TECHNICKÁ UNIVERZITA V KOŠICIACH - VÝZVA pre tvoju BUDÚCNOSŤ

Moderné vzdelávanie pre vedomostnú spoločnosť/
Projekt je spolufinancovaný zo zdrojov EÚ

# MATEMATIKA 1

# MATHEMATICS 1

*Fakulta elektrotechniky a informatiky*

*Štefan Berežný*

Európska únia
Európsky sociálny fond

Agentúra
Ministerstva školstva, vedy, výskumu a športu SR
pre štrukturálne fondy EÚ

Operačný program
VZDELÁVANIE

Táto publikácia vznikla za finančnej podpory z **Európskeho sociálneho fondu** v rámci Operačného programu **VZDELÁVANIE.**

Prioritná os 1 Reforma vzdelávania a odbornej prípravy

Opatrenie 1.2 Vysoké školy a výskum a vývoj ako motory rozvoja vedomostnej spoločnosti.

Názov projektu: Balík prvkov pre skvalitnenie a inováciu vzdelávania na TUKE

NÁZOV: Mathematics 1 (Matematika 1)
AUTOR: RNDr. Štefan Berežný, PhD.
RECENZENTI: Mgr. Ján Buša, Ph.D.,
RNDr. Kristína Budajová, PhD.,
prof. RNDr. Jozef Džurina, PhD.
VYDAVATEĽ: Technická univerzita v Košiciach
ROK: 2014
ROZSAH: 170 strán
NÁKLAD: 100 ks
VYDANIE: prvé
ISBN: 978-80-553-1788-5

Rukopis neprešiel jazykovou úpravou.
Za odbornú a obsahovú stránku zodpovedajú autori.

TECHNICKÁ UNIVERZITA V KOŠICIACH - VÝZVA pre tvoju BUDÚCNOSŤ

Moderné vzdelávanie pre vedomostnú spoločnosť/
Projekt je spolufinancovaný zo zdrojov EÚ

# MATEMATIKA 1

# MATHEMATICS 1

*Fakulta elektrotechniky a informatiky*

*Štefan Berežný*

Európska únia
Európsky sociálny fond

Agentúra
Ministerstva školstva, vedy, výskumu a športu SR
pre štrukturálne fondy EÚ

Operačný program
VZDELÁVANIE

Technical University of Košice
Faculty of Electrical Engineering and Informatics

# MATHEMATICS 1

University Textbook

Košice 2014

Technical University of Košice
Faculty of Electrical Engineering and Informatics
Department of Mathematics and Theoretical Informatics

# Mathematics 1

University Textbook

Štefan Berežný

Košice 2014

# Mathematics 1

First Edition

**Author:** © **RNDr. Štefan BEREŽNÝ, PhD., 2014**

Reviewed by: Mgr. Ján BUŠA, Ph.D.
RNDr. Kristína BUDAJOVÁ, PhD.
prof. RNDr. Jozef DŽURINA, PhD.

Editor: Technical University of Košice
Faculty of Electrical Engineering and Informatics

ISBN: 978-80-553-1788-5

Za odbornú a jazykovú stránku tejto vysokoškolskej učebnice zodpovedá autor.
Rukopis neprešiel redakčnou ani jazykovou úpravou.

# Preface

The textbook Mathematics 1 contains an overview of the theory, solved examples and unsolved tasks for subject Mathematics 1 for bachelor's degrees students at Applied Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice. The textbook consists of five chapters. Each chapter is divided into sub-chapters in particular areas of Mathematics. At the end of each chapter are subsections Solved Examples, Unsolved Tasks and results of them.

In addition to numerical mathematics are in this textbook the theory and examples of basic information from mathematical analysis and linear algebra, as this required course of study Applied Informatics.

This textbook is available on CD and on the web site DMTI FEEI TUKE (KMTI FEI TU) and Moodle system, which is managed by the FEEI TUKE.

Košice, $31^{st}$ of August 2014 Author

# Contents

**Preface** **5**

**Contents** **9**

**List of abbreviations and symbols** **11**

**List of Figures** **13**

**List of Tables** **15**

**1 Real Function** **17**

1.1 Real Number . . . . . . . . . . . . . . . . . . . . . . . . . . . 17
1.2 The Real Function of One Real Variable . . . . . . . . . . . . 24
1.3 Elementary Functions . . . . . . . . . . . . . . . . . . . . . . 28
1.3.1 Constant Function . . . . . . . . . . . . . . . . . . . . 28
1.3.2 Linear Function . . . . . . . . . . . . . . . . . . . . . . 28
1.3.3 Quadratic Function . . . . . . . . . . . . . . . . . . . 29
1.3.4 Power Function . . . . . . . . . . . . . . . . . . . . . . 32
1.3.5 Exponential Function . . . . . . . . . . . . . . . . . . 33
1.3.6 Logarithmic Function . . . . . . . . . . . . . . . . . . 34
1.3.7 Trigonometric Functions . . . . . . . . . . . . . . . . . 36
1.3.8 Cyclometric Functions . . . . . . . . . . . . . . . . . . 38
1.3.9 Polynomial Function . . . . . . . . . . . . . . . . . . . 41
1.4 Limit of Function . . . . . . . . . . . . . . . . . . . . . . . . . 42
1.5 The Derivative . . . . . . . . . . . . . . . . . . . . . . . . . . 49
1.6 Solved Examples . . . . . . . . . . . . . . . . . . . . . . . . . 59
1.7 Unsolved Tasks . . . . . . . . . . . . . . . . . . . . . . . . . . 63
1.8 Results of Unsolved Tasks . . . . . . . . . . . . . . . . . . . . 66

**2 Approximate Solution of a Nonlinear Equation 67**
2.1 Separation of Roots . . . 67
2.2 Bisection Method . . . 68
2.3 Fixed Point Iteration Method . . . 71
2.4 Newton's Method . . . 74
2.5 Solved Examples . . . 77
2.6 Unsolved Tasks . . . 82
2.7 Results of Unsolved Tasks . . . 83

**3 Approximation of Functions 85**
3.1 Interpolation . . . 85
3.2 The Least Squares Method . . . 87
3.3 Solved Examples . . . 90
3.4 Unsolved Tasks . . . 94
3.5 Results of Unsolved Tasks . . . 95

**4 The Calculation of Definite Integrals 97**
4.1 Antiderivative . . . 97
4.2 Indefinite Integral . . . 98
4.2.1 Integration Rules . . . 98
4.2.2 Integration Formulas . . . 98
4.2.3 Integration by Parts . . . 99
4.2.4 Integration by the Substitution Method . . . 100
4.3 The Definite Integral . . . 100
4.3.1 Properties of Definite Integrals . . . 102
4.3.2 Aplication of Definite Integrals . . . 103
4.4 Numerical Integration . . . 105
4.4.1 Newton-Cotes Formulas . . . 105
4.4.2 Rectangular Method . . . 107
4.4.3 Trapezoidal Method . . . 108
4.4.4 Simpson's Method . . . 109
4.5 Solved Examples . . . 111
4.6 Unsolved Tasks . . . 115
4.7 Results of Unsolved Tasks . . . 118

**5 Linear Algebra 119**
5.1 Vector Space . . . 119
5.2 Matrix . . . 123

5.3 Determinant . . . 129
5.4 Systems of Linear Equations . . . 132
5.5 Numerical Methods for Solving Systems of Linear Equations . 138
5.5.1 Jacobi's Method . . . 138
5.5.2 Gauss-Seidel's Method . . . 143
5.6 Solved Examples . . . 144
5.7 Unsolved Tasks . . . 146
5.8 Results of Unsolved Tasks . . . 151

**Register** **153**

**Lexicon** **159**

**Bibliography** **169**

# List of Abbreviations and Symbols

SLE – system of linear algebraic equations

HSLE – homogeneous system of linear algebraic equations

$\mathbb{N}$ – set of positive integers (natural numbers)

$\mathbb{Z}$ – set of integers

$\mathbb{Q}$ – set of rational numbers

$\mathbb{R}$ – set of real numbers

$\mathbb{C}$ – set of complex numbers

$\mathcal{D}(f)$ – domain of definition of a function $f$

$\mathcal{R}(f)$ – range of a function $f$

$\det(A)$ – determinant of a matrix $A$

$\operatorname{rank}(A)$ – rank of a matrix $A$

$A^\top$ – matrix transposed to a matrix $A$

$\vec{v}$ – vector $v$

$\Omega(S)$ – set of solutions of SLE $S$

# List of Figures


1.1 The Real Number Line . . . . . . . . . . . . . . . . . . . . . . . 20
1.2 The Real Number Line: Relation “$<$” and “$>$” . . . . . . . . . . 20
1.3 Intervals . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 24
1.4 Constant Function . . . . . . . . . . . . . . . . . . . . . . . . . 29
1.5 Linear function . . . . . . . . . . . . . . . . . . . . . . . . . . 30
1.6 Quadratic Function . . . . . . . . . . . . . . . . . . . . . . . . 31
1.7 Square Root Function . . . . . . . . . . . . . . . . . . . . . . . 34
1.8 Exponential Function . . . . . . . . . . . . . . . . . . . . . . . 35
1.9 Logarithmic Function . . . . . . . . . . . . . . . . . . . . . . . 36
1.10 Trigonometric Function: $\sin x$ . . . . . . . . . . . . . . . . . . 38
1.11 Trigonometric Function: $\cos x$ . . . . . . . . . . . . . . . . . . 38
1.12 Trigonometric Function: $\operatorname{tg} x$ . . . . . . . . . . . . . . . . . . . 39
1.13 Trigonometric Function: $\operatorname{cotg} x$ . . . . . . . . . . . . . . . . . 40
1.14 Domain of Function . . . . . . . . . . . . . . . . . . . . . . . . 59
1.15 Domain of Function . . . . . . . . . . . . . . . . . . . . . . . . 60

2.1 Bisection Method . . . . . . . . . . . . . . . . . . . . . . . . . 70
2.2 Newton’s Method . . . . . . . . . . . . . . . . . . . . . . . . . 75
2.3 Bisection Method - Separating the Roots . . . . . . . . . . . . 77
2.4 Fixed Point Iteration Method - Separating the Roots . . . . . 80

# List of Tables


2.1 Table for the bisection method. . . . . . . . . . . . . . . . . . . 71
2.2 Solving equation by bisection method. . . . . . . . . . . . . . . 78
2.3 Newton's method for solving of equation. . . . . . . . . . . . . 79
2.4 Iterative methods for solving of the equation. . . . . . . . . . . 81

3.1 The Least Squares Method for $y(x) \sim \frac{a}{x} + b \cdot \ln(x)$. . . . . . . 93

4.1 Trapezoidal method. . . . . . . . . . . . . . . . . . . . . . . . . 114

5.1 Solving of SLE by the Jacobi's iteration method. . . . . . . . . 145

# Chapter 1

# Real Function

## 1.1 Real Number

**Natural** (denoted by $\mathbb{N}$) *Natural numbers* are the counting numbers $\mathbb{N} = \{1, 2, 3, 4, \dots\}$. Natural numbers are called also *positive integers*. The whole numbers $\{0, 1, 2, 3, 4, \dots\}$ are denoted by $\mathbb{N}_0$ and are also called *non-negative integers*. These numbers represent the cardinality of sets.[1] For natural numbers applies the principle of complete mathematical induction. If $N$ is any set of positive integers that contains the number 1, and that with each natural number $n$ contains the number $n + 1$, then set $N$ contains all natural numbers, i. e. $N = \mathbb{N}$.

**Integer** (denoted by $\mathbb{Z}$) *Integers* are the natural numbers, zero and their negatives $\mathbb{Z} = \{\dots - 4, -3, -2, -1, 0, 1, 2, 3, 4, \dots\}$.

**Rational** (denoted by $\mathbb{Q}$) *Rational number* is a number that can be expressed as a fraction of an integer numerator $m$ and a non-zero integer number denominator $n$. Fractions are written as two numbers, the numerator and the denominator, with a dividing bar between them. In the fraction written

$$\frac{m}{n} = \frac{\text{numerator}}{\text{denominator}},$$

$m$ represents equal parts, where $n$ equal parts of that size make up $m$ wholes. Two different fractions may represent the same rational

[1]Mathematicians use the term "natural" in both cases.

number. For example:

$$\frac{2}{3} = \frac{4}{6} = -\frac{-6}{9}.$$

If the absolute value of $m$ is greater than the absolute value of $n$, then the absolute value of the fraction is greater than 1. Fractions can be greater than, less than, or equal to 1 and can also be positive, negative, or 0. The set of all rational numbers includes the integers, since every integers can be written as a fraction with denominator 1. Formally

$$\mathbb{Q} = \left\{ q : q = \frac{m}{n}; m, n \in \mathbb{Z} \text{ and } n \neq 0 \right\}.$$

A rational number is a real number that is the quotient of integers. Each rational number can be written in lowest terms, that is, as a quotient of integers with no common factor larger than 1.

**Irrational** (denoted by $\mathbb{I}$ or $\mathbb{I}_r$) *Irrational numbers* are real numbers that cannot be written as a simple fraction. Irrational means not rational. These are numbers that can be written as decimals, but not as fractions. They are non-repeating, non-terminating decimals. Famous irrational numbers are $\pi$, e, $\sqrt{2}$.

**Real** (denoted by $\mathbb{R}$) *Real numbers* is the set which is denoted by the letter $\mathbb{R}$. Every number (except complex numbers) is contained in the set of real numbers. When the general term "number" is used, it refers to a real number. Real numbers are all the numbers on the continuous number line with no gaps. Every decimal expansion is a real number. Real numbers may be rational or irrational, and algebraic or non-algebraic (transcendental). Real numbers $\pi = 3.14159\ldots$ and $\mathrm{e} = 2.71828\ldots$ are transcendental. A transcendental number can be defined by an infinite series. Real numbers can also be positive, negative or zero. The square root of minus 1 ($\sqrt{-1} = \mathrm{i}$) is not a real number, it is an *imaginary number*. Infinity ($\infty$) is not a real number.

**Complex** (denoted by $\mathbb{C}$)[2] Definition of *complex numbers* you can see in [1, 4, 11].

Properties of real numbers:

[2]Each of the number set mentioned is a proper subset of the next number set. Symbolically, $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

(a) Commutative laws: $\forall a, b \in \mathbb{R} : a + b = b + a$,
$\forall a, b \in \mathbb{R} : a \cdot b = b \cdot a$,
(b) Associative laws: $\forall a, b, c \in \mathbb{R} : (a + b) + c = a + (b + c)$,
$\forall a, b, c \in \mathbb{R} : (a \cdot b) \cdot c = a \cdot (b \cdot c)$,
(c) Distributive laws: $\forall a, b, c \in \mathbb{R} : a \cdot (b + c) = a \cdot b + a \cdot c$,
$\forall a, b, c \in \mathbb{R} : (a + b) \cdot c = a \cdot c + b \cdot c$,
(d) $\forall a \in \mathbb{R}$: $a + 0 = a$,
(e) $\forall a \in \mathbb{R}$: $a \cdot 1 = a$,
(f) For each real number $a$ there exists a real number $-a$, such that $a + (-a) = 0$,
(g) For each real number $a \neq 0$ there exists a real number $a^{-1} \in \mathbb{R}$, such that $a \cdot a^{-1} = 1$.

**The real number line** is like an actual geometric line. A point is chosen on the line to be the "origin" (zero = 0), points to the right will be positive, and points to the left will be negative. A distance is chosen to be "1". We choose a starting point, a unit length and a positive direction (the direction from 0 toward 1). Then we mark the points $2, 3, 4, \ldots$ to the right and the points $-1, -2, -3, -4, \ldots$ to the left of the starting point. To complete the labelling, we must make a fundamental assumption. We take it as an axiom that there is a one-to-one correspondence between the points on the line and the system $\mathbb{R}$ of real numbers. Any point on the line is a real number.

More precisely:

- Each real number $a$ corresponds to a point $P_a$ on the line.
- Each point $P$ on the line is correspondent of a single real number $a$.
- If $P_a$ is left of $P_b$, then $b - a$ is the distance between the points $P_a$ and $P_b$.

Because of the close association of the real number system $\mathbb{R}$ with the set of points on a line, it is common to refer to a line as the real number system and to the real number system as a line: *the real line* or *the number line* (see Figure 1.1).

**Remark 1.1** In general, the rules for order of operations are as follows:

- Parentheses – perform operations that are in parentheses first, working from the inside out.

- Exponents – evaluate all powers.
- Multiplication and Division – multiply and divide from left to right.
- Addition and Subtraction – add and subtract from left to right.

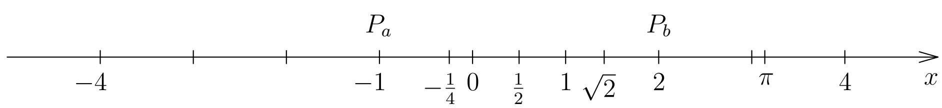


Figure 1.1: The real number line.

**Definition 1.1** If a real number $a$ is positive, then we write $a > 0$. If a real number $a$ lies to the left of real number $b$ (or real number $b - a > 0$ is positive), then we say that "$a$ is *less than* $b$" and we write:

$$a < b.$$

Alternatively, if a real number $a$ lies to the right of real number $b$ (or real number $a - b > 0$ is positive), then we say that "$a$ is *greater than* $b$" and we write:

$$a > b.$$

The notation $a \leqq b$, which is read "$a$ is *less than or equal* to $b$", means that either $a < b$ or $a = b$. The same condition can be written $b \geqq a$, then it is read "$b$ is *greater than or equal* to $a$". The relations *less than* and *greater than* and so on impose what is called *order* on the real number system (see Figure 1.2).

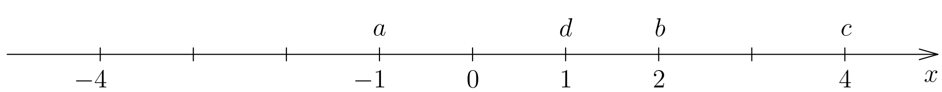


Figure 1.2: The relations: $a < b$ and $c > d$.

**Theorem 1.1 (Properties of Order)**

(a) Reflexivity: $\forall a \in \mathbb{R}$: $a \leqq a$,
(b) Anti-symetry: $\forall a, b \in \mathbb{R}$: if $a \leqq b$ and $b \leqq a$, then $a = b$,
(c) Transitivity: $\forall a, b, c \in \mathbb{R}$: if $a \leqq b$ and $b \leqq c$, then $a \leqq c$,
(d) Trichotomy: If $a$ and $b$ are real numbers, then exactly one of these three relations holds:
(1) $a < b$
(2) $a > b$
(3) $a = b$.
(e) The arithmetic operations $+$, $-$, $\cdot$ and $\div$ are closely linked to the order relations
(1) $\forall a, b \in \mathbb{R}$: $a + b \in \mathbb{R}$
(2) $\forall a, b \in \mathbb{R}$: $a - b \in \mathbb{R}$
(3) $\forall a, b \in \mathbb{R}$: $a \cdot b \in \mathbb{R}$
(4) $\forall a, b \in \mathbb{R}$: $a \div b \in \mathbb{R}$.
(f) If $a < b$ and if $c$ is any real number, then $a + c < b + c$ and $a - c < b - c$.
(g) $\forall a, b, c \in \mathbb{R}$: $a < b \wedge c > 0 \Rightarrow ac < bc$.
(h) $\forall a, b, c \in \mathbb{R}$: $a < b \wedge c < 0 \Rightarrow ac > bc$.
(i) $\forall a, b, c, d \in \mathbb{R}$: $a < b \wedge c \leq d \Rightarrow a + c < b + d$.
(j) If $0 < a < b$ or $a < b < 0$, then $1/b < 1/a$.
(k) The statement $a < b$ is equivalent to the statement $b - a > 0$.
(l) If $x$ is any real number, then $x^2 \geqq 0$. If $x^2 = 0$, then $x = 0$.

**Theorem 1.2** We can order rational numbers as follows: If $q, r \in \mathbb{Q}$, $q = \dfrac{a}{b}$ and $r = \dfrac{c}{d}$, where $a, c \in \mathbb{Z}$ and $b, d \in \mathbb{N}$, then holds:

(1) $q < r$, if $a \cdot d < b \cdot c$,

(2) $q = r$, if $a \cdot d = b \cdot c$,

(3) $q > r$, if $a \cdot d > b \cdot c$.

**Definition 1.2** *Algebraic numbers* are those that can be expressed as the solution to a polynomial equation with integer coefficients. The complement of the algebraic numbers are the transcendental numbers. Real number $\alpha$ is called algebraic, if $\alpha$ is root of some algebraic equation of the form $a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + a_{n-2} \cdot x^{n-2} + a_{n-3} \cdot x^{n-3} + \cdots + a_1 \cdot x + a_0 = 0$ with rational coefficients $a_1, a_2, a_3, \dots, a_n$. If number $\alpha$ is not algebraic, then it called *transcendental*. Transcendental numbers are for example $\pi$ or e and so on.

**Definition 1.3** The set $K \subseteq \mathbb{R}$ of real numbers is called *bounded from above*, if there exists real number $U$ such that, $U$ is greater then or equal to every numbers of set $K$. Number $U$ is called *upper bound of set $K$*. A set $K$ with an upper bound $U$ is said to be bounded from above by that bound. The set $K \subseteq \mathbb{R}$ of real numbers is called *bounded from below*, if there exists real number $L$ such that, $L$ is less then or equal to every numbers of set $K$. Number $L$ is called *lower bound of set $K$*. A set $K$ with a lower bound $L$ is said to be bounded from below by that bound. The set $K$ is called *bounded (bounded set)*, if there exist upper and lower bound simultaneously. Let $K \subseteq \mathbb{R}$. *Upper bound of set $K$* is called every real number $U \in \mathbb{R}$: $\forall x \in K$ is $x \leqq U$. Let $K \subseteq \mathbb{R}$. *Lower bound of set $K$* is called every real number $L \in \mathbb{R}$: $\forall x \in K$ is $x \geqq L$.

**Definition 1.4** Least upper bound of set $K \subseteq \mathbb{R}$ is called *suprema* of set $K$. It is denoted by $\sup(K)$. Greatest lower bound of set $K \subseteq \mathbb{R}$ is called *infima* of set $K$. We will denote it by $\inf(K)$.

**Theorem 1.3** Every nonempty set of real numbers bounded from above has suprema. Every nonempty set of real numbers bounded from below has infima.

**Definition 1.5** If set $K$ contains an upper bound then that upper bound is unique and is called the *greatest element of set $K$* (*maximum*). It is denoted by $\max(K)$.

**Definition 1.6** If set $K$ contains a lower bound then that lower bound is unique and is called the *least element of set $K$* (*minimum*). It is denoted by $\min(K)$.

**Theorem 1.4** The greatest element of $K$ (if it exists) is also the least upper bound of $K$ (suprema of $K$). Let $K \subseteq \mathbb{R}$. Maximum of set $K$ is real number $M \in K$, for which holds: $\forall x \in K$ is $x \leqq M$. If $M = \max(K)$ exists, then $\sup(K) = \max(K)$.

The least element of $K$ (if it exists) is also the greatest lower bound of $K$ (infima of $K$). Let $K \subseteq \mathbb{R}$. Minimum of set $K$ is real number $m \in K$, for which holds: $\forall x \in K$ is $x \geqq m$. If $m = \min(K)$ exists, then $\inf(K) = \min(K)$.

**Theorem 1.5** Between two different real numbers lie infinitely many rational numbers and infinitely many irrational numbers.

**Remark 1.2** Let set $K \subseteq \mathbb{R}$ is given. For set $K$ holds: $\inf(K)$ and $\sup(K)$ may or may not belong to the set $K$. This follows from the Theorem 1.4.

**Definition 1.7** The *absolute value* of real number $a$, written $|a|$, is the distance on the real number line from $a$ to 0. It comes to this, that:

$$|a| = \begin{cases} a & \text{if } a \geqq 0, \\ -a & \text{if } a < 0. \end{cases}$$

For any real numbers $a$ and $b$, we say that $b$ is *larger in magnitude* than $a$ if $|b| > |a|$.

**Theorem 1.6** Rules for absolute value are:

(1) $|-a| = |a|$,

(2) $|a \cdot b| = |a| \cdot |b|$,

(3) $\left|\frac{a}{b}\right| = \frac{|a|}{|b|}$; where $b \neq 0$,

(4) $|a + b| \leqq |a| + |b|$,

(5) $|a - b| \leqq |a| + |b|$,

(6) $|a| - |b| \leqq |a - b|$.

**Definition 1.8** The set of all numbers between two fixed numbers is called an *interval* on the number line. The interval may include one or both of its end points, or neither. The interval that includes both end points is called a *closed interval*. We use the notation: $\langle a, b\rangle$. The interval that excludes both end points is called an *open interval*. We use the notation: $(a, b)$. We also need notation for hybrid intervals called *half-open* or *half-closed* intervals. Let $a, b \in \mathbb{R}$, $a < b$, and then the suggestive notation for these intervals is:

- $\langle a, b\rangle$, for all points $x \in \langle a, b\rangle$ satisfying: $a \leqq x \leqq b$,
- $(a, b)$, for all points $x \in (a, b)$ satisfying: $a < x < b$,
- $(a, b\rangle$, for all points $x \in (a, b\rangle$ satisfying $a < x \leqq b$,

- $\langle a, b)$, for all points $x \in \langle a, b)$ satisfying $a \leqq x < b$.

The intervals that we have defined so far are *bounded*. We also deal with *unbounded* intervals, these are, intervals that go off indefinitely in one direction or the other. We use one of the following notations: $(-\infty, b)$, $(-\infty, b\rangle$, $\langle a, \infty)$, $(a, \infty)$ and $(-\infty, \infty)$. This is shown on the Figure 1.3.

**Remark 1.3** Using absolute values and inequalities, we can develop nice shorthand to express geometrical facts about distances and intervals.

- The inequality $|x - a| < r$ describes the open interval $(a - r, a + r)$,
- The inequality $|x - a| \leqq r$ describes the closed interval $\langle a - r, a + r\rangle$.

We can think of $|x-a| < r$ as representing the open interval with "center" at $a$ and "radius" equal to positive real number $r$. In calculus, the Greek letter $\varepsilon$ (epsilon) generally denotes a small positive real number, so $|x - a| < \varepsilon$ describes a small interval centered at $a$ (neighbourhood or $\varepsilon$-neighbourhood of point $a$, $\mathcal{O}(a)$ or $\mathcal{O}_\varepsilon(a)$). These terms can be used to express a simple, but important principle:

$$\text{If } |x| < \varepsilon \text{ for every } \varepsilon > 0, \text{then } x = 0.$$

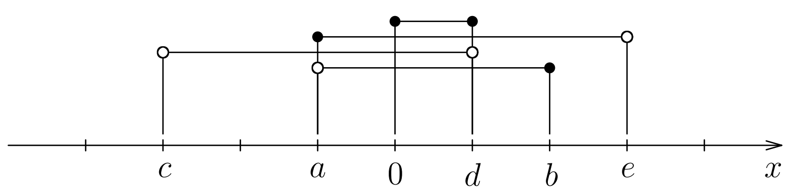


Figure 1.3: Display of the intervals on the number line: $(a, b\rangle$, $(c, d)$,, $\langle a, e)$, $\langle 0, d\rangle$.

## 1.2 The Real Function of One Real Variable

**Definition 1.9** Let the symbol $x$ represents any real number belonging to a certain set $D \subseteq \mathbb{R}$ of real numbers. Suppose there is a rule $f$ that associates with each such $x$ a real number $y$. Then this rule $f$ is called a *function* whose *domain* is the set $D = \mathcal{D}(f)$. We can think of a function as a "black box", that is a machine whose inner workings are hidden from us. We input an $x$ from $D$, the "black box" outputs a $y$. The set of all numbers $y$ that a function assigns to the numbers $x$ in its domain is called the *range* of the function ($R = \mathcal{R}(f)$). We sometimes say that a function maps its domain onto its range ($f\colon\ D \longrightarrow R$). The symbol used to denote a typical real number in the domain of a function is sometimes called the *independent variable*. The symbol used to denote the typical real number in the range is called the *dependent variable*. Generally, but not always, variables are denoted by lower-case letters such as $x, y, z, t, \ldots$ Functions are generally denoted by $f, g, h$, and capital letters. If $f$ denotes a function, $x$ the independent variable, and $y$ the dependent variable, then it is common practice to write $y = f(x)$, which reads: "$y$ equals $f$ of $x$" or "$y$ equals $f$ at $x$". This means that the function $f$ assigns to each $x$ in its domain a number $f(x)$, which is also written $y$. Given a function $f$, we can construct a geometric picture of the function. For each number $x$ in the domain of $f$, we find the corresponding number $y = f(x)$ in the range and then we plot the point $[x, y]$. The set of all such points is called the *graph* of function $f(x)$.

