (1, -2)$, which is not on the open circle int$Q$, that is, f has no local extreme value inside the circle. Since Q is a compact set (bounded and closed), therefore the continuous function f has a minimum and a maximum in Q. So, the extreme value points are on the boundary of Q.

Find the conditional minimum and maximum of the function f under the condition $g(x,y) := x^2+y^2-1 = 0$.

Prepare the function $F(x,y) := f(x,y) + \lambda g(x,y) = x^2+y^2-2x+4y-1+\lambda(x^2+y^2-1)$.

$$\begin{aligned} \partial_1 F(x,y) &= 2x-2+2\lambda x = 0, \\ \partial_2 F(x,y) &= 2y+4+2\lambda y = 0, \\ x^2+y^2-1 &= 0. \end{aligned}$$

Solving the system of equations in three unknowns (we have exactly as many equations as unknowns...), we obtain that $x = \frac{1}{1+\lambda}$, $y = -\frac{2}{1+\lambda}$, from which

$$\frac{1}{(1+\lambda)^2} + \frac{4}{(1+\lambda)^2} = 1.$$

There are two solutions: $\lambda_1 = \sqrt{5}-1$ and $\lambda_2 = -\sqrt{5}-1$. The points corresponding to these solutions are $P_1(\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}})$ and $P_2(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})$.

First let $\lambda_1 = \sqrt{5}-1$ and $P_1(\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}})$.

$$\begin{aligned} F_1(x,y) &= f(x,y) + \lambda_1 g(x,y), \\ F_1'(x,y) &= \begin{bmatrix} 2x-2+2(\sqrt{5}-1)x & 2y+4+2(\sqrt{5}-1)y \end{bmatrix}, \\ F_1''(x_1,y_1) &= \begin{bmatrix} 2+2(\sqrt{5}-1) & 0 \\ 0 & 2+2(\sqrt{5}-1) \end{bmatrix} = \begin{bmatrix} 2\sqrt{5} & 0 \\ 0 & 2\sqrt{5} \end{bmatrix} \end{aligned}$$

is the matrix of a positive definite quadratic form, so for any vector $h \in \mathbb{R}^2, h \neq 0$: $\langle F_1''(x_1,y_1)h, h\rangle > 0$. Therefore, at the point $P_1(\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}})$ the function f has a minimum under the condition $g = 0$.

$\lambda_2 = -\sqrt{5} - 1$ and $P_2(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})$ define a function $F_2(x, y) = f(x, y) + \lambda_2 g(x, y)$ as well.

$$F_2'(x, y) = \begin{bmatrix} 2x - 2 + 2(-\sqrt{5} - 1)x & 2y + 4 + 2(-\sqrt{5} - 1)y \end{bmatrix},$$
$$F_2''(x_2, y_2) = \begin{bmatrix} 2 + 2(-\sqrt{5} - 1) & 0 \\ 0 & 2 + 2(-\sqrt{5} - 1) \end{bmatrix} = \begin{bmatrix} -2\sqrt{5} & 0 \\ 0 & -2\sqrt{5} \end{bmatrix}$$

is the matrix of a negative quadratic form, so for any vector $h \in \mathbb{R}^2$, $h \neq 0$: $\langle F_2''(x_2, y_2)h, h\rangle < 0$. Therefore, at the point $P_2(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})$ the function f has a maximum under the condition $g = 0$.

Chapter 13

Line integrals

We generalize the integral of a function $f : [a,b] \to \mathbb{R}$. The role of the interval $[a,b]$ will be played by a curve, while the role of the function f by a multivariable vector-valued function. The following topics will be discussed.

- Line integral and its properties
- Potential and the relation between its existence and the line integral
- The differentiability of parametric integrals
- A sufficient condition for the existence of potential

13.1 Line integrals

13.1.1 The concept and properties of line integral

When a sleigh is pulled from point A to point B along a displacement s with a force F parallel with the path, then the work done is $W = F \cdot s$ (Fig. 13.1).

When the force F encloses an angle α with the displacement (Fig. 13.2), then the work is approximately

$$W = F \cos\alpha = \langle \underline{F}, \underline{s} \rangle.$$

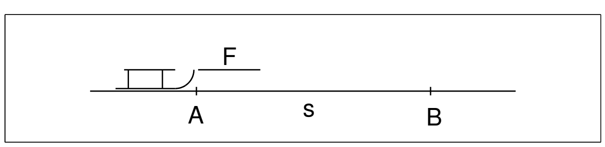


Figure 13.1

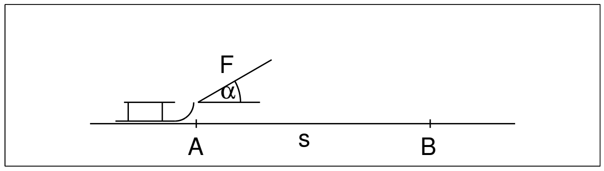


Figure 13.2

The work of a force $F \in \mathbb{R}^3 \to \mathbb{R}^3$, changing from point to point (i.e., a force field) along a space curve $r : [\alpha, \beta] \to \mathbb{R}^3$ can be approximated by dividing the interval $[\alpha, \beta]$ by the division points $\alpha = t_0 < t_1 < \ldots < t_{i-1} < t_i < \ldots < t_n = \beta$ and taking further points

$$t_{i-1} \le \xi_i \le t_i \quad (i = 1, \ldots, n).$$

Then the elementary work is

$$\Delta W_i := \langle F(r(\xi_i)), r(t_i) - r(t_{i-1})\rangle,$$

and the work done by the force field F along the curve is

$$W \approx \sum \Delta W_i = \sum_{i=1}^{n} \langle F(r(\xi_i)), r(t_i) - r(t_{i-1})\rangle = \sum \langle F(r(\xi_i)), \Delta r_i\rangle.$$

If r is sufficiently smooth (differentiable), then

$$r(t_i) - r(t_{i-1}) = \begin{bmatrix} x(t_i) - x(t_{i-1}) \\ y(t_i) - y(t_{i-1}) \\ z(t_i) - z(t_{i-1}) \end{bmatrix} = \begin{bmatrix} \dot{x}(\eta_i)(t_i - t_{i-1}) \\ \dot{y}(\vartheta_i)(t_i - t_{i-1}) \\ \dot{z}(\zeta_i)(t_i - t_{i-1}) \end{bmatrix} \approx \dot{r}(\xi_i)(t_i - t_{i-1}),$$

provided that $\dot{r}$ is continuous. One can see that

$$W \approx \sum_{i=1}^{n} \langle F(r(\xi_i)), \dot{r}(\xi_i)\rangle (t_i - t_{i-1}),$$

which looks like a sum approximation of an integral. This consideration is the ground for the further concepts.

Let $\Omega \subset \mathbb{R}^n$ be a connected (meaning that any two points of Ω can be connected by a continuous curve running in Ω) and open set, that is, a **domain**. Let $f \in \mathbb{R}^n \rightarrowtail \mathbb{R}^n$, $D(f) := \Omega$ be a continuous vector function, $f \in C(\Omega)$. Let $r : [\alpha, \beta] \to \Omega$ be a smooth space curve, that is, $r \in C[\alpha, \beta]$, $r \in D(\alpha, \beta)$, $\dot{r}(t) \neq 0$ $(t \in (\alpha, \beta))$, and for all $t_1, t_2 \in (\alpha, \beta)$, $t_1 \neq t_2$: $r(t_1) \neq r(t_2)$.

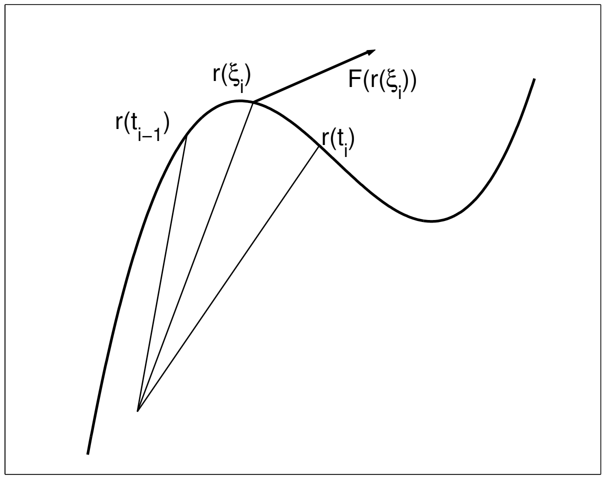


Figure 13.3

Definition 13.1. The **line integral** of the function f along r is

$$\int_r f := \int_\alpha^\beta \langle f(r(t)), \dot{r}(t)\rangle \mathrm{d}t.$$

For example,

$$f : \mathbb{R}^3 \to \mathbb{R}^3, \quad f(x,y,z) := \begin{bmatrix} x+y \\ x-y \\ z \end{bmatrix}$$

(here $\Omega = \mathbb{R}^3$) and for

$$r : [0,1] \to \mathbb{R}^3, \quad r(t) := \begin{bmatrix} t \\ 2t \\ 3t \end{bmatrix}$$

we have $\dot{r}(t) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, thus,

$$\int_r f = \int_0^1 \left\langle \begin{bmatrix} t+2t \\ t-2t \\ 3t \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\rangle \mathrm{d}t = \int_0^1 [(t+2t) + 2(t-2t) + 9t]\mathrm{d}t$$

$$= \int_0^1 10t\mathrm{d}t = \left[10\frac{t^2}{2}\right]_0^1 = 5.$$

The properties of line integral

1^o If $r_1 : [\alpha, \beta] \to \Omega$, $r_2 : [\beta, \gamma] \to \Omega$ and $r_1(\beta) = r_2(\beta)$, then $r_1 \cup r_2 : [\alpha, \gamma] \to \Omega$, where $r_1 \cup r_2|_{[\alpha,\beta]} = r_1$ and $r_1 \cup r_2|_{[\beta,\gamma]} = r_2$ is called **union of the two curves**. Then

$$\int_{r_1 \cup r_2} f = \int_{r_1} f + \int_{r_2} f.$$

2^o If $r : [\alpha, \beta] \to \Omega$, then we call $\overleftarrow{r} : [\alpha, \beta] \to \Omega, \overleftarrow{r}(t) := r(\alpha + \beta - t)$ the **reverse curve** to r. Then

$$\int_{\overleftarrow{r}} f = -\int_r f.$$

3^o If f is bounded on the domain Ω, that is, there is a number $K > 0$ such that for all $x \in \Omega$, $\|f(x)\| \leq K$, and the length of $r : [\alpha, \beta] \to \Omega$ is equal to L, then

$$\left| \int_r f \right| \leq K \cdot L.$$

13.1.2 Potential

Line integral is basically related to the "work of a force field". When work (energy) is considered, one may be interested whether there is energy loss or perhaps energy gain along a closed curve. It may also be important to know what path is worth following from the point of view of the work done. Such questions are answered by the following theorem.

Theorem 13.1. *Let $\Omega \subset \mathbb{R}^n$ be a domain, $f : \Omega \to \mathbb{R}^n$, $f \in C(\Omega)$. The following three properties of the function (or force field) f are equivalent:*

1^o *For any smooth* **closed curve** $r : [\alpha, \beta] \to \Omega$, $r(\alpha) = r(\beta)$ $\oint f = 0$ *(here the symbol $\oint$ emphasizes that we integrate along a closed curve).*

2^o *For any fixed $a, b \in \Omega$ and any arbitrary curves r_1 and r_2 "connecting a and b" (that is, $r_1 : [\alpha_1, \beta_1] \to \Omega$ and $r_2 : [\alpha_2, \beta_2] \to \Omega$ for which $r_1(\alpha_1) = r_2(\alpha_2) = a$ and $r_1(\beta_1) = r_2(\beta_2) = b$),*

$$\int_{r_1} f = \int_{r_2} f.$$

3^o *There exists a so-called* **potential function** $\Phi : \Omega \to \mathbb{R}$, $\Phi \in D(\Omega)$ *such that for all $x \in \Omega$ and $i = 1, 2, \ldots, n$*

$$\partial_i \Phi(x) = f_i(x),$$

or, more briefly, $\mathrm{grad}\Phi = f$.

This theorem states that if the force field f has a potential, then the work done along any closed curve is zero, moreover, the work done between two points is independent of the curve that connects them.

Obviously, it would be interesting to know how we can decide whether a force field f has a potential.