**Remark 1.4** Suppose $D \subseteq \mathbb{R}$ and mapping $f : D \longrightarrow R$ is a real function of one real variable, then domain, range and graph are the following sets:

- $\mathcal{D}(f) = \{x \in \mathbb{R} : \exists! y \in \mathbb{R} : y = f(x)\}$.
- $\mathcal{R}(f) = \{y \in \mathbb{R} : \exists x \in \mathbb{R} : y = f(x)\}$.
- $\mathcal{G}(f) = \{[x, y] \in \mathbb{R}^2 :\ y = f(x);\ x \in \mathcal{D}(f)\}$.

**Definition 1.10** Let $f$ be a function and $\mathcal{D}(f)$ is its domain and $M \subseteq \mathcal{D}(f)$. If for each of two real numbers $x_1, x_2 \in M$ such that $x_1 < x_2$ holds:

(a) $f(x_1) < f(x_2)$, then the function $f$ is called *increasing* on $M$,

(b) $f(x_1) > f(x_2)$, then the function $f$ is called *decreasing* on $M$,

(c) $f(x_1) \leqq f(x_2)$, then the function $f$ is called *non-decreasing* on $M$,

(d) $f(x_1) \geqq f(x_2)$, then the function $f$ is called *non-increasing* on $M$.

If function $f$ is non-increasing or non-decreasing on $M$, then $f$ is called *monotonic* on $M$. If $f$ is increasing or decreasing on set $M$, then $f$ is called *strictly monotonic* on $M$.

**Remark 1.5** We can write this more formally:

(a) $f$ is increasing on $M$ if $\forall x_1, x_2 \in M\colon\ x_1 < x_2 \Longrightarrow f(x_1) < f(x_2)$,

(b) $f$ is decreasing on $M$ if $\forall x_1, x_2 \in M\colon\ x_1 < x_2 \Longrightarrow f(x_1) > f(x_2)$,

(c) $f$ is non-decreasing on $M$ if $\forall x_1, x_2 \in M\colon\ x_1 < x_2 \Longrightarrow f(x_1) \leqq f(x_2)$,

(d) $f$ is non-increasing on $M$ if $\forall x_1, x_2 \in M\colon\ x_1 < x_2 \Longrightarrow f(x_1) \geqq f(x_2)$.

A function which is increasing on its whole domain is called shortly increasing (without a specification where). Similarly, one can introduce the notions of a decreasing, non-increasing, non-decreasing, monotonic and strictly monotonic function.

**Definition 1.11** Function $f\colon\ y = f(x)$ is said to be *one-to-one* (*unique*) if for all $x_1, x_2 \in \mathcal{D}(f)$ such that $x_1 \neq x_2$ holds: $f(x_1) \neq f(x_2)$.

**Definition 1.12** Let $f$ be a function and $\mathcal{D}(f)$ be its domain. If for all $x \in \mathcal{D}(f)$ is also $-x \in \mathcal{D}(f)$, then function $f$ is called:

(a) *even*, if for all $x \in \mathcal{D}(f)$ holds: $f(-x) = f(x)$,

(b) *odd*, if for all $x \in \mathcal{D}(f)$ holds: $f(-x) = -f(x)$.

**Remark 1.6** The graph of an even function is symmetric with respect to the $y$-axis and the graph of an odd function is symetric with respect to the origin, i. e. point $O = [0, 0]$.

**Definition 1.13** Let $f$ be a function and $\mathcal{D}(f)$ be a domain of $f$. Let $p$ be a positive real number. Function $f$ is called *periodic* with *period* $p$ if

(1) $\forall x \in \mathcal{D}(f)\colon\ x + p \in \mathcal{D}(f)$ and

(2) $\forall x \in \mathcal{D}(f)\colon\ f(x + p) = f(x)$.

The smallest positive real number $p$ with the above properties is called a *period of the function* $f$.

**Definition 1.14** Let functions $g$: $z = g(x)$ $g : A \longrightarrow C$ and $f$: $y = f(x)$ $f : C \longrightarrow B$ (where $A$, $B$ and $C$ are subsets of real number set) are given. If functions $f$ and $g$ are such functions that $\mathcal{R}(g) \subset \mathcal{D}(f)$, we can define a function $F$: $y = F(x)$ by the equation $y = f(g(x))$ for $x \in \mathcal{D}(g)$. The function $F : A \longrightarrow B$ is called the *composite function* of functions $f$ and $g$. We use the notation $F = f \circ g$. The function $f$ is called the *outside function* and the function $g$ the *inside function* of the function $F$.

**Definition 1.15** Let $f$: $y = f(x)$ be a one-to-one function and $\mathcal{D}(f)$ be a domain and $\mathcal{R}(f)$ be a range of function $f$. Function that assigns to each real number $y \in \mathcal{R}(f)$, such that for real number $x \in \mathcal{D}(f)$ applies that $y = f(x)$ is called the *inverse function* of a function $f$, and denoted by the symbol $f^{-1}$. There are inverse mappings to one-to-one mappings, there are also called inverse function to one-to-one functions.

**Remark 1.7** The graphs of the functions $f^{-1}$ and $f$ are symmetric with respect to the axis of the $1^{\text{st}}$ and $3^{\text{rd}}$ quadrant (i. e. straight line $p$: $y = x$). For all $x \in \mathcal{D}(f)$ applies that $f^{-1}(f(x)) = x$ and for all $y \in \mathcal{D}(f^{-1})$ applies that $f(f^{-1}(y)) = y$. For functions $f^{-1}$ and $f$ applies that $\mathcal{D}(f^{-1}) = \mathcal{R}(f)$ and $\mathcal{R}(f^{-1}) = \mathcal{D}(f)$. Inverse function exists only for one-to-one function. A strictly monotonic function is one-to-one and so an inverse function exists.

**Definition 1.16** Function $f$ is called *bounded from above* (*upper bounded*) if there exists a real number $K \in \mathbb{R}$ such that ($\forall x \in \mathcal{D}(f)$: $f(x) \leqq K$). We can define analogously a function *bounded from below* (*lower bounded*). Function $f$ is called *bounded* if $f$ is bounded from above and below. Assume further that $M \subseteq \mathcal{D}(f)$. Function $f$ is called *upper bounded on the set* $M$ if there exists a real number $K \in \mathbb{R}$ such that ($\forall x \in M$: $f(x) \leqq K$). We can similarly define the notion of a function *lower bounded on the set* $M$ and the notion of a function *bounded on the set* $M$.

**Definition 1.17** We say that function $f$ has its *maximum* at the point $x_0 \in \mathcal{D}(f)$ if ($\forall x \in \mathcal{D}(f)$: $f(x) \leqq f(x_0)$). Analogously, function $f$ has its *minimum* at the point $x_0$. The maximum and minimum of function $f$ are both called *extreme value* of $f$. Suppose that $M \subseteq \mathcal{D}(f)$. We say that function $f$ has its *maximum on the set* $M$ at the point $x_0 \in M$ if $\forall x \in M$: $f(x) \leqq f(x_0)$. We can also define the *minimum* of function $f$ *on the set* $M$ at the point $x_0 \in M$ if ($\forall x \in M$: $f(x) \geqq f(x_0)$).

**Definition 1.18** The supremum of the set of values of function $f$ (i. e. of the set $\mathcal{R}(f)$) is called the *supremum of the function* $f$ and it is the least upper bound of $f$. The infimum of the set of values of the function $f$ (i. e. of the set $\mathcal{R}(f)$) is called the *infimum of the function* $f$ and it is the greatest lower bound of $f$.

## 1.3 Elementary Functions

### 1.3.1 Constant Function

The general form of a *constant function* $f$ is $f\colon\ y = k$, where $k \in \mathbb{R}$. The domain of the constant function is the set of all real numbers, $\mathcal{D}(f) = \mathbb{R}$. Range is a single element set with element $k$, $\mathcal{R}(f) = \{k\}$. Graph of the constant function is a straight line parallel to the axis of $x$, which intersects the axis of $y$ in point $[0, k]$. See the Figure 1.4.

### 1.3.2 Linear Function

The general form of a *linear function* $f$ is $f\colon\ y = a \cdot x + b$, where $a, b \in \mathbb{R}$ and $a \neq 0$. The domain of the function $f$ and the range of the function $f$ is the real number set, $\mathcal{D}(f) = \mathcal{R}(f) = \mathbb{R}$. The graph of the linear function is a straight line. The coefficients of the linear function $a$ and $b$ have the following meanings:[3]

| | | |
|---|---|---|
| $a = \operatorname{tg} \varphi$ | – | slope of the line, which is a graph of a linear function, |
| $a > 0$ | – | linear function is increasing, |
| $a < 0$ | – | linear function is decreasing, |
| $b$ | – | section which is cut off by the straight line on the $y$-axis, |
| $b = 0$ | – | the graph of linear function is passing through origin of the coordinate system, |
| $a = 0$ | – | straight line is parallel to the axis $x$. |

In this case ($a = 0$) it is not a linear function, but the function is constant.

[3] $\varphi$ is the angle formed by a straight line (linear function graph) with positive orientation of axis $x$.

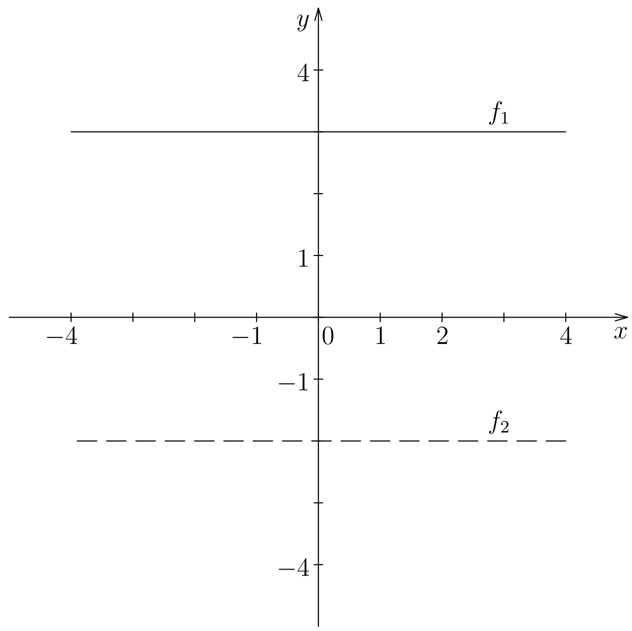


Figure 1.4: Function $f_1(x)$ is given by formula $f_1 : y = 3$ and function $f_2(x)$ is given by formula $f_2 : y = -2$.

Functions shown in the Figure 1.5 are examples of linear functions. The function $f_1$ is given by the rule $y = x$. Graph of the function $f_1$ is a straight line that intersects the $y$ axis at the point $[0, 0]$ and the function $f_1$ is increasing. The function $f_2$ is given by the rule $y = x + 2$. The graph of the function $f_2$ was created by moving the graph of the function $f_1$ in the positive direction with respect the $y$ axis by 2. The graph of this function is a straight line that intersects the $y$ axis at the point $[0, 2]$ and the function $f_2$ is increasing. The function $f_3$ is given by the rule $y = -x - 1$. The graph is a straight line that intersects the $y$ axis at the point $[0, -1]$ and the function $f_3$ is decreasing.

### 1.3.3 Quadratic Function

*A quadratic function* $f$ has the form $f\colon\ y = a \cdot x^2 + b \cdot x + c$, where $a, b, c \in \mathbb{R}$ and $a \neq 0$. The domain of the function $f$ is the real numbers set, $\mathcal{D}(f) = \mathbb{R}$ and the range of the quadratic function is set:

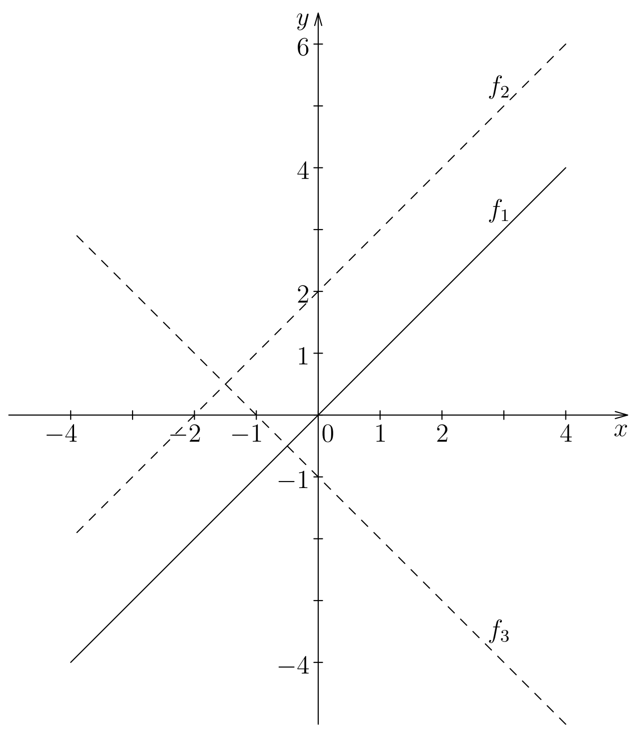


Figure 1.5: Linear function: graphs of functions $f_1(x) = x$, $f_2(x) = x + 2$, and $f_3(x) = -x - 1$.

(a) $\mathcal{R}(f) = \left\langle c - \dfrac{b^2}{4 \cdot a}, \infty \right)$, if $a > 0$ or

(b) $\mathcal{R}(f) = \left( -\infty, c - \dfrac{b^2}{4 \cdot a} \right\rangle$, if $a < 0$.

The graph of a quadratic function is a parabola whose axis is parallel to the $y$ axis. For the positive values of the parameter $a$ is the parabola open upwards and for the negative values of the parameter $a$ is the parabola open downwards. The peak $V$ of parabola has coordinates:

$$V = \left[ \frac{-b}{2 \cdot a}, c - \frac{b^2}{4 \cdot a} \right].$$

The special forms of quadratic function:

$f\colon\ y = a \cdot x^2$ – parabola with the peak at the point $V = [0, 0]$,
$f\colon\ y = a \cdot x^2 + c$ – parabola with the peak at the point $V = [0, c]$,
$f\colon\ y = a \cdot (x + d)^2$ – parabola with the peak at the point $V = [-d, 0]$,
$f\colon\ y = a \cdot (x + d)^2 + k$ – parabola with the peak at the point $V = [-d, k]$,
$f\colon\ y = x^2 + p \cdot x + q$ – standardized parabola with the peak at the point $V = \left[-\frac{p}{2}, q - \left(\frac{p}{2}\right)^2\right]$.

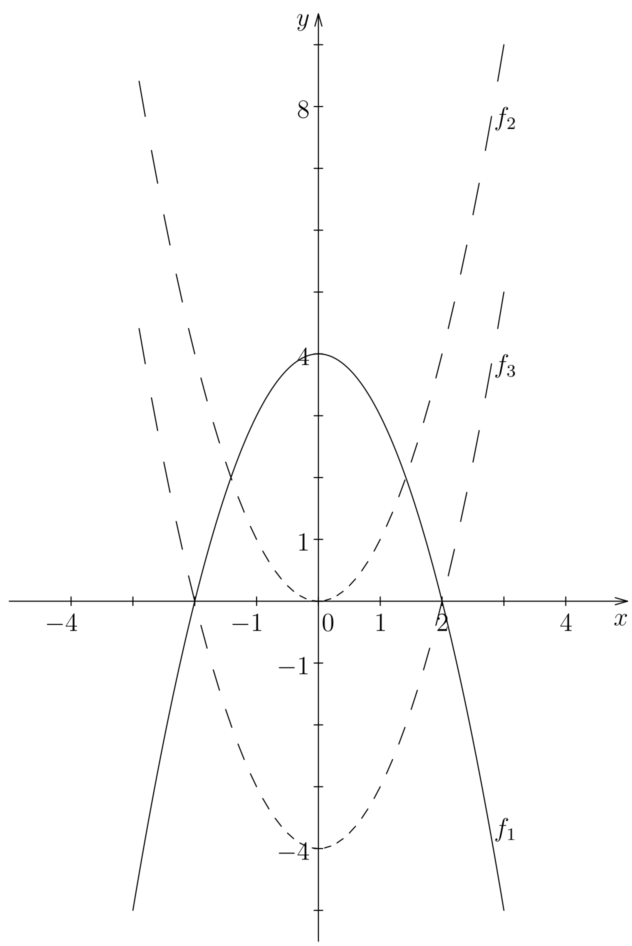


Figure 1.6: Quadratic Function: graphs of functions $f_1(x) = 4 - x^2$, $f_2(x) = x^2$, and $f_3(x) = x^2 - 4$.

Three graphs of quadratic functions (parabolas) $f_1(x)$, $f_2(x)$ and $f_3(x)$ are shown in the Figure 1.6. The function $f_2(x)$ is given by the rule $f_2\colon\ y = x^2$ and the peak is at the point $[0, 0]$, which is the minimum of the function $f_2$. The function $f_3(x)$ is given by the rule $f_3\colon\ y = x^2 - 4$ and the peak is at the point $[0, -4]$. This peak is the minimum of the function $f_3$. The function $f_1(x)$ is given by the rule $f_1\colon\ y = 4 - x^2$. The function $f_1$ has the maximum at the point $[0, 4]$.

### 1.3.4 Power Function

A *power function* $f$ has the form $f\colon\ y = a \cdot x^k$, where $k \in \mathbb{N}$, and $a \in \mathbb{R} - \{0\}$. The domain of the power function is the set of all real numbers, $\mathcal{D}(f) = \mathbb{R}$.

- for $a > 0$ and $k = 2n$, $n \in \mathbb{N}$; the graph of the function $f$ is a parabola of $k$-th degree with the peak at the beginning of the coordinate system $V = [0, 0]$, which is open upward and the range is the set of non-negative real numbers, $\mathcal{R}(f) = \langle 0, \infty)$,

- for $a < 0$ and $k = 2n$, $n \in \mathbb{N}$; the graph of the function $f$ is a parabola of $k$-th degree with the peak at the beginning of the coordinate system $V = [0, 0]$, which is open downwards and the range is the set of non-positive real numbers, $\mathcal{R}(f) = (-\infty, 0\rangle$,

- for $a > 0$ and $k = 2n + 1$, $n \in \mathbb{N}$; the graph of the function $f$ is a parabola of $k$-th degree, which lies in the first and the third quadrant, where the center of symmetry is the origin of the coordinate system, point $V = [0, 0]$. The range of the function is the set $\mathcal{R}(f) = \mathbb{R}$,

- for $a < 0$ and $k = 2n + 1$, $n \in \mathbb{N}$; the graph of the function $f$ is a parabola of $k$-th degree, which lies in the second and the fourth quadrant, where the center of symmetry is the origin of the coordinate system, point $V = [0, 0]$. The range of the function is the set $\mathcal{R}(f) = \mathbb{R}$.

If the exponent $k$ can have a negative value, we obtain a function in the form $f\colon\ y = a \cdot x^{-k}$, where $k \in \mathbb{N}$ and $a \in \mathbb{R} - \{0\}$. The domain of this function is the set of all real numbers excluding zero, $\mathcal{D}(f) = \mathbb{R} - \{0\}$.

- for $a > 0$ and $k = 2n$, $n \in \mathbb{N}$; the graph of the function $f$ is a hyperbola of $k$-th degree, which lies in the first and the second quadrant. The range of the function is the set of positive real numbers, $\mathcal{R}(f) = (0, \infty)$,
- for $a < 0$ and $k = 2n$, $n \in \mathbb{N}$; the graph of the function $f$ is a hyperbola of $k$-th degree, which lies in the third and the fourth quadrant. The range of the function is the set of negative real numbers, $\mathcal{R}(f) = (-\infty, 0)$,
- for $a > 0$ and $k = 2n + 1$, $n \in \mathbb{N}$; the graph of the function $f$ is a hyperbola of $k$-th degree, which lies in the first and the third quadrant, where the center of symmetry is the origin of the coordinate system, point $V = [0, 0]$. The range of the function is the set $\mathcal{R}(f) = \mathbb{R} - \{0\}$. For $n = 0$ we obtain the function $y = \frac{1}{x}$ (inversely proportional = lowers with raising), for which it holds: $f = f^{-1}$. Its graph is a equiaxed hyperbola,
- for $a < 0$ and $k = 2n + 1$, $n \in \mathbb{N}$; the graph of the function $f$ is a hyperbola of $k$-th degree, which lies in the second and the fourth quadrant, where the center of symmetry is the origin of the coordinate system, point $V = [0, 0]$. The range of the function is the set $\mathcal{R}(f) = \mathbb{R} - \{0\}$.

If the exponent $k$ is chosen from the set of rational numbers, we get graphs of functions that correspond to the square roots of real numbers. Example of such functions can be functions shown in the Figure 1.7. These functions have the following forms: $f_1$: $y = \sqrt{x}$, $f_2$: $y = \sqrt{x} - 2$, and $f_3$: $y = 2 - \sqrt{x}$.

### 1.3.5 Exponential Function

*An exponential function* $f$ has the form $f$: $y = a^x$, where $a > 0$ and $a \neq 1$. The domain of the exponential function is the set of all real numbers, $\mathcal{D}(f) = \mathbb{R}$. The range of the exponential function $f$ is the set of positive real numbers, $\mathcal{R}(f) = (0, \infty)$. The graph of the exponential function is an exponential curve that passes through the point $[0, 1]$.The function is one-to-one. The function is increasing for values of parameter $a > 1$ and the function is decreasing for values of parameter $a \in (0, 1)$. The most important of exponential functions is the function $y = \mathrm{e}^x$, where $\mathrm{e} = 2{,}7182818\ldots$ (well-known Euler's number),

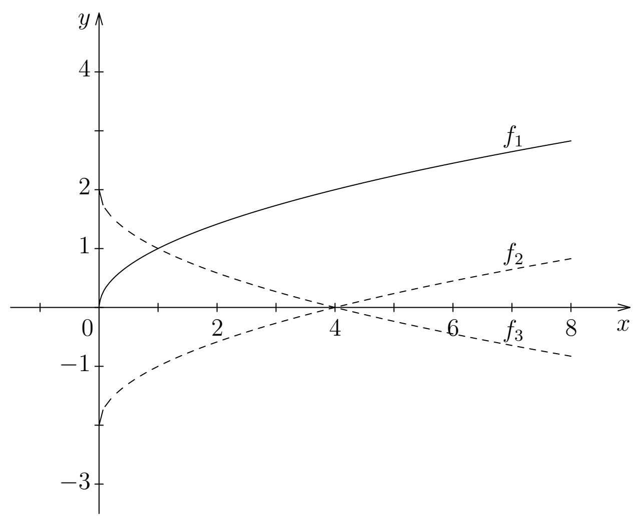


Figure 1.7: Function of the square root: the graphs of the functions $f_1(x) = \sqrt{x}$, $f_2(x) = \sqrt{x} - 2$, and $f_3(x) = 2 - \sqrt{x}$.

which is called the natural exponential function. Exponential function is one of the transcendent functions.

Examples of graphs of exponential functions are shown in the Figure 1.8. The function $f_1$ has the form $y = \mathrm{e}^x$. It intersects at the point $[0, 1]$ the $y$ axis and it is increasing. The function $f_2$ has the form $y = \mathrm{e}^{\frac{x}{5}}$. The function $f_2$ intersect the $y$ axes at the point $[0, 1]$ and it is also increasing and the one-to-one function. The function $f_3$ has the form $y = 4 - \mathrm{e}^x$. The function $f_3$ intersect the axis of $y$ at the point $[0, 3]$. The function $f_3$ is decreasing and it is the one-to-one function.

### 1.3.6 Logarithmic Function

*A logarithmic function* is the inverse function of the corresponding exponential function. The general form of the logarithmic function is $f\colon\ y = \log_a x$, where $a > 0$ and $a \neq 1$. The domain of the logarithmic function is the set of all positive real numbers, $\mathcal{D}(f) = (0, \infty)$. The range of the logarithmic function is the real numbers set, $\mathcal{R}(f) = \mathbb{R}$. The graph of the logarithmic function is a logarithmic curve passing through the point $[1, 0]$. The function

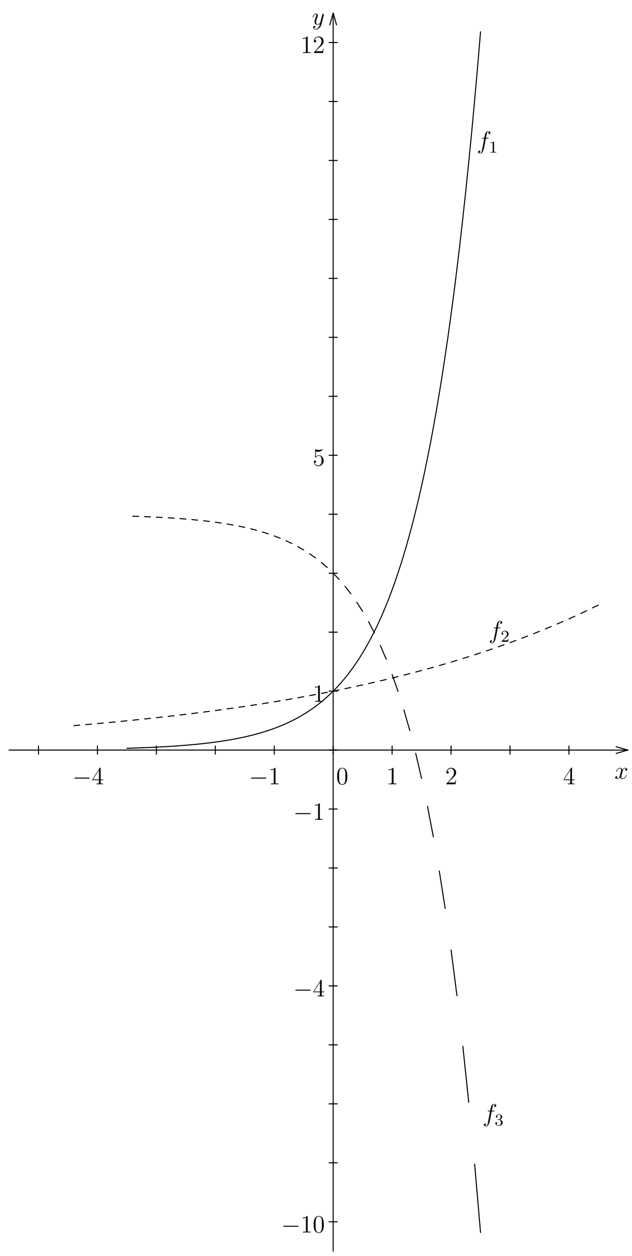


Figure 1.8: Exponential function: graphs of functions $f_1(x) = \mathrm{e}^x$, $f_2(x) = \mathrm{e}^{\frac{x}{5}}$, and $f_3(x) = 4 - \mathrm{e}^x$.

is one-to-one. The logarithmic function is increasing for $a > 1$ and decreasing for $a \in (0, 1)$. The most important logarithmic functions are functions $y = \log_e x = \ln x$, where $e = 2{,}7182818\ldots$, which is called the natural logarithmic function (natural logarithm) and $y = \log_{10} x = \log x$, which is called the decadic logarithm. Logarithmic function also belongs to the transcendent functions.

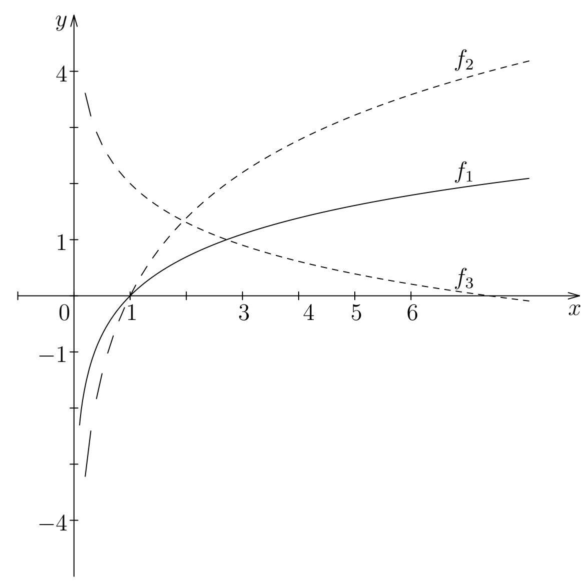


Figure 1.9: Logarithmic function: the graphs of the functions $f_1(x) = \ln x$, $f_2(x) = 2 \cdot \ln x$, and $f_3(x) = 2 - \ln x$

Examples of logarithmic functions are shown in the Figure 1.9. The function $f_1$: $y = \ln x$ intersects the $x$ axis at the point $[1, 0]$ and it is increasing function. The graph of the function $f_2$: $y = 2 \cdot \ln x$ intersects the $x$ axis at the point $[1, 0]$. It is increasing and one-to-one function. The function $f_3$: $y = 2 - \ln x$ is decreasing and one-to-one function.

### 1.3.7 Trigonometric Functions

*Trigonometric functions* is the common name for the functions sine (formally $y = \sin x$), cosine ($y = \cos x$), tangent ($y = \operatorname{tg} x = \frac{\sin x}{\cos x}$)[4], and cotangent

[4]In English texts is used such labelling of tangent function: $y = \tan x$

($y = \operatorname{cotg} x = \frac{\cos x}{\sin x}$)[5]. The domain of the sine and the cosine functions is the set of all real numbers, $\mathcal{D}(f) = \mathbb{R}$ and the range is the closed interval between $-1$ and 1, $\mathcal{R}(f) = \langle -1, 1\rangle$. The domain of the tangent function is the set $\mathcal{D}(f) = \mathbb{R} - \{\frac{\pi}{2} + k\pi, k \in \mathbb{Z}\}$ and the domain of the cotangent function is the set $\mathcal{D}(f) = \mathbb{R} - \{k\pi, k \in \mathbb{Z}\}$. The range of functions tangent and cotangent is the real numbers set, $\mathcal{R}(f) = \mathbb{R}$. Trigonometric functions are periodic. Functions sine and cosine have the period $2\pi$ and the tangent and cotangent functions have the period $\pi$. For these functions holds, the following equality for all $x \in \mathcal{D}(f)$:

$$\begin{aligned}
&\sin x = \sin(x + 2\pi \cdot k) && \text{for } \forall k \in \mathbb{Z},\\
&\cos x = \cos(x + 2\pi \cdot k) && \text{for } \forall k \in \mathbb{Z},\\
&\operatorname{tg} x = \operatorname{tg}(x + k \cdot \pi) && \text{for } \forall k \in \mathbb{Z},\\
&\operatorname{cotg} x = \operatorname{cotg}(x + k \cdot \pi) && \text{for } \forall k \in \mathbb{Z}.
\end{aligned}$$

Graphs of trigonometric functions are:

$$\begin{aligned}
&f_1 = \{[x, y] : y = \sin x, x \in \mathbb{R}, y \in \langle -1, 1\rangle\} && \text{(sinusoid)},\\
&f_2 = \{[x, y] : y = \cos x, x \in \mathbb{R}, y \in \langle -1, 1\rangle\} && \text{(cosinusoid)},\\
&f_3 = \{[x, y] : y = \operatorname{tg} x, x \in \mathbb{R} - \{\tfrac{\pi}{2} + k\pi, k \in \mathbb{Z}\}, y \in \mathbb{R}\} && \text{(tangentoid)},\\
&f_4 = \{[x, y] : y = \operatorname{cotg} x, x \in \mathbb{R} - \{k\pi, k \in \mathbb{Z}\}, y \in \mathbb{R}\} && \text{(cotangentoid)}.
\end{aligned}$$

Graphs of trigonometric functions sine, cosine, tangent, and cotangent are shown in Figures 1.10, 1.11, 1.12, and 1.13. The function $\sin x$ is shown in the Figure 1.10 as a function $f_1$. The function $f_2$ has the form $y = 1 + \sin x$. The graph of the function $f_3$ has the form $y = -2 \cdot \sin x$.

The function $\cos x$ is shown in the Figure 1.11. The graph of the function $f_1$ has the form $y = \cos x$. the function $f_2$ has the form $y = 1 + \cos x$ and the graph of the function $f_3$ is given by the form $y = 3 \cdot \cos x$. The graph of the function $f$: $y = \operatorname{tg} x$ is shown in the Figure 1.12 and the graph of the function $f$: $y = \operatorname{cotg} x$ is shown in the Figure 1.13. Points that do not belong to the domain of these two functions are shown by dashed lines perpendicular to the axis $x$.

[5] In English texts is used such labelling of cotangent function: $y = \cot x$

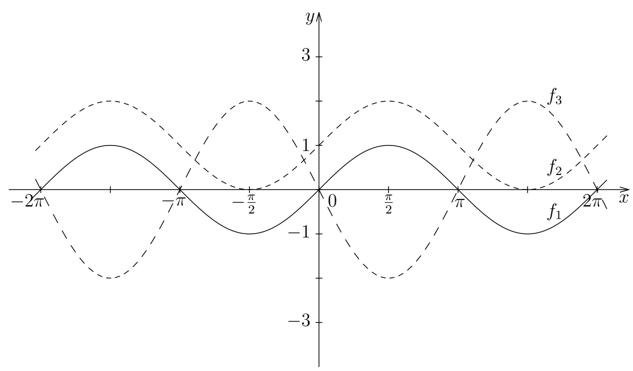


Figure 1.10: Function sine: the graphs of the functions $f_1(x) = \sin x$, $f_2(x) = 1 + \sin x$, and $f_3(x) = -2 \cdot \sin x$.

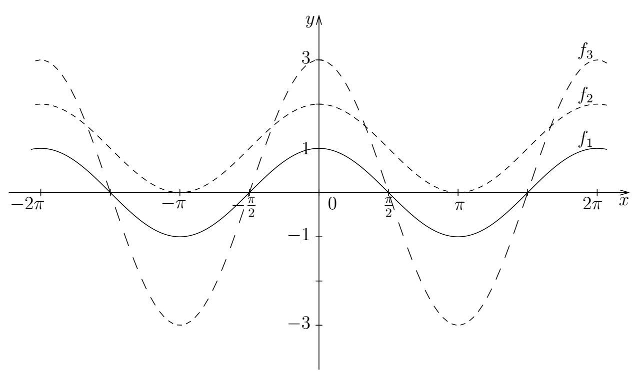


Figure 1.11: Function cosine: the graphs of the functions $f_1(x) = \cos x$, $f_2(x) = 1 + \cos x$, and $f_3(x) = 3 \cdot \cos x$.

### 1.3.8 Cyclometric Functions

Inverse trigonometric functions:
Trigonometric functions are not one-to-one in their domain of definition, therefore, to them, there are not inverse functions. If we turn our attention to the domain of a suitable interval so that it was a one-to-one function, then we can define the inverse function of it. Thus created inverse functions of trigonometric functions are called inverse trigonometric functions.

**1.** ***Function arcsine:*** The function $y = \sin x$ is increasing and one-to-

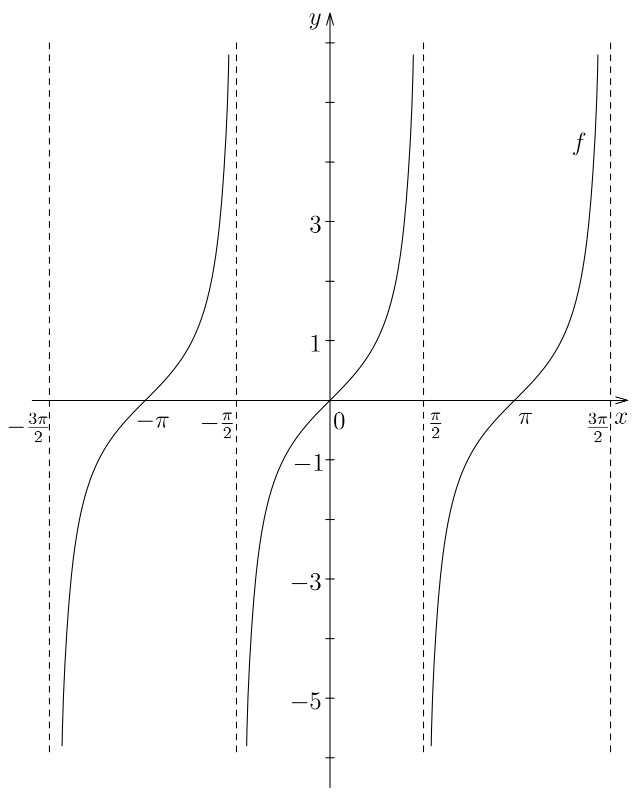


Figure 1.12: Function tangent: the graph of the function $f$: $y = \operatorname{tg} x$.

one on the closed interval $\langle -\frac{\pi}{2}, \frac{\pi}{2} \rangle$ and is mapping this interval to a closed interval $\langle -1, 1 \rangle$. The inverse function to the function $\sin x$, for $x \in \langle -\frac{\pi}{2}, \frac{\pi}{2} \rangle \subseteq \mathcal{D}(f)$ is the function arcsine, $y = \arcsin x$. The domain of the function $y = \arcsin x$ is the interval $\mathcal{D}(f) = \langle -1, 1 \rangle$ and the range of the function is the interval $\mathcal{R}(f) = \langle -\frac{\pi}{2}, \frac{\pi}{2} \rangle$. The function is increasing and is one-to-one on the interval $\langle -1, 1 \rangle$.

**2.** ***Function arccosine*:** The function $y = \cos x$ is decreasing and one-to-one on the closed interval $\langle 0, \pi \rangle$ and is mapping this interval to a closed interval $\langle -1, 1 \rangle$. The inverse function to the function $\cos x$, for $x \in \langle 0, \pi \rangle \subseteq \mathcal{D}(f)$ is the function arccosine, $y = \arccos x$. The domain of the function $y = \arccos x$ is the interval $\mathcal{D}(f) = \langle -1, 1 \rangle$ and the range of the function is the interval $\mathcal{R}(f) = \langle 0, \pi \rangle$. The function is decreasing and one-to-one on the interval $\langle -1, 1 \rangle$.

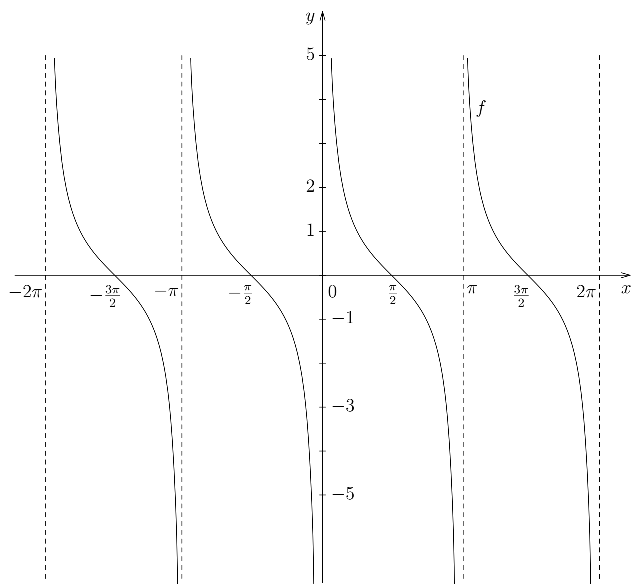


Figure 1.13: Function cotangent: the graph of the function $f$: $y = \operatorname{cotg} x$.

**3.** ***Function arctangent:*** The function $y = \operatorname{tg} x$ is increasing and one-to-one on the opened interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ and is mapping this interval to the set of all real numbers $(-\infty, \infty)$. The inverse function to the function $\operatorname{tg} x$ is the function arctangent, $y = \operatorname{arctg} x$, for $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \subseteq \mathcal{D}(f)$. The domain of the function $y = \operatorname{arctg} x$ is the real numbers set $\mathcal{D}(f) = \mathbb{R}$ and the range of the function is the interval $\mathcal{R}(f) = \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. The function is increasing and one-to-one on the set $(-\infty, \infty)$.

**4.** ***Function arccotangent:*** The function $y = \operatorname{cotg} x$ is decreasing and one-to-one on the opened interval $(0, \pi)$ and is mapping this interval to the real numbers set $(-\infty, \infty)$. The inverse function to the function $\operatorname{cotg} x$ is the function arccotangent, $y = \operatorname{arccotg} x$, for $x \in (0, \pi) \subseteq \mathcal{D}(f)$. The domain of the function $y = \operatorname{arccotg} x$ is the set of all real numbers $\mathcal{D}(f) = \mathbb{R}$ and the range of the function is the interval $\mathcal{R}(f) = (0, \pi)$. The function is decreasing and one-to-one on the set $(-\infty, \infty)$.

### 1.3.9 Polynomial Function

A polynomial function (polynomial):
The polynomial of the $n$-th degree is the function:

$$P(x) = a_0 + a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3 + \cdots + a_{n-1} \cdot x^{n-1} + a_n \cdot x^n,$$

where $a_0, a_1, a_2, a_3, \ldots, a_n$ are real numbers and $a_n \neq 0$. Specially, if $n = 0$, then $P(x)$ is called the *constant function*, if $n = 1$, then $P(x)$ is called the *linear polynomial* (*linear function*), if $n = 2$, then $P(x)$ is called the *quadratic polynomial* (*quadratic function*) and if $n = 3$, then $P(x)$ is the *cubic polynomial* (*cubic function*).

## 1.4 Limit of Function

**Definition 1.19** We say that the function $f\colon\ y = f(x)$ has a *right side limit* (Right Hand Limit or Limit from the right) at the point $a$ equal to real number $L$, if following holds: $(\forall\varepsilon > 0)(\exists\delta > 0)(\forall x \in (a, a+\delta)\colon (0 < |x - a| < \delta \Longrightarrow |f(x) - L| < \varepsilon))$. We write:

$$\lim_{x\to a^+} f(x) = L$$

and it is said the limit of the function $f$ as $x$ approaches $a$ from the right (above) is $L$.

**Definition 1.20** We say that the function $f\colon\ y = f(x)$ has a *left side limit* (Left Hand Limit or Limit from the left) at the point $a$ equal to real number $L$, if following holds: $(\forall\varepsilon > 0)(\exists\delta > 0)(\forall x \in (a-\delta, a)\colon (0 < |x - a| < \delta \Longrightarrow |f(x) - L| < \varepsilon))$. We write:

$$\lim_{x\to a^-} f(x) = L$$

and it is said the limit of the function $f$ as $x$ approaches $a$ from the left (below) is $L$.

**Definition 1.21** Let $f\colon\ y = f(x)$ be a function with a domain $\mathcal{D}(f)$. Let $a$ be a point of the real axis such that $\mathcal{D}(f)$ contains points different from $a$ that are arbitrarily close to $a$. Let $L$ be a real number. Then the limit as $x$ approaches $a$ of $f(x)$ is $L$:

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

if for each number $\varepsilon > 0$ there exists a number $\delta > 0$ such that $|f(x) - L| < \varepsilon$ for all $x$ in $\mathcal{D}(f)$ that the inequality $0 < |x - a| < \delta$ holds. Formally, it is written as follows: $(\forall\varepsilon > 0)(\exists\delta > 0)(\forall x \in (a-\delta, a+\delta)\colon (0 < |x - a| < \delta \Longrightarrow |f(x) - L| < \varepsilon))$.This means that there are the same two one-sided limits.

**Remark 1.8** It does not mater whether $a$ is in the domain $\mathcal{D}(f)$ or not. Even if $a$ is in $\mathcal{D}(f)$, the value of $f(x)$ at $a$ itself does not enter into the definition, because we consider only points $x$ for which $0 < |x - a|$. The definition depends only on the values of $f(x)$ for $x$ near $a$.

**Remark 1.9** The domain $\mathcal{D}(f)$ must contain points different from $a$ but arbitrarily close to $a$. The most important case are when $\mathcal{D}(f)$ is an interval and $a$ is the point of $\mathcal{D}(f)$, and when $a$ is an end point of $\mathcal{D}(f)$ but not necessarily in $\mathcal{D}(f)$. The expression "arbitrarily close to $a$" means that, given any positive $\delta$, no mater how small, there is a point $x$ in $\mathcal{D}(f)$ such that $0 < |x - a| < \delta$.

**Remark 1.10** In practice, the definition works this way. If provided with a arbitrary $\varepsilon$, you must determine a suitable $\delta$. You must be able to do so for every $\varepsilon > 0$, not just a particular $\varepsilon$. Note that the $\delta$ you produce will depend on the $\varepsilon$ you get. A $\delta$ that works for one particular $\varepsilon$ will not work for a different (much smaller) $\varepsilon$, generally.

**Remark 1.11** Limits involving infinity: Suppose a function $f(x)$ is defined on a domain that includes arbitrarily large real numbers, and suppose $L$ is a real number. We define $\lim\limits_{x\to\infty} f(x) = L$ to mean that for any $\varepsilon > 0$, there exists a real number $N$, such that $|f(x) - L| < \varepsilon$ for all $x$ in the domain of $f(x)$, such that $x > N$.

If $a = \infty$ or $a = -\infty$, then for a real function $f(x)$, the limit of $f$ as $x$ approaches infinity (the limit of $f(x)$ as $x$ approaches negative infinity) is $L$, denoted:

$$\lim_{x\to\infty} f(x) = L$$

or

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

**Theorem 1.7** If $\lim\limits_{x\to a} g(x) = M$ and $g(x) \neq M$ in a certain neighbourhood of the point $a$ and $\lim\limits_{z\to M} f(z) = L$, then $\lim\limits_{x\to a} f(g(x)) = L$.

**Theorem 1.8** If a function $f(x)$ is real-valued, then the limit of $f(x)$ at the point $a$ is $L$ if and only if both the right-hand limit and the left-hand limit of $f(x)$ at $a$ exist and are equal to $L$. i. e.

$$\lim_{x\to a^-} f(x) = \lim_{x\to a^+} f(x) = \lim_{x\to a} f(x) = L$$

**Theorem 1.9** (*Basic Limit Rules*) Let functions $f$: $y = f(x)$ and $g$: $y = g(x)$ be given. Let $\lim\limits_{x\to a} f(x) = L \in \mathbb{R}$ and $\lim\limits_{x\to a} g(x) = M \in \mathbb{R}$. Then:

(1) $\lim\limits_{x\to a} |f(x)| = |L|$,

(2) $\lim\limits_{x\to a} (f(x) + g(x)) = \lim\limits_{x\to a} f(x) + \lim\limits_{x\to a} g(x) = L + M$,

(3) $\lim\limits_{x\to a} (f(x) - g(x)) = \lim\limits_{x\to a} f(x) - \lim\limits_{x\to a} g(x) = L - M$,

(4) $\lim\limits_{x\to a} (f(x) \cdot g(x)) = \lim\limits_{x\to a} f(x) \cdot \lim\limits_{x\to a} g(x) = L \cdot M$,

(5) If for all $x$ from the neighbourhood of the point $a$ is $g(x) \neq 0$ and $\lim\limits_{x\to a} g(x) \neq 0$, then $\lim\limits_{x\to a} \left(\frac{f(x)}{g(x)}\right) = \frac{\lim\limits_{x\to a} f(x)}{\lim\limits_{x\to a} g(x)} = \frac{L}{M}$.

**Theorem 1.10** Let functions $f$: $y = f(x)$ and $g$: $y = g(x)$ be given. Let $\lim\limits_{x\to a} f(x) = 0$ and the function $g$ be bounded. Then $\lim\limits_{x\to a} (f(x) \cdot g(x)) = 0$.

**Theorem 1.11** Let functions $f$: $y = f(x)$ and $g$: $y = g(x)$ be given. Then:

(1) if $\lim\limits_{x\to a} f(x) = \infty$, then $\lim\limits_{x\to a} (-f(x)) = -\infty$,

(2) if $\lim\limits_{x\to a} f(x) = -\infty$, then $\lim\limits_{x\to a} (-f(x)) = \infty$,

(3) if $\lim\limits_{x\to a} f(x) = -\infty$ or $\lim\limits_{x\to a} (-f(x)) = \infty$, then $\lim\limits_{x\to a} |f(x)| = \infty$,

(4) if $\lim\limits_{x\to a} f(x) = \infty$ and the set $\mathcal{R}(g)$ (range of the function $g$) is bounded from below, then $\lim\limits_{x\to a} (f(x) + g(x)) = \infty$,

(5) if $f(x) > 0$ and $\lim\limits_{x\to a} f(x) = 0$, then $\lim\limits_{x\to a} \frac{1}{f(x)} = \infty$,

(6) if $\lim\limits_{x\to a} |f(x)| = \infty$, then $\lim\limits_{x\to a} \frac{1}{f(x)} = 0$.

**The basic formulas for calculating limits:** [6]

| | |
|---|---|
| (1) $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = 1$ | $\left[\frac{0}{0}\right]$ |
| (2) $\lim\limits_{x\to 0} \dfrac{e^x - 1}{x} = 1$ | $\left[\frac{0}{0}\right]$ |
| (3) $\lim\limits_{x\to \infty} \left(1 + \dfrac{1}{x}\right)^x = e$ | $[1^{+\infty}]$ |
| (4) $\lim\limits_{x\to -\infty} \left(1 + \dfrac{1}{x}\right)^x = e$ | $[1^{-\infty}]$ |
| (5) $\lim\limits_{x\to 0} \dfrac{1}{x} =$ does not exist | $\left[\frac{1}{0}\right]$ |
| (6) $\lim\limits_{x\to 0^-} \dfrac{1}{x} = -\infty$ | $\left[\frac{1}{0^-}\right]$ |
| (7) $\lim\limits_{x\to 0^+} \dfrac{1}{x} = \infty$ | $\left[\frac{1}{0^+}\right]$ |
| (8) $\lim\limits_{x\to \infty} a^x = \infty$, for $a > 1$ | $[a^\infty]$ |
| (9) $\lim\limits_{x\to \infty} a^x = 1$, for $a = 1$ | $[a^\infty]$ |
| (10) $\lim\limits_{x\to \infty} a^x = 0$, for $a \in (0,1)$ | $[a^\infty]$ |
| (11) $\lim\limits_{x\to -\infty} a^x = 0$, for $a > 1$ | $[a^{-\infty}]$ |
| (12) $\lim\limits_{x\to -\infty} a^x = 1$, for $a = 1$ | $[a^{-\infty}]$ |
| (13) $\lim\limits_{x\to -\infty} a^x = \infty$, for $a \in (0,1)$ | $[a^{-\infty}]$ |
| (14) $\lim\limits_{x\to \infty} e^x = \infty$ | $[e^\infty]$ |
| (15) $\lim\limits_{x\to -\infty} e^x = 0$ | $[e^{-\infty}]$ |
| (16) $\lim\limits_{x\to 0^+} \ln x = -\infty$ | $[\ln 0]$ |
| (17) $\lim\limits_{x\to \infty} \ln x = \infty$ | $[\ln \infty]$ |

[6] In square brackets is written the type of the limit.

(18) $\lim_{x\to\infty} x^n = \infty$, for $n \in \mathbb{N}$ $[\infty^n]$

(19) $\lim_{x\to-\infty} x^n = \infty$, for $n \in \mathbb{N}$, $n$ even $[(-\infty)^n]$

(20) $\lim_{x\to-\infty} x^n = -\infty$, for $n \in \mathbb{N}$, $n$ odd $[(-\infty)^n]$

(21) $\lim_{x\to\pm\infty} \frac{1}{x^n} = 0$, for $n \in \mathbb{N}$ $\left[\frac{1}{(\pm\infty)^n}\right]$

(22) $\lim_{x\to 0} \frac{1}{x^n} = \infty$, for $n \in \mathbb{N}$, $n$ even $\left[\frac{1}{(0)^n}\right]$

(23) $\lim_{x\to 0} \frac{1}{x^n} =$ does not exist, for $n \in \mathbb{N}$, $n$ odd $\left[\frac{1}{(0)^n}\right]$

(24) $\lim_{x\to 0^-} \frac{1}{x^n} = -\infty$, for $n \in \mathbb{N}$, $n$ odd $\left[\frac{1}{(0^-)^n}\right]$

(25) $\lim_{x\to 0^+} \frac{1}{x^n} = \infty$, for $n \in \mathbb{N}$, $n$ odd $\left[\frac{1}{(0^+)^n}\right]$

(26) $\lim_{x\to\left(\frac{\pi}{2}\right)^-} \operatorname{tg} x = \infty$ $\left[\operatorname{tg} \frac{\pi}{2}\right]$

(27) $\lim_{x\to\left(\frac{\pi}{2}\right)^+} \operatorname{tg} x = -\infty$ $\left[\operatorname{tg} \frac{\pi}{2}\right]$

(28) $\lim_{x\to 0^-} \operatorname{cotg} x = -\infty$ $[\operatorname{cotg} 0]$

(29) $\lim_{x\to 0^+} \operatorname{cotg} x = \infty$ $[\operatorname{cotg} 0]$

(30) $\lim_{x\to\infty} \operatorname{arctg} x = \frac{\pi}{2}$ $[\operatorname{arctg} \infty]$

(31) $\lim_{x\to-\infty} \operatorname{arctg} x = -\frac{\pi}{2}$ $[\operatorname{arctg} -\infty]$

(32) $\lim_{x\to\infty} \operatorname{arccotg} x = 0$ $[\operatorname{arccotg} \infty]$

(33) $\lim_{x\to-\infty} \operatorname{arccotg} x = \pi$ $[\operatorname{arccotg} -\infty]$

(34) $\lim_{x\to\pm\infty} \sin x =$ does not exist $[\sin \pm\infty]$

(35) $\lim_{x\to\pm\infty} \cos x =$ does not exist $[\cos \pm\infty]$

**Remark 1.12** When calculating the limit of the function $\lim_{x\to a} f(x)$ we can get the following results:

(1) $\lim_{x\to a} f(x) = b$, $b \in \mathbb{R}$ – i. e. there exists the real limit $b$ of the function $f(x)$ (bounded limit),

(2) $\lim_{x\to a} f(x) = \pm\infty$ – i. e. there exists infinite limit of function, but rather we say that the limit is infinity, the proper thing is to say that the function "diverges" or "grows without bound". It is said the limit of $f(x)$ as $x$ approaches $a$ is infinity (negative infinity) (unbounded limit).

(3) The limit $\lim_{x\to a} f(x)$ does not exist, but exist one-side limits, such that the following property holds: $\lim_{x\to a^-} f(x) = L$ and $\lim_{x\to a^+} f(x) = M$ and $L \neq M$.

(4) The limit does not make sense (is not defined), because the function $f(x)$ is not defined in the neighbourhood of the point $a$, the neighbourhood on the right, or the neighbourhood on the left side of the point $a$, respectively.

**Definition 1.22** Let the function $f$ be defined in some neighbourhood of the point $a \in I \subseteq \mathcal{D}(f)$. It is said that a *function $f$ is continuous at the point $a$*, if the following property holds: $\lim_{x\to a} f(x) = f(a)$, i. e. $(\forall\varepsilon > 0)(\exists\delta > 0)(\forall x \in I : |x - a| < \delta)(|f(x) - f(a)| < \varepsilon)$. It is said that a *function $f$ is continuous at the point $a$ from the right*, if $\lim_{x\to a^+} f(x) = f(a)$. It is said that a *function $f$ is continuous at the point $a$ from the left*, if $\lim_{x\to a^-} f(x) = f(a)$.

Therefore a function $f$ is said to be *continuous at the point $a$* if it is both defined at $a$ and its value at $a$ equals to the limit of $f$ as $x$ approaches $a$.

**Definition 1.23** A *function $f$ is continuous* if it is continuous at each point of the domain of the function $f$. The function $f$ is *continuous on set $I \subseteq \mathcal{D}(f)$*, if it is continuous at each point of the set $I$.

**Definition 1.24** It is said that a function $f$ is *continuous on an open interval* $(a, b)$, if it is continuous at each point of this interval. It is said that a function $f$ is *continuous on a closed interval $\langle a, b\rangle$*, if it is continuous at each point of this interval $(a, b)$ and, moreover, is continuous at the point $a$ from the right and continuous at the point $b$ from the left.

**Remark 1.13** A function $f$ is continuous at the point $a$ if and only if the limit of $f(x)$ as $x$ approaches $a$ exists and is equal to $f(a)$. If $f : M \longrightarrow N$ is a function between metric spaces $M$ and $N$, then it is equivalent that $f$ transforms every sequence in $M$ which converges towards $a$ into a sequence in $N$ which converges towards $f(a)$.

**Remark 1.14** If $N$ is a normed vector space, then the limit operation is linear in the following sense: if the limit of $f(x)$ as $x$ approaches $a$ is $L$ and the limit of $g(x)$ as $x$ approaches $a$ is $M$, then the limit of $f(x) + g(x)$ as $x$ approaches $a$ is $L + M$. If $\alpha$ is a scalar from the base field, then the limit of $\alpha \cdot f(x)$ as $x$ approaches $a$ is $\alpha \cdot L$.

**Theorem 1.12** If a function $f$ is continuous at a point $a$ from the right and also from the left, then the function $f$ is continuous at the point $a$.

**Theorem 1.13** Let functions $f$: $y = f(x)$ and $g$: $y = g(x)$ be continuous at the point $a \in \mathcal{D}(f)$ and let $\alpha \in \mathbb{R}$. Then the functions $f + g$, $f - g$, $\alpha \cdot f$, $f \cdot g$, $|f|$ are also continuous at the point $a$. If it is true, that $g(a) \neq 0$, then the function $\frac{f}{g}$ is continuous at the point $a$.

**Theorem 1.14** Let a function $f$: $y = f(x)$ be continuous at the point $a$ and a function $g$: $y = g(x)$ be continuous at the point $f(a)$, then the function $y = f(g(x))$ is continuous at the point $a$.