Theorem 13.2. *If $\Omega \subset \mathbb{R}^n$ is a domain that has a point $a \in \Omega$ such that for all $x \in \Omega$*

$$[a, x] := \{a + t(x - a) \in \mathbb{R}^n \mid t \in [0, 1]\} \subset \Omega$$

*($[a, x]$ is called the "section connecting the point a with the point x", such a set Ω is called a "**starlike domain**" with respect to the point a), and $f : \Omega \to \mathbb{R}^n$, $f \in C(\Omega)$ is such that for all $i, j = 1, 2, \dots, n$ $\partial_i f_j \in C(\Omega)$ (each partial derivative of each coordinate function of f is continuous at all points of Ω), moreover,*

$$\partial_i f_j(x) = \partial_j f_i(x) \quad \text{for all } x \in \Omega, i, j = 1, 2, \dots, n$$

(which means that the derivative matrix $f'(x)$ is symmetric), then the function f has a potential $\Phi : \Omega \to \mathbb{R}$.

When $\Omega = \mathbb{R}^3$, then this domain is starlike. If the force field

$$f := \begin{bmatrix} P \\ Q \\ R \end{bmatrix} : \mathbb{R}^3 \to \mathbb{R}^3$$

is sufficiently smooth with coordinate functions $P, Q, R : \mathbb{R}^3 \to \mathbb{R}$ (also mentioned as "components" of the force field), then the condition $\partial_i f_j = \partial_j f_i$ means that the derivative matrix

$$f'(x) = \begin{bmatrix} \partial_1 P(x) & \partial_2 P(x) & \partial_3 P(x) \\ \partial_1 Q(x) & \partial_2 Q(x) & \partial_3 Q(x) \\ \partial_1 R(x) & \partial_2 R(x) & \partial_3 R(x) \end{bmatrix} \quad (x \in \Omega)$$

is symmetric. If we introduce the concept of rotation, then

$$\mathrm{rot} f := \nabla \times f := \begin{vmatrix} e_1 & e_2 & e_3 \\ \partial_1 & \partial_2 & \partial_3 \\ P & Q & R \end{vmatrix}$$

$$:= (\partial_2 R - \partial_3 Q)e_1 - (\partial_1 R - \partial_3 P)e_2 + (\partial_1 Q - \partial_2 P)e_3 = 0 \in \mathbb{R}^3$$

on the whole $\mathbb{R}^3$ space. In physics this theorem is often mentioned in the following form: "A rotation-free force field has a potential."

Finally, let us examine how we can find the potential of a force field f in case it exists, and what else we gain from knowing it.

For simplicity, let

$$f : \mathbb{R}^2 \to \mathbb{R}^2,\ f(x, y) := \begin{bmatrix} x + y \\ x - y \end{bmatrix}.$$

Since $\partial_y(x + y) = 1$ and $\partial_x(x - y) = 1$, therefore $f'(x, y)$ is symmetric, so f has a potential Φ, and this – temporarily unknown – potential is a function $\Phi : \mathbb{R}^2 \to \mathbb{R}$ for which

$$\begin{aligned} \partial_x\Phi(x, y) &= x + y, \\ \partial_y\Phi(x, y) &= x - y. \end{aligned}$$

If $\partial_x\Phi(x, y) = x + y$, then Φ is of the form $\Phi(x, y) = \frac{x^2}{2} + xy + \phi(y)$, where $\phi : \mathbb{R} \to \mathbb{R}$ is differentiable, but otherwise arbitrary. Then $\partial_y\Phi(x, y) = x + \phi'(y) = x - y$, so $\phi'(y) = -y$, from which $\phi(y) = -\frac{y^2}{2} + c$, where $c \in \mathbb{R}$ is arbitrary.

Consequently, the potential of f can only be a function of the form $\Phi : \mathbb{R}^2 \to \mathbb{R}$, $\Phi(x, y) = \frac{x^2}{2} + xy - \frac{y^2}{2} + c$.

If after all this we would like to calculate the work of the force field f along any smooth curve $r : [\alpha, \beta] \to \mathbb{R}^2$, then (in view of the chain rule of differentiating a composition of functions)

$$\begin{aligned} \int_r f = \int_\alpha^\beta \langle f(r(t)), \dot{r}(t)\rangle \mathrm{d}t = \int_\alpha^\beta \langle \mathrm{grad}\Phi(r(t)), \dot{r}(t)\rangle \mathrm{d}t = \int_\alpha^\beta \Phi'(r(t)) \cdot \dot{r}(t) \mathrm{d}t \\ = \int_\alpha^\beta (\Phi(r(t)))' \mathrm{d}t = [\Phi(r(t))]_\alpha^\beta = \Phi(r(\beta)) - \Phi(r(\alpha)), \end{aligned}$$

which shows that the value of the line integral only depends on the two "end points" of the curve, and is independent of the curve that connects the points $r(\alpha)$ and $r(\beta)$ (a fact also suggested by the theorem). Particularly, if $r(\alpha) = r(\beta)$, then we have a closed curve, and then $\Phi(r(\beta)) = \Phi(r(\alpha))$, so $\oint_r f = 0$, as it was also stated by the theorem.

13.2 Exercises

1. Let

$$f:\mathbb{R}^3\to\mathbb{R}^3,\quad f(x,y,z):=\begin{bmatrix} x+y+z\\ y-z\\ x+z\end{bmatrix}\ \text{and}\ \ r:[0,6\pi]\to\mathbb{R}^3,$$
$$r(t):=\begin{bmatrix} 2\cos t\\ 2\sin t\\ t\end{bmatrix}.$$

Calculate the line integral $\int_r f$.

2. Let $f:\mathbb{R}^2\setminus\{(0,0)\}\to\mathbb{R}^2,\quad f(x,y):=\begin{bmatrix}\frac{x}{x^2+y^2}\\ -\frac{y}{x^2+y^2}\end{bmatrix}$.

Calculate the line integral of the tangent space f along a closed circle of radius 1, centered at the origin and with positive (counterclockwise) direction.

3. Show that the force field

$$F:\mathbb{R}^3\setminus\{(0,0,0)\}\to\mathbb{R}^3,\quad F(x_1,x_2,x_3):=\begin{bmatrix} -\frac{x_1}{(x_1^2+x_2^2+x_3^2)^{3/2}}\\ -\frac{x_2}{(x_1^2+x_2^2+x_3^2)^{3/2}}\\ -\frac{x_3}{(x_1^2+x_2^2+x_3^2)^{3/2}}\end{bmatrix}$$

has a potential. Calculate the potential Φ.

Solution:

Let $i,j=1,2,3$ and $i\neq j$. Then

$$\begin{aligned}\partial_i\left(-\frac{x_j}{(x_1^2+x_2^2+x_3^2)^{3/2}}\right)&=-x_j\left(-\frac{3}{2}(x_1^2+x_2^2+x_3^2)^{-5/2}\right)\cdot 2x_i\\ &=-x_i\left(-\frac{3}{2}(x_1^2+x_2^2+x_3^2)^{-5/2}\right)\cdot 2x_j=\partial_j\left(-\frac{x_i}{(x_1^2+x_2^2+x_3^2)^{3/2}}\right),\end{aligned}$$

therefore the force field F has a potential. If $\Phi:\mathbb{R}^3\setminus\{(0,0,0)\}\to\mathbb{R}$, then

$$\partial_i\Phi(x_1,x_2,x_3)=-\frac{x_i}{(x_1^2+x_2^2+x_3^2)^{3/2}},$$

and then

$$\Phi(x_1,x_2,x_3)=\frac{1}{(x_1^2+x_2^2+x_3^2)^{1/2}}+c,$$

since

$$\partial_i \Phi(x_1, x_2, x_3) = -\frac{1}{2}(x_1^2 + x_2^2 + x_3^2)^{-3/2}(2x_i)$$
$$= -\frac{x_i}{(x_1^2 + x_2^2 + x_3^2)^{3/2}}, \quad i = 1, 2, 3.$$

We remark that F can be considered as the gravitational force field of a mass point $M = 1$ located at the origin, since the force acting on the mass $m = 1$ at the point with location vector $\underline{r}$ (apart from the multiplication factor arising from the choice of the unit system) is

$$\underline{F}(\underline{r}) = -\frac{1}{\|\underline{r}\|^2} \cdot \frac{\underline{r}}{\|\underline{r}\|} \quad (\underline{r} \neq \underline{0}).$$

The potential of this force field is

$$\Phi(\underline{r}) = \frac{1}{\|\underline{r}\|} \quad (\underline{r} \neq \underline{0}).$$

4. Let

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \quad f(x, y) := \begin{bmatrix} xy^2 \\ x^2 y \end{bmatrix},$$

and let the curve be a lemniscate, for example

$$L := \left\{(x, y) \in \mathbb{R}^2 \mid \sqrt{(x-2)^2 + y^2} \cdot \sqrt{(x+2)^2 + y^2} = 8\right\}.$$

(The curve L is the locus of points on the plane whose product of distances from the points $C_1(2, 0)$ and $C_2(-2, 0)$ equals 8.) Calculate the line integral of f along this lemniscate.

Chapter 14

Differential equations

Differential equations are suitable for the description of several natural and social phenomena. The following topics will be discussed.

- The concept of differential equation
- Separable differential equations
- Applications

14.1 Differential equations

14.1.1 Basic concepts

Let $\Omega \subset \mathbb{R}^2$ be a domain, and $f : \Omega \to \mathbb{R}$ a continuous function. We look for such functions $y : \mathbb{R} \rightarrowtail \mathbb{R}$ whose range $D(y)$ is an open interval, y is continuously differentiable, and for all $x \in D(y)$: $(x, y(x)) \in \Omega$ and

$$y'(x) = f(x, y(x)).$$

This problem is called **first order differential equation**, and we denote it by the symbol $y' = f(x, y)$. We will see that such a problem usually has infinitely many solutions. If, however, at some point x_0 we prescribe the value of the solution $y(x_0)$, then, as a rule, we will obtain one solution. The relation

$$y(x_0) = y_0$$

is called **initial condition**. As the exercises will reveal, such types of conditions naturally belong to the differential equations.

We can raise the question of whether the continuity of the function f is sufficient for the differential equation $y' = f(x, y)$ to have a solution, moreover, if it has one, how can we calculate it.

14.1.2 Separable differential equations

As a first step, let us deal with the special case where the function $f : \mathbb{R}^2 \rightarrowtail \mathbb{R}$ is of the form

$$f(x, y) = g(x)h(y),$$

where $g, h : \mathbb{R} \rightarrowtail \mathbb{R}$ are continuous functions, $D(g)$ is an interval, and the function h is nowhere equal to zero. This type of differential equations is called **separable** differential equation, and can be denoted more briefly as

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

Assume that some function $y : I \to \mathbb{R}$ solves the problem, that is, for all $x \in I$, $y'(x) = g(x)h(y(x))$. Then

$$\frac{y'(x)}{h(y(x))} = g(x) \qquad (x \in I). \tag{14.1}$$

Let $H := \int 1/h$ and $G = \int g$ be primitive functions of $1/h$ and g, respectively. For any arbitrary $x \in I$

$$(H \circ y)'(x) = H'(y(x))y'(x) = \frac{y'(x)}{h(y(x))} \quad \text{and} \quad G'(x) = g(x).$$

Since the derivative functions $(H \circ y)'$ and G' are equal on the interval I by (14.1), therefore the functions $H \circ y$ and G can only differ by a constant. So, there exists a number $c \in \mathbb{R}$ such that for all $x \in I$

$$H(y(x)) = G(x) + c.$$

If H has an inverse function, H^{-1}, then

$$H^{-1}(H(y(x))) = H^{-1}(G(x) + c),$$

that is, for all $x \in I$

$$y(x) = H^{-1}(G(x) + c). \tag{14.2}$$

Consequently, the solution of the problem $y' = g(x)h(y)$ can be given in the form (14.2). (One can check by substitution that these functions are really solutions.) We remark that following this train of thought just formally, we can get to a solution procedure that is easy to memorize:

$$\begin{aligned} y' &= g(x)h(y), \\ \frac{\mathrm{d}y}{\mathrm{d}x} &= g(x)h(y), \\ \frac{\mathrm{d}y}{h(y)} &= g(x)\mathrm{d}x. \end{aligned}$$

Integrating this equation and introducing the primitive functions $H := \int 1/h$ and $G = \int g$, we are led to the equation

$$H(y) = G(x) + c.$$

Applying the inverse function H^{-1} to both sides, we obtain the solution (14.2).

Let us look at what kinds of phenomena can be described by such a simple type of differential equations.