**Theorem 1.15** Each elementary function is continuous on its domain of definition.

**Theorem 1.16** Let a function $f : y = f(x)$ be continuous on a closed interval $\langle a, b\rangle$. Then the function $f$ attains its minimum and maximum at interval $\langle a, b\rangle$ and the function $f$ attains every value between the minimum and the maximum.

**Theorem 1.17** Let a function $f$ $y = f(x)$ be continuous on a closed interval $\langle a, b\rangle$ and let $f(a) \cdot f(b) < 0$. Then there exists a point $c \in (a, b)$ such that $f(c) = 0$.

**Remark 1.15** If $\lim\limits_{x\to a} f(x) = L$, then one of the following statements is true:

(1) If $f(a) = L$, then the function $f$ is continuous at the point $a$.

(2) If $f(a) \neq L$, then the function $f$ is not continuous at the point $a$, but the function $f$ is defined at the point $a$.

(3) If exists $\lim_{x \to a} f(x) = L$, but $f(a)$ is not defined, then the function $f$ is not continuous at the point $a$ and also the function $f$ is not defined at the point $a$.

**Remark 1.16** For infinite limits are used the following rules:

(1) $L + \infty = \infty$ for $L \neq -\infty$,

(2) $L \cdot \infty = \infty$ if $L > 0$,

(3) $L \cdot \infty = -\infty$ if $L < 0$,

(4) $\dfrac{L}{\infty} = 0$ for $L \neq \pm\infty$.

## 1.5 The Derivative

**Definition 1.25** Let a function $f\colon\ y = f(x)$ be given and defined in the neighbourhood of the point $x_0 \in \mathcal{D}(f)$. *Derivative* of the function $f$ at the point $x_0$ is the real number:

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}. \tag{1.1}$$

resp.

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}. \tag{1.2}$$

**Definition 1.26** We say that a function $f\colon\ y = f(x)$ has at the point $x_0 \in \mathbb{R}$ *left-hand derivation* if it is defined in the same left neighbourhood of the point $x_0 \in \mathcal{D}(f)$ and exists the limit:

$$f'_-(x_0) = \lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0} \tag{1.3}$$

resp.

$$f'_-(x_0) = \lim_{h \to 0^-} \frac{f(x_0 + h) - f(x_0)}{h} \tag{1.4}$$

We say that the function $f\colon\ y = f(x)$ has at the point $x_0 \in \mathbb{R}$ *right-hand derivation* if it is defined in the same right neighbourhood of the point $x_0 \in \mathcal{D}(f)$ and exists the limit:

$$f'_+(x_0) = \lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \tag{1.5}$$

resp.

$$f'_+(x_0) = \lim_{h \to 0^+} \frac{f(x_0 + h) - f(x_0)}{h} \tag{1.6}$$

**Remark 1.17** We also usually use the terms: *one-sided derivative*, *left derivative*, *right derivative*, *left-hand derivative*, *right-hand derivative*, *one-sided limit*.

**Theorem 1.18** If a function $f$ has derivative at a point $x_0$, then the function $f$ is continuous at this point.

**Theorem 1.19** Let a function $f\colon\ y = f(x)$ be given and the point $x_0 \in \mathcal{D}(f)$ be an internal point of the domain of the function $f$. The function $f$ has at the point $x_0$ derivative $f'(x_0)$ if and only if the function $f$ has at the point $x_0$ left derivative $f'_-(x_0)$, right derivative $f'_+(x_0)$ and holds equality: $f'_-(x_0) = f'_+(x_0)$.

**Theorem 1.20** Let a function $f\colon\ y = f(x)$ and a point $T = [x_0, y_0]$, where $x_0 \in \mathcal{D}(f)$, $y_0 \in \mathcal{H}(f)$ and $y_0 = f(x_0)$ be given. If there exists derivative of the function $f$ at the point $x_0$ $(f'(x_0))$, then the *tangent line* (*tangent*) $t$ to the graph of the function $f$ at the point $T$ has the equation:

$$t: \quad y - y_0 = f'(x_0) \cdot (x - x_0). \tag{1.7}$$

**Theorem 1.21** Let a function $f\colon\ y = f(x)$ and a point $T = [x_0, y_0]$, where $x_0 \in \mathcal{D}(f)$, $y_0 \in \mathcal{H}(f)$ and $y_0 = f(x_0)$ be given. If there exists derivative of the function $f$ at the point $x_0$ ($f'(x_0)$) and $f'(x_0) \neq 0$, then the *normal line* $n$ to the graph of the function $f$ at the point $T$ has the equation: [7]

$$n: \quad y - y_0 = \frac{-1}{f'(x_0)} \cdot (x - x_0). \tag{1.8}$$

**Theorem 1.22** (*Fundamental Differentiation Rules*) Let functions $f$: $y = f(x)$ and $g$: $y = g(x)$ have derivatives $f'(x_0)$ and $g'(x_0)$ at the point $x_0$ . Let $c \in \mathbb{R}$. Then holds:

(1) $(c \cdot f(x_0))' = c \cdot f'(x_0)$,

(2) $(f(x_0) + g(x_0))' = f'(x_0) + g'(x_0)$,

(3) $(f(x_0) - g(x_0))' = f'(x_0) - g'(x_0)$,

(4) $(f(x_0) \cdot g(x_0))' = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0)$,

(5) $\left(\dfrac{f(x_0)}{g(x_0)}\right)' = \dfrac{f'(x_0) \cdot g(x_0) - f(x_0) \cdot g'(x_0)}{g^2(x_0)}$.

**Theorem 1.23** (*The Derivative of a Composite Function* or *Chain Rule*) Let a composite function $h$: $y = f\Big(g(x)\Big)$ be defined on an interval $(a, b)$ and $x_0 \in (a, b)$. Let a function $g$ has derivation $g'(x_0)$ at the point $x_0$ and a function $f$ has derivation $f'(z_0)$ at the point $z_0 = g(x_0)$. Then the function $h$ has derivation $h'(x_0) = f'(z_0) \cdot g'(x_0)$ at the point $x_0$.

**Fundamental Formulas for Differentiation:**

(1) $(c)' = 0$, where $c$ is real constant $c \in \mathbb{R}$,

(2) $(x)' = 1$,

(3) $(x^n)' = n \cdot x^{n-1}$, for $n \in \mathbb{R}$,

[7] Normal line to a curve: The line perpendicular to the tangent line to the curve at the point $x_0$ of tangency is called the normal line to the curve at that point $x_0$. The slopes of perpendicular lines have product $-1$, so if the equation of the curve is $y = f(x)$ then the slope of the normal line is $\frac{-1}{f'(x_0)}$.

(4) $(\sin x)' = \cos x$, for $x \in \mathbb{R}$,

(5) $(\cos x)' = -\sin x$, for $x \in \mathbb{R}$,

(6) $(\operatorname{tg} x)' = \dfrac{1}{\cos^2 x}$, for $x \in \mathbb{R} - \{\frac{(2k+1)\pi}{2}; k \in \mathbb{Z}\}$,

(7) $(\operatorname{cotg} x)' = \dfrac{-1}{\sin^2 x}$, for $x \in \mathbb{R} - \{k\pi; k \in \mathbb{Z}\}$,

(8) $(\arcsin x)' = \dfrac{1}{\sqrt{1-x^2}}$, for $x \in (-1, 1)$,

(9) $(\arccos x)' = \dfrac{-1}{\sqrt{1-x^2}}$, for $x \in (-1, 1)$,

(10) $(\operatorname{arctg} x)' = \dfrac{1}{1+x^2}$, for $x \in \mathbb{R}$,

(11) $(\operatorname{arccotg} x)' = \dfrac{-1}{1+x^2}$, for $x \in \mathbb{R}$,

(12) $(\ln x)' = \dfrac{1}{x}$, for $x \in (0, \infty)$,

(13) $(\log_a x)' = \dfrac{1}{x \cdot \ln a}$, where $a > 0$ and $a \neq 1$, for $x \in (0, \infty)$,

(14) $(\mathrm{e}^x)' = \mathrm{e}^x$, for $x \in \mathbb{R}$,

(15) $(a^x)' = a^x \cdot \ln a$, where $a > 0$ and $a \neq 1$, for $x \in \mathbb{R}$.

We show how to make derivative of the function that has the form $y = f(x)^{g(x)}$, where $f(x) > 0$ for all $x \in \mathcal{D}(f)$.

$$y' = \left(f(x)^{g(x)}\right)' = \left(\mathrm{e}^{\ln\left(f(x)^{g(x)}\right)}\right)' = \left(\mathrm{e}^{g(x)\cdot\ln f(x)}\right)' = \left(\mathrm{e}^{g(x)\cdot\ln f(x)}\right)\cdot\left[g(x)\cdot\ln f(x)\right]' =$$

$$= \left(f(x)^{g(x)}\right) \cdot \left[g'(x) \cdot \ln f(x) + g(x) \cdot (\ln f(x))'\right] =$$

$$= \left(f(x)^{g(x)}\right) \cdot \left[g'(x) \cdot \ln f(x) + g(x) \cdot \frac{1}{f(x)} \cdot f'(x)\right].$$

We got additional differentiation formula of the form:

$$\left(f(x)^{g(x)}\right)' = f(x)^{g(x)} \cdot \left[g'(x) \cdot \ln f(x) + \frac{g(x)}{f(x)} \cdot f'(x)\right]. \tag{1.9}$$

**Theorem 1.24** Let $f\colon\ y = f(x)$ be given and $\langle a, b\rangle \subseteq \mathcal{D}(f)$. Suppose $f$ has in an internal point $c$ of the interval $\langle a, b\rangle$ the greatest or the smallest value, respectively. If the function $f$ has the first derivative at the point $c$, then $f'(c) = 0$.

**Theorem 1.25** (*Rolle's Theorem*) Let a function $f\colon\ y = f(x)$ be continuous on a closed interval $\langle a, b\rangle$ and has the first derivative ($f$ be differentiable) on the open interval $(a, b)$, and $f(a) = f(b)$. Then there exists at least one point $\xi$ in the open interval $(a, b)$ for which $f'(\xi) = 0$.

**Remark 1.18** The tangent to a graph of the function $f$, where the derivative is "hidden", is parallel to the $x$-axis, and also with the straight line joining the two "end" points $(a, f(a))$ and $(b, f(b))$ on a curve on the graph of the function $f$. Thus Rolle's theorem claims the existence of a point at which the tangent to the graph is parallel to the given horizontal straight line.

**Theorem 1.26** (*Mean Value Theorem, Lagrange's Theorem*) Let a function $f\colon\ y = f(x)$ be continuous on a closed interval $\langle a, b\rangle$ and has the first derivative ($f$ be differentiable) on the open interval $(a, b)$. Then there exists at least one point $\xi$ in the open interval $(a, b)$ for which

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

**Remark 1.19** The Mean Value Theorem (MVT) claims the existence of a point at which the tangent is parallel to the straight line joining $(a, f(a))$ and $(b, f(b))$. Rolle's theorem is clearly a particular case of the MVT in which $f$ satisfies an additional condition: $f(a) = f(b)$.

**Theorem 1.27** (*Cauchy's Theorem*) Let functions $f\colon\ y = f(x)$ and $g\colon\ y = g(x)$ be continuous on a closed interval $\langle a, b\rangle$ and have the first derivative ($f$, $g$ be differentiable) on an open interval $(a, b)$. Suppose $g'(x) \neq 0$ for $\forall x \in (a, b)$. Then there exists at least one point $\xi$ in the open interval $(a, b)$, such that

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

**Theorem 1.28** (*L'Hospital's rule*) Let functions $f\colon\ y = f(x)$ and $g\colon\ y = g(x)$ have the derivatives over the neighbourhood of the point $a \in \mathbb{R} \cup \{\pm\infty\}$. Suppose $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = 0$ or $\lim\limits_{x\to a} |g(x)| = +\infty$. If $\lim\limits_{x\to a} \frac{f'(x)}{g'(x)}$ has a finite value or if the limit is $\pm\infty$, then exists also the limit $\lim\limits_{x\to a} \frac{f(x)}{g(x)}$ and holds:

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

**Remark 1.20** L'Hospital's rule is also called Bernoulli's rule, uses derivatives to help evaluate limits involving the indeterminate forms. L'Hospital's rule tells us that if we have an indeterminate form $\frac{0}{0}$ or $\frac{\pm\infty}{\pm\infty}$ all we need to do is differentiate the numerator and differentiate the denominator and then take the limit.

**Theorem 1.29** Let a function $f\colon\ y = f(x)$ be continuous on an interval $I$ and has a derivative at all internal points of the interval $I$. Then:

(1) If the function $f$ is on the interval $I$ non-decreasing, then $f'(x) \geqq 0$ for each inner point of the interval $I$.

(2) If the function $f$ is on the interval $I$ non-increasing, then $f'(x) \leqq 0$ for each inner point of the interval $I$.

(3) If the function $f$ is on the interval $I$ increasing, then $f'(x) > 0$ for each inner point of the interval $I$ and $f'$ is nonzero on each open subinterval of the interval $I$.

(4) If the function $f$ is on the interval $I$ decreasing, then $f'(x) < 0$ for each inner point of the interval $I$ and $f'$ is nonzero on each open subinterval of the interval $I$.

**Theorem 1.30** Let a function $f\colon\ y = f(x)$ be continuous on an interval $I$ and has a derivative at all internal points of an interval $I$. Then:

(1) if $f'(x) > 0$ for each inner point of the interval $I$, then the function $f$ is increasing on the interval $I$,

(2) if $f'(x) < 0$ for each inner point of the interval $I$, then the function $f$ is decreasing on the interval $I$,

(3) if $f'(x) \geqq 0$ for each inner point of the interval $I$, then the function $f$ is non-decreasing on the interval $I$,

(4) if $f'(x) \leqq 0$ for each inner point of the interval $I$, then the function $f$ is non-increasing on the interval $I$.

**Definition 1.27** We say that a function $f\colon\ y = f(x)$ has in the inner point $x_0 \in I$, where $I \subseteq \mathcal{D}(f)$ the *local maximum*, if there exists a neighbourhood of the point $x_0$, such that for all points of given neighbourhood the following applies: $f(x) \leqq f(x_0)$.

We say that a function $f\colon\ y = f(x)$ has in the inner point $x_0 \in I$, where $I \subseteq \mathcal{D}(f)$ the *local minimum*, if there exists a neighbourhood of the point $x_0$, such that for all points of given neighbourhood the following applies: $f(x) \geqq f(x_0)$.

We say that a function $f\colon\ y = f(x)$ has in the inner point $x_0 \in I$, where $I \subseteq \mathcal{D}(f)$ the *strictly local maximum*, if there exists a neighbourhood of the point $x_0$, such that for all points of given neighbourhood the following applies: $f(x) < f(x_0)$.

We say that a function $f\colon\ y = f(x)$ has in the inner point $x_0 \in I$, where $I \subseteq \mathcal{D}(f)$ the *strictly local minimum*, if there exists a neighbourhood of the point $x_0$, such that for all points of given neighbourhood the following applies: $f(x) > f(x_0)$.

A function $f$ has an *absolute maximum* at a point $x_0 \in \mathcal{D}(f)$, if $f(x_0) \geqq f(x)$ for all $x$ in the domain of the function $f$.

A function $f$ has an *absolute minimum* at a point $x_0 \in \mathcal{D}(f)$, if $f(x_0) \leqq f(x)$ for all $x$ in the domain of the function $f$.

Together, the absolute minimum and the absolute maximum are known as the *absolute extrema* of the function $f$.

The point $x_0$ is called the *stationary point* or the *turning point* of the function $f\colon\ y = f(x)$, if there exists $f'(x_0)$ and holds: $f'(x_0) = 0$.[8]

**Theorem 1.31** Let $f'(x_0)$ exists. If the function $f$ has local extrema at the point $x_0$, then $f'(x_0) = 0$.[9]

[8]Maxima and minima are points where a function reaches the highest or the lowest value, respectively. There are two kinds of extrema: global and local (absolute and relative), respectively.

[9]Condition $f'(x_0) = 0$ is only a necessary condition for the existence of local extremum. From this condition does not automatically follow that the function $f$ has at the point $x_0$

**Theorem 1.32** (*Extreme Value Theorem*) If a function $f$ is a continuous function on the closed interval $\langle a, b\rangle$, then the function $f$ attains both an absolute maximum and an absolute minimum on the closed interval $\langle a, b\rangle$.

**Definition 1.28** The function $f\colon\ y = f(x)$ is called *convex* (*concave up*) on the interval $I \subseteq \mathcal{D}(f)$, if for each set of three points $x_1, x_2, x_3 \in I$, such that $x_1 < x_2 < x_3$, the point $[x_2, f(x_2)]$ lies below or on the straight line, which is determined by the points $[x_1, f(x_1)]$ and $[x_3, f(x_3)]$.

The function $f\colon\ y = f(x)$ is called *concave* (*concave down*) on the interval $I \subseteq \mathcal{D}(f)$, if for each set of three points $x_1, x_2, x_3 \in I$, such that $x_1 < x_2 < x_3$, the point $[x_2, f(x_2)]$ lies above or on the straight line, which is determined by the points $[x_1, f(x_1)]$ and $[x_3, f(x_3)]$.

The function $f\colon\ y = f(x)$ is called *strictly convex* (*purely concave up*) on the interval $I \subseteq \mathcal{D}(f)$, if for each set of three points $x_1, x_2, x_3 \in I$, such that $x_1 < x_2 < x_3$, the point $[x_2, f(x_2)]$ lies below the straight line, which is determined by the points $[x_1, f(x_1)]$ and $[x_3, f(x_3)]$.

The function $f\colon\ y = f(x)$ is called *strictly concave* (*purely concave down*) on the interval $I \subseteq \mathcal{D}(f)$, if for each set of three points $x_1, x_2, x_3 \in I$, such that $x_1 < x_2 < x_3$, the point $[x_2, f(x_2)]$ lies above the straight line, which is determined by the points $[x_1, f(x_1)]$ and $[x_3, f(x_3)]$.

**Theorem 1.33** Suppose the function $f\colon\ y = f(x)$ has derivative $f'$ in all internal points of the interval $I \subseteq \mathcal{D}(f)$. If for each pair of points $x_0, x_1 \in I$, such that $x_0 \neq x_1$, the point $[x_1, f(x_1)]$ lies above the tangent line to the graph of the function $f$ at the point $T = [x_0, f(x_0)]$, then the function $f$ is strictly convex.

**Theorem 1.34** Suppose the function $f\colon\ y = f(x)$ has derivative $f'$ in all internal points of the interval $I \subseteq \mathcal{D}(f)$. If for each pair of points $x_0, x_1 \in I$, such that $x_0 \neq x_1$, the point $[x_1, f(x_1)]$ lies below the tangent line to the graph of the function $f$ at the point $T = [x_0, f(x_0)]$, then the function $f$ is strictly concave.

**Theorem 1.35** Let a function $f\colon\ y = f(x)$ be continuous on an interval $I$ and the second derivative be defined in all internal points of the interval $I$. Then:

---

local extremum.

The function $f$ can have a local extrema at the points which are not stationary points of the function $f$, i. e. even at the points where the function $f$ has no derivative.

(1) if $f''(x) > 0$ for each inner point of the interval $I$, then the function $f$ is strictly convex on the interval $I$,

(2) if $f''(x) < 0$ for each inner point of the interval $I$, then the function $f$ is strictly concave on the interval $I$,

(3) if $f''(x) \geqq 0$ for each inner point of the interval $I$, then the function $f$ is convex on the interval $I$,

(4) if $f''(x) \leqq 0$ for each inner point of the interval $I$, then the function $f$ is concave on the interval $I$.

**Definition 1.29** Let a function $f\colon\ y = f(x)$ be continuous on an interval $I \subseteq \mathcal{D}(f)$. A point $x_0 \in I$ is called the *inflection point* (the *point of inflection*, the*inflexion*) of the function $f$, if the function $f$ is strictly concave (strictly convex) in some left neighbourhood of the point $x_0$ and is strictly convex (strictly concave) in some right neighbourhood of the point $x_0$.

**Theorem 1.36** Suppose $f''(x_0)$ is defined. If $x_0$ is the point of inflection of the function $f$, then holds $f''(x_0) = 0$. [10]

**Theorem 1.37** Let $f''(x_0) = 0$ and $f'''(x_0) \neq 0$, then the point $x_0$ is the point of inflection of the function $f$.

**Theorem 1.38** Suppose that a function $f\colon\ y = f(x)$ has the nonzero $n$-th derivative $f^{(n)}(x_0) \neq 0$, for $n \geqq 2$ at an inner point $x_0$ of an interval $I \subseteq \mathcal{D}(f)$ and $f'(x_0) = f''(x_0) = \cdots = f^{(n-2)}(x_0) = f^{(n-1)}(x_0) = 0$. Then holds:

(1) If $n$ is an even number and $f^{(n)}(x_0) > 0$, then the function $f$ has the strictly local minimum at the point $x_0$.

(2) If $n$ is an even number and $f^{(n)}(x_0) < 0$, then the function $f$ has the strictly local maximum at the point $x_0$.

(3) If $n$ is an odd number, then the function $f$ has a point of inflexion $x_0$.

[10]Condition $f''(x_0) = 0$ is a necessary condition for the existence of an inflection point, because from this condition does not automatically follows that the point $x_0$ is the inflection point of the function $f$.

**Remark 1.21** A global maximum of a function is a point in which the function takes the largest value on the entire range of the function, while a global minimum is the point in which the function takes the smallest value on the range of the function. On the other hand, local extrema are the largest or smallest values of the function in the small neighbourhood of extrema.

A global extremum is always a local extremum too. It is also possible to have a function with no extrema.

For each extremum, the slope of the graph is necessarily zero. The graph must stop rising or decreasing at an extremum, and begin to continue in the opposite direction. Because of this, extrema are also commonly called stationary points or turning points. Therefore, the first derivative of a function is equal to zero at extrema.

However, a slope of zero ($f'(x) = 0$) does not guarantee existence of a maximum or a minimum. It could be a stationary point which is called a point of inflection (inflection point).

Method to classify a stationary point is called the extremum test, or the $2^{nd}$ *derivative test*. The second derivative of the function tells us the rate of change of the first derivative. If the second derivative is positive at the stationary point, then the gradient is increasing and it is a minimum. Conversely, if the second derivative is negative at that point, then it is a maximum.

If the second derivative is zero, we have a problem. It could be a point of inflexion, or it could still be an extremum. We differentiate a function $f$ until we get, at the $(n+1)$-st derivative a non-zero result at the stationary point. If $n$ is odd, then the stationary point is a true extremum. If the $(n+1)$-st derivative is positive, it is a minimum; if the $(n+1)$-st derivative is negative, it is a maximum. If $n$ is even, then the stationary point is the point of inflection.

Critical points are the points where a function's derivative is 0 or not defined, or the endpoints on given closed interval of a function $f$ that is continuous on that interval.

## 1.6 Solved Examples

**Example 1.1** Find the domain of the function $f$ defined by

$$f : y = \frac{x+3}{x^2-9} + \ln(x+4) - \sqrt{16-x^2}, \; x \in \mathbb{R}.$$

*Solution:*
For this function we have three conditions. In the first summand the denominator must not be zero. In the second summand must be the argument of a logarithmic function positive and the third summand must have nonnegative expression under the square root. Formally, it is written as follows: $x^2 - 9 \neq 0 \wedge x + 4 > 0 \wedge 16 - x^2 \geqq 0$. Resolving these inequalities we get the following sets: $x \in \mathbb{R} - \{-3, 3\} \wedge x \in (-4, \infty) \wedge x \in \langle -4, 4 \rangle$. Intersection of these sets is the resulting domain of the function $f$ i. e. $\mathcal{D}(f) = (-4, -3) \cup (-3, 3) \cup (3, 4\rangle$. See Figure 1.14.

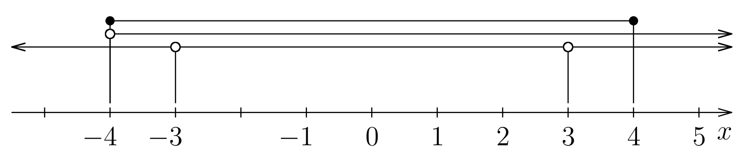


Figure 1.14: Domain of the function:$f : y = \frac{x+3}{x^2-9} + \ln(x+4) - \sqrt{16-x^2}$.

√

**Example 1.2** Find the domain of the function:

$$f : y = \frac{x+3}{(x^2-9) \cdot \sqrt{16-x^2}} + \arcsin\left(\frac{x+2}{6}\right), \; x \in \mathbb{R}.$$

*Solution:*
We will write all conditions that must be satisfied for the function $f$ to be defined. $x^2 - 9 \neq 0 \wedge -1 \leqq \frac{x+2}{6} \leqq 1 \wedge 16 - x^2 \geqq 0 \wedge 16 - x^2 \neq 0$. We rewrite these conditions to the following form: $x^2 - 9 \neq 0 \wedge -6 \leqq x + 2 \wedge x + 2 \leqq 6 \wedge 16 - x^2 > 0$. Resolving these inequalities we get the following sets: $x \in \mathbb{R} - \{-3, 3\} \wedge x \in \langle -8, \infty) \wedge x \in (-\infty, 4\rangle \wedge x \in (-4, 4)$. Intersection of these sets is the resulting domain of the function $f$, i. e. $\mathcal{D}(f) = (-4, -3) \cup (-3, 3) \cup (3, 4)$. See Figure 1.15. √

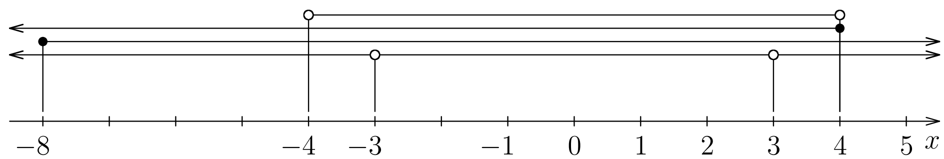


Figure 1.15: Domain of the function: $f : y = \frac{x+3}{(x^2-9)\cdot\sqrt{16-x^2}} + \arcsin\left(\frac{x+2}{6}\right)$.

**Example 1.3** Calculate the derivative of the given functions:

$$f_1 : y = x^3 - 2x^2 + \frac{x}{5} + e^3 - \frac{1}{x^4} + \frac{2}{5x}$$
$$f_2 : y = \sqrt{x} + \sqrt[5]{x} - \frac{4}{\sqrt{x}} + \frac{x^2}{\sqrt[3]{x}} - \frac{1}{e^x}$$
$$f_3 : y = x^2 \cdot (3 - 2 \cdot \ln x) + \frac{2-3x}{x-1}$$
$$f_4 : y = \ln(e - e^{2x})$$
$$f_5 : y = \arccos\left(\sqrt{1-x^2}\right)$$

*Solution:*
We successively calculate the derivative of the each function:

$$f_1' : y' = \left(x^3 - 2x^2 + \frac{x}{5} + e^3 - \frac{1}{x^4} + \frac{2}{5x}\right)' =$$
$$= (x^3)' - 2(x^2)' + \frac{1}{5} \cdot x' + e^3 \cdot (1)' - (x^{-4})' + \frac{2}{5} \cdot (x^{-1})' =$$
$$= 3x^2 - 4x + \frac{1}{5} + e^3 \cdot 0 - (-4)x^{-5} + \frac{2}{5} \cdot (-1)x^{-2} =$$
$$= 3x^2 - 4x + \frac{1}{5} + \frac{4}{x^5} - \frac{2}{5x^2}$$

---

$$f_2' : y' = \left(\sqrt{x} + \sqrt[5]{x} - \frac{4}{\sqrt{x}} + \frac{x^2}{\sqrt[3]{x}} - \frac{1}{e^x}\right)' =$$
$$= \left(\sqrt{x}\right)' + \left(\sqrt[5]{x}\right)' - \left(\frac{4}{\sqrt{x}}\right)' + \left(\frac{x^2}{\sqrt[3]{x}}\right)' - \left(\frac{1}{e^x}\right)' =$$
$$= \left(x^{\frac{1}{2}}\right)' + \left(x^{\frac{1}{5}}\right)' - \left(4x^{-\frac{1}{2}}\right)' + \left(x^2 \cdot x^{-\frac{1}{3}}\right)' - \left(e^{-x}\right)' =$$

$$= \frac{1}{2}x^{\frac{1}{2}-1} + \frac{1}{5}x^{\frac{1}{5}-1} - 4\left(-\frac{1}{2}\right)x^{-\frac{1}{2}-1} + \left(x^{\frac{5}{3}}\right)' - \mathrm{e}^{-x}\,(-x)' =$$
$$= \frac{1}{2}x^{-\frac{1}{2}} + \frac{1}{5}x^{-\frac{4}{5}} + 2x^{-\frac{3}{2}} + \frac{5}{3}x^{\frac{2}{3}} - \mathrm{e}^{-x}(-1) =$$
$$= \frac{1}{2\sqrt{x}} + \frac{1}{5\sqrt[5]{x^4}} + \frac{2}{x\sqrt{x}} + \frac{5}{3}\sqrt[3]{x^2} + \frac{1}{\mathrm{e}^x}$$

---

The function $f$ is the sum of two functions $g$ and $h$, where $f : y = g + h$. The function $g$ is the product of two functions $x^2$ and $(3 - 2 \cdot \ln x)$, and the function $h$ is the ratio of two functions $2 - 3x$ and $x - 1$. We use the product rule (4) in Theorem 1.22 for $g$ and the quotient rule (5) in Theorem 1.22 for $h$ to differentiate $f$ as follows:

$$f_3' : y' = \left(x^2 \cdot (3 - 2 \cdot \ln x) + \frac{2-3x}{x-1}\right)' =$$
$$= \left(x^2 \cdot (3 - 2 \cdot \ln x)\right)' + \left(\frac{2-3x}{x-1}\right)' =$$
$$= (x^2)' \cdot (3 - 2 \cdot \ln x) + x^2 \cdot (3 - 2 \cdot \ln x)' + \frac{(2-3x)'(x-1) - (2-3x)(x-1)'}{(x-1)^2} =$$
$$= 2x \cdot (3 - 2 \cdot \ln x) + x^2 \cdot \left(0 - 2 \cdot \frac{1}{x}\right) + \frac{(-3)(x-1) - (2-3x)(1)}{(x-1)^2}$$

We expand and group the results to obtain $f'(x)$ as follows:

$$f_3' : y' = 4x - 4x \cdot \ln x + \frac{1}{(x-1)^2}$$

---

$$f_4' : y' = \left(\ln(\mathrm{e} - \mathrm{e}^{2x})\right)' = \frac{1}{(\mathrm{e} - \mathrm{e}^{2x})} \cdot (\mathrm{e} - \mathrm{e}^{2x})' = \frac{1}{(\mathrm{e} - \mathrm{e}^{2x})} \cdot (0 - \mathrm{e}^{2x}(2x)') =$$
$$= \frac{-2\mathrm{e}^{2x}}{(\mathrm{e} - \mathrm{e}^{2x})}$$

---

$$f_5' : y' = \left[\arccos\left(\sqrt{1-x^2}\right)\right]' = \frac{-1}{\sqrt{1 - \left(\sqrt{1-x^2}\right)^2}} \cdot \left(\sqrt{1-x^2}\right)' =$$
$$= \frac{-1}{\sqrt{1 - (1-x^2)}} \cdot \left((1-x^2)^{\frac{1}{2}}\right)' = \frac{-1}{\sqrt{x^2}} \cdot \frac{1}{2} \cdot (1-x^2)^{\frac{1}{2}-1}(1-x^2)' =$$
$$= \frac{-1}{2|x|} \cdot \frac{1}{\sqrt{1-x^2}}(0 - 2x) = \frac{\pm 1}{\sqrt{1-x^2}}$$

$\surd$

**Example 1.4** Calculate the first and the second derivative of the function $f$ given by

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

*Solution:*
We use the quotient rule (5) in Theorem 1.22 on the page 51 to differentiate $f$ as follows:

$$f'(x) = \left(\frac{(2x-3)^2}{(1-2x)^3}\right)' = \frac{((2x-3)^2)' \cdot (1-2x)^3 - (2x-3)^2 \cdot ((1-2x)^3)'}{((1-2x)^3)^2} =$$

$$= \frac{2 \cdot (2x-3)^1 \cdot (2x-3)' \cdot (1-2x)^3 - (2x-3)^2 \cdot 3 \cdot (1-2x)^2 \cdot (1-2x)'}{(1-2x)^6} =$$

$$= \frac{4 \cdot (2x-3) \cdot (1-2x)^3 - (2x-3)^2 \cdot (-6) \cdot (1-2x)^2}{(1-2x)^6} =$$

$$= \frac{(1-2x)^2 \cdot [4 \cdot (2x-3) \cdot (1-2x) + 6 \cdot (2x-3)^2]}{(1-2x)^6} =$$

$$= \frac{4 \cdot (2x - 4x^2 - 3 + 6x) + 6 \cdot (4x^2 - 12x + 9)}{(1-2x)^4} = = \frac{8x^2 - 40x + 42}{(1-2x)^4}.$$

$$f''(x) = \left(\frac{8x^2 - 40x + 42}{(1-2x)^4}\right)' = \frac{32x^2 - 224x + 296}{(1-2x)^5}. \qquad \surd$$

**Example 1.5** Find the derivative of a function $f$ given by

$$f : y = x^{\ln x}.$$

*Solution:*
We use the formula (1.9) on the page 53.

$$f' : y' = \left(x^{\ln x}\right)' = \left(x^{\ln x}\right) \cdot \left[(\ln x)' \cdot \ln x + \frac{\ln x}{x}(x)'\right] =$$

$$= \left(x^{\ln x}\right) \cdot \left[\frac{1}{x} \cdot \ln x + \frac{\ln x}{x} \cdot 1\right] = \left(x^{\ln x}\right) \cdot \left[\frac{2 \ln x}{x}\right].$$

$\surd$

## 1.7 Unsolved Tasks

**1.1** Find the domain of the following functions:

a) $f : y = \dfrac{3-x}{4x^2-1}$

b) $f : y = \dfrac{(5x+1)\cdot \sin x}{2x^3+3x}$

c) $f : y = \sqrt{x+x^2} - \sqrt{4-x^2} + \dfrac{\mathrm{e}^x - 2x}{1-\sqrt{1-\frac{x}{2}}}$

d) $f : y = \ln\left(\dfrac{x-1}{1+x}\right)$

e) $f : y = \dfrac{x^2-5x+6}{\sqrt{1+\ln x}-1}$

f) $f : y = \ln\left(\dfrac{5x}{8x^2+1}\right) - \ln(1-x)$

g) $f : y = \sqrt{1-\dfrac{1}{x}} - \sqrt{\dfrac{1}{x}+1} - \arccos\dfrac{10x}{16+x^2}$

h) $f : y = \dfrac{\ln(1-x)\cdot \ln(x+1)}{3x^2+2}$

i) $f : y = \dfrac{\ln[\ln(\ln x)]}{3\mathrm{e}^{3x}+2\mathrm{e}^{2x}+\mathrm{e}^x+1}$

j) $f : y = \dfrac{\mathrm{e}^x \cdot \arccos(1+x)}{x+\sqrt{3+2x}}$

k) $f : y = \ln(3x-6)\cdot\sqrt{4x-x^2}$

l) $f : y = \dfrac{\sqrt{(\ln x)\cdot \arccos(6x-5)}}{1+x^2}$

m) $f : y = \sqrt{\dfrac{x^2-1}{2+3x+x^2}}$

n) $f : y = \dfrac{\sqrt{\log_{\frac{1}{3}}\left(x^2 - 2x + 1\right)}}{8 + x^3}$

**1.2** Find the first derivative of each of the following functions.

a) $f : y = 3x^5 - \dfrac{x^4}{2} + 7x - 6 + 2x^3 \cdot \ln 2 - \cos 1$

b) $f : y = \dfrac{5}{2x^3} - \dfrac{\sqrt{3}}{x} + \dfrac{1}{\sqrt[3]{8x}} - 4\sqrt{x}$

c) $f : y = \left(1 - \sqrt{x}\right) \cdot (1 + x)$

d) $f : y = \dfrac{1 - x^2}{\sqrt{x}}$

e) $f : y = \dfrac{x^5}{5} \cdot \left(\ln x - \dfrac{1}{5}\right) - \dfrac{(x-2)^2}{x}$

f) $f : y = \mathrm{e}^x \cdot \left(x^3 - 3x^2 + 6x - 6\right)$

g) $f : y = \dfrac{x - \cos x \cdot \sin x}{2}$

h) $f : y = \dfrac{x}{2} - \dfrac{1 + x^2}{2} \cdot \operatorname{arctg} x$

i) $f : y = \dfrac{\mathrm{e}^x + 1}{\mathrm{e}^x - 1}$

j) $f : y = \mathrm{e}^x \cdot x^3 \cdot \cos x$

k) $f : y = \ln\left(\dfrac{2 + x}{2 - x}\right)$

l) $f : y = (2x^3 - 4)^5$

m) $f : y = \operatorname{arctg} \dfrac{x - 1}{x + 1} - \operatorname{arctg} \dfrac{1}{x}$

n) $f : y = -\dfrac{1}{\sqrt{2}} \cdot \arcsin \dfrac{\sqrt{2} \cdot x}{1 + x^2}$

**1.3** Find the second derivative of each of the following functions.

a) $f : y = 4x^3 - x^4$

b) $f : y = \dfrac{x^2 - 2\ln(x-1)}{2}$

c) $f : y = 2x + \ln(\cos x)$

d) $f : y = \dfrac{x^2}{x-1}$

e) $f : y = \dfrac{7+x^2}{3+x^2}$

f) $f : y = \dfrac{x^2+1}{1-x^2}$

g) $f : y = x^2 \cdot \mathrm{e}^{-x}$

h) $f : y = x + \operatorname{arctg} x$

**1.4** Write the equation of the tangent $t$ to the function $f$ at the given point $T = [x_0, y_0]$:

a) $f : y = \dfrac{x^2}{x-1}$, ak $x_0 = 3$

b) $f : y = \operatorname{arctg} x$, ak $x_0 = -1$

c) $f : y = \dfrac{\ln x}{x}$, ak $x_0 = \mathrm{e}$

d) $f : y = \dfrac{2x+1}{x^2}$, ak $x_0 = -2$

e) $f : y = \sqrt{1-x^2}$, ak $x_0 = -\frac{\sqrt{2}}{2}$

Find all points on the graph of $y = x3 - 3x$ where the tangent line is parallel to the $x$ axis (or the horizontal tangent line).

Find $a$ and $b$ so that the line $y = -3x + 4$ is tangent to the graph of $y = ax^3 + bx$ at $x = 1$.

## 1.8 Results of Unsolved Tasks

**1.1** **a)** $\mathcal{D}(f) = (-\infty, -2) \cup (-2, 2) \cup (2, \infty)$ **b)** $\mathcal{D}(f) = \mathbb{R} - \{0\}$ **c)** $\mathcal{D}(f) = (-2, -1\rangle \cup (0, 2\rangle$ **d)** $\mathcal{D}(f) = (-\infty, -1) \cup (1, \infty)$ **e)** $\mathcal{D}(f) = \langle \frac{1}{\mathrm{e}}, 1) \cup (1, \infty)$ **f)** $\mathcal{D}(f) = (0, 1)$ **g)** $\mathcal{D}(f) = (-\infty, -8\rangle \cup \langle -2, -1\rangle \cup \langle 1, 2\rangle \cup \langle 8, \infty)$ **h)** $\mathcal{D}(f) = (-1, 1)$ **i)** $\mathcal{D}(f) = (\mathrm{e}, \infty)$ **j)** $\mathcal{D}(f) = \langle -\frac{3}{2}, -1) \cup (-1, 0)$ **k)** $\mathcal{D}(f) = (2, 4\rangle$ **l)** $\mathcal{D}(f) = \langle -\frac{1}{2}, 0\rangle$ **m)** $\mathcal{D}(f) = (-\infty, -2) \cup (1, \infty)$ **n)** $\mathcal{D}(f) = \langle 0, 1) \cup (1, 2\rangle$

**1.2** **a)** $y' = 15x^4 - 2x^3 + 6x^2 \cdot \ln 2 + 7$ **b)** $y' = -\frac{15}{2x^4} + \frac{\sqrt{3}}{x^2} - \frac{1}{6x\sqrt[3]{x}} - \frac{2}{\sqrt{x}}$ **c)** $y' = 1 - \frac{3}{2}\sqrt{x} - \frac{1}{2\sqrt{x}}$ **d)** $y' = -\frac{1}{2x\sqrt{x}} - \frac{3}{2}\sqrt{x}$ **e)** $y' = x^4 \cdot \ln x - 1 + \frac{4}{x^2}$ **f)** $y' = x^3 \cdot \mathrm{e}^x$ **g)** $y' = \sin^2 x$ **h)** $y' = -x \cdot \operatorname{arctg} x$ **i)** $y' = \frac{-2\mathrm{e}^x}{(\mathrm{e}^x - 1)^2}$ **j)** $y' = x^2 \cdot \mathrm{e}^x \cdot (x \cdot \cos x + 3\cos x - x \sin x)$ **k)** $y' = \frac{4}{4 - x^2}$ **l)** $y' = 30x^2(2x^3 - 4)^4$ **m)** $y' = \frac{2}{1 + x^2}$ **n)** $y' = \frac{x^4 - 1}{\sqrt{x^4 + 1}}$

**1.3** **a)** $y'' = 12x \cdot (2 - x)$ **b)** $y'' = \frac{(x-1)^2 - 1}{(x-1)^2}$ **c)** $y'' = \frac{-1}{(\cos x)^2}$ **d)** $y'' = \frac{2}{(x-1)^3}$ **e)** $y'' = 24 \cdot \frac{x^2 - 1}{(3 + x^2)^3}$ **f)** $y'' = 4 \cdot \frac{3x^2 + 1}{(1 - x^2)^3}$ **g)** $y'' = (2 - 4x + x^2) \cdot \mathrm{e}^{-x}$ **h)** $y'' = \frac{-2x}{(x^2 + 1)^2}$

**1.4** **a)** $T = [3, \frac{9}{2}]$, $t : 3x - 4y + 9 = 0$ **b)** $T = [-1, \frac{\pi}{4}]$, $t : 2x - 2y + 2 - \pi = 0$ **c)** $T = [\mathrm{e}, \frac{1}{\mathrm{e}}]$, $t : \mathrm{e} \cdot y - 1 = 0$ **d)** $T = [-2, -\frac{3}{4}]$, $t : x + 4y + 5 = 0$ **e)** $T = [-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}]$, $t : x - y + \sqrt{2} = 0$

# Chapter 2

# Approximate Solution of a Nonlinear Equation

In this chapter we will show some basic numerical methods for solving of the equation of one real variable $f(x) = 0$. We show whether the methods converges to the solution always, or only under certain conditions. We will insinuate the speed of convergence of these methods.

## 2.1 Separation of Roots

Let be given a non-linear equation $f(x) = 0$. We are trying to find such points $c \in \mathbb{R}$, for which applies $f(c) = 0$. These points $c$ are called the roots of the equation $f(x) = 0$. When solving the equation $f(x) = 0$ we try to determine how many roots does this equation have and we are determining intervals in which there is exactly one root of the equation. The process of finding these intervals is called the *separation of the roots* of the equation $f(x) = 0$. We say that the roots of the equation $f(x) = 0$ are separated, if holds:

(1) $\mathcal{D}(f) = \langle a_1, b_1\rangle \cup \langle a_2, b_2\rangle \cup \cdots \cup \langle a_{n-1}, b_{n-1}\rangle \cup \langle a_n, b_n\rangle$, [1]

(2) $(a_i, b_i) \cap (a_j, b_j) = \emptyset$, for $i, j \in \{1, 2, \ldots, n\}$, $i \neq j$

(3) Each interval $\langle a_i, b_i\rangle$, $i \in \{1, 2, \ldots, n\}$ contains not more than one root of the equation $f(x) = 0$.

[1] We assume that $b_i = a_{i+1}$ for $i \in \{1, 2, \ldots, n-1\}$ and $a_i \leqq a_j$ for $i < j$.

Once we have roots separated, we approximate them by one of the approximations methods described below. For finding the roots of equations is useful following theorem.

**Theorem 2.1** Let a function $f\colon\ y = f(x)$ be given and let $\langle a, b\rangle \subseteq \mathcal{D}(f)$. If the function $f$ is continuous on an interval $\langle a, b\rangle$ and holds:

$$f(a) \cdot f(b) < 0, \tag{2.1}$$

then on the interval $\langle a, b\rangle$ is situated at least one root of the equation $f(x) = 0$.

**Remark 2.1** Condition (2.1) in the Theorem 2.1 means that the signs of functional values at points $a$ and $b$ are opposite. On a given interval $\langle a, b\rangle$ can also be more than one root. But if the condition (2.1) in the Theorem 2.1 is not satisfied, then the interval $\langle a, b\rangle$ may contain roots of the equation $f(x) = 0$.[2]

To be able to separate the roots of the equation, it is appropriate to examine the characteristics and properties of functions $f$ and based on these characteristics determine the intervals of separation. For easier to find the roots we can use the basic properties of the elementary functions, therefore it is appropriate to modify the equation $f(x) = 0$ to the form $g(x) = h(x)$, so that the graphs of the functions $g$ and $h$ are easier to draw. We draw both graphs of functions $g$ and $h$ into the same coordinate system. Searched roots of functions $f$ are the points where the graphs of functions $g$ and $h$ intersect.

## 2.2 Bisection Method

The *Bisection method* (or *the method of the half-partitioning of interval*) is the simplest numerical method for solving non-linear equations. In this section we describe the algorithm how to find an approximate solution of non-linear equations using the bisection method.

Let a continuous function $f\colon\ y = f(x)$ be given on an interval $\langle a, b\rangle$, which is the interval of separation of the equation $f(x) = 0$ and it contains one root $c$ exactly. Our task is to find the root $c$ or find such approximate $c_k$, which is sufficiently close to the root $c$.

[2] For example, the equation $x^2 = 0$ has the root $c = 0$, but at any interval $\langle a, b\rangle$ condition (2.1) can not be satisfied.

We suppose that on the interval $\langle a,b\rangle$ holds: $f(a)\cdot f(b)<0$. We denote this initial interval as the interval $\langle a_0,b_0\rangle$. We divide this interval $\langle a_0,b_0\rangle$ in a half. We denote its centre as $c_0$. For the middle point from the interval $\langle a_0,b_0\rangle$ holds: $c_0=\frac{a_0+b_0}{2}$. We created two new intervals $\langle a_0,c_0\rangle$ and $\langle c_0,b_0\rangle$. From these two intervals we choose one in which is the root $c$. That we can find out by conditions for endpoints of both intervals. If $f(a_0)\cdot f(c_0)<0$, then we denote this interval as $\langle a_1,b_1\rangle$, i. e. $a_1=a_0$ and $b_1=c_0$. If $f(a_0)\cdot f(c_0)>0$ and $f(c_0)\cdot f(b_0)<0$, then we denote the interval $\langle c_0,b_0\rangle$ as $\langle a_1,b_1\rangle$, where $a_1=c_0$ and $b_1=b_0$. If $f(a_0)\cdot f(c_0)=0$ resp. $f(c_0)\cdot f(b_0)=0$, then we know, that $f(c_0)=0$. We have found the root $c=c_0$ and we end the iterative process. If we did not found the root $c$, then the new interval $\langle a_1,b_1\rangle$ will be devided in two again and we proceed by the same way in marking new elements and in next decision-making. This procedure creates a sequence of intervals $\langle a_0,b_0\rangle$, $\langle a_1,b_1\rangle$, $\langle a_2,b_2\rangle$, $\ldots$, $\langle a_k,b_k\rangle$, $\ldots$ Each subsequent interval is obtained dividing the interval $\langle a_k,b_k\rangle$ to half and denoting the middle point $c_k=\frac{a_k+b_k}{2}$, for some $k\in\mathbb{N}$. If $f(a_k)\cdot f(c_k)<0$, then we denote this interval as $\langle a_{k+1},b_{k+1}\rangle$, i. e. $b_{k+1}=c_k$. If $f(a_k)\cdot f(c_k)>0$ and $f(c_k)\cdot f(b_k)<0$, then we denote the interval $\langle c_k,b_k\rangle$ as $\langle a_{k+1},b_{k+1}\rangle$, where $a_{k+1}=c_k$. If $f(a_k)\cdot f(c_k)=0$ resp. $f(c_k)\cdot f(b_k)=0$, then we know that $f(c_k)=0$. We have found the root $c=c_k$ and we can end the iterative process. If we did not found the root $c$, then we create a new interval $\langle a_{k+1},b_{k+1}\rangle$ of half-length until its length is less than the number $2\cdot\varepsilon$, for given small real number $\varepsilon>0$, i. e.[3]

$$b_k-a_k<2\cdot\varepsilon. \tag{2.2}$$

Approximate solution of an equation $f(x)=0$ corresponds to the value of the middle point of the last $k$-th interval:

$$c\approx c_k=\frac{a_k+b_k}{2}. \tag{2.3}$$

We estimate the error of the obtained solution of our equation. We know that the root $c$ of the equation $f(x)=0$ is inside the interval $\langle a_k,b_k\rangle$. Therefore, the approximate value $c_k$ can be from the exact value $c$ at a distance at most half of the length of the interval $\langle a_k,b_k\rangle$, i. e. about the value $\varepsilon$. Therefore, for the estimation of the error of the *bisection method* applies that the form

[3] If the real root $c$ lies within the interval $\langle a,b\rangle$ and $c_1=\frac{a+b}{2}$, then $|c-c_1|<\frac{b-a}{2}$, i. e. $c_1$ is the approximate value of the root $c$ with an accuracy $\frac{b-a}{2}$.
If for some point $x$ holds $f(x-\varepsilon)\cdot f(x+\varepsilon)<0$, then $|x-c|<\varepsilon$.

of the error estimation after the $k$-th iteration is:

$$|c_k - c| \leqq \frac{b-a}{2^{k+1}} \leqq \varepsilon. \tag{2.4}$$

The bisection method converges to the root of the equation whenever the interval $\langle a, b\rangle$ contains the root $c$. If, in the interval are more roots, then this method finds one of them. The disadvantage of the bisection method is that this method converges to the root slowly. It is therefore appropriate to use this method to reduce the length of initial interval, where is the root $c$, and then use one of the faster methods. Four iterations of the bisection method are shown in Figure 2.1.

We can use the Table 2.1 on the page 71, to visualize the process of solving an equation $f(x) = 0$ by the bisection method. In this table are entered four iterative steps of the bisection method.

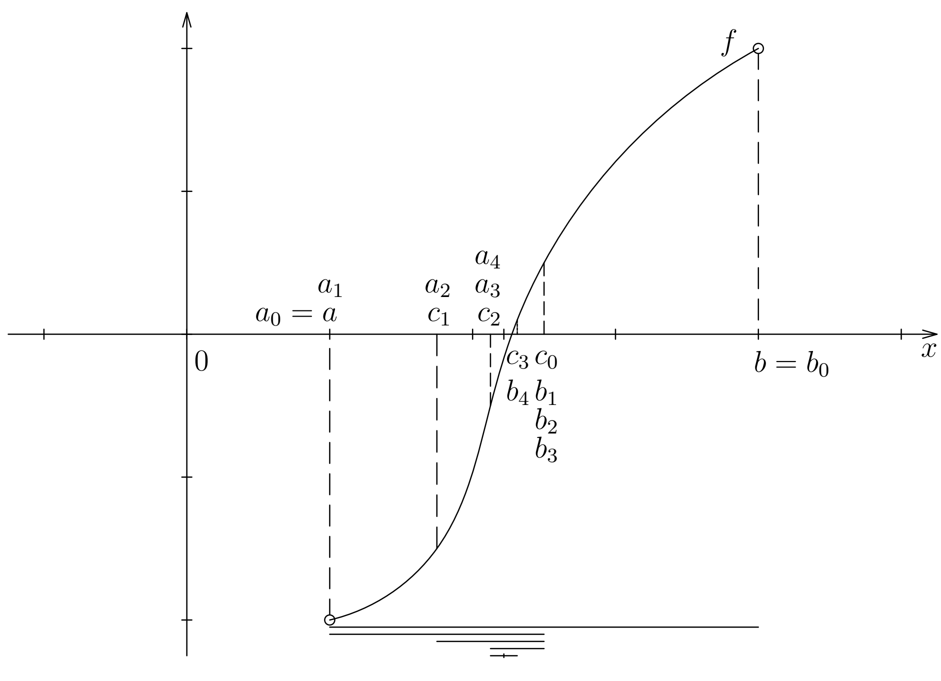


Figure 2.1: Four iteration of bisection method.

We fill the Table 2.1 on the page 71 until the value in the last column will be less than $2 \cdot \varepsilon$. Finding the approximate solutions by the bisection method

Table 2.1: Table for the bisection method.

| $k$ | $a_k$ | $b_k$ | $c_k$ | $\lvert b_k - a_k \rvert$ | |
|---|---|---|---|---|---|
| 0 | $a_0 = a$ | $b_0 = b$ | $c_0 = \frac{a_0+b_0}{2}$ | $b_0 - a_0$ | $\text{sgn} f(c_0) = \text{sgn} f(b_0)$ |
| 1 | $a_1 = a_0$ | $b_1 = c_0$ | $c_1 = \frac{a_1+b_1}{2}$ | $b_1 - a_1$ | $\text{sgn} f(c_1) = \text{sgn} f(a_1)$ |
| 2 | $a_2 = c_1$ | $b_2 = b_1$ | $c_2 = \frac{a_2+b_2}{2}$ | $b_2 - a_2$ | $\text{sgn} f(c_2) = \text{sgn} f(a_2)$ |
| 3 | $a_3 = c_2$ | $b_3 = b_2$ | $c_3 = \frac{a_3+b_3}{2}$ | $b_3 - a_3$ | $\text{sgn} f(c_3) = \text{sgn} f(b_3)$ |
| 4 | $a_4 = a_3$ | $b_4 = c_3$ | $c_4 = \frac{a_4+b_4}{2}$ | $b_4 - a_4$ | |
| 5 | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | |

with a given accuracy does not depend on the form of the function $f(x)$. It can be shown that improving the result on one decimal place requires always next three to four steps of this method.

## 2.3 Fixed Point Iteration Method

The *fixed point iteration method* for solving one nonlinear equation of one real variable is the application of the general method of successive approximations, which we describe now.

**Definition 2.1** We say that $g$ is a *mapping* from the set $A$ to the set $B$ (we write $g\colon\ A \longrightarrow B$), if for each element $x \in A$ it assigns just one element $y \in B$ using $g$, such that the following applies: $y = g(x)$.

We are interested in sets which are mapping to themselves and in elements that are mapping them to themselves.

**Definition 2.2** An element $x \in A$ is called the *fixed point* (*fixpoint*) of the mapping $g\colon\ A \longrightarrow A$, iff

$$g(x) = x. \tag{2.5}$$

If the set $A = \mathbb{R}$, then the mapping $g\colon\ \mathbb{R} \longrightarrow \mathbb{R}$ is a real function of one real variable. For example, the quadratic function $f\colon\ y = x^2$ has two fixed points. These are the points at which the graph of $f$ intersects the line $y = x$. This is true for points $x = 0$ and $x = 1$, or $0^2 = 0$ and $1^2 = 1$.

We ask whether each mapping has a fixed point and if it does how to find it. It can be proven that some mappings have a fixed point always, and there exists a procedure to find the fixed point.

**Definition 2.3** Let $A \subset \mathbb{R}$ be a metric space with metric $d$. We say that the mapping $g\colon\ A \longrightarrow A$ is *contractive* (*contraction*) if there exists a real number $\alpha \in \langle 0, 1)$, such that for every two points $x, y \in A$ holds:

$$d(g(x), g(y)) \leqq \alpha \cdot d(x, y). \tag{2.6}$$

Number $\alpha$ is called the *coefficient of contraction*.

Contraction, (narrowing) may be freely interpreted in the way that contraction mappings have images (functional values) closer than were their patterns.

**Theorem 2.2** Let $A$ be a complete metric space and let mapping $g\colon\ A \longrightarrow A$ be contraction. Then there exists exactly one fixed point of mapping $g$ (denoted by $x_p$), for which the following applies:

$$x_p = \lim_{n\to\infty} x_n, \tag{2.7}$$

where $\{x_n\}_{n=0}^{\infty}$ is a sequence of approximations, which is defined as follows: $x_0$ is any element of a set $A$ and the other members of the sequence are defined by:

$$x_{k+1} = g(x_k), \qquad k = 0, 1, 2, \ldots \tag{2.8}$$

and, moreover, for all non negative integers $n$ holds:

$$d(x_p, x_n) \leqq \frac{\alpha}{1-\alpha} \cdot d(x_n, x_{n-1}), \tag{2.9}$$

$$d(x_p, x_n) \leqq \frac{\alpha^n}{1-\alpha} \cdot d(x_0, x_1), \tag{2.10}$$

where $\alpha$ is coefficient of the contraction of the mapping $g$.

The Theorem 2.2 gives us instructions how to find a fixed point of given mapping $g$. We choose an arbitrary point $x_0 \in A$. The point $x_0$ is called the initial approximation. Then we calculate the other members of the sequence according to the relation (2.8) in the Theorem 2.2. This calculation is called an *iterative process*, and the $k$-th member of a sequence $\{x_n\}_{n=0}^{\infty}$ is called the $k$-th approximation.