14.1.3 Application

Assume that a radioactive substance has mass m_0 at time t_0. As time passes, at $t > 0$ we denote the mass $m(t)$, while at time $t + \Delta t$ by $m(t + \Delta t)$. We suppose that the change of mass $\Delta m := m(t + \Delta t) - m(t)$ between the time instances t and $t + \Delta t$ is proportional to the mass $m(t)$ at time t and the time Δt that has passed: $\Delta m \sim m(t)\Delta t$. Due to the radioactive decay, $\Delta m < 0$ if $\Delta t > 0$. Introducing a proportionality factor $k > 0$, we have $\Delta m = -km(t)\Delta t$, and so we are led to the equality

$$\frac{\Delta m}{\Delta t} = -km(t).$$

In the limit $\Delta t \to 0$ we obtain the differential equation

$$\lim_{\Delta t \to 0} \frac{\Delta m}{\Delta t} = \frac{\mathrm{d}m}{\mathrm{d}t} = -km(t).$$

Here the variable is denoted by t instead of x, and the unknown function by $m(t)$ instead of $y(x)$. The differential equation is separable (now k does not even depend on the variable t). Let us solve the equation by using the previously introduced method. By separating the variables we have

$$\frac{\mathrm{d}m}{m} = -k\mathrm{d}t.$$

Integrating both sides yields

$$\ln m = -kt + c,$$

from which

$$m = \mathrm{e}^{-kt+c} = \mathrm{e}^{-kt}\mathrm{e}^{c}.$$

Since from the initial condition $m_0 = m(0) = \mathrm{e}^0\mathrm{e}^c = \mathrm{e}^c$, therefore the solution for any $t > 0$ is

$$m(t) = m_0\mathrm{e}^{-kt}.$$

An important characteristic of radioactive substances is the half-life T, which is the time required for one half of the amount of radioactive material to degrade. The decay constant k can be expressed in terms of T, since

$$\frac{m_0}{2} = m(T) = m_0 \mathrm{e}^{-kT},$$

from which

$$k = \frac{\ln 2}{T}.$$

Research has shown that the concentration of the carbon-14 isotope is constant in living plants, since the radiated C^{14} is replenished from the atmosphere during the assimilation. However, when a tree dies, no more C^{14} will infiltrate, therefore, its concentration decreases in the material of the tree. A decayed trunk has been found in which the amount of C^{14} per unit volume is just 90% of the usual amount. When did the tree die if we know that the half-life of C^{14} is 5370 years?

Since the amount of C^{14} at time t after the death of the tree is given by the formula

$$m(t) = m_0 \mathrm{e}^{-\frac{\ln 2}{5370} t},$$

and now the amount of C^{14} is $0,9m_0$ in the wood, therefore the time in question is given by the equation

$$0,9m_0 = m_0 \mathrm{e}^{-\frac{\ln 2}{5370} t}.$$

Dividing both sides by m_0 and then taking their logarithms we have

$$\ln 0,9 = -\frac{\ln 2}{5370} t,$$

from which

$$t = -5370 \frac{\ln 0,9}{\ln 2} = 816 \text{ years}.$$

So, the tree died 816 years ago. This example illustrates the method of absolute dating, for which W. Libby, American chemist was awarded the Nobel Prize in 1960...

14.2 Exercises

1. (The model of unlimited reproduction) Assume that the mass of virus in the population of a city at time $t = 0$ is m_0. Describe the formation of the epidemic (if there is no remedy of the disease...).

2. (The model of limited growth) On an island the grass can support at most a population of rabbits of total amount (e.g., mass) M. Assume that a rabbit population of mass m_0 are settled on the island. Describe how the amount of rabbits changes in time.

   Solution: Denote by $m(t)$ the amount in question at time t. We may assume that the change of this amount in time Δt at t is proportional to the time Δt that has passed, the amount $m(t)$ of the rabbits and the remaining carrying capacity of the island. Thus,

$$m(t+\Delta t) - m(t) \sim m(t)(M - m(t))\Delta t.$$

   By introducing the reproduction factor k,

$$m(t+\Delta t) - m(t) = km(t)(M - m(t))\Delta t.$$

   Dividing the equation by Δt, and then taking the limit as $\Delta t \to 0$, we obtain the separable differential equation

$$\frac{\mathrm{d}m}{\mathrm{d}t} = m' = km(M-m).$$

   Separating the variables we have

$$\frac{\mathrm{d}m}{m(M-m)} = k\mathrm{d}t.$$

   Exploiting the equality

$$\frac{1}{m(M-m)} = \frac{1}{M}\left(\frac{1}{m} + \frac{1}{M-m}\right),$$

   we obtain that

$$\int \frac{1}{m(M-m)}\mathrm{d}m = \frac{1}{M}\left(\ln m - \ln(M-m)\right) = \frac{1}{M}\ln\frac{m}{M-m}.$$

   Integrating both sides of the equation yields

$$\begin{aligned} \frac{1}{M}\ln\frac{m}{M-m} &= kt + c, \\ \ln\frac{m}{M-m} &= Mkt + Mc, \\ \frac{m}{M-m} &= \mathrm{e}^{Mkt}\mathrm{e}^{Mc}. \end{aligned}$$

   From here

$$m(t) = M\frac{\mathrm{e}^{Mkt}}{\mathrm{e}^{-Mc} + \mathrm{e}^{Mkt}}.$$

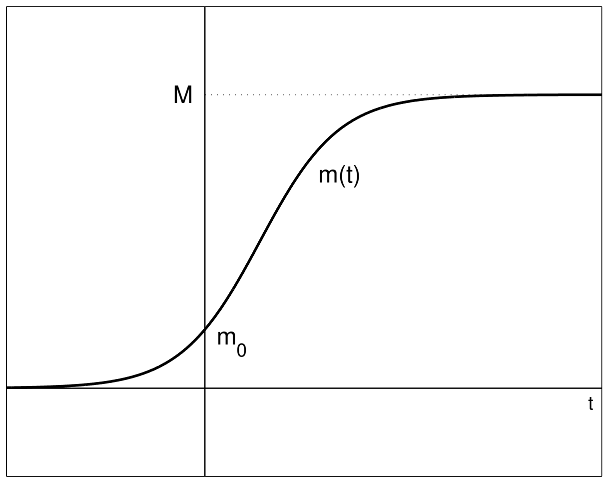


Figure 14.1

Due to the initial condition $m(0) = m_0$,

$$m_0 = M\frac{1}{e^{-Mc} + 1},$$

and so

$$e^{Mc} = \frac{m_0}{M - m_0}.$$

So, the solution is

$$m(t) = M\frac{e^{Mkt}}{\frac{M-m_0}{m_0} + e^{Mkt}}.$$

One can see that

$$\lim_{t\to\infty} m(t) = M\lim_{t\to\infty}\frac{1}{\frac{M-m_0}{m_0}e^{-Mkt} + 1} = M.$$

3. Solve the following differential equations:

   a) $y' = xy$, $x \in \mathbb{R}$,

   b) $y' = -y\mathrm{tg}x$, $x \in (-\pi/2, \pi/2)$,

   c) $y' = \frac{1}{2x}\sqrt{1+y^2}$, $x > 0$.

Chapter 15

Integration of multivariable functions

Now we generalize the integration of real functions in another direction. We will get to the volume of space under a surface, the calculation of which will be traced back to integrals of real functions. The following topics will be discussed.

- The definition of the Riemann integral for multivariable functions
- Calculating the integral on a rectangle by Fubini's theorem
- Calculating the integral on a normal domain
- Calculating the integral on other domains by integral transformation

15.1 Multiple integrals

15.1.1 The concept of multiple integral

Let $T := [a, b] \times [c, d] \subset \mathbb{R}^2$ be a closed rectangle. Let $f : \mathbb{R}^2 \supset\!\!\to \mathbb{R}$ be a continuous function of two variables for which $T \subset D(f)$. Prepare a partition $a = x_0 < x_1 < \ldots < x_{i-1} < x_i < \ldots < x_n = b$ of the interval $[a, b]$ and $c = y_0 < \ldots < y_{k-1} < y_k < \ldots < y_m = b$ of the interval $[c, d]$. At each subinterval $[x_{i-1}, x_i]$ we set a point $\xi_i \in [x_{i-1}, x_i]$, and at each sub-interval $[y_{k-1}, y_k]$ a point $\eta_k \in [y_{k-1}, y_k]$ $(i = 1, \ldots, n,\ k = 1, \ldots, m)$. Prepare the sum approximation

$$\sigma_{n,m} := \sum_{i=1}^{n} \sum_{k=1}^{m} f(\xi_i, \eta_k)(x_i - x_{i-1})(y_k - y_{k-1}).$$

(Here $\sigma_{n,m}$ can be illustrated as the sum of "signed" volumes of prisms with a rectangular base of area $[x_{i-1}, x_i] \times [y_{k-1}, y_k]$ and "height" $f(\xi_i, \eta_k)$ (which can be negative, too!).)

Under the condition that f is continuous, one can verify that the sum approximations have a limit in the sense that there exists a number $I \in \mathbb{R}$ such that for all $\varepsilon > 0$ there is a number $\delta > 0$ such that for all partitions satisfying the property

$$\max\{x_i - x_{i-1} \mid i = 1, 2, \ldots, n\} < \delta$$

and

$$\max\{y_k - y_{k-1} \mid k = 1, 2, \ldots, m\} < \delta$$

and for arbitrarily chosen values $\xi_i \in [x_{i-1}, x_i]$ $(i = 1, \ldots, n)$ and $\eta_k \in [y_{k-1}, y_k]$ $(k = 1, \ldots, m)$

$$|\sigma_{n,m} - I| < \varepsilon.$$

Such a number $I \in \mathbb{R}$ is called **integral of the function f on the rectangle** T and denoted as

$$\int_T f := I.$$

This concept is often referred to as

$$\begin{aligned}\int_T f &= \lim_{x_i - x_{i-1} \to 0,\ y_k - y_{k-1} \to 0} \sum_{i,k} f(\xi_i, \eta_k)(x_i - x_{i-1})(y_k - y_{k-1}) \\ &= \lim_{\Delta x_i \to 0,\ \Delta y_k \to 0} \sum_{i,k} f(\xi_i, \eta_k) \Delta x_i \Delta y_k = \int_{[a,b]\times[c,d]} f(x, y)\mathrm{d}x\mathrm{d}y.\end{aligned}$$

The number $\int_T f \in \mathbb{R}$ is called "signed" volume of the space

$$\begin{aligned}H := \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in T,\ 0 \le z \le f(x, y) \text{ if } f(x, y) \ge 0 \\ \text{or } f(x, y) \le z \le 0 \text{ if } f(x, y) < 0\}\end{aligned}$$

under the surface f.

15.1.2 Integration on rectangular and normal domains

Obviously, performing the procedure just introduced would make it rather complicated to calculate the integral of a function f on a rectangle T.