According to the relation (2.7) in the Theorem 2.2 a fixed point of mapping $g$ is the limit of the sequence $\{x_n\}_{n=0}^{\infty}$. Successive approximations are approaching to the fixed point $x_p$. So if we could perform this iterative process indefinitely, we would obtain by this procedure the exact value of the fixed point $x_p$. In real conditions it is not possible, because after a certain number of steps we stop the iterative process and we approximate the fixed point by the last member of the calculated sequence $\{x_n\}_{n=0}^{\infty}$.

When to stop the iterative process depends on the accuracy with which we have a fixed point $x_p$ to calculate. For this we can use the relation (2.9) or the relation (2.10) in the Theorem 2.2, which bounds the distance of the $n$-th approximation from the fixed point $x_p$.

Let us return to solving the equation $f(x) = 0$, with that we will use above-mentioned facts. We modify the equation $f(x) = 0$ to the form:

$$x = g(x). \tag{2.11}$$

The function $g$ is called the *iteration function*. After this modification, searching for the root $c$ of the equation $f(x) = 0$ will be the same as finding a fixed point $x_p$ of a function $g(x)$. We choose the initial approximation $c_0 \in \mathcal{D}(g)$. We calculate the following approximations of the fixed point (or solve the equation) according to the formula:

$$c_{k+1} = g(c_k), \tag{2.12}$$

for $k = 1, 2, 3, \ldots$ We create a sequence of approximations $\{c_n\}_{n=0}^{\infty}$, which in general case does not need to converge. Therefore, we will show when fixed point iteration method converges to the solution.

**Theorem 2.3** Let the function $g$ maps the interval $\langle a, b\rangle$ into itself and has the first derivative of this function on this interval. Then, if there exists a real number $\alpha \in \langle 0, 1)$, such that

$$|g'(x)| \leqq \alpha \qquad \forall x \in \langle a, b\rangle, \tag{2.13}$$

then the mapping $g(x)$ is a contraction with coefficient $\alpha$ and there exists a fixed point $x_p$ of a function $g$ in the interval $\langle a, b\rangle$ and sequence of approximations obtained by relation (2.12) on the page 73 converges to this fixed point for any initial approximation $x_0 \in \langle a, b\rangle$. For the error estimates holds:

$$|c_k - x_p| \leqq \frac{\alpha}{1-\alpha} \cdot |c_k - c_{k-1}|, \tag{2.14}$$

$$|c_k - x_p| \leqq \frac{\alpha^k}{1-\alpha} \cdot |c_0 - c_1|, \tag{2.15}$$

We can use the error estimates (2.14) and (2.15) from the Theorem 2.3 when deciding about termination of the iterative process. However, the verification of condition (2.13) in the Theorem 2.3 may be in general case complicated. In such case we can use as stopping condition:

$$|c_k - c_{k-1}| < \varepsilon, \tag{2.16}$$

which does not mean that holds: $|c_k - x_p| < \varepsilon$. It is therefore often used criterion:

$$f(c_k - \varepsilon) \cdot f(c_k + \varepsilon) < 0. \tag{2.17}$$

## 2.4 Newton's Method

The *Newton's method* (also called *tangential method*) is the method that approximates the solution of the equation $f(x) = 0$ using the tangents to the graph of the function $f$: $y = f(x)$.

The basic principle of the Newton's method is the construction of tangents to the chosen, respectively calculated point. We choose the initial approximation $x_0$ of the solution $c$ of the equation $f(x) = 0$. We calculate the value of the function $f(x_0)$. We construct a tangent line $t_0$ to the graph of the function $f$ at the point $[x_0, f(x_0)]$. We denot the intersection of the tangent line $t_0$ with $x$-axis as the point $x_1$. The point $x_1$ is a further approximation of the solution of the equation $f(x) = 0$. Again we calculate the value of the function $f(x_1)$ and construct a tangent line $t_1$ to the graph of the function $f$ at the point $[x_1, f(x_1)]$. In this way we continue in creation of the sequence of approximations $x_0, x_1, x_2, \ldots$ Let the function $f$ has a first derivative. Then the equation of the tangent line $t_k$ to the graph of the functions $f$ at the point $T = [x_k, y_k]$ reads (see 2.2):

$$t_k : y - y_k = f'(x_k) \cdot (x - x_k). \tag{2.18}$$

The intersection of the tangent line $t_k$ at the point $[x_k, f(x_k)]$ with $x$-axis is computed using the formula:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}. \tag{2.19}$$

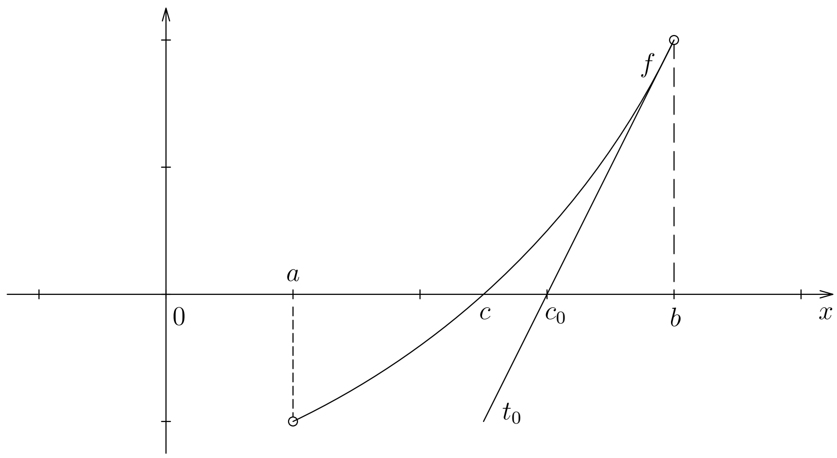


Figure 2.2: First iteration of Newton's method.

The Newton's method also can be derived using the Taylor formula. Suppose that we have the $k$-th approximation $x_k$ of the root $c$. Then we can write the following relation:

$$f(c) = f(x_k) + f'(x_k) \cdot (c - x_k) + ER,$$

where $ER$ is the residue in Taylor's formula.

If we neglect residue $ER$ and we realize that $f(c) = 0$ because $c$ is the root of the equation $f(x) = 0$, then from the previous equation we can express approximately the root $c$ as follows:

$$c \doteq x_k - \frac{f(x_k)}{f'(x_k)},$$

which corresponds to the approximation $x_{k+1}$.

From Taylor's formula we can also derive the error estimates for the $k$-th approximation of the root $c$ obtained by Newton's method. If the function $f$ on the interval $I$ has the second derivative, where $x_k \in I$ and $c \in I$, then the following relations hold:

$$|c - x_k| \leqq \frac{M_2}{2 \cdot m_1} \cdot (x_k - x_{k-1})^2, \quad |c - x_k| \leqq \frac{M_2}{2 \cdot m_1} \cdot (c - x_{k-1})^2, \qquad (2.20)$$

where $M_2 = \max_{x \in I} |f''(x)|$ and $m_1 = \min_{x \in I} |f'(x)|$.

Based on the Lagrange's Theorem 1.26 on page 53 between points $x_k$ and $c$ lies such a point $\xi$, for which holds $f'(\xi) = \frac{f(x_k)-f(c)}{x_k-c}$. Whereas for the root $c$ is true $f(c) = 0$, we have $|x_k - c| = \frac{|f(x_k|}{|f'(\xi)|}$, then for each number $m$, $0 < m \leqq |f'(\xi)|$ applies estimate:

$$|c - x_k| \leqq \frac{|f(x_k)|}{m}. \tag{2.21}$$

The best estimation we get for $m = m_1$.

Another way to express the relation of a calculation of the $k$-th approximation of the root is to combine knowledge from the theory of fixed point and Newton's method. We try to express $x$ from the equation $f(x) = 0$, which is the basic principle of the fixed point iteration method. Let us divide this equation by derivative of a function $f$. We get the equation in the form $\frac{f(x)}{f'(x)}$. We multiply this equality with $-1$ and we add $x$ to the both sides of the equation. Then we get the relation:

$$x = x - \frac{f(x)}{f'(x)},$$

which corresponds to Newton's method and it is a special case of fixed point iteration method.

The Newton's method is from by us considered methods for solving nonlinear equations most effective, but it does not always converge. Convergence of the Newton's method often depends on the choice of the initial approximation $x_0$. Therefore we describe the conditions that guarantee us the convergence of Newton's method.

**Theorem 2.4** (*Fourier's condition*) Let a function $f$ be continuous on an interval $\langle a, b\rangle$ and there be exactly one root $c$ of an equation $f(x) = 0$ in this interval. Let the first and the second derivative of the function $f$ be continuous on the interval $\langle a, b\rangle$ and let the function do not change sign on this interval.[4] If we choose for the initial approximation a point $x_0$ satisfying:

$$f(x_0) \cdot f''(x_0) > 0, \tag{2.22}$$

then the Newton's method will converge to the root $c$, where $x_0 \in \langle a, b\rangle$.

[4]The condition that the first derivative of the function $f$ does not change sign on the interval $\langle a, b\rangle$ means that the function $f$ is growing or decreasing all over the given interval. The condition that the second derivative of the function $f$ does not change sign throughout the interval $\langle a, b\rangle$ indicates, that the function $f$ is concave up on the interval, or concave down.

## 2.5 Solved Examples

**Example 2.1** Solve the given equation by bisection method, perform 8 iterations and estimate the error:

$$\frac{x}{8} - 2 + \ln 3x = 0.$$

*Solution:*
Separating the roots () $\ln(3x) = 2 - \frac{x}{8} \iff g(x) = h(x)$) is easy to find the interval in which lies the root, see Figure 2.3.

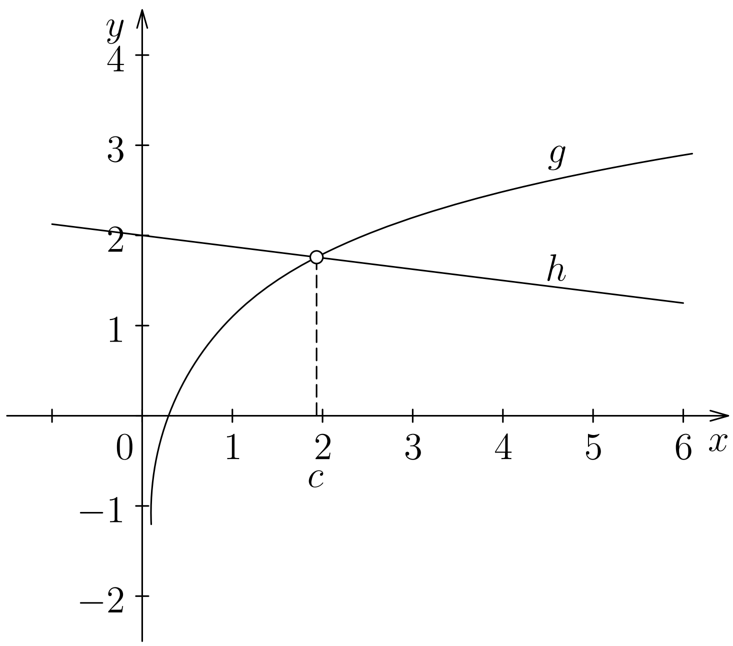


Figure 2.3: Separating the roots of the equation $\ln(3x) = 2 - \frac{x}{8}$.

Let us denote as the function $f$ the left side of the equation, $f\colon\ y = \frac{x}{8} - 2 + \ln 3x$. We verify that the root $c$ lies in the interval $\langle 1, 2\rangle$. The function $f$ is continuous on the interval $\langle 1, 2\rangle$ and $f(1) \cdot f(2) < 0$. Thus the root $c \in \langle 1, 2\rangle$. We will create the Table 2.2 based on the Table 2.1 on page 71, using which we find approximation of solution $c$ of the given equation.[5]

[5]Sign in circles in Tab 2.2 is the sign of the functional value of the function $f$ at those points. For example, 1,75 $\ominus$ means that $f(1{,}75) < 0$.

Table 2.2: Solving equation by bisection method.

| $k$ | $a_k$ $\ominus$ | $b_k$ $\oplus$ | $c_k$ | | $\lvert b_k - a_k \rvert$ |
|---|---|---|---|---|---|
| 0 | 1,0 | 2,0 | 1,5 | $\ominus$ | 1,0 |
| 1 | 1,5 | 2,0 | 1,75 | $\ominus$ | 0,5 |
| 2 | 1,75 | 2,0 | 1,875 | $\ominus$ | 0,25 |
| 3 | 1,875 | 2,0 | 1,9375 | $\oplus$ | 0,125 |
| 4 | 1,875 | 1,9375 | 1,90625 | $\ominus$ | 0,0625 |
| 5 | 1,90625 | 1,9375 | 1,921875 | $\ominus$ | 0,03125 |
| 6 | 1,921875 | 1,9375 | 1,9296875 | $\ominus$ | 0,015625 |
| 7 | 1,9296875 | 1,9375 | 1,93359375 | $\ominus$ | 0,0078125 |
| 8 | 1,93359375 | 1,9375 | 1,935546875 | $\oplus$ | 0,00390625 |

Based on the last row in the Table 2.2, we get the approximation of the solution to the equation $c_8 = 1{,}935546875$. The error estimate of the solution reads:

$$|c_8 - c| \leqq \frac{b-a}{2^{k+1}} = \frac{2-1}{2^{8+1}} = \frac{1}{2^9} \doteq 0{,}001953125.$$

$\checkmark$

**Example 2.2** Solve the given equation by Newton's method with the accuracy $\varepsilon = 5 \cdot 10^{-6}$:

$$\frac{x}{8} - 2 + \ln 3x = 0.$$

*Solution:*
Similarly to the previous example, we see easily that the root of the equation $c \in \langle 1, 2 \rangle$. Let us denote as the function $f$: $y = f(x)$ the left side of the equation, $f$: $y = \frac{x}{8} - 2 + \ln 3x$. In order to investigate the conditions for the convergence of Newton's method we need to know the first and the second derivative of the function $f$.

$$f: \quad y = \frac{x}{8} - 2 + \ln 3x,$$

$$f': \quad y' = \left(\frac{x}{8} - 2 + \ln 3x\right)' = \frac{1}{8} + \frac{1}{x},$$

$$f'' : \quad y'' = \left(\frac{1}{8} + \frac{1}{x}\right)' = -\frac{1}{x^2}.$$

The second derivative of the function $f$ is negative (and thus $f''$ does not change the sign) on the interval $\langle 1, 2\rangle$. It is true that $f(1)\cdot f''(1) > 0$, therefore we choose as the initial approximation of the solution the point $c_0 = 1$. We write the iterative relation whereby we fill a table with approximations of the solution of our equation.

$$c_{k+1} = c_k - \frac{f(c_k)}{f'(c_k)} = c_k - \frac{\frac{c_k}{8} - 2 + \ln(3c_k)}{\frac{1}{8} + \frac{1}{c_k}}.$$

We create the table:

Table 2.3: Newton's method for solving of equation.

| $k$ | $c_k$ | $f(c_k - \varepsilon) \cdot f(c_k + \varepsilon)$ |
|---|---|---|
| 0 | 1 | — |
| 1 | 1,690122410 | $\ominus \cdot \ominus = \oplus$ |
| 2 | 1,920801430 | $\ominus \cdot \ominus = \oplus$ |
| 3 | 1,934036886 | $\ominus \cdot \ominus = \oplus$ |
| 4 | 1,934073692 | $\ominus \cdot \oplus = \ominus$ |

Whereas the stopping criterion has been met (the second column in Table 2.3), the approximate solution of our equation for a given accuracy is the number $c_4 = 1{,}934073692$. For this solution we will estimate the error.

Holds:

$$m_1 = \min_{x\in\langle 1,2\rangle} |f'(x)| = 0{,}625,$$

$$|c - c_4| \leqq \frac{|f(c_4)|}{m_1} = 2{,}8974 \cdot 10^{-10} < \varepsilon.$$

$\checkmark$

**Example 2.3** Separate the roots of the equation $x \cdot \sqrt{x+1} = 1$, verify that the conditions of convergence of fixed point iteration method are satisfied. Make 3 steps of this method and estimate the error of the approximate solution after the third step.

*Solution:*
To separate the roots we can modify our equation to the form:

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

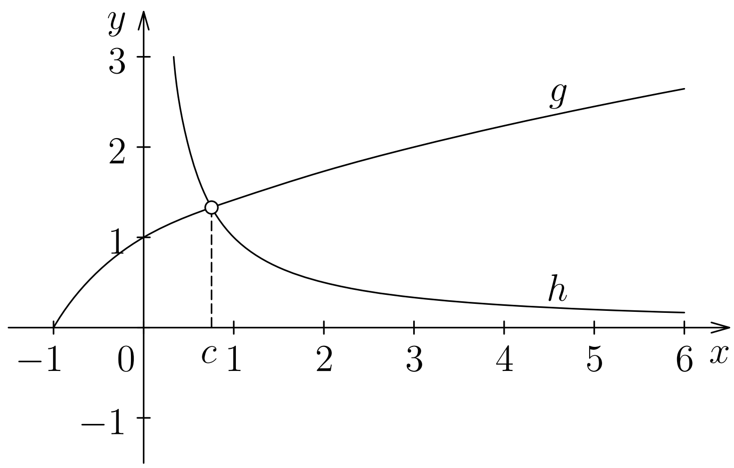


Figure 2.4: Separating the roots of equation $\sqrt{x+1} = \frac{1}{x}$.

From the graphs of elementary functions that make up the right and left side of the equation, we can easily guess that the root $c$ lies in the interval $\langle \frac{1}{2}, 1\rangle$, see Figure 2.4. We will verify that the root $c$ is really within that interval. Let us denote $f$: $y = x \cdot \sqrt{x+1} - 1$. Function $f$ is continuous on interval $\langle \frac{1}{2}, 1\rangle$ and applies that $f(\frac{1}{2}) \cdot f(1) = (\frac{1}{2} \cdot \sqrt{\frac{3}{2}} - 1) \cdot (\sqrt{2} - 1) < 0$. We have shown that $c \in \langle \frac{1}{2}, 1\rangle$. Let us now express the function $\varphi$ modifying our equation to the form $x = \varphi(x)$.

$$x = \frac{1}{\sqrt{x+1}} \quad \Longrightarrow \quad \varphi(x) = \frac{1}{\sqrt{x+1}} \qquad \text{for } x \in \left\langle \frac{1}{2}, 1 \right\rangle$$

We calculate the first derivative of the function $\varphi$:

$$\varphi'(x) = \left(\frac{1}{\sqrt{x+1}}\right)' = -\frac{1}{2} \cdot (x+1)^{-\frac{3}{2}} = \frac{-1}{2\sqrt{(x+1)^3}}.$$

Function $|\varphi'(x)|$ is a decreasing on the interval $\langle \frac{1}{2}, 1\rangle$. Therefore, we can bound it:

$$|\varphi'(x)| = \left|\frac{-1}{2\sqrt{(x+1)^3}}\right| \leqq \frac{3}{10} < 1 \qquad \text{for } x \in \left\langle \frac{1}{2}, 1 \right\rangle.$$

The mapping $\varphi$ is contractive on the interval $\langle \frac{1}{2}, 1\rangle$ and $\alpha = \frac{3}{10}$, so the iterative method will converge. As an initial approximation we choose the point $c_0 = \frac{3}{4}$ and we express iterative relation:

$$c_{k+1} = \frac{1}{\sqrt{c_k + 1}}.$$

We create a table into which we write the results the above iterative relation.

Table 2.4: Iterative methods for solving of the equation.

| $k$ | $c_k$ | $\lvert c_k - c_{k-1}\rvert$ |
|---|---|---|
| 0 | 0 | — |
| 1 | 1 | 1 |
| 2 | 0,7071 | 0,2929 |
| 3 | 0,7654 | 0,0583 |

We still need to estimate the error of our solution. We know the value $\alpha$ and we use the relation (2.14), which is shown on page 73.

$$|c_k - c| \leqq \frac{\alpha}{1-\alpha} \cdot |c_k - c_{k-1}|,$$

$$|c_3 - c| \leqq \frac{\alpha}{1-\alpha} \cdot |c_3 - c_2|,$$

$$|c_3 - c| \leqq \frac{\frac{1}{2}}{1-\frac{1}{2}} \cdot 0{,}0583 = 0{,}0583.$$

We got the approximate solution $c_3 = 0{,}7654$, where the upper bound of the error of this solution is 0,0583. $\surd$

## 2.6 Unsolved Tasks

**2.1** Separate the real roots of the given equation and solve the given equation by bisection method for the specified root. Make $n$ steps of the iterative method and estimate the upper bound of the error.

a) $x^5 = 6x^2 - 1$, where $n = 10$ and choose the smaller positive root,

b) $x^5 = 4x^4 + 2$, where $n = 14$ and choose the positive root,

c) $x^4 = 7 - 8x$, where $n = 15$ and choose the greatest root,

d) $x^3 = 2 - 4x^2$, where $n = 10$ and choose the greater negative root,

e) $\mathrm{e}^x + 2x = 2$, where $n = 12$ and choose the positive root,

f) $x^2 + \ln x = 4$, where $n = 12$ and choose the positive root.

**2.2** Separate the real roots of the given equation and solve the given equation by fixed point iteration method for the specified root with an accuracy of $\varepsilon$. Estimate the upper bound of the error calculation.

a) $x^3 - 1 = 12x$, where $\varepsilon = 10^{-3}$ and choose the lowest root,

b) $x^3 - 1 = 12x^2$, where $\varepsilon = 10^{-2}$ and choose lower root,

c) $\ln x = 4 - 2x$, where $\varepsilon = 10^{-3}$ and choose the positive root,

d) $\mathrm{e}^{2x} - 9 = x$, where $\varepsilon = 10^{-4}$ and choose the positive root,

e) $4x^3 + 1 - x^2 = 0$, where $\varepsilon = 10^{-3}$ and choose the lowest negative root.

**2.3** Separate the real roots of the given equation and solve the given equation by Newton's method for the specified root with an accuracy of $\varepsilon$. Estimate the upper bound of the error calculation.

a) $x^3 + x = 3$, where $\varepsilon = 10^{-3}$ and choose the positive root,

b) $x^4 + 3 = 5x^3$, where $\varepsilon = 10^{-4}$ and choose the greatest positive root,

c) $\ln x = x - 2$, where $\varepsilon = 10^{-4}$ and choose the greatest negative root,

d) $\mathrm{e}^{2x} = 8x$, where $\varepsilon = 10^{-4}$ and choose the lowest negative root,

e) $x^5 + 4 - x^2 = 0$, where $\varepsilon = 10^{-4}$ and choose the negative root.

## 2.7 Results of Unsolved Tasks

**2.1** **a)** $0{,}4111328125 \in \langle 0, 1\rangle$ **b)** $4{,}0076904296875 \in \langle 3, 5\rangle$ **c)** $0{,}81881713867188 \in \langle 0, 1\rangle$ **d)** $-0{,}7900390625 \in \langle -1, 0\rangle$ **e)** $0{,}314697265625 \in \langle 0, 1\rangle$ **f)** $1{,}841064453125 \in \langle 1, 2\rangle$

**2.2** **a)** $-3{,}505 \in \langle -4, -3\rangle$ **b)** $0{,}29 \in \langle 0, 1\rangle$ **c)** $1{,}727 \in \langle 1, 2\rangle$ **d)** $1{,}0374 \in \langle 0, 2\rangle$ **e)** $-0{,}24490799644825 \in \langle -1, 0\rangle$

**2.3** **a)** $1{,}213 \in \langle 1, 2\rangle$ **b)** $4{,}9756 \in \langle 4, 6\rangle$ **c)** $3{,}1462 \in \langle 3, 4\rangle$ **d)** $0{,}17870 \in \langle 0, 1\rangle$ **e)** $-1{,}2056 \in \langle -2, 0\rangle$

# Chapter 3

# Approximation of Functions

It is often the case, that the function is not specified by any functional expression and we know only its functional values in some of its points, which are often received from different measurements. For such functions it is difficult to obtain the functional values in other than the given points, find its derivative, or integrate it. Therefore it is appropriate to replace this function by another function, which is similar to given function and it has the functional expression and it is easy to do calculations with it.

The most commonly used in approximation as the polynomial function (polynomial of the $n$-th degree), which is defined on the set $\mathbb{R}$. A polynomial function is easily differentiable and integrable on the set $\mathbb{R}$. The requirements on the function, by which we want to approximate the given function may be various. If we use *interpolation* for approximation of function, we demand that approximate function has the same functional values (and/or derivatives) at selected points as original function. Using the *least squares method* is not necessary for approximative function to directly pass through given points, just that in a certain sense, to be as close as possible to given points of original function.

## 3.1 Interpolation

Formulation of the problem: Let the function $f$ be given by $n+1$ each other different points $x_0, x_1, \ldots, x_n$. These points are called *nodes* (nodes of interpolation). We denote the function values at those points as $y_0, y_1, \ldots, y_n$, where $y_0 = f(x_0)$, $y_1 = f(x_1)$, $y_2 = f(x_2)$, $\ldots$, $y_n = f(x_n)$. Let us find the polynomial $P_n(x)$ of degree at most $n$, such that the node points take the same functional values as the function $f$ i. e. $P_n(x_i) = y_i$ for $i = 0, 1, 2, \ldots, n$.

**Theorem 3.1** Let nodes $[x_i, y_i]$ for $i = 0, 1, 2, \ldots, n$ be given and $x_i \neq x_j$ for $i \neq j$. Then there exists a polynomial $P_n(x)$ of degree at most $n$, such that $P_n(x_i) = y_i$ for $i = 0, 1, 2, \ldots, n$.

### Construction of the interpolation polynomial

Below we describe the construction of a polynomial, which satisfies the conditions of Theorem 3.1, and is called *Lagrange's interpolating polynomial.*

The nodes $[x_i, y_i]$ for $i = 0, 1, 2, \ldots, n$ are given and $x_i \neq x_j$ for $i \neq j$. We construct polynomials $p_i(x)$ for $i = 0, 1, 2, \ldots, n$, such that:

$$p_i(x_j) = \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j. \end{cases}$$

Then the Lagrange's interpolating polynomial has the form:

$$L_n(x) = p_0(x) \cdot y_0 + p_1(x) \cdot y_1 + p_2(x) \cdot y_2 + \cdots + p_{n-1}(x) \cdot y_{n-1} + p_n(x) \cdot y_n.$$

It is easy to see that so defined function $L_n(x)$ satisfies allegation of the Theorem 3.1. It only remains to show the construction of the polynomial $p_i(x)$ itself.

$$p_0(x) = \frac{(x - x_1) \cdot (x - x_2) \cdot (x - x_3) \cdot \cdots \cdot (x - x_n)}{(x_0 - x_1) \cdot (x_0 - x_2) \cdot (x_0 - x_3) \cdot \cdots \cdot (x_0 - x_n)},$$

$$p_1(x) = \frac{(x - x_0) \cdot (x - x_2) \cdot (x - x_3) \cdot \cdots \cdot (x - x_n)}{(x_1 - x_0) \cdot (x_1 - x_2) \cdot (x_1 - x_3) \cdot \cdots \cdot (x_1 - x_n)},$$

$$p_2(x) = \frac{(x - x_0) \cdot (x - x_1) \cdot (x - x_3) \cdot \cdots \cdot (x - x_n)}{(x_2 - x_0) \cdot (x_2 - x_1) \cdot (x_2 - x_3) \cdot \cdots \cdot (x_2 - x_n)},$$

$$p_{n-1}(x) = \frac{(x - x_0) \cdot (x - x_1) \cdot \cdots \cdot (x - x_{n-2}) \cdot (x - x_n)}{(x_{n-1} - x_0) \cdot (x_{n-1} - x_1) \cdot \cdots \cdot (x_{n-1} - x_{n-2}) \cdot (x_{n-1} - x_n)},$$

$$p_n(x) = \frac{(x - x_0) \cdot (x - x_1) \cdot \cdots \cdot (x - x_{n-2}) \cdot (x - x_{n-1})}{(x_n - x_0) \cdot (x_n - x_1) \cdot \cdots \cdot (x_n - x_{n-2}) \cdot (x_n - x_{n-1})}.$$

It is easy to see that these functional expressions satisfy the definition of the polynomial $p_i(x)$ at node points $x_j$ for $i$, $j = 0, 1, 2, \ldots, n$. After substituting into the expression of the $L_n(x)$ we obtain Lagrange's interpolating polynomial:

$$L_n(x) = \sum_{j=0}^{n} \prod_{\substack{i=0 \\ i \neq j}}^{n} \frac{(x - x_i)}{(x_j - x_i)} \cdot y_j. \tag{3.1}$$

In a similar way we can define the *Inverse Lagrange's interpolating polynomial.* Let nodes $[x_i, y_i]$ for $i = 0, 1, 2, \ldots, n$ be given and $y_i \neq y_j$ for $i \neq j$. We construct polynomials $p_i(y)$:

$$p_i(y_j) = \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j. \end{cases}$$

Then the Inverse Lagrange's interpolating polynomial has the form:

$$L_n^{\text{Inv}}(y) = p_0(y) \cdot x_0 + p_1(y) \cdot x_1 + \cdots + p_{n-1}(y) \cdot x_{n-1} + p_n(y) \cdot x_n.$$

By similar procedure as in the construction of the Lagrange's interpolating polynomial we get for the node points $[x_i, y_i]$ for $i = 0, 1, 2, \ldots, n$, and $y_i \neq y_j$ for $i \neq j$: $L_n^{\text{Inv}}(y)$ – Inverse Lagrange's interpolating polynomial:

$$L_n^{Inv}(y) = \sum_{j=0}^{n} \prod_{\substack{i=0 \\ i \neq j}}^{n} \frac{(y - y_i)}{(y_j - y_i)} \cdot x_j. \tag{3.2}$$

## 3.2 The Least Squares Method

In experiments measurements are often carried out several times under the same conditions, which is contrary to the assumptions of interpolation, requiring that all nodes are different from each other. Also, in measurements we receive data containing errors. This date is therefore not appropriate to interpolate as this would propagate these errors. Therefore, if we know at least a little about the functional relationship (linear, quadratic, logarithmic, exponential, etc.), we can approximate that function so that from the supposed type of functional dependency (the set of all linear function or the set of all quadratic functions, etc.) we choose such a function, which is to given points in a sense, as close as possible.