Let us remember how the integrals of real functions can be applied to the calculation of volumes. Let the plane section of H at an arbitrary point $x \in [a, b]$ be denoted by $S(x)$ (Fig. 15.1). This area is the integral of the function $[c, d] \ni y \mapsto f(x, y)$ on $[c, d]$:

$$S(x) = \int_c^d f(x, y)\mathrm{d}y.$$

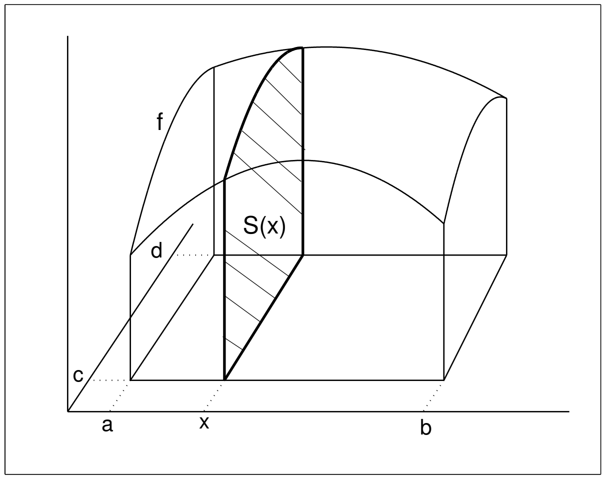


Figure 15.1

If this function $[a, b] \ni x \mapsto S(x)$ (which is continuous due to the continuity of f) is integrated on the interval $[a, b]$, then

$$\int_T f = \int_{[a,b]\times[c,d]} f(x, y)\mathrm{d}x\mathrm{d}y = \int_a^b S(x)\mathrm{d}x = \int_a^b \left(\int_c^d f(x, y)\mathrm{d}y \right) \mathrm{d}x.$$

By a similar consideration,

$$\int_T f = \int_c^d \left(\int_a^b f(x, y)\mathrm{d}x \right) \mathrm{d}y.$$

Theorem 15.1 (Fubini's threorem)**.** *Let $f : \mathbb{R}^2 \rightarrowtail \mathbb{R}$, f continuous and $[a, b] \times [c, d] \subset D(f)$. Then*

$$\int_{[a,b]\times[c,d]} f = \int_a^b \left(\int_c^d f(x, y)\mathrm{d}y \right) \mathrm{d}x = \int_c^d \left(\int_a^b f(x, y)\mathrm{d}x \right) \mathrm{d}y.$$

For example, let $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = xy$. $T := [0, 1] \times [2, 3]$. Then

$$\int_T f = \int_2^3 \left(\int_0^1 xy\mathrm{d}x \right) \mathrm{d}y = \int_2^3 \left[\frac{x^2}{2} y \right]_0^1 \mathrm{d}y = \int_2^3 \frac{y}{2}\mathrm{d}y = \left[\frac{y^2}{4} \right]_2^3 = \frac{9}{4} - 1 = \frac{5}{4}.$$

The definition of the integral of f on the rectangle T did not require the function f to be continuous. If f is not continuous, then it may happen that the number $I \in \mathbb{R}$ does not exist. If, however, the number I with the desired property exists, then the function f is called **integrable** on the rectangle T, and then

$$\int_T f := I.$$

With this remark we move on to the integrability and integration of a function $f : \mathbb{R}^2 \supset\!\!\to \mathbb{R}$ on domains that are not rectangular.

Let $\alpha, \beta : [a,b] \to \mathbb{R}$ be a continuous function, such that for all $x \in [a,b]$, $\alpha(x) \leq \beta(x)$. Let

$$N_x := \{(x,y) \in \mathbb{R}^2 \mid x \in [a,b] \text{ and } \alpha(x) \leq y \leq \beta(x)\}$$

be a **normal domain with respect to the x axis**. Let $f : N_x \to \mathbb{R}$ be a continuous function. Since $\alpha, \beta \in C[a,b]$, therefore there exist $c, d \in \mathbb{R}$ such that for all $x \in [a,b]$, $c \leq \alpha(x) \leq \beta(x) \leq d$. Let us extend the function f to the rectangle

$$T := [a,b] \times [c,d]$$

as follows:

$$\hat{f} : T \to \mathbb{R}, \quad \hat{f}(x,y) := \begin{cases} f(x,y) & \text{if } (x,y) \in N_x, \\ 0 & \text{if } (x,y) \in T \setminus N_x. \end{cases}$$

This function f is such that $\hat{f}_{|N_x}$ is continuous, while it is constant zero on the set $T \setminus N_x$. It can be verified that such a function $\hat{f}$ is integrable, and we define the integral of the function f on the normal domain N_x as the integral of the function $\hat{f}$ on the rectangle T:

$$\int_{N_x} f := \int_T \hat{f}.$$

By Fubini's theorem

$$\begin{aligned}
\int_{N_x} f = \int_T \hat{f} &= \int_a^b \left(\int_c^d \hat{f}(x,y)\mathrm{d}y \right) \mathrm{d}x \\
&= \int_a^b \left(\int_c^{\alpha(x)} \hat{f}(x,y)\mathrm{d}y + \int_{\alpha(x)}^{\beta(x)} \hat{f}(x,y)\mathrm{d}y + \int_{\beta(x)}^d \hat{f}(x,y)\mathrm{d}y \right) \mathrm{d}x \\
&= \int_a^b \left(\int_{\alpha(x)}^{\beta(x)} f(x,y)\mathrm{d}y \right) \mathrm{d}x,
\end{aligned}$$

since $[c, \alpha(x)] \ni y \mapsto \hat{f}(x,y) = 0$, $[\alpha(x), \beta(x)] \ni y \mapsto \hat{f}(x,y) = f(x,y)$ and $[\beta(x), d] \ni y \mapsto \hat{f}(x,y) = 0$ for all $x \in [a,b]$.

For example, let $f : \mathbb{R}^2 \to \mathbb{R},\ f(x,y) = xy$ and

$$N_x := \left\{(x,y) \in \mathbb{R}^2 \mid x \in [-1,1] \text{ és } x^2 - 1 \le y \le 1 - x^2\right\}.$$

Then

$$\begin{aligned}\int_{N_x} f &= \int_{-1}^{1} \left(\int_{x^2-1}^{1-x^2} xy\mathrm{d}y \right) \mathrm{d}x = \int_{-1}^{1} \left[x\frac{y^2}{2} \right]_{x^2-1}^{1-x^2} \mathrm{d}x \\ &= \int_{-1}^{1} \frac{x}{2}(1-x^2)^2 - \frac{x}{2}(x^2-1)^2 \mathrm{d}x = \int_{-1}^{1} 0\,\mathrm{d}x = 0.\end{aligned}$$

We obtain the integral of f on a normal domain N_y with respect to y with obvious modifications.

We can similarly construct the integral of a function $f : \mathbb{R}^3 \supset\!\to \mathbb{R}$ on the brick $T := [a_1,b_1] \times [a_2,b_2] \times [a_3,b_3]$, the corresponding Fubini's theorem, and then the integral on a normal domain N_{xy} with respect to the plane xy.

The set $N_{xy} \subset \mathbb{R}^3$ is a **normal domain** with respect to the plane xy if there exists a closed interval $[a,b] \subset \mathbb{R}$ and continuous functions $\alpha, \beta : [a,b] \to \mathbb{R}$ for which $\alpha(x) \le \beta(x)$ $(x \in [a,b])$, and there exist continuous functions $\lambda, \mu : \mathbb{R}^2 \supset\!\to \mathbb{R}$ for which $\lambda(x,y) \le \mu(x,y)$ $(x \in [a,b],\ \alpha(x) \le y \le \beta(x))$ such that

$$N_{xy} = \left\{(x,y,z) \in \mathbb{R}^3 \mid x \in [a,b],\ \alpha(x) \le y \le \beta(x),\ \lambda(x,y) \le z \le \mu(x,y)\right\}.$$

Assume that $f : \mathbb{R}^3 \supset\!\to \mathbb{R}$, it is continuous, and $N_{xy} \subset D(f)$. Then

$$\int_{N_{xy}} f = \int_a^b \left\{ \int_{\alpha(x)}^{\beta(x)} \left(\int_{\lambda(x,y)}^{\mu(x,y)} f(x,y,z)\mathrm{d}z \right) \mathrm{d}y \right\} \mathrm{d}x.$$

15.1.3 The transformation of integrals

The integration by substitution of real functions has its equivalent in multiple integration. According to the integration by substitution in the real case, if $\phi : [\alpha,\beta] \to [a,b]$ is a strictly monotonically increasing one-to-one function, and $\phi \in D$, then

$$\int_a^b f(x)\mathrm{d}x = \int_\alpha^\beta f(\phi(t)) \cdot \phi'(t)\mathrm{d}t.$$

Assume that we would like to integrate the function $f : \mathbb{R}^2 \supset\!\to \mathbb{R}$ on the set $Q \subset D(f)$. If we are lucky, then we can find such a one-to-one function $\Phi = (\phi,\psi) : T \to Q$, where $T = [\alpha,\beta] \times [\gamma,\delta] \subset \mathbb{R}^2$ is a rectangle, Φ is continuously differentiable, and for all $(u,v) \in T$

$$\det \Phi'(u,v) = \begin{vmatrix} \partial_u \phi(u,v) & \partial_v \phi(u,v) \\ \partial_u \psi(u,v) & \partial_v \psi(u,v) \end{vmatrix} \ne 0.$$

One can prove that

$$\int_Q f = \int_T f(\phi(u,v),\psi(u,v)) \cdot |\det \Phi'(u,v)| \mathrm{d}u\mathrm{d}v.$$

For example, let $Q := \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 \le 4\}$, which is a closed circle of radius 2, centered at the origin. Then $f : \mathbb{R}^2 \to \mathbb{R},\ f(x,y) := x^2+y^2$.

Since

$$(\phi,\psi) := [0,2]\times[0,2\pi] \to Q,\ \ \phi(u,v) := u\cos v,\ \ \psi(u,v) := u\sin v$$

is a one-to-one function (known as polar transformation), and

$$\det(\phi,\psi)'(u,v) = \begin{vmatrix} \cos v & -u\sin v \\ \sin v & u\cos v \end{vmatrix} = u\cos^2 v + u\sin^2 v = u,$$

therefore

$$\begin{aligned}\int_Q x^2+y^2\mathrm{d}x\mathrm{d}y &= \int_{[0,2]\times[0,2\pi]} \left\{(u\cos v)^2 + (u\sin v)^2\right\} u\mathrm{d}u\mathrm{d}v \\ &= \int_0^2 \left(\int_0^{2\pi} u^3\mathrm{d}v\right)\mathrm{d}u = \int_0^2 [u^3 v]_0^{2\pi}\mathrm{d}u = \left[2\pi\frac{u^4}{4}\right]_0^2 = 8\pi.\end{aligned}$$

It makes the integration of a function $f : \mathbb{R}^3 \supset\!\!\to \mathbb{R}$ easier if we notice that by the transformation

$$\begin{aligned} X(r,\phi,\vartheta) &:= r\sin\vartheta\cos\phi, \\ Y(r,\phi,\vartheta) &:= r\sin\vartheta\sin\phi, \\ Z(r,\phi,\vartheta) &:= r\cos\phi \end{aligned}$$

the brick

$$[R_1,R_2]\times[\phi_1,\phi_2]\times[\vartheta_1,\vartheta_2] =: T$$

is mapped to the domain of integration $Q \subset \mathbb{R}^3$ by the function $\Phi := (X,Y,Z) : T \to Q$ in a one-to-one manner, and $\det\Phi' \ne 0$ on the brick T. Then

$$\det\Phi'(r,\phi,\vartheta) = \begin{vmatrix} \sin\vartheta\cos\phi & -r\sin\vartheta\sin\phi & r\cos\vartheta\cos\phi \\ \sin\vartheta\sin\phi & r\sin\vartheta\cos\phi & r\cos\vartheta\sin\phi \\ \cos\vartheta & 0 & -r\sin\vartheta \end{vmatrix} = -r^2\sin\vartheta.$$

Then

$$\int_Q f(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z = \int_T f(X(r,\phi,\vartheta),Y(r,\phi,\vartheta),Z(r,\phi,\vartheta)) \cdot r^2\sin\vartheta \cdot \mathrm{d}r\mathrm{d}\phi\mathrm{d}\vartheta.$$

The polar transformation, applied here, is suitable for domains Q that are part of a sphere (hemisphere, spherical layer etc.).

Chapter 16

Vector analysis

We define and calculate the characteristics of space curves (curvature, torsion, arc length). We introduce the concept of integral on a surface. As a generalization of the Newton–Leibniz theorem, we formulate theorems for integral transforms (Gauss, Stokes). The following topics well be discussed.

- Tangent, binormal and principal normal vectors of a curve
- Rectifiable curves and their lengths
- Curvature, osculating circle, torsion
- Parametric definition of a surface
- The definition of surface
- Surface integral of a scalar field
- Surface integral of a vector field
- Gradient, divergence, rotation
- Stokes' theorem, Gauss' theorem

16.1 Vector analysis

16.1.1 Space curves

Let $r : [\alpha, \beta] \to \mathbb{R}^3$ be a sufficiently smooth space curve ($\dot{r}, \ddot{r}, \dddot{r} \in C$ and for all $t \in (\alpha, \beta)$ $\dot{r}(t), \ddot{r}(t), \dddot{r}(t) \neq 0$). As we have seen, $\dot{r}(t_0)$ is a **tangent vector** to the curve at the point corresponding to the parameter value t_0. Denote by $\underline{t}$ the unit vector pointing to the direction of $\dot{r}(t_0)$:

$$\underline{t} := \frac{\dot{r}(t_0)}{\|\dot{r}(t_0)\|}.$$

This is called **tangential** vector.

Now let the point P_0 on the curve be the end point of the vector $r(t_0)$. Take two arbitrary points P_1 and P_2 ($P_1, P_2 \neq P_0$) on the curve. If P_1, P_0, P_2 do not fall on one straight line, then they determine a plane. Let P_1 and P_2 both approach P_0. Assume that the planes determined by them also approach a limit position, which is a plane, too. This plane is called **osculating plane** of the curve belonging to the point P_0 (in fact, this plane contains a small piece of the curve close to P_0). It is possible to show that the osculating plane is spanned by the vectors $\dot{r}(t_0)$ and $\ddot{r}(t_0)$, therefore $\dot{r}(t_0) \times \ddot{r}(t_0)$ is a normal vector of the plane. Therefore, for the position vector $\underline{r}$ of any point of the osculating plane we have

$$\langle (\dot{r}(t_0) \times \ddot{r}(t_0)), \underline{r} - r(t_0) \rangle = 0 \quad \text{(the equation of the osculating plane).}$$

The unit vector derived from a normal vector of the osculating plane is called **binormal vector**:

$$\underline{b} := \frac{\dot{r}(t_0) \times \ddot{r}(t_0)}{\|\dot{r}(t_0) \times \ddot{r}(t_0)\|}.$$

Clearly, the binormal vector $\underline{b}$ is perpendicular to the tangential vector $\underline{t}$. The plane spanned by $\underline{b}$ and $\underline{t}$ is called **rectifying plane**. A normal vector to this plane is $\underline{t} \times \underline{b}$, and the unit vector $\underline{f}$ derived from this vector is called **principal normal vector**:

$$\underline{f} := \frac{(\dot{r}(t_0) \times \ddot{r}(t_0)) \times \dot{r}(t_0)}{\|(\dot{r}(t_0) \times \ddot{r}(t_0)) \times \dot{r}(t_0)\|}.$$

The plane spanned by $\underline{f}$ and $\underline{b}$ is called **normal plane** (a normal vector of which is the tangent vector $\dot{r}(t_0)$).

The vectors $\underline{t}$, $\underline{f}$ and $\underline{b}$ are pairwise orthogonal unit vectors, which form a right-handed system in this order. The vector system $\underline{t}$, $\underline{f}$, $\underline{b}$ fit to the point of the curve with position vector $r(t_0)$ is called **accompanying trieder** (if t_0 changes, the accompanying trieder changes as well, but this system seems quite natural for the curve).

The length of a path is important to know even our in everyday life. We will clarify when a space curve $r : [\alpha, \beta] \to \mathbb{R}^3$ has an arc length, and if it does, how we can define it.

Let τ be an arbitrary partition of $[\alpha, \beta]$:

$$\tau : \alpha = t_0 < t - 1 < \ldots < t_{i-1} < t_i < \ldots < t_n = \beta.$$

The vectors $r(t_{i-1})$ and $r(t_i)$ are position vectors of two points on the curve, so $\|r(t_i) - r(t_{i-1})\|$ is the length of the section connecting them. Let

$$L(\tau) := \sum_{i=1}^{n} \|r(t_i) - r(t_{i-1})\|,$$

that is, the broken line corresponding to the points $r(t_0), r(t_1), \ldots r(t_n)$. Let us prepare the set containing the lengths of all these broken lines:

$$\{L(\tau) \mid \tau \text{ is a partition of the interval } [\alpha, \beta]\}.$$

If this set is bounded above, then the space curve r is called **rectifiable** ("straightenable"), and then the real number

$$\sup\{L(\tau) \mid \tau \text{ is a partition of the interval } [\alpha, \beta]\} =: L$$

is called length of the space curve. If this set is not bounded above, then the curve has no length (or infinite length).

We have already seen that in case of a smooth curve

$$r(t_i) - r(t_{i-1}) \approx \dot{r}(\xi_i) \cdot (t_i - t_{i-1}) \quad (\xi_i \in [t_{i-1}, t_i]),$$

thus

$$L(\tau) = \sum_{i=1}^{n} \|r(t_i) - r(t_{i-1})\| \approx \sum_{i=1}^{n} \|\dot{r}(\xi_i) \cdot (t_i - t_{i-1})\|,$$

which is the sum approximation of an integral. One can show that in case of a smooth curve this leads us to the arc length of the curve: a smooth curve is rectifiable, and

$$L = \int_\alpha^\beta \|\dot{r}(t)\| \mathrm{d}t.$$

If $r : [\alpha, \beta] \to \mathbb{R}$ is a smooth curve, then for all $t \in [\alpha, \beta]$ let $s(t) := \int_\alpha^t \|\dot{r}(u)\| \mathrm{d}u$, that is, the arc length of the curve between the parameter point α and t (which also exists obviously). From the definition one can see that

$$s'(t) = \frac{\mathrm{d}}{\mathrm{d}t} \int_\alpha^t \|\dot{r}(u)\| \mathrm{d}u = \|\dot{r}(t)\|.$$

If $t, t + \Delta t \in [\alpha, \beta]$, then

$$\Delta s := s(t + \Delta t) - s(t) = \int_t^{t+\Delta t} \|\dot{r}(u)\| \mathrm{d}u \approx \|\dot{r}(t)\| \Delta t \text{ if} \Delta t \approx 0.$$

From this we can derive the **orbital** (tangential) velocity of a point moving along the space curve r:

$$v(t) = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t} = \frac{\mathrm{d}s}{\mathrm{d}t} = \|\dot{r}(t)\|.$$

The concept of angular velocity is also frequently used. Assume that the angle enclosed by the vectors $r(t)$ and $r(t+\Delta t)$ is $\Delta\phi$. We define the **angular velocity** at the point t of the space curve r as

$$\omega(t) := \lim_{\Delta t \to 0} \frac{\Delta \phi}{\Delta t}.$$

For a sufficiently smooth curve we calculate the angular velocity $\omega(t)$. It is known that

$$\|r(t) \times r(t+\Delta t)\| = \|r(t)\| \cdot \|r(t+\Delta t)\| \cdot \sin \Delta\phi.$$

By exploiting the fact that $r(t) \times r(t) = \underline{0}$,

$$\sin \Delta\phi = \frac{\|r(t) \times (r(t+\Delta t) - r(t))\|}{\|r(t)\| \cdot \|r(t+\Delta t)\|}.$$

We remember that $\lim_{\Delta\phi \to 0} \frac{\sin \Delta\phi}{\Delta\phi} = 1$, and in case $\Delta t \to 0$, $\Delta\phi \to 0$, we continue the calculation as

$$\frac{\sin \Delta\phi}{\Delta\phi} \cdot \frac{\Delta\phi}{\Delta t} = \frac{\|r(t) \times \frac{r(t+\Delta t)-r(t)}{\Delta t}\|}{\|r(t)\| \cdot \|r(t+\Delta t)\|},$$

$$\lim_{\Delta t \to 0} \frac{r(t+\Delta t) - r(t)}{\Delta t} = \dot{r}(t),$$

$$\omega(t) = \lim_{\Delta t \to 0} \frac{\Delta\phi}{\Delta t} = \frac{\|r(t) \times \dot{r}(t)\|}{\|r(t)\|^2}.$$

When sitting in a car we arrive at a bend, it is important to know how steep the bend is. Denote by Δs the arc length of a sufficiently smooth curve between the parameter points t and $t + \Delta t$. Let the angle of the tangent vectors at these two points be $\Delta\alpha$. The **curvature** at point t is defined as

$$G(t) := \lim_{\Delta s \to 0} \frac{\Delta\alpha}{\Delta s}.$$

The curvature informs us about the change of the angle along the displacement Δs. If G is big, then the bend is “steep”, if G is close to zero, then the path is practically straight.

We calculate the curvature in case of a sufficiently smooth curve ($\dot{r}, \ddot{r} \in C$). If $\Delta s \approx 0$, then $\Delta t \approx 0$, so

$$\frac{\Delta\alpha}{\Delta s} = \frac{\Delta\alpha}{\Delta t} : \frac{\Delta s}{\Delta t},$$

from which

$$G(t) = \lim_{\Delta s \to 0} \frac{\Delta\alpha}{\Delta s} = \lim_{\Delta t \to 0} \frac{\Delta\alpha}{\Delta t} : \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t} = \Omega(t) : \|\dot{r}(t)\|,$$

where $\Omega(t)$ is the angular velocity of $\dot{r}$ (this is also a space curve!) at the point belonging to the parameter t. So,

$$G(t) = \frac{\|\dot{r}(t) \times \ddot{r}(t)\|}{\|\dot{r}(t)\|^2} : \|\dot{r}(t)\| = \frac{\|\dot{r}(t) \times \ddot{r}(t)\|}{\|\dot{r}(t)\|^3}.$$

Let

$$R(t) := \frac{1}{G(t)} > 0$$

in case $G(t) \neq 0$. It is possible to verify that the circle of radius $R(t)$ that fits the point of the curve belonging to the parameter t, which is in the osculating plane, and whose center lies on the principal normal vector $\underline{f}$, approaches the curve most tightly. It is called **osculating circle**. Any movement along a small section of the curve can be replaced by movement along the osculating circle.

Curvature measures how much a curve deviates from the straight line. Another characteristic property, the so-called torsion informs us about the deviation of a space curve from a plane curve.

A normal vector of the osculating plane of a curve is called binormal vector. If the osculating plane changes with the parameter, then this is shown by the deflection of the binormal vector. The change in the angle of the binormal vector whereas the arc length changes by Δs is characterized by the torsion (twist), that is, the **torsion** at the location of parameter value t is defined as

$$T(t) := \lim_{\Delta s \to 0} \frac{\Delta \beta}{\Delta s},$$

where $\Delta\beta$ is the angle enclosed by the normal vectors at the points $r(t)$ and $r(t+\Delta t)$ of the curve (that is, the angle of $\underline{b}(t)$ and $\underline{b}(t+\Delta t)$), and Δs is the arc length between the two points. Similarly to our previous consideration, for a sufficiently smooth curve ($\dot{r}, \ddot{r}, \dddot{r} \in C$)

$$\begin{aligned} T(t) = \lim_{\Delta s \to 0} \frac{\Delta \beta}{\Delta s} &= \lim_{\Delta t \to 0} \frac{\Delta \beta}{\Delta t} : \frac{\Delta s}{\Delta t} = \lim_{\Delta t \to 0} \frac{\Delta \beta}{\Delta t} : \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t} \\ &= \frac{\|\underline{b}(t) \times \dot{\underline{b}}(t)\|}{\|\underline{b}(t)\|^2} : \|\dot{r}(t)\|, \end{aligned}$$

where the dividend is the angle velocity of the binormal $\underline{b}$ as a space curve. Substituting the already known form of the binormal vector, after some simplifications we obtain that

$$T(t) = \frac{|\langle \dot{r}(t) \times \ddot{r}(t), \dddot{r}(t)\rangle|}{\|\dot{r}(t) \times \ddot{r}(t)\|^2}.$$

We remark that if we do not take the absolute value of the "mixed product" in the numerator, then the curve makes a right-handed screw if $T(t) > 0$, and a left-handed screw if $T(t) < 0$. In case of screws and winding staircases this may also be of great importance ...

16.1.2 Surfaces

Let S be a plane in the space. Let the vector $\underline{r}_0$ be directed to a point of the plane S. Assume that $\underline{a}$ and $\underline{b}$ are two non-parallel plane vectors. It is known that a vector $\underline{r}$ directed to an arbitrary point of the plane S can be given in the form

$$\underline{r} = \underline{r}_0 + u\underline{a} + v\underline{b},$$

where $u, v \in \mathbb{R}$ are suitable numbers. To put it coordinate-wise,

$$\begin{aligned} x &= x_0 + ua_1 + vb_1, \\ y &= y_0 + ua_2 + vb_2, \\ z &= z_0 + ua_3 + vb_3. \end{aligned}$$

So, we could give any point of the plane S with the aid of three functions of two variables.

This is generally true. Let

$$\Phi : \mathbb{R}^2 \supset\!\!\to \mathbb{R}^3, \quad \Phi = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix},$$

where $X, Y, Z : \mathbb{R}^2 \supset\!\!\to \mathbb{R}$. If $\Omega := D(\Phi) \subset \mathbb{R}^2$, then for all $(u, v) \in \Omega$

$$\begin{bmatrix} X(u,v) \\ Y(u,v) \\ Z(u,v) \end{bmatrix} \in \mathbb{R}^3$$

defines a position vector of a point of the plane. These points define a (two-parameter) surface. For example, the function $\Phi : [0, 2\pi] \times [0, \pi] \to \mathbb{R}^3$,

$$\Phi(u,v) := \begin{bmatrix} 3\sin v \cos u \\ 3\sin v \sin u \\ 3\cos v \end{bmatrix}$$

is the two-parameter representation of the surface of a sphere with radius $R = 3$, centered at the point $(0, 0, 0)$ (Fig. 16.1).