Formulation of the problem: Let nodes $x_0, x_1, x_2, \ldots, x_n$ and corresponding function values $y_0, y_1, y_2, \ldots, y_n$ be given . Let functions $\varphi_0, \varphi_1, \ldots, \varphi_m$, where $m < n + 1$ be given. Then, from all functions of the form (linear combination of functions $\varphi_0, \varphi_1, \varphi_2, \ldots, \varphi_m$):

$$f_m(x) = a_0 \cdot \varphi_0(x) + a_1 \cdot \varphi_1(x) + \cdots + a_m \cdot \varphi_m(x) = \sum_{j=0}^{m} a_j \cdot \varphi_j(x)$$

we are looking for such a function, for which the quadratic error:

$$g(a_0, a_1, \ldots, a_m) = \sum_{i=0}^{n} (y_i - f_m(x_i))^2 \tag{3.3}$$

obtains the minimum value. Unknown variables $a_0, a_1, \ldots, a_m$ are from the set $\mathbb{R}$. The quadratic error $g$ expresses the sum of the squares with a length of side $|y_i - f_m(x_i)|$, for $i = 0, 1, 2, \ldots, n$. The function for which is obtains the minimum is called the *best approximation of the experimental data* $y_0, \ldots, y_n$ in a given class of functions in the sense of the least squares method.

**Determination of the best approximation:**

The points $[x_i, y_i]$, $i = 0, 1, 2, \ldots, n$ and functions $\varphi_j$, $j = 0, 1, 2, \ldots, m$ are given, therefore the quadratic error:

$$g(a_0, a_1, \ldots, a_m) = \sum_{i=0}^{n} (y_i - f_m(x_i))^2 =$$

$$= \sum_{i=0}^{n} \left[y_i - (a_0 \cdot \varphi_0(x_i) + a_1 \cdot \varphi_1(x_i) + \cdots + a_m \cdot \varphi_m(x_i))\right]^2$$

depends only on the coefficients $a_0, a_1, \ldots, a_m$ of the function $g$. From the differential calculus of functions of several variables we know, that a necessary condition for the function $g(a_0, a_1, \ldots, a_m)$ to obtain a minimum, is fulfilment of the conditions expressed by the following equations:

$$\frac{\partial g}{\partial a_j} = \frac{\partial}{\partial a_j}\left[\sum_{i=0}^{n} \left[y_i - (a_0 \cdot \varphi_0(x_i) + a_1 \cdot \varphi_1(x_i) + \cdots + a_m \cdot \varphi_m(x_i))\right]^2\right] = 0,$$ [1]

for $j = 0, 1, \ldots, m$. After differentiation, we get a system of $m + 1$ linear equations:

$$\left[\sum_{i=0}^{n} 2 \cdot \left[y_i - (a_0 \cdot \varphi_0(x_i) + a_1 \cdot \varphi_1(x_i) + \cdots + a_m \cdot \varphi_m(x_i))\right] \cdot (-\varphi_j(x_i))\right] = 0,$$

[1] $\frac{\partial g}{\partial a_j}$ means $g'(a_j)$.

for $j = 0, 1, \ldots, m$. We divide all equations by number $-2$ and we split them into individual sums. Then we get the following system of $m+1$ linear equations with $m+1$ unknown variables:

$$\left[\sum_{i=0}^{n}(y_i \cdot \varphi_j(x_i)) - \sum_{i=0}^{n}(a_0 \cdot \varphi_0(x_i) \cdot \varphi_j(x_i)) - \cdots - \sum_{i=0}^{n}(a_m \cdot \varphi_m(x_i) \cdot \varphi_j(x_i))\right] = 0,$$

for $j = 0, 1, \ldots, m$. In each sum we can move the corresponding coefficient $a_k$ in front of the sum and after this modification we receive so-called *normal equations* for unknown variables $a_0, a_1, \ldots, a_m$:

$$a_0 \cdot \sum_{i=0}^{n}(\varphi_0(x_i) \cdot \varphi_j(x_i)) + \cdots + a_m \cdot \sum_{i=0}^{n}(\varphi_m(x_i) \cdot \varphi_j(x_i)) = \sum_{i=0}^{n}(y_i \cdot \varphi_j(x_i)), \tag{3.4}$$

for $j = 0, 1, \ldots, m$.

In this way we obtain system of $m+1$ linear equations with $m+1$ unknowns $a_0, a_1, \ldots, a_m$, which we already know easily to solve.

We will show a special case of approximation by the least squares method using algebraic polynomials. Let us choose the functions $\varphi_i$ as: $\varphi_i(x) = x^i$, for $i = 0, 1, 2, \ldots, m$, i. e.

$$\varphi_0(x) = 1, \quad \varphi_1(x) = x, \quad \varphi_2(x) = x^2, \quad \ldots, \quad \varphi_m(x) = x^m.$$

We denote the approximate function $f_m$ as polynomial $P_m$:

$$P_m(x) = a_0 + a_1 \cdot x + a_2 \cdot x^2 + \cdots + a_m \cdot x^m. \tag{3.5}$$

We receive the sums from the system of normal equations:

$$\sum_{i=0}^{n}(\varphi_0(x_i) \cdot \varphi_0(x_i)) = \sum_{i=0}^{n}(\varphi_0(x_i))^2 = \sum_{i=0}^{n}\left(x_i^0\right)^2 = \sum_{i=0}^{n}1^2 = (n+1),$$

$$\sum_{i=0}^{n}(\varphi_0(x_i) \cdot \varphi_1(x_i)) = \sum_{i=0}^{n}\left(x_i^0 \cdot x_i^1\right) = \sum_{i=0}^{n}x_i,$$

$$\sum_{i=0}^{n}(\varphi_0(x_i) \cdot \varphi_2(x_i)) = \sum_{i=0}^{n}\left(x_i^0 \cdot x_i^2\right) = \sum_{i=0}^{n}x_i^2,$$

$$\vdots$$

$$\sum_{i=0}^{n}(\varphi_r(x_i) \cdot \varphi_s(x_i)) = \sum_{i=0}^{n}(x_i^r \cdot x_i^s) = \sum_{i=0}^{n}x_i^{r+s}, \qquad r, s \in \{0, 1, \ldots, m\}$$

The system of normal equations is in the form:

$$\begin{aligned}
a_0 \cdot (n+1) + a_1 \cdot \textstyle\sum_{i=0}^n x_i + \cdots + a_m \cdot \sum_{i=0}^n x_i^m &= \textstyle\sum_{i=0}^n y_i \\
a_0 \cdot \textstyle\sum_{i=0}^n x_i + a_1 \cdot \sum_{i=0}^n x_i^2 + \cdots + a_m \cdot \sum_{i=0}^n x_i^{m+1} &= \textstyle\sum_{i=0}^n x_i \cdot y_i \\
a_0 \cdot \textstyle\sum_{i=0}^n x_i^2 + a_1 \cdot \sum_{i=0}^n x_i^3 + \cdots + a_m \cdot \sum_{i=0}^n x_i^{m+2} &= \textstyle\sum_{i=0}^n x_i^2 \cdot y_i \\
\vdots \qquad & \quad \vdots \\
a_0 \cdot \textstyle\sum_{i=0}^n x_i^m + a_1 \cdot \sum_{i=0}^n x_i^{m+1} + \cdots + a_m \cdot \sum_{i=0}^n x_i^{2m} &= \textstyle\sum_{i=0}^n x_i^m \cdot y_i
\end{aligned}$$

For the approximation by linear function (straight line) $P_1(x) = a_0 + a_1 \cdot x$ we get system of normal equations in the form:

$$\begin{aligned}
a_0 \cdot (n+1) &+ a_1 \cdot \textstyle\sum_{i=0}^n x_i &= \textstyle\sum_{i=0}^n y_i \\
a_0 \cdot \textstyle\sum_{i=0}^n x_i &+ a_1 \cdot \textstyle\sum_{i=0}^n x_i^2 &= \textstyle\sum_{i=0}^n x_i \cdot y_i
\end{aligned}$$

and for the approximation by quadratic function (parabola) $P_2(x) = a_0 + a_1 \cdot x + a_2 \cdot x^2$ we get system of normal equations in the form:

$$\begin{aligned}
a_0 \cdot (n+1) &+ a_1 \cdot \textstyle\sum_{i=0}^n x_i &+ a_2 \cdot \textstyle\sum_{i=0}^n x_i^2 &= \textstyle\sum_{i=0}^n y_i \\
a_0 \cdot \textstyle\sum_{i=0}^n x_i &+ a_1 \cdot \textstyle\sum_{i=0}^n x_i^2 &+ a_2 \cdot \textstyle\sum_{i=0}^n x_i^3 &= \textstyle\sum_{i=0}^n x_i \cdot y_i \\
a_0 \cdot \textstyle\sum_{i=0}^n x_i^2 &+ a_1 \cdot \textstyle\sum_{i=0}^n x_i^3 &+ a_2 \cdot \textstyle\sum_{i=0}^n x_i^4 &= \textstyle\sum_{i=0}^n x_i^2 \cdot y_i
\end{aligned}$$

Besides polynomials the trigonometric functions also are used, if we assume periodic behaviour of the approximated data. For example, we can choose the functions $\varphi_i$ as follows (assuming period is $2\pi$):

$$\varphi_0(x) = 1,\ \varphi_1(x) = \cos x,\ \varphi_2(x) = \sin x,\ \varphi_3(x) = \cos 2x,\ \varphi_4(x) = \sin 2x, \ldots$$

## 3.3 Solved Examples

**Example 3.1** Write the Lagrange's interpolating polynomial for the function $f$, which is given by node points:

| $x_i$ | $-1$ | $0$ | $1$ | $2$ |
|---|---|---|---|---|
| $y_i$ | $3$ | $2$ | $1$ | $6$ |

*Solution:*
From the assignment we see, that $n = 3$. The Lagrange's interpolating polynomial has the form:

$$L_3(x) = \sum_{j=0}^{3} \prod_{\substack{i=0 \\ i\neq j}}^{3} \frac{(x - x_i)}{(x_j - x_i)} \cdot y_j =$$

$$= \frac{(x-0)\cdot(x-1)\cdot(x-2)}{((-1)-0)\cdot((-1)-1)\cdot((-1)-2)} \cdot 3 + \\ + \frac{(x-(-1))\cdot(x-1)\cdot(x-2)}{(0-(-1))\cdot(0-1)\cdot(0-2)} \cdot 2 + \\ + \frac{(x-(-1))\cdot(x-0)\cdot(x-2)}{(1-(-1))\cdot(1-0)\cdot(1-2)} \cdot 1 + \\ + \frac{(x-(-1))\cdot(x-0)\cdot(x-1)}{(2-(-1))\cdot(2-0)\cdot(2-1)} \cdot 6 = \\ = \frac{3x\cdot(x^2-3x+2)}{(-1)\cdot(-2)\cdot(-3)} + \frac{(x^2-1)\cdot(2x-4)}{1\cdot(-1)\cdot(-2)} + \\ + \frac{x\cdot(x^2-1)}{3\cdot 2\cdot 1} + \frac{x\cdot(x^2-x-2)}{2\cdot 1\cdot(-1)} = x^3 - 2x + 2.$$

√

**Example 3.2** Solve the equation $f(x) = 0$ using the inverse Lagrange's interpolation polynomial if the function $f$ is given by node points:

| $x_i$ | $0$ | $1$ | $3$ | $4$ |
|---|---|---|---|---|
| $y_i$ | $5$ | $3$ | $-1$ | $-2$ |

*Solution:*
To solve the equation $f(x) = 0$ using the inverse Lagrange interpolating polynomial, it means to determine the value of this polynomial at point zero i. e. $L_n^{\mathrm{Inv}}(0)$.[2] From the assignment we see, that $n = 3$. The inverse Lagrange's interpolating polynomial has the form:

$$L_3^{\mathrm{Inv}}(y) = \sum_{j=0}^{3} \prod_{\substack{i=0 \\ i\neq j}}^{3} \frac{(y - y_i)}{(y_j - y_i)} \cdot x_j =$$

[2] We suppose that $f(x) = 0$ if and only if $x = f^{-1}(0) \doteq L_n^{\mathrm{Inv.}}(0)$.

$$= \frac{(y-3)\cdot(y-(-1))\cdot(y-(-2))}{(5-3)\cdot(5-(-1))\cdot(5-(-2))}\cdot 0+$$
$$+\frac{(y-5)\cdot(y-(-1))\cdot(y-(-2))}{(3-5)\cdot(3-(-1))\cdot(3-(-2))}\cdot 1+$$
$$+\frac{(y-5)\cdot(y-3)\cdot(y-(-2))}{((-1)-5)\cdot((-1)-3)\cdot((-1)-(-2))}\cdot 3+$$
$$+\frac{(y-5)\cdot(y-3)\cdot(y-(-1))}{((-2)-5)\cdot((-2)-3)\cdot((-2)-(-1))}\cdot 4$$

In this case it is not necessary to know the functional formula of this polynomial, it is enough to calculate the value of the polynomial at the point zero, i. e. we calculate:

$$L_3^{\mathrm{Inv}}(0) = \sum_{j=0}^{3}\prod_{\substack{i=0\\ i\neq j}}^{3}\frac{(0-y_i)}{(y_j-y_i)}\cdot x_j =$$
$$= \frac{(0-3)\cdot(0-(-1))\cdot(0-(-2))}{(5-3)\cdot(5-(-1))\cdot(5-(-2))}\cdot 0+$$
$$+\frac{(0-5)\cdot(0-(-1))\cdot(0-(-2))}{(3-5)\cdot(3-(-1))\cdot(3-(-2))}\cdot 1+$$
$$+\frac{(0-5)\cdot(0-3)\cdot(0-(-2))}{((-1)-5)\cdot((-1)-3)\cdot((-1)-(-2))}\cdot 3+$$
$$+\frac{(0-5)\cdot(0-3)\cdot(0-(-1))}{((-2)-5)\cdot((-2)-3)\cdot((-2)-(-1))}\cdot 4=$$
$$= \frac{-6}{84}\cdot 0+\frac{-10}{-40}\cdot 1+\frac{30}{24}\cdot 3+\frac{15}{-35}\cdot 4 = \frac{1}{4}+\frac{90}{24}-\frac{60}{35} = \frac{16}{7} = 2{,}29571.$$

We get $L_3^{\mathrm{Inv}}(0) = 2{,}29571$, therefore $L_3(2{,}29571) = 0$, therefore the solution of the equation $f(x) = 0$ is $x = 2{,}29571$. $\surd$

**Example 3.3** For the values which are given in the table

| $x$ | 0,1 | 0,5 | 0,7 | 1,2 |
|---|---|---|---|---|
| $y$ | 14,61 | 3,39 | 2,14 | 0,47 |

determine by the least squares method the optimal coefficients $a$ and $b$ if the function $f$: $y = f(x)$ is given by $y(x) \sim \frac{a}{x} + b\cdot\ln(x)$. Write out a theoretical and numerical matrix of the system.

*Solution:*
The function $f_2(x)$ is given by $f_2(x) = a \cdot \varphi_0(x) + b \cdot \varphi_1(x) = a \cdot \frac{1}{x} + b \cdot \ln(x)$, where $\varphi_0(x) = \frac{1}{x}$ and $\varphi_1(x) = \ln(x)$. The theoretical matrix is:

$$\left( \begin{array}{cc|c} \sum_{i=1}^4 \varphi_0(x_i) \cdot \varphi_0(x_i) & \sum_{i=1}^4 \varphi_0(x_i) \cdot \varphi_1(x_i) & \sum_{i=1}^4 y_i \cdot \varphi_0(x_i) \\ \sum_{i=1}^4 \varphi_1(x_i) \cdot \varphi_0(x_i) & \sum_{i=1}^4 \varphi_1(x_i) \cdot \varphi_1(x_i) & \sum_{i=1}^4 y_i \cdot \varphi_1(x_i) \end{array} \right) =$$

$$= \left( \begin{array}{cc|c} \sum_{i=1}^4 \left(\frac{1}{x_i}\right)^2 & \sum_{i=1}^4 \frac{1}{x_i} \cdot \ln(x_i) & \sum_{i=1}^4 y_i \cdot \frac{1}{x_i} \\ \sum_{i=1}^4 \ln(x_i) \cdot \frac{1}{x_i} & \sum_{i=1}^4 (\ln(x_i))^2 & \sum_{i=1}^4 y_i \cdot \ln(x_i) \end{array} \right)$$

Creation of the calculation table:

Table 3.1: The Least Squares Method for $y(x) \sim \frac{a}{x} + b \cdot \ln(x)$.

| $i$ | $x_i$ | $y_i$ | $\left(\frac{1}{x_i}\right)^2$ | $(\ln(x_i))^2$ | $\frac{1}{x_i} \cdot \ln(x_i)$ | $y_i \cdot \frac{1}{x_i}$ | $y_i \cdot \ln(x_i)$ |
|---|---|---|---|---|---|---|---|
| 1 | 1,0 | 14,61 | 1,000 | 0,000 | 0,000 | 14,610 | 0,000 |
| 2 | 0,5 | 3,39 | 4,000 | 0,480 | −1,386 | 6,780 | −2,349 |
| 3 | 0,7 | 2,14 | 2,041 | 0,127 | −0,509 | 3,057 | −0,763 |
| 4 | 1,2 | 0,47 | 0,694 | 0,033 | 0,152 | 0,392 | 0,086 |
| SUM | 3,4 | 20,16 | 7,735 | 0,641 | $-1,744$ | 24,839 | −3,027 |

Based on the values in Table 3.1 (last row), we can write the numerical matrix of the system of normal equations:

$$\left( \begin{array}{cc|c} 7{,}735 & -1{,}744 & 24{,}839 \\ -1{,}744 & 0{,}641 & -3{,}027 \end{array} \right)$$

The solution of this system of normal equations is $a = \frac{10{,}64007}{1{,}91644} = 5{,}551985$ and $b = \frac{19{,}89885}{1{,}91644} = 10{,}383214$. We can approximate the function $f$: $y = f(x)$ by formula:

$$y(x) \approx \frac{5{,}551985}{x} + 10{,}383214 \cdot \ln(x).$$

$\surd$

## 3.4 Unsolved Tasks

**3.1** Construct the Lagrange interpolating polynomials for the function $f(x)$, which are given by the following tables:

a)

| $x$ | $-1$ | $0$ | $1$ | $2$ |
|---|---|---|---|---|
| $y$ | $2$ | $-1$ | $2$ | $4$ |

,

b)

| $x$ | 0,1 | 0,5 | 0,7 | 1,2 |
|---|---|---|---|---|
| $y$ | 7,2 | 3,6 | 2,4 | 0,8 |

,

c)

| $x$ | $-2$ | $0$ | $2$ | $3$ |
|---|---|---|---|---|
| $y$ | $4$ | $-1$ | $-2$ | $1$ |

,

d)

| $x$ | $-2$ | $1$ | $3$ |
|---|---|---|---|
| $y$ | $5$ | $-3$ | $5$ |

,

e)

| $x$ | $-1$ | $0$ | $2$ | $3$ | $5$ |
|---|---|---|---|---|---|
| $y$ | $12$ | $10$ | $0$ | $-4$ | $15$ |

,

f)

| $x$ | $-3$ | $-2$ | $1$ | $3$ |
|---|---|---|---|---|
| $y$ | $12$ | $0$ | $4$ | $-15$ |

.

**3.2** The function $f(x)$ is given by table. Approximate the function $f$ by linear function $g$: $y = ax + b$ by the least squares method.

a)

| $x$ | $-1$ | $2$ | $3$ | $5$ |
|---|---|---|---|---|
| $y$ | $4$ | $4$ | $5$ | $7$ |

,

b)

| $x$ | $-0{,}1$ | 0,5 | 0,7 | 1,2 |
|---|---|---|---|---|
| $y$ | 7,2 | 3,6 | 2,4 | 0,8 |

,

c)

| $x$ | $-1$ | $0$ | $2$ | $3$ | $5$ |
|---|---|---|---|---|---|
| $y$ | $6$ | $2$ | $0$ | $-6$ | $-12$ |

.

**3.3** The function $f(x)$ is given by the table:

| $x$ | 0 | 0,5 | 1,1 | 2,0 |
|---|---|---|---|---|
| $y$ | $-1{,}75$ | 0,53 | 0,95 | 1,25 |

.

By the least squares method approximate the function $f$ by the function $g$: $y = a + b \cdot \mathrm{e}^x$.

## 3.5 Results of Unsolved Tasks

**3.1** **a)** $L_3(x) = -\frac{7}{6}x^3 + 3x^2 + \frac{7}{6}x - 1$ **b)** $L_3(x) = -\frac{10}{11}x^3 + \frac{68}{11}x^2 - \frac{1367}{110}x + \frac{461}{55}$ **c)** $L_3(x) = \frac{1}{5}x^3 + \frac{1}{2}x^2 - \frac{23}{10}x - 1$ **d)** $L_2(x) = \frac{4}{3}x^2 - \frac{4}{3}x - 3$ **e)** $L_4(x) = \frac{1}{12}x^4 - \frac{5}{4}x^2 - \frac{19}{6}x + 10$ **f)** $L_3(x) = -\frac{11}{12}x^3 - \frac{1}{3}x^2 + \frac{15}{4}x + \frac{3}{2}$

**3.2** **a)** $g\colon\ y = 0{,}48x + 3{,}92$ **b)** $g\colon\ y = -2{,}89473684210526x + 3{,}21052631578947$ **c)** $g\colon\ y = -4{,}99135446685879x + 6{,}3700288184438$

**3.3** $g\colon\ y = -0{,}8085 + 0{,}3231 \cdot \mathrm{e}^x$

# Chapter 4

# The Calculation of Definite Integrals

## 4.1 Antiderivative

With the concept of derivation we met in section 1.5. In this chapter we will discuss the inverse process to differentiation of a function. We are looking for such a function, which will be equal to the derivative of the given function.

**Definition 4.1** Let a function $f$ with a domain $\mathcal{D}(f) \subseteq \mathbb{R}$ be given. A function $F$ is called an *antiderivative* of the function $f$ on an interval $I \subseteq \mathcal{D}(f)$, if for all $x \in I$ holds: $F'(x) = f(x)$. We assume that in the end points of the interval $I$, the function $F$ has at least one-sided derivatives.

The consequence of the definition of antiderivative of the function $f$ is the fact that each antiderivative of the function $f$ is the continuous function $F$ on the interval $I$, because it has the first derivative on the interval $I$.

**Theorem 4.1** Let a function $F : y = F(x)$ be antiderivative to a function $f : y = f(x)$ on the interval $I \subseteq \mathcal{D}(f) \subseteq \mathbb{R}$ and let $c \in \mathbb{R}$ be a real constant. Then also the function $G = F + c$ is an antiderivative to the function $f$ on the interval $I$.

**Theorem 4.2** Let the functions $F$ and $G$ be antiderivatives to a function $f$ on the interval $I \subseteq \mathcal{D}(f) \subseteq \mathbb{R}$. Then the function $H = F - G$ is a constant function on the interval $I$.

## 4.2 Indefinite Integral

**Definition 4.2** The set $\{F + c : c \in \mathbb{R} \wedge F' = f\}$ is called the *indefinite integral* and it is denoted by symbol $\int f(x)\,\mathrm{d}x = F(x) + c$. Finding the antiderivative to the function $f$ is called *integration* (integration of the function $f$, integrating the function $f$), the argument $x$ is called the *integration variable* and the constant $c$ is called the *constant of integration* (integration constant).

**Remark 4.1** Following relations are valid:

a) $\displaystyle\int f'(x)\,\mathrm{d}x = f(x) + c,$

b) $\displaystyle\left(\int f(x)\,\mathrm{d}x\right)' = f(x).$

**Theorem 4.3** Let a function $f : y = f(x)$ be continuous on an interval $I \subseteq \mathcal{D}(f) \subseteq \mathbb{R}$. Then for the function $f$ on the interval $I$ exists the antiderivative $F : y = F(x)$.

### 4.2.1 Integration Rules

(1) $\displaystyle\int (\alpha \cdot f(x))\,\mathrm{d}x = \alpha \cdot \int f(x)\,\mathrm{d}x$, for $\alpha \in \mathbb{R}$

(2) $\displaystyle\int (f(x) + g(x))\,\mathrm{d}x = \int f(x)\,\mathrm{d}x + \int g(x)\,\mathrm{d}x$

(3) $\displaystyle\int (f(x) - g(x))\,\mathrm{d}x = \int f(x)\,\mathrm{d}x - \int g(x)\,\mathrm{d}x$

(4) $\displaystyle\int \frac{f'(x)}{f(x)}\,\mathrm{d}x = \ln|f(x)| + c$, $c \in \mathbb{R}$

### 4.2.2 Integration Formulas

(1) $\displaystyle\int 1\,\mathrm{d}x = x + c, \quad c \in \mathbb{R}$

(2) $\displaystyle\int k\,\mathrm{d}x = k \cdot x + c, \quad \text{for } k, c \in \mathbb{R}$

(3) $\displaystyle\int x^n\,\mathrm{d}x = \frac{1}{n+1} \cdot x^{n+1} + c, \quad \text{for } n \neq -1,\ c \in \mathbb{R}$

$$(4)\quad \int a^x \,\mathrm{d}x = \frac{1}{\ln a}\cdot a^x + c, \qquad \text{for } a>0 \wedge a\neq 1,\ c\in\mathbb{R}$$

$$(5)\quad \int \mathrm{e}^x \,\mathrm{d}x = \mathrm{e}^x + c, \qquad c\in\mathbb{R}$$

$$(6)\quad \int \frac{1}{x} \,\mathrm{d}x = \ln|x| + c, \qquad c\in\mathbb{R}$$

$$(7)\quad \int \sin x \,\mathrm{d}x = -\cos x + c, \qquad c\in\mathbb{R}$$

$$(8)\quad \int \cos x \,\mathrm{d}x = \sin x + c, \qquad c\in\mathbb{R}$$

$$(9)\quad \int \frac{1}{\cos^2 x} \,\mathrm{d}x = \operatorname{tg} x + c, \qquad c\in\mathbb{R}$$

$$(10)\quad \int \frac{1}{\sin^2 x} \,\mathrm{d}x = -\operatorname{cotg} x + c, \qquad c\in\mathbb{R}$$

$$(11)\quad \int \frac{1}{\sqrt{1-x^2}} \,\mathrm{d}x = \arcsin x + c, \qquad c\in\mathbb{R}$$

$$(12)\quad \int \frac{1}{1+x^2} \,\mathrm{d}x = \operatorname{arctg} x + c, \qquad c\in\mathbb{R}$$

$$(13)\quad \int \frac{1}{1-x^2} \,\mathrm{d}x = \frac{1}{2}\cdot\ln\left|\frac{1+x}{1-x}\right| + c, \qquad c\in\mathbb{R}$$

$$(14)\quad \int \frac{1}{\sqrt{x^2+a^2}} \,\mathrm{d}x = \ln\left|x+\sqrt{x^2+a^2}\right| + c, \qquad c\in\mathbb{R}$$

$$(15)\quad \int \frac{1}{\sqrt{x^2-a^2}} \,\mathrm{d}x = \ln\left|x+\sqrt{x^2-a^2}\right| + c, \qquad c\in\mathbb{R}$$

### 4.2.3 Integration by Parts

As we may notice between the rules for integrating we lack rules for integrating the product and the division of two functions. Unlike by derivatives, by integration there does not exist universal rule for integrating such functions. *Integration by parts*[1] allows us to integrate a product of some functions and the elementary functions that are not among the formulas for integration.