Let $\Phi : \mathbb{R}^2 \supset\!\!\to \mathbb{R}^3$, $(u_0, v_0) \in D(\Phi)$. The curve $p : \mathbb{R} \supset\!\!\to \mathbb{R}^3$, $p(u) := \Phi(u, v_0)$, lying on the surface is called **u-parameter curve**, while $q : \mathbb{R} \supset\!\!\to \mathbb{R}^3$, $q(v) := \Phi(u_0, v)$ is called **v-parameter curve** (Fig. 16.2).

If Φ is a smooth function (the partial derivatives of the coordinate functions X, Y, Z are continuous), then $\dot{p}(u_0)$ and $\dot{q}(v_0)$ are tangent vectors of the az u-parameter curve and the v-parameter curve, and then $\underline{n} := \dot{p}(u_0) \times \dot{q}(v_0)$ is

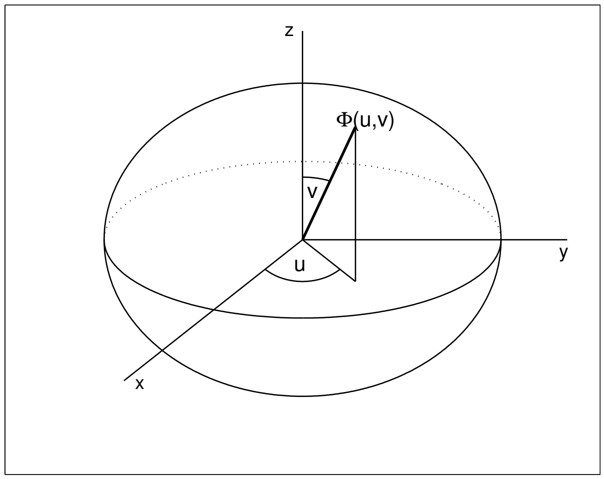


Figure 16.1

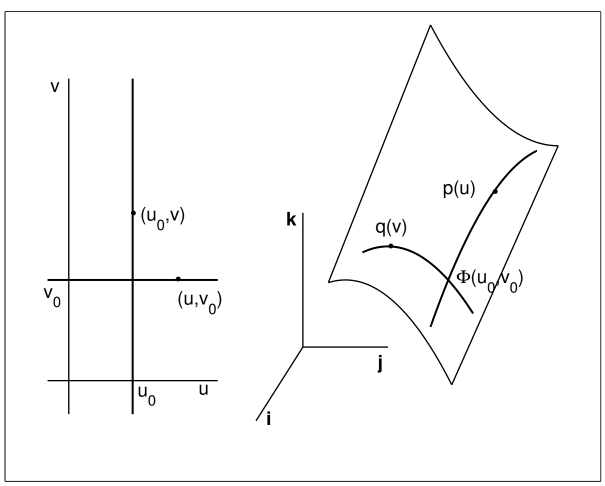


Figure 16.2

a normal vector to the tangent plane of the surface Φ at the point $\Phi(u_0, v_0)$. Since

$$\dot{p}(u_0) = \begin{bmatrix} \partial_u X(u_0, v_0) \\ \partial_u Y(u_0, v_0) \\ \partial_u Z(u_0, v_0) \end{bmatrix} \quad \text{and} \quad \dot{q}(v_0) = \begin{bmatrix} \partial_v X(u_0, v_0) \\ \partial_v Y(u_0, v_0) \\ \partial_v Z(u_0, v_0) \end{bmatrix},$$

therefore the normal vector of the tangent plane is the determinant

$$\underline{n} = \begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ \partial_u X & \partial_u Y & \partial_u Z \\ \partial_v X & \partial_v Y & \partial_v Z \end{vmatrix}$$

where the partial derivatives are to be evaluated at the point (u_0, v_0).

Let Φ be a smooth surface, and consider the rectangle defined by the points $(u, v),\ (u + \Delta u, v),\ (u + \Delta u, v + \Delta v),\ (u, v + \Delta v) \in D(\Phi)$. The area of this rectangle is $\Delta u \cdot \Delta v$. Then

$$\Phi(u + \Delta u, v) - \Phi(u, v) \approx \begin{bmatrix} \partial_u X(u, v) \\ \partial_u Y(u, v) \\ \partial_u Z(u, v) \end{bmatrix} \Delta u =: \underline{a},$$

$$\Phi(u, v + \Delta v) - \Phi(u, v) \approx \begin{bmatrix} \partial_v X(u, v) \\ \partial_v Y(u, v) \\ \partial_v Z(u, v) \end{bmatrix} \Delta v =: \underline{b},$$

therefore, the area of the piece of surface, enclosed by the u- and v-parameter curves, characterized by the "vertices" $\Phi(u, v),\ \Phi(u + \Delta u, v),\ \Phi(u + \Delta u, v + \Delta v),\ \Phi(u, v + \Delta v)$ can be approximated with the area of a parallelogram, lying on the tangent plane of the surface at the point $\Phi(u, v)$, the side vectors of which are $\underline{a}$ and $\underline{b}$. This area can be expressed by the vector product

$$\|\underline{a} \times \underline{b}\| = \|\underline{n}\| \Delta u \Delta v = \| \begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ \partial_u X & \partial_u Y & \partial_u Z \\ \partial_v X & \partial_v Y & \partial_v Z \end{vmatrix} \| \cdot \Delta u \Delta v.$$

Partitioning the parameter domain finely enough with straight lines parallel with the u and v axes, cells of area $\Delta u \Delta v$ are obtained (Fig. 16.3).

The image of a cell is a "curvilinear cell" on the surface, the area of which has just been calculated. Summing these up, we obtain a sum approximation of the area of the surface Φ:

$$\sum_u \sum_v \| \begin{bmatrix} \partial_u X(u, v) \\ \partial_u Y(u, v) \\ \partial_u Z(u, v) \end{bmatrix} \times \begin{bmatrix} \partial_v X(u, v) \\ \partial_v Y(u, v) \\ \partial_v Z(u, v) \end{bmatrix} \| \Delta u \Delta v,$$

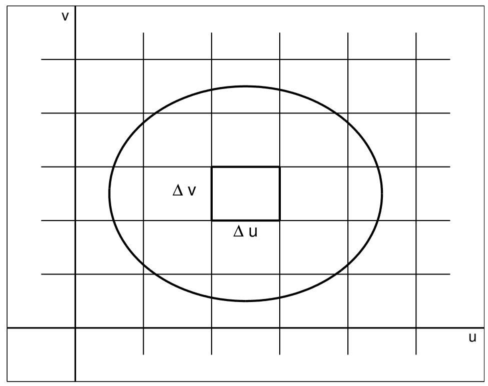


Figure 16.3

which, by refining the partition beyond all bounds, tends to the integral of the surface Φ over its domain of definition Ω. So, the area of the surface Φ is

$$S := \int_\Omega \|\partial_u \Phi \times \partial_v \Phi\| \mathrm{d}u \mathrm{d}v,$$

where

$$\partial_u \Phi = \begin{bmatrix} \partial_u X \\ \partial_u Y \\ \partial_u Z \end{bmatrix} \quad \text{and} \quad \partial_v \Phi = \begin{bmatrix} \partial_v X \\ \partial_v Y \\ \partial_v Z \end{bmatrix}.$$

We can simplify the integrand. Since for the vectors $\underline{a}$, $\underline{b}$

$$\begin{aligned} \|\underline{a} \times \underline{b}\|^2 &= \|\underline{a}\|^2 \cdot \|\underline{b}\|^2 \sin^2 \alpha = \|\underline{a}\|^2 \|\underline{b}\|^2 (1 - \cos^2 \alpha) \\ &= \|\underline{a}\|^2 \|\underline{b}\|^2 - \|\underline{a}\|^2 \|\underline{b}\|^2 \cos^2 \alpha = \|\underline{a}\|^2 \|\underline{b}\|^2 - (\langle \underline{a}, \underline{b} \rangle)^2, \end{aligned}$$

therefore

$$S = \int_\Omega (\|\partial_u \Phi\|^2 \|\partial_v \Phi\|^2 - (\langle \partial_u \Phi, \partial_v \Phi \rangle)^2)^{1/2} \mathrm{d}u \mathrm{d}v.$$

Surface integral of a scalar field

Let $\Phi : \mathbb{R}^2 \supset\!\!\to \mathbb{R}^3,\ \Omega := D(\Phi)$ be a smooth surface. Assume that to each point of the surface a real number is assigned, so, let $U : \mathbb{R}^3 \supset\!\!\to \mathbb{R},\ D(U) := \Phi(\Omega)$. Assume that U is continuous. The integral of the "scalar function" U on the surface Φ is defined in the usual manner:

1^o We divide the parameter domain Ω into cells of area $\Delta u \Delta v$.

2^o We take arbitrary points (u', v') in the cells.

3^o We prepare the product $U(\Phi(u',v')) \cdot \|\partial_u\Phi \times \partial_v\Phi\|\Delta u\Delta v$ (we multiplied the value of the function U at a surface point by the approximate area of the image of the cell of area $\Delta u\Delta v$).

4^o The sum $\sum_u \sum_v U(\Phi(u',v')) \cdot \|\partial_u\Phi \times \partial_v\Phi\|\Delta u\Delta v$ is the approximation of an integral.

5^o The integral of the scalar function U on the surface Φ is defined as the limit of the sum approximations:

$$\begin{aligned}\int_\Phi U &:= \lim_{\Delta u\to 0, \Delta v\to 0} \sum_u \sum_v U(\Phi(u',v')) \cdot \|\partial_u\Phi \times \partial_v\Phi\|\Delta u\Delta v \\ &= \int_\Omega U(\Phi(u,v)) \cdot \|\partial_u\Phi(u,v) \times \partial_v\Phi(u,v)\|\mathrm{d}u\mathrm{d}v.\end{aligned}$$

Surface integral of a vector field

Let $\Phi : \mathbb{R}^2 \supset\!\!\to \mathbb{R}^3$, $\Omega := D(\Phi)$ be a smooth surface. Assume that to each point of the surface a vector is assigned, that is $F : \mathbb{R}^3 \supset\!\!\to \mathbb{R}^3$, $D(F) = \Phi(\Omega)$. Assume that F is continuous. The integral of the "vector-valued" function F on the surface Φ can be defined in the usual manner:

1^o We divide the parameter domain Ω into cells of area $\Delta u\Delta v$.

2^o We take arbitrary points (u',v') in the cells.

3^o We prepare the vector $F(\Phi(u',v'))$.

4^o To the point $\Phi(u,v)$, belonging to the "vertex" (u,v) of the cell, corresponds a tangent plane, a normal vector of which is

$$\partial_u\Phi(u,v) \times \partial_v\Phi(u,v).$$

Since the area of the surface element corresponding to the cell of area $\Delta u\Delta v$ is

$$\Delta S \approx \|\partial_u\Phi(u,v)\Delta u \times \partial_v\Phi(u,v)\Delta v\| = \|\partial_u\Phi(u,v) \times \partial_v\Phi(u,v)\|\Delta u\Delta v,$$

therefore the vector

$$\underline{\Delta S} := (\partial_u\Phi(u,v) \times \partial_v\Phi(u,v))\Delta u\Delta v$$

is called **surface vector**. (The length of $\underline{\Delta S}$ is exactly the area of the surface element, and its direction is orthogonal to the surface so that it forms a right-handed system with the vectors $\dot{p}(u)$, $\dot{q}(v)$.)

5^o We prepare the sum of scalar products

$$\sum_u \sum_v \langle F(\Phi(u', v')), \underline{\Delta S}\rangle.$$

This is the sum approximation of an integral.

6^o The integral of the vector-valued function F on the surface Φ is defined as the limit of the sum approximations:

$$\int_\Phi F := \lim_{\Delta u \to 0, \Delta v \to 0} \sum_u \sum_v \langle F(\Phi(u', v')), \partial_u \Phi(u, v) \times \partial_v \Phi(u, v)\rangle \Delta u \Delta v$$
$$= \int_\Omega \langle F(\Phi(u, v)), \partial_u \Phi(u, v) \times \partial_v \Phi(u, v)\rangle \mathrm{d}u\mathrm{d}v. \tag{16.1}$$

16.1.3 The nabla symbol

Nabla is a symbolic vector, which expresses partial differentiation with respect to x, y and z. It is denoted as ∇. It can be used as a vector and can be employed as a factor in scalar and vector products.

$$\nabla := \begin{bmatrix} \partial_x \\ \partial_y \\ \partial_z \end{bmatrix}$$

If $f : \mathbb{R}^3 \supset\!\!\to \mathbb{R}$ is a smooth function, then

$$\mathrm{grad} f = \nabla f = \begin{bmatrix} \partial_x f \\ \partial_y f \\ \partial_z f \end{bmatrix}.$$

(Here f behaves like a scalar multiplier of a vector, however, in an unusual way it is located behind the vector.)

If $f : \mathbb{R}^3 \supset\!\!\to \mathbb{R}^3$ is a smooth function, then let the sum of the elements in the main diagonal of the derivative matrix $f'(x, y, z) \in \mathbb{R}^{3\times 3}$ be

$$\mathrm{div} f(x, y, z) = \partial_x f_1(x, y, z) + \partial_y f_2(x, y, z) + \partial_z f_3(x, y, z).$$

The **divergence** of f in terms of the nabla vector reads as

$$\mathrm{div} f = \langle \nabla, f\rangle$$

(the scalar product of nabla and the vector f).

If $f : \mathbb{R}^3 \supset\!\!\to \mathbb{R}^3$ is a smooth function, then the vector consisting of the differences of the elements, symmetrical to the main diagonal of the derivative matrix $f'(x,y,z) \in \mathbb{R}^{3\times 3}$ is called **rotation** of f, and

$$\mathrm{rot} f(x,y,z) := \begin{bmatrix} \partial_y f_3(x,y,z) - \partial_z f_2(x,y,z) \\ \partial_z f_1(x,y,z) - \partial_x f_3(x,y,z) \\ \partial_x f_2(x,y,z) - \partial_y f_1(x,y,z) \end{bmatrix}.$$

The rotation of f can be expressed in terms of the ∇ vector as

$$\mathrm{rot} f = \nabla \times f$$

(the vector product of nabla and the vector f).

The meaning of $\mathrm{grad} f$ was discovered in connection with the directional derivative, while $\mathrm{rot} f$ arose in the topic of the line integral, where the sufficient condition of the existence of a potential was investigated. ($\mathrm{div} f$ will soon appear, too.) One can see that grad, div and rot are differentiation operations, which become more transparent with the aid of the ∇ symbol.

∇ really behaves like a vector. For example, in case of a sufficiently smooth function $f : \mathbb{R}^3 \supset\!\!\to \mathbb{R}$

$$\mathrm{rot}(\mathrm{grad} f) = \nabla \times (\nabla f) = 0,$$

since ∇ and ∇f are "parallel". (After performing the tedious derivations we would get the same result.)

Note that the scalar product of ∇ by itself is known as the Laplacian operator:

$$\triangle := \langle \nabla, \nabla \rangle,$$

that is, if $f : \mathbb{R}^3 \supset\!\!\to \mathbb{R}$ is a sufficiently smooth scalar function, then

$$\triangle f := \mathrm{div}(\mathrm{grad} f) = \partial^2_{xx} f + \partial^2_{yy} f + \partial^2_{zz} f$$

is called "Laplacian of f".

16.1.4 Theorems for integral transforms

We are going to generalize the Newton–Leibniz theorem of real functions. As a consequence of this theorem, if a function $f : \mathbb{R} \supset\!\!\to \mathbb{R}$ is continuously differentiable, then

$$\int_a^b f' = f(b) - f(a),$$

which means that the integral of the function f on the set $[a,b]$ is equal to the change of the function on the boundary of the set.

The first generalization is as follows:

Let $\Phi : \mathbb{R}^2 \supset\!\!\to \mathbb{R}^3$ be a smooth surface, bounded by an oriented curve $r : \mathbb{R} \supset\!\!\to \mathbb{R}^3$ (r should be positively oriented from the viewpoint of the surface vectors).

Theorem 16.1 (Stokes' theorem)**.** *If $F : \mathbb{R}^3 \supset\!\!\to \mathbb{R}^3$ is a smooth vector function, then*

$$\int_\Phi \mathrm{rot}F = \int_r F,$$

that is, the surface integral of $\mathrm{rot}F$ *(here a differential operator has been applied to the function F) is equal to the line integral of F on the boundary of the surface.*

The second generalization is as follows:

Let $\Phi : \mathbb{R}^2 \supset\!\!\to \mathbb{R}^3$ be a closed, smooth surface, enclosing the space $V \subset \mathbb{R}^3$ (the surface vectors are oriented "outwards").

Theorem 16.2 (Gauss' theorem)**.** *If $F : \mathbb{R}^3 \supset\!\!\to \mathbb{R}^3$ is a smooth vector function, then*

$$\int_V \mathrm{div}F = \int_\Phi F,$$

that is, when integrating the function $\mathrm{div}F$ *on the space domain V (here another differential operator has been applied to the function F), this integral can be given as a surface integral on the boundary of the space domain.*

16.2 Exercises

1. Consider the helix $r : [0, 6\pi] \to \mathbb{R}^3,\ r(t) := \begin{bmatrix} \cos t \\ \sin t \\ t \end{bmatrix}$. At the parameter point $t_0 := \frac{\pi}{2}$ calculate

   a) the vectors $\underline{t}, \underline{f}, \underline{b}$ of the accompanying trieder,

   b) the curvature $G(t_0)$, and the radius $R(t_0)$ of the osculating circle,

   c) the torsion $T(t_0)$.

   Calculate the length of the helix.

2. Consider the spherical surface

$$\Phi : [0, 2\pi] \times [0, \pi] \to \mathbb{R}^3,\ \Phi(u, v) := \begin{bmatrix} 3 \sin v \cos u \\ 3 \sin v \sin u \\ 3 \cos v \end{bmatrix}.$$

Give the u- and v-parameter curves $p : [0, 2\pi] \to \mathbb{R}^3,\ p(u) := \Phi(u, \frac{\pi}{2})$ and $q : [0, \pi] \to \mathbb{R}^3,\ q(v) := \Phi(0, v)$ for the parameter values $(u_0, v_0) := (0, \frac{\pi}{2})$. (What would they mean on the surface of the Earth?)

Give the equation of the tangent plane corresponding to the parameter values $(u_0, v_0) := (\frac{\pi}{3}, \frac{\pi}{3})$. (What angle does it enclose with the plane of the Equator?)

3. Calculate the area of the previous surface $\Phi(u, v)$ over the parameter domain
$$\Omega := \left\{(u, v) \mid 0 \le u \le \frac{\pi}{4},\ \frac{\pi}{3} \le v \le \frac{\pi}{2}\right\}.$$

4. Show that the area of the smooth function $f : [a, b] \times [c, d] \to \mathbb{R}$ as a surface is given by the integral
$$\int_c^d \left(\int_a^b \sqrt{1 + [\partial_x f(x, y)]^2 + [\partial_y f(x, y)]^2}\mathrm{d}x\right)\mathrm{d}y.$$

Solution: Define the surface as the two-parameter function
$$\Phi : [a, b] \times [c, d] \to \mathbb{R}^3,\ \Phi(x, y) := \begin{bmatrix} x \\ y \\ f(x, y) \end{bmatrix}.$$
$$\partial_x \Phi(x, y) = \begin{bmatrix} 1 \\ 0 \\ \partial_x f(x, y) \end{bmatrix}, \quad \partial_y \Phi(x, y) = \begin{bmatrix} 0 \\ 1 \\ \partial_y f(x, y) \end{bmatrix},$$
$$\|\partial_x \Phi(x, y)\|^2 = 1 + [\partial_x f(x, y)]^2, \quad \|\partial_y \Phi(x, y)\|^2 = 1 + [\partial_y f(x, y)]^2,$$
$$(\langle \partial_x \Phi(x, y), \partial_y \Phi(x, y)\rangle)^2 = (\partial_x f(x, y) \cdot \partial_y f(x, y))^2,$$
$$\|\partial_x \Phi\|^2 \cdot \|\partial_y \Phi\|^2 - \langle \partial_x \Phi, \partial_y \Phi\rangle^2 = 1 + (\partial_x f)^2 + (\partial_y f)^2.$$
The latter implies the statement.

5. Let $\Phi : [0, 1] \times [0, 1] \to \mathbb{R}^3,\ \Phi(u, v) := \begin{bmatrix} u + v \\ u - v \\ u \end{bmatrix}$, and $U : \mathbb{R}^3 \to \mathbb{R}$, $U(x, y, z) := x + y + z$. Calculate the surface integral $\int_\Phi U$.

6. Let
$$\Phi : [0, 1] \times [0, 1] \to \mathbb{R}^3, \quad \Phi(u, v) := \begin{bmatrix} u + v \\ u - v \\ u \end{bmatrix},$$

and

$$F : \mathbb{R}^3 \to \mathbb{R}^3, \quad F(x,y,z) := \begin{bmatrix} y \\ x \\ z \end{bmatrix}.$$

Calculate the surface integral $\int_\Phi F$.

7. Consider the "upper hemisphere"

$$\Phi : [0, 2\pi] \times [0, \pi/2] \to \mathbb{R}^3, \ \Phi(u,v) := \begin{bmatrix} \cos v \cos u \\ \cos v \sin u \\ \sin v \end{bmatrix}$$

and its boundary curve

$$r : [0, 2\pi] \to \mathbb{R}^3, \ r(t) := \begin{bmatrix} \cos t \\ \sin t \\ 0 \end{bmatrix}.$$

Let $F : \mathbb{R}^3 \to \mathbb{R}^3, \ F(x,y,z) := \begin{bmatrix} x^2 y \\ yz \\ z \end{bmatrix}$ be a vector function. Check the validity of Stokes' theorem.

Solution:

$$\mathrm{rot} F(x,y,z) = \nabla \times F(x,y,z) = \begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ \partial_x & \partial_y & \partial_z \\ x^2 & yz & z \end{vmatrix}$$

$$= \underline{i}(0-y) - \underline{j}(0-0) + \underline{k}(0-x^2),$$

so $\mathrm{rot} F(x,y,z) = \begin{bmatrix} -y \\ 0 \\ -x^2 \end{bmatrix}, \ \mathrm{rot} F(\Phi(u,v)) = \begin{bmatrix} -\cos v \sin u \\ 0 \\ -(\cos v \cos u)^2 \end{bmatrix}.$

$$\partial_u \Phi(u,v) = \begin{bmatrix} -\cos v \sin u \\ \cos v \cos u \\ 0 \end{bmatrix}, \ \partial_v \Phi(u,v) = \begin{bmatrix} -\sin v \cos u \\ -\sin v \sin u \\ \cos v \end{bmatrix},$$

$$\partial_u \Phi(u,v) \times \partial_v \Phi(u,v) = \begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ -\cos v \sin u & \cos v \cos u & 0 \\ -\sin v \cos u & -\sin v \sin u & \cos v \end{vmatrix}$$

$$= \underline{i}(\cos^2 v \cos u) - \underline{j}(-\cos^2 v \sin u) + \underline{k}(\cos v \sin v \sin^2 u + \cos v \sin v \cos^2 u)$$

$$= \begin{bmatrix} \cos^2 v \cos u \\ \cos^2 v \sin u \\ \cos v \sin v \end{bmatrix},$$

and these vectors are oriented “outwards”.

The left-hand side of Stokes’ theorem:

$$\begin{aligned}
\int_{\Phi} \operatorname{rot} F &= \int_{[0,2\pi]\times[0,\pi/2]} \langle \operatorname{rot} F(\Phi(u,v)), \partial_u \Phi(u,v) \times \partial_v \Phi(u,v) \rangle \mathrm{d}u\mathrm{d}v \\
&= \int_0^{2\pi} \left(\int_0^{\pi/2} \left\langle \begin{bmatrix} -\cos v \sin u \\ 0 \\ -\cos^2 v \cos^2 u \end{bmatrix}, \begin{bmatrix} \cos^2 v \cos u \\ \cos^2 v \sin u \\ \cos v \sin v \end{bmatrix} \right\rangle \mathrm{d}v \right) \mathrm{d}u \\
&= \int_0^{2\pi} \left(\int_0^{\pi/2} (-\cos^3 v \sin u \cos u - \cos^3 v \sin v \cos^2 u) \mathrm{d}v \right) \mathrm{d}u.
\end{aligned}$$

Since

$$\begin{aligned}
\int -\cos^3 v \mathrm{d}v &= -\int \cos v (1 - \sin^2 v) \mathrm{d}v \\
&= -\int \cos v \mathrm{d}v + \int \sin^2 v \cos v \mathrm{d}v = -\sin v + \frac{\sin^3 v}{3},
\end{aligned}$$

therefore

$$\int_0^{\pi/2} -\cos^3 v \mathrm{d}v = \left[-\sin v + \frac{\sin^3 v}{3} \right]_0^{\pi/2} = -\frac{2}{3}.$$

On the other hand,

$$\int_0^{\pi/2} \cos^3 v (-\sin v) \mathrm{d}v = \left[\frac{\cos^4 v}{4} \right]_0^{\pi/2} = -\frac{1}{4}.$$

By exploiting all this,

$$\begin{aligned}
\int_{\Phi} \operatorname{rot} F &= \int_0^{2\pi} \sin u \cos u \cdot \left(-\frac{2}{3} \right) + \cos^2 u \cdot \left(-\frac{1}{4} \right) \mathrm{d}u \\
&= -\frac{2}{3} \left[\frac{\sin^2 u}{2} \right]_0^{2\pi} - \frac{1}{4} \int_0^{2\pi} \frac{1 + \cos 2u}{2} \mathrm{d}u \\
&= 0 - \frac{1}{4} \left[\frac{1}{2} u + \frac{\sin 2u}{4} \right]_0^{2\pi} = -\frac{\pi}{4}.
\end{aligned}$$

The right-hand side of Stokes' theorem is a line integral:

$$\int_r F = \int_0^{2\pi} \langle F(r(t)), \dot{r}(t)\rangle \mathrm{d}t = \int_0^{2\pi} \left\langle \begin{bmatrix} \cos^2 t \sin t \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -\sin t \\ \cos t \\ 0 \end{bmatrix} \right\rangle \mathrm{d}t$$
$$= \int_0^{2\pi} -\cos^2 t \sin^2 t \mathrm{d}t = \int_0^{2\pi} -\frac{\sin^2 2t}{4}\mathrm{d}t = \int_0^{2\pi} -\frac{1-\cos 4t}{8}\mathrm{d}t$$
$$= \left[-\frac{1}{8}t + \frac{\sin 4t}{32}\right]_0^{2\pi} = -\frac{\pi}{4}.$$

So, in this example
$$\int_\Phi \mathrm{rot} F = \int_r F = -\frac{\pi}{4}.$$

8. Consider the vector function $F : \mathbb{R}^3 \to \mathbb{R}^3,\ F(x,y,z) := \begin{bmatrix} x^2 y \\ yz \\ z \end{bmatrix}$.

Let
$$V := \left\{(x,y,z) \in \mathbb{R}^3 \ \mid x^2 + y^2 + z^2 \le R^2\right\}$$
be a sphere of radius R, centered at the origin, which is bounded by the surface
$$\Phi : [0,2\pi] \times [-\pi/2, \pi/2] \to \mathbb{R}^3,\ \Phi(u,v) := \begin{bmatrix} R\cos v \cos u \\ R \cos v \sin u \\ R \sin v \end{bmatrix}.$$

Check the validity of Gauss' theorem.

Solution:
$$\mathrm{div} F(x,y,z) = \langle \nabla, F\rangle(x,y,z) = \partial_x(x^2y) + \partial_y(yz) + \partial_z(z) = 2xy + z + 1.$$

The left-hand side of Gauss' theorem:
$$\int_V \mathrm{div} F = \int_V (2xy + z + 1)\mathrm{d}x\mathrm{d}y\mathrm{d}z.$$

For the calculation of the integral it is convenient to perform polar transformation. Let
$$\Psi : [0,2\pi] \times [\pi/2, \pi/2] \times [0, R] \to \mathbb{R}^3,\ \Psi(u,v,r) := \begin{bmatrix} r\cos v \cos u \\ r \cos v \sin u \\ r \sin v \end{bmatrix}.$$

Ψ is a one-to-one correspondence between $T := [0, 2\pi] \times [\pi/2, \pi/2] \times [0, R]$ and the sphere V. Calculate the determinant of the derivative of the substituting function:

$$\begin{aligned}
\det \Psi'(u, v) &= \begin{vmatrix} -r\cos v \sin u & -r \sin v \cos u & \cos v \cos u \\ r \cos v \cos u & -r \sin v \sin u & \cos v \sin u \\ 0 & r \cos v & \sin v \end{vmatrix} \\
&= -r\cos v(-r\cos^2 v \sin^2 u - r\cos^2 v \cos^2 u) \\
&\quad + \sin v(r^2 \cos v \sin v \sin^2 u + r^2 \cos v \sin v \cos^2 u) \\
&= r^2 \cos v.
\end{aligned}$$

Since $v \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, therefore $|\det \Psi'(u, v)| = r^2 \cos v$. By exploiting this,

$$\begin{aligned}
&\int_V (2xy + z + 1)\mathrm{d}x\mathrm{d}y\mathrm{d}z \\
&= \int_0^{2\pi}\left(\int_{-\pi/2}^{\pi/2}\left(\int_0^R (2(r\cos v\cos u)(r\cos v\sin u) + r\sin v + 1)r^2\cos v\mathrm{d}r\right)\mathrm{d}v\right)\mathrm{d}u \\
&= \int_0^{2\pi}\left(\int_{-\pi/2}^{\pi/2}\left(\int_0^R (2r^4\cos^3 v\cos u\sin u + r^3\sin v\cos v + r^2\cos v)\mathrm{d}r\right)\mathrm{d}v\right)\mathrm{d}u \\
&= \int_0^{2\pi}\left(\int_{-\pi/2}^{\pi/2}\left(\frac{2}{5}R^5\cos^3 v\cos u\sin u + \frac{1}{4}R^4\sin v\cos v + \frac{1}{3}R^3\cos v\right)\mathrm{d}v\right)\mathrm{d}u.
\end{aligned}$$

By Exercise 7,

$$\int_{-\pi/2}^{\pi/2} \cos^3 v\mathrm{d}v = \left[\sin v - \frac{\sin^3 v}{3}\right]_{-\pi/2}^{\pi/2} = \frac{4}{3}.$$

$$\int_{-\pi/2}^{\pi/2} \sin v\cos v\mathrm{d}v = \left[\frac{\sin^2 v}{2}\right]_{-\pi/2}^{\pi/2} = 0 \quad \text{és} \quad \int_{-\pi/2}^{\pi/2} \cos v\mathrm{d}v = 2.$$

By continuing the integration:

$$\begin{aligned}
\int_0^{2\pi}\left[\frac{4}{3}\cdot\frac{2}{5}R^5\cos u\sin u + \frac{1}{4}R^4 + \frac{2}{3}R^3\right]\mathrm{d}u &= \frac{8}{15}R^5\left[\frac{\sin^2 u}{2}\right]_0^{2\pi} + \frac{4\pi}{3}R^3 \\
&= \frac{4\pi}{3}R^3.
\end{aligned}$$

The right-hand side of Gauss' theorem is a surface integral:

$$\int_\Phi F = \int_{[0,2\pi]\times[-\frac{\pi}{2},\frac{\pi}{2}]} \langle F(\Phi(u, v)), \partial_u\Phi(u, v) \times \partial_v\Phi(u, v)\rangle\mathrm{d}u\mathrm{d}v.$$

$$\partial_u \Phi(u,v) = \begin{bmatrix} -R\cos v \sin u \\ R\cos v\cos u \\ 0 \end{bmatrix}, \quad \partial_v \Phi(u,v) = \begin{bmatrix} -R\sin v\cos u \\ -R\sin v\sin u \\ R\cos v \end{bmatrix}$$

$$\begin{aligned} \partial_u \Phi(u,v) \times \partial_v \Phi(u,v) &= \begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ -R\cos v\sin u & R\cos v\cos u & 0 \\ -R\sin v\cos u & -R\sin v\sin u & R\cos v \end{vmatrix} \\ &= \underline{i}(R^2\cos^2 v\cos u) - \underline{j}(-R^2\cos^2 v\sin u) \\ &\quad + \underline{k}(R^2\cos v\sin v\sin^2 u + R^2\cos v\sin v\cos^2 u) \\ &= \begin{bmatrix} R^2\cos^2 v\cos u \\ R^2\cos^2 v\sin u \\ R^2\cos v\sin v \end{bmatrix}. \end{aligned}$$

$$F(\Phi(u,v)) = \begin{bmatrix} (R\cos v\cos u)^2(R\cos v\sin u) \\ (R\cos v\sin u)(R\sin v) \\ R\sin v \end{bmatrix}.$$

By using these functions,

$$\begin{aligned} &\int_\Phi F = \int_{[0,2\pi]\times[-\frac{\pi}{2},\frac{\pi}{2}]} \left\langle \begin{bmatrix} R^3\cos^3 v\cos^2 u\sin u \\ R^2\cos v\sin v\sin u \\ R\sin v \end{bmatrix}, \begin{bmatrix} R^2\cos^2 v\cos u \\ R^2\cos^2 v\sin u \\ R^2\cos v\sin v \end{bmatrix} \right\rangle \mathrm{d}u\mathrm{d}v \\ &= \int_0^{2\pi}\left(\int_{-\pi/2}^{\pi/2} (R^5\cos^5 v\cos^3 u\sin u + R^4\cos^3 v\sin v\sin^2 u + R^3\sin^2 v\cos v)\mathrm{d}v\right)\mathrm{d}u. \end{aligned}$$

We calculate the integral of each term:

$$\int \cos^5 v\mathrm{d}v = \int (1-\sin^2 v)\cos^3 v\mathrm{d}v = \int \cos^3 v\mathrm{d}v - \int \cos^3 v\sin^2 v\mathrm{d}v.$$

From Exercise 7 we have $\int \cos^3 v\mathrm{d}v = \sin v - \frac{\sin^3 v}{3}$.

$$\begin{aligned} \int \cos^3 v\sin^2 v\mathrm{d}v &= \int (\cos^3 v\sin v)\sin v\mathrm{d}v \\ &= -\frac{\cos^4 v}{4}\sin v - \int -\frac{\cos^4 v}{4}\cos v\mathrm{d}v \\ &= -\frac{1}{4}\cos^4 v\sin v + \frac{1}{4}\int \cos^5 v\mathrm{d}v. \end{aligned}$$

After rearrangement of the results we obtain

$$
\begin{aligned}
\int \cos^5 v \mathrm{d}v &= \sin v - \frac{\sin^3 v}{3} + \frac{1}{4}\cos^4 v \sin v - \frac{1}{4}\int \cos^5 v \mathrm{d}v, \\
\frac{5}{4}\int \cos^5 v \mathrm{d}v &= \sin v - \frac{\sin^3 v}{3} + \frac{1}{4}\cos^4 v \sin v, \\
\int_{-\pi/2}^{\pi/2} \cos^5 v \mathrm{d}v &= \frac{4}{5}\left[\sin v - \frac{\sin^3 v}{3} + \frac{1}{4}\cos^4 v \sin v\right]_{-\pi/2}^{\pi/2} = \frac{4}{5}\cdot\frac{4}{3} = \frac{16}{15}.
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
&\int_{-\pi/2}^{\pi/2} (R^5 \cos^5 v \cos^3 u \sin u + R^4 \cos^3 v \sin v \sin^2 u + R^3 \sin^2 v \cos v)\mathrm{d}v \\
&\quad = \frac{16}{15}R^5 \cos^3 u \sin u + R^4 \sin^2 u \left[-\frac{\cos^4 v}{4}\right]_{-\pi/2}^{\pi/2} + R^3\left[\frac{\sin^3 v}{3}\right]_{-\pi/2}^{\pi/2} \\
&\quad = \frac{16}{15}R^5 \cos^3 u \sin u + \frac{2}{3}R^3.
\end{aligned}
$$

Finally,

$$
\begin{aligned}
&\int_0^{2\pi} \left(\frac{16}{15}R^5 \cos^3 u \sin u + \frac{2}{3}R^3\right)\mathrm{d}u \\
&= \left[\frac{16}{15}R^5\left(-\frac{\cos^4 u}{4}\right) + \frac{2}{3}R^3 u\right]_0^{2\pi} = \frac{4\pi}{3}R^3.
\end{aligned}
$$

So,

$$
\int_V \mathrm{div} F = \int_\Phi F = \frac{4\pi}{3}R^3.
$$