[1] It also usually be called per-partes method.

**Theorem 4.4** Let the functions $u : y = u(x)$ and $v : y = v(x)$ have continuous the first derivative on the interval $I \subseteq \mathbb{R}$. Then holds:

$$\int u(x) \cdot v'(x)\,\mathrm{d}x = u(x) \cdot v(x) - \int u'(x) \cdot v(x)\,\mathrm{d}x, \tag{4.1}$$

$$\int u'(x) \cdot v(x)\,\mathrm{d}x = u(x) \cdot v(x) - \int u(x) \cdot v'(x)\,\mathrm{d}x. \tag{4.2}$$

### 4.2.4 Integration by the Substitution Method

The *Substitution method* allows us to integrate many composite functions, the product of some functions and often fraction of functions.

**Theorem 4.5** Let a function $g : y = g(x)$ be defined on an opened interval $I = (a; b) \subseteq \mathcal{D}(g) \subseteq \mathbb{R}$ with range $I_H = (c; d) \subseteq \mathcal{H}(g) \subseteq \mathbb{R}$ (i. e. $g : (a; b) \longrightarrow (c; d)$). Let the function $g$ be continuous and differentiable. Let a function $F : y = F(x)$ be antiderivative to a function $f : y = f(x)$, which is defined on an interval $J = I_H = (c; d) \subseteq \mathcal{D}(f) \subseteq \mathbb{R}$. Then the function $F(g(x))$ is antiderivative to the function $f(g(x)) \cdot g'(x)$ on the interval $I = (a; b)$.

We can formally rewrite Theorem 4.5 into the form:

$$\int f(g(x)) \cdot g'(x)\,\mathrm{d}x \left| \begin{array}{rl} g(x) & = t \\ g'(x)\,\mathrm{d}x & = \mathrm{d}t \end{array} \right| = \int f(t)\,\mathrm{d}t. \tag{4.3}$$

## 4.3 The Definite Integral

Definite integral is one of the base instruments that we can use to calculate the area of surface (plane figures).

**Definition 4.3** Let be given a non-negative and continuous function $f : y = f(x)$, where $f : \langle a; b\rangle \longrightarrow \mathbb{R}$, $\langle a; b\rangle \subseteq \mathcal{D}(f) \subseteq \mathbb{R}$. The *partition* $D_n$ of the interval $\langle a; b\rangle$ is given, which is given by $n$ separate points $D_n = \{x_0; x_1; x_2; \ldots; x_n\}$, where $a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$ and $\langle a; b\rangle = \bigcup_{i=1}^{n} \langle x_{i-1}; x_i\rangle$. Denote $\Delta x_i = x_i - x_{i-1}$ the length of the $i$-th interval for $i = 1, 2, \ldots, n$. For each positive integer $n$ we can create some partitioning $D_n$ of the interval $\langle a; b\rangle$ and we get a sequence of partitioning $\{D_n\}_{n=1}^{\infty}$ of interval $\langle a;b\rangle$. Denote the *norm of the partition* $D_n$ as a number $||D_n|| = \max\{\Delta x_1, \Delta x_2, \ldots, \Delta x_{n-1}, \Delta x_n\}$. Assume that the sequence of partitioning $\{D_n\}_{n=1}^{\infty}$ is *normal* (i. e. $\lim_{x\to\infty} ||D_n|| = 0$). Let $\xi_i$ be an arbitrary point of the interval $\langle x_{i-1}; x_i\rangle$ for $i = 1, 2, 3, \ldots, n$. Denote $S(f, D_n)$ the *integral sum* of the function $f$ for partition $D_n$ of the interval $\langle a;b\rangle$ and for any choice of points $\xi_1; \xi_2; \ldots; \xi_n$, where $S(f, D_n) = \sum_{i=1}^{n} f(\xi_i) \cdot \Delta x_i$.[2] Then if exists the number

$$S = \lim_{n\to\infty} S(f, D_n) \tag{4.4}$$

(the limit of a sequence of integral sums of the function $f$ on the interval $\langle a;b\rangle$), it is called the *definite integral* (Riemann integral) of function $f$ on the interval $\langle a;b\rangle$. The definite integral is denoted by the symbol $\int_a^b f(x)\,\mathrm{d}x$.[3] The number $a$ is called the *lower limit* and the number $b$ is called the *upper limit* of the definite integral.

**Definition 4.4** Let a function $f$ be defined and bounded on a closed interval $\langle a;b\rangle$. The function $f$ is *integrable* on the interval $\langle a;b\rangle$, if for every normal sequence $\{D_n\}_{n=1}^{\infty}$ on the interval $\langle a;b\rangle$, the sequence of integral sums $\{S(f, D_n)\}_{n=1}^{\infty}$ of the function $f$ for partitioning $D_n$ of the interval $\langle a;b\rangle$ and arbitrary selection $\xi_1; \xi_2; \ldots; \xi_n$ for the partition $D_n$ is convergent.

**Remark 4.2** If we mark a planar shape – curvilinear trapezoid as a set of points $K = \{[x,y] \in \mathbb{R}^2 : a \leqq x \leqq b \wedge 0 \leqq y \leqq f(x)\}$, then the surface area of this region is equal to the real number $S$, i. e. value of the definite integral $\int_a^b f(x)\,\mathrm{d}x$.

**Theorem 4.6** The function $f$ is integrable on $\langle a,b\rangle$ with integral $S$ if and only if $\lim_{n\to\infty} S(f, D_n) = S$ for every sequence $\{S(f, D_n)\}_{n=1}^{\infty}$ of Riemann sums associated with a sequence of partitions $\{D_n\}_{n=1}^{\infty}$ of $\langle a,b\rangle$ such that $||D_n|| \longrightarrow 0$ as $n \longrightarrow \infty$.

**Remark 4.3** Based on the Theorem 4.6 we can express the definite integral $S$ as the limit:

$$S = \lim_{||D_n||\to 0} S(f, D_n)$$

[2]The set of points $\{\xi_1; \xi_2; \ldots; \xi_n\}$ is called a selection for the partition $D_n$.

[3]The summing definition of integral introduced Cauchy for continuous functions and Riemann for discontinuous functions. Title definite integral introduced Bernoulli and symbol $\int$ is actually the letter S as the first letter of the word sum which used Leibniz, Bernoulli's teacher. Therefore, the definite integral is called the Riemann integral or the Cauchy-Riemann integral.

**Theorem 4.7** (*The Newton-Leibniz Formula*) Let a function $f : y = f(x)$ be integrable on a closed interval $\langle a, b\rangle$ and let it has on the interval $\langle a, b\rangle$ antiderivative $F : y = F(x)$. Then holds:

$$\int\limits_a^b f(x)\,\mathrm{d}x = \Big[F(x)\Big]_a^b = F(b) - F(a). \tag{4.5}$$

### 4.3.1 Properties of Definite Integrals

**Theorem 4.8** Suppose that functions $f$ and $g$ are integrable. Then holds:

(1) $\int_a^b f(x)\,\mathrm{d}x = -\int_b^a f(x)\,\mathrm{d}x$,

(2) $\int_a^a f(x)\,\mathrm{d}x = 0$,

(3) $\int_a^b (f(x) + g(x))\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x$,

(4) $\int_a^b \alpha \cdot f(x)\,\mathrm{d}x = \alpha \cdot \int_a^b f(x)\,\mathrm{d}x$, where $\alpha \in \mathbb{R}$,

(5) $\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x$, where $c \in \langle a, b\rangle$.

**Theorem 4.9** (*Existence of the Integral*) If a function $f : y = f(x)$ is continuous on an interval $\langle a, b\rangle$, then the function $f$ is integrable on this interval.

**Theorem 4.10** Let a function $f : y = f(x)$ be bounded and has a finite number of points of discontinuity on an interval $\langle a, b\rangle$. Then the function $f : y = f(x)$ is integrable on the interval $\langle a, b\rangle$.

**Theorem 4.11** Let a function $f : y = k$ be the constant function ($k \in \mathbb{R}$). Then holds:

$$\int\limits_a^b k\,\mathrm{d}x = k \cdot (b - a).$$

**Theorem 4.12** Let a function $f: y = f(x)$ be integrable and non-negative on an interval $\langle a, b\rangle$. Then holds:

$$\int_a^b f(x)\,\mathrm{d}x \geqq 0.$$

**Theorem 4.13** Let functions $f: y = f(x)$ and $g: y = g(x)$ be integrable on an interval $\langle a, b\rangle$. Let $f(x) \leqq g(x)$ for all values $x \in \langle a, b\rangle$. Then holds:

$$\int_a^b f(x)\,\mathrm{d}x \leqq \int_a^b g(x)\,\mathrm{d}x.$$

**Theorem 4.14** (*Integration by Parts*) Let functions $u: y = u(x)$ and $v: y = v(x)$ be continuously differentiable on a closed interval $\langle a, b\rangle$. Then holds:

$$\int_a^b u(x)\cdot v'(x)\,\mathrm{d}x = \Big[u(x)\cdot v(x)\Big]_a^b - \int_a^b u'(x)\cdot v(x)\,\mathrm{d}x. \tag{4.6}$$

**Theorem 4.15** (*Substitution*) Let a function $f: y = f(x)$ be continuous on a closed interval $\langle a, b\rangle$. Let a function $g: y = g(x)$ has a continuous the first derivative on bounded and the closed interval $\langle c, d\rangle$, while it maps the interval $\langle c, d\rangle$ to the interval $\langle a, b\rangle$. Then holds:

$$\int_a^b f(x)\,\mathrm{d}x = \int_c^d f(g(t))\cdot g'(t)\,\mathrm{d}t, \tag{4.7}$$

where $g(c) = a$ and $g(d) = b$.

### 4.3.2 Aplication of Definite Integrals

**Definition 4.5** Let functions $f: \langle a, b\rangle \longrightarrow \mathbb{R}$ and $g: \langle a, b\rangle \longrightarrow \mathbb{R}$ be continuous and let for each $x \in \langle a, b\rangle$: $f(x) \leqq g(x)$. Then the set $D = \{[x, y] \in \mathbb{R}^2 : a \leqq x \leqq b \wedge f(x) \leqq y \leqq g(x)\}$ is called the *elemental area* (region) in $\mathbb{R}^2$ with respect to the $x$ axis $o_x$.

**Definition 4.6** Let functions $\phi : \langle c, d\rangle \longrightarrow \mathbb{R}$ and $\psi : \langle c, d\rangle \longrightarrow \mathbb{R}$ be continuous and for each $y \in \langle c, d\rangle$ is $\phi(y) \leqq \psi(y)$. Then the set $H = \{[x, y] \in \mathbb{R}^2 : c \leqq y \leqq d \wedge \phi(y) \leqq x \leqq \psi(y)\}$ is called the *elemental area* (region) in $\mathbb{R}^2$ with respect to the $y$ axis $o_y$.

**Theorem 4.16** A *surface area* of the elemental area $D$ from Definition 4.5 is calculated by the formula:

$$P = \int_a^b \Big(g(x) - f(x)\Big)\,\mathrm{d}x. \tag{4.8}$$

**Theorem 4.17** A *surface area* of the elemental area $H$ from Definition 4.6 is calculated by the formula:

$$P = \int_c^d (\psi(y) - \phi(y))\,\mathrm{d}y. \tag{4.9}$$

**Theorem 4.18** If the *curve* $\gamma$ is a graph of the function $f : \langle a, b\rangle \longrightarrow \mathbb{R}$, having a continuous the first derivative, then for its *length* $d$ holds formula:

$$d = \int_a^b \sqrt{1 + (f'(x))^2}\,\mathrm{d}x. \tag{4.10}$$

**Theorem 4.19** Let a function $f : y = f(x)$ be continuous and non-negative on a closed interval $\langle a, b\rangle$, which creates in a plane curvilinear trapezoid $L$ over the interval $\langle a, b\rangle$. By rotation of curvilinear trapezium $L$ in the space $\mathbb{R}^3$ with axes $o_x$, $o_y$, and $o_z$ around an axis $o_x$ arises *rotating shape* whose volume $V$ is computed using the volume formula:

$$V = \lim_{x\to\infty} \sum_{i=1}^{n} \pi \cdot f^2(\xi_i) \cdot \Delta x_i = \pi \cdot \int_a^b f^2(x)\,\mathrm{d}x. \tag{4.11}$$

**Theorem 4.20** Let a function $f : y = f(x)$ be continuous, differentiable and non-negative on the closed interval $\langle a, b\rangle$, which creates in a plane curvilinear trapezoid $L$ over the interval $\langle a, b\rangle$. By rotation of curvilinear trapezium $L$ in the space $\mathbb{R}^3$ with axes $o_x$, $o_y$, and $o_z$ about the axis $o_x$ arises *rotating shape* whose surface area of rotating area $S$ is computed using the formula:

$$S = 2 \cdot \pi \cdot \int_a^b f(x) \cdot \sqrt{1 + (f'(x))^2}\, \mathrm{d}x. \tag{4.12}$$

## 4.4 Numerical Integration

As we already know from previous chapters, not every function must have a primitive function (antiderivative), or to find the primitive function can be very complicated. Unfamiliarity of primitive function does not allow us to use the Newton-Leibniz formula to calculate the definite integral. We can still use the definition of the definite integral, but its use is not very practical for calculations. It is therefore appropriate to suitably adjust the definition of the definite integral so we can approximate it more easily. For this we use the so-called Newton-Cotes formulas.

### 4.4.1 Newton-Cotes Formulas

We want to calculate the definite integral of a function $f$: $y = f(x)$ on an interval $\langle a, b\rangle$:

$$\int_a^b f(x)\, \mathrm{d}x. \tag{4.13}$$

We obtain the *Newton-Cotes quadrature formula* (quadrature formula) by integrating the interpolating polynomial with equidistant nodes, which approximates the integrand $f$ (the expression to be integrated). These formulas can be divided into two basic groups:

a) *closed formulas*, where we take the interval endpoints for quadrature nodes,

b) *open formulas*, where we do not take the endpoints for quadrature nodes, but the nodes are symmetrical according to the centre of the interval.

In this chapter we will discuss one open formula (rectangular method) and two closed formulas (trapezoidal method and Simpson's method). The difference between these methods is by which polynomial function is the integrand $f$ approximated. When we apply the rectangular method we use the polynomial of the zero degree (constant function). When we apply the trapezoidal method we use the polynomial of the first degree (linear function – straight line), and when we apply the Simpson's method we use the polynomial of the second degree (quadratic function – a parabola).

We know that the definite integral of a positive function $f$ on the interval $\langle a, b\rangle$ corresponds to the surface area of the plane, which is surrounded, from above by the function $f$, from below by the $x$ axis $o_x$ and on the both sides by the straight lines $x = a$ and $x = b$. With the rectangular method we calculate the area of the rectangles that approximate that area. With the trapezoidal method we calculate the surface area of trapezoids that approximate the given area, and with Simpson's method we calculate surface of areas that are bound from above by parabola.

Let the integral (4.13) be given. We divide the interval $\langle a, b\rangle$ into the $n$ equal parts, using $n+1$ nodal points $x_0, x_1, x_2, \ldots, x_n$. We say the partition $D_n = \{x_0, x_1, x_2, \ldots, x_n\}$ is *regular*. Let $h \in \mathbb{R}$, for which holds:

$$h = \frac{(b-a)}{n}. \tag{4.14}$$

Then we can express the nodal points using the constant $h$ as follows:

$$x_{i+1} = x_i + h, \tag{4.15}$$

where $i = 0, 1, 2, \ldots, n-1$ and $x_0 = a$, $x_n = b$.

**Remark 4.4** As explained in Definition 4.3 , the choice of $\xi_i$ can be anything within the interval $\langle x_{i-1}, x_i\rangle$. Depending on our choices, we can have a vast variety of sums. However, there are three main sums that relate and are usable to our cases:

(1) *Left Riemann Sum*: $S_{left}(f, D_n) = \sum\limits_{i=1}^{n} f(x_{i-1}) \cdot \Delta_i$, we choice the left endpoint $\xi_i = x_{i-1}$ (the left endpoint sum),

(2) *Right Riemann Sum*: $S_{right}(f, D_n) = \sum\limits_{i=1}^{n} f(x_i) \cdot \Delta_i$, we choice the right endpoint $\xi_i = x_i$ (the right endpoint sum),

(3) *Middle Riemann Sum*: $S_{mid}(f, D_n) = \sum\limits_{i=1}^{n} f(\xi_i) \cdot \Delta_i$, we choice the midpoint $\xi_i = \frac{x_{i-1}+x_i}{2}$ (the midpoint sum).

### 4.4.2 Rectangular Method

Let the integral (4.13) on page 105 be given. We approximate the function $f$ by the constant function $f_i$: $y = f\left(\frac{x_i+x_{i+1}}{2}\right)$ for $i = 0, 1, 2, \ldots, n-1$ at each interval $\langle x_i, x_{i+1}\rangle$, for $i = 0, 1, 2, \ldots, n-1$. The area of the rectangle over the interval $\langle x_i, x_{i+1}\rangle$, for $i \in \{0, 1, 2, \ldots, n-1\}$ can be expressed by the formula:

$$s_i = (x_{i+1} - x_i) \cdot f\left(\frac{x_i + x_{i+1}}{2}\right). \tag{4.16}$$

Then for the sum of the of all $n$ areas of rectangles applies:

$$s(n) = \sum_{i=0}^{n-1} s_i = \sum_{i=0}^{n-1} (x_{i+1} - x_i) \cdot f\left(\frac{x_i + x_{i+1}}{2}\right).$$

Given this relations (4.14) and (4.15) we can modify the sum in to the following form:

$$s(n) = \sum_{i=0}^{n-1} h \cdot f\left(\frac{x_i + x_i + h}{2}\right) = h \cdot \sum_{i=0}^{n-1} f\left(x_i + \frac{h}{2}\right). \tag{4.17}$$

The resulting relation for calculation of the integral by the *rectangular method* with number of subintervals $n$ has the form:

$$\int_a^b f(x)\,\mathrm{d}x \approx s(n) = h \cdot \sum_{i=1}^{n} f\left(a + (2i-1) \cdot \frac{h}{2}\right). \tag{4.18}$$

For estimation of the error holds:

$$\left| s(n) - \int_a^b f(x)\,\mathrm{d}x \right| \leqq \frac{(b-a)^3}{24n^2} \cdot M_2, \tag{4.19}$$

where $M_2 \geqq \max_{x \in \langle a,b\rangle} |f''(x)|$.

**Remark 4.5** The formula (4.18) is also called *Midpoint Rule*, if we use the Middle Riemann Sum. If we use the Left-handed Riemann sum, then we say *Rectangular Rule* (the left endpoint of the subinterval is used). Then the formula (4.18) is written as:

$$\int_a^b f(x)\,\mathrm{d}x \approx s(n) = h \cdot \sum_{i=0}^{n-1} f\left(a + i \cdot h\right).$$

For estimation of the error holds:

$$\left| s(n) - \int_a^b f(x)\,\mathrm{d}x \right| \leqq \frac{(b-a)^2}{2n} \cdot M_1,$$

where $M_1 \geqq \max_{x \in \langle a,b \rangle} |f'(x)|$.

### 4.4.3 Trapezoidal Method

Let the integral (4.13) on page 105 be given. We divide the interval $\langle a, b \rangle$ on $n$ equal parts, by $n+1$ nodal points $x_0, x_1, x_2, \ldots, x_n$. The length of each interval $\langle x_i, x_{i+1} \rangle$ is equal to the number $h \in \mathbb{R}$, for which the relation (4.14) holds. Then we can express the nodal points by using the relation (4.15), where $i \in \{0, 1, 2, \ldots, n-1\}$, $x_0 = a$ and $x_n = b$. We approximate the function $f$ by a linear function $f_i$: $y = f_i(x)$, which passes through the points $[x_i, f(x_i)]$ and $[x_{i+1}, f(x_{i+1})]$ at each interval $\langle x_i, x_{i+1} \rangle$ for $i = 0, 1, 2, \ldots, n-1$. We approximate the surface of the original area over the interval $\langle x_i, x_{i+1} \rangle$ using the area of trapezoid over the interval $\langle x_i, x_{i+1} \rangle$ for $i \in \{0, 1, 2, \ldots, n-1\}$, which can be expressed by the formula:

$$s_i = \frac{1}{2} \cdot h \cdot \left(f(x_i) + f(x_{i+1})\right). \tag{4.20}$$

Then for the integral (4.13) holds:

$$\int_a^b f(x)\,\mathrm{d}x = \int_{x_0}^{x_1} f(x)\,\mathrm{d}x + \int_{x_1}^{x_2} f(x)\,\mathrm{d}x + \int_{x_2}^{x_3} f(x)\,\mathrm{d}x + \cdots$$

$$\cdots + \int_{x_{n-2}}^{x_{n-1}} f(x)\,\mathrm{d}x + \int_{x_{n-1}}^{x_n} f(x)\,\mathrm{d}x \approx$$

$$\approx s_0 + s_1 + s_2 + \cdots + s_{n-2} + s_{n-1} =$$

$$= \frac{h}{2} \cdot (f(x_0) + f(x_1)) + \frac{h}{2} \cdot (f(x_1) + f(x_2)) + \frac{h}{2} \cdot (f(x_2) + f(x_3)) + \cdots$$

$$\cdots + \frac{h}{2} \cdot (f(x_{n-2}) + f(x_{n-1})) + \frac{h}{2} \cdot (f(x_{n-1}) + f(x_n)) =$$

$$= \frac{h}{2} \cdot \Big[ (f(x_0) + f(x_1)) + (f(x_1) + f(x_2)) + (f(x_2) + f(x_3)) + \cdots$$

$$\cdots + (f(x_{n-2}) + f(x_{n-1})) + (f(x_{n-1}) + f(x_n)) \Big] =$$

$$= \frac{h}{2} \cdot \Big[ f(x_0) + 2 \cdot f(x_1) + 2 \cdot f(x_2) + 2 \cdot f(x_3) + \cdots + 2 \cdot f(x_{n-2}) + 2 \cdot f(x_{n-1}) + f(x_n) \Big].$$

Thus:

$$\int\limits_a^b f(x)\,\mathrm{d}x \approx \frac{h}{2} \cdot \left[ f(x_0) + f(x_n) + 2 \cdot \sum_{i=1}^{n-1} f(x_i) \right]. \tag{4.21}$$

The following holds for the error estimation:

$$\left| s(n) - \int_a^b f(x)\,\mathrm{d}x \right| \leqq \frac{(b-a)^3}{12n^2} \cdot M_2, \tag{4.22}$$

where $M_2 \geqq \max_{x \in \langle a,b \rangle} |f''(x)|$.
For estimation of errors we also use the relation in the form:

$$\left| s(n) - \int_a^b f(x)\,\mathrm{d}x \right| \leqq \frac{(b-a)}{12} \cdot h^2 \cdot M_2, \tag{4.23}$$

where $M_2 \geqq \max_{x \in \langle a,b \rangle} |f''(x)|$.

### 4.4.4 Simpson's Method

In *Simpson's method* we proceed similarly to the trapezoidal method. We determine the approximate value of the integral (4.13) on page 105. Let us divide the interval $\langle a, b \rangle$ to the even number $n = 2m$, $m \in \mathbb{N}$ of equal parts, by $n + 1$ nodal points $x_0, x_1, x_2, \ldots, x_n$. The length of each interval $\langle x_i, x_{i+1} \rangle$ is equal to the number $h \in \mathbb{R}$, for which the relation (4.14) holds. Then we can express the nodal points by using the relation (4.15), where $i \in \{0, 1, 2, \ldots, n-1\}$, $x_0 = a$ and $x_n = b$. We approximate the function

$f$ by the quadratic function $f_i$: $y = f_i(x)$, which passes through the points $[x_{2i}, f(x_{2i})]$, $[x_{2i+1}, f(x_{2i+1})]$, and $[x_{2i+2}, f(x_{2i+2})]$ at each interval $\langle x_{2i}, x_{2i+2}\rangle$, for $i = 0, 1, 2, \ldots, m-1$. We replace the area of the original region over the interval $\langle x_{2i}, x_{2i+2}\rangle$ by the area of a curvilinear trapezoid over the interval $\langle x_{2i}, x_{2i+2}\rangle$, for $i \in \{0, 1, 2, \ldots, m-1\}$, which can be expressed by the formula:

$$s_i = \frac{1}{3} \cdot h \cdot \left(f(x_i) + 4 \cdot f(x_{i+1}) + f(x_{i+2})\right). \tag{4.24}$$

Then for the integral (4.13) holds:

$$\int_a^b f(x)\,\mathrm{d}x = \int_{x_0}^{x_2} f(x)\,\mathrm{d}x + \int_{x_2}^{x_4} f(x)\,\mathrm{d}x + \int_{x_4}^{x_6} f(x)\,\mathrm{d}x + \cdots$$

$$\cdots + \int_{x_{n-4}}^{x_{n-2}} f(x)\,\mathrm{d}x + \int_{x_{n-2}}^{x_n} f(x)\,\mathrm{d}x \approx$$

$$\approx s_0 + s_1 + s_2 + \cdots + s_{m-2} + s_{m-1} =$$

$$= \frac{h}{3} \cdot (f(x_0) + 4f(x_1) + f(x_2)) + \frac{h}{3} \cdot (f(x_2) + 4f(x_3) + f(x_4)) +$$

$$+ \frac{h}{3} \cdot (f(x_4) + 4f(x_5) + f(x_6)) + \cdots +$$

$$+ \frac{h}{3} \cdot (f(x_{n-4}) + 4f(x_{n-3}) + f(x_{n-2})) + \frac{h}{3} \cdot (f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)) =$$

$$= \frac{h}{3} \cdot \Big[ (f(x_0) + 4f(x_1) + f(x_2)) + (f(x_2) + 4f(x_3) + f(x_4)) +$$

$$+ (f(x_4) + 4f(x_5) + f(x_6)) + \cdots +$$

$$+ (f(x_{n-4}) + 4f(x_{n-3}) + f(x_{n-2})) + (f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)) \Big] =$$

$$= \frac{h}{3} \cdot \Big[f(x_0) + 4 \cdot f(x_1) + 2 \cdot f(x_2) + 4 \cdot f(x_3) + \cdots + 2 \cdot f(x_{n-2}) + 4 \cdot f(x_{n-1}) + f(x_n)\Big].$$

Thus:

$$\int_a^b f(x)\,\mathrm{d}x \approx \frac{h}{3} \cdot \left[ f(x_0) + f(x_n) + 4 \cdot \sum_{i=1}^{\frac{n}{2}} f\left(x_{2i-1}\right) + 2 \cdot \sum_{i=1}^{\frac{n-2}{2}} f\left(x_{2i}\right) \right]. \tag{4.25}$$

For the error estimate holds:

$$\left| s(n) - \int_a^b f(x)\,\mathrm{d}x \right| \leqq \frac{(b-a)^5}{180n^4} \cdot M_4, \tag{4.26}$$

where $M_4 \geqq \max_{x\in\langle a,b\rangle} |f^{(IV)}(x)|$.

For estimation of errors we also use the relation in the form:

$$\left| s(n) - \int_a^b f(x)\,\mathrm{d}x \right| \leqq \frac{(b-a)}{180} \cdot h^4 \cdot M_4, \tag{4.27}$$

where $M_4 \geqq \max_{x\in\langle a,b\rangle} |f^{(IV)}(x)|$.

When we want to use the Simpson's method we must not forget that the number of dividing points of the interval $\langle a, b\rangle$ must be always the odd (i. e. $n$ is the even). This follows from the relation (4.24) on the page 110. We must pay attention, when we estimate the number of dividing points $n$ for a given $\varepsilon$, because the nearest larger non-negative integer can be also the odd but we need the even $n$.

## 4.5 Solved Examples

**Example 4.1** Calculate the following indefinite integrals:

(a) $\displaystyle\int \left(4x^2 - 2x + 3 - \sqrt{x} - \frac{1}{\sqrt[4]{x^3}}\right) \mathrm{d}x,$

(b) $\displaystyle\int \frac{(2-x)^2}{x^3}\,\mathrm{d}x,$

(c) $\displaystyle\int x^2 \cdot \cos x\,\mathrm{d}x,$

(d) $\displaystyle\int \frac{(\operatorname{arctg} x)^5}{x^2+1}\,\mathrm{d}x.$

*Solution:*
(a) Integral is rewritten using the rules for integration so that we can use the integration formula.

$$\int \left(4x^2 - 2x + 3 - \sqrt{x} - \frac{1}{\sqrt[4]{x^3}}\right) \mathrm{d}x =$$
$$= 4 \cdot \int x^2 \,\mathrm{d}x - 2 \cdot \int x \,\mathrm{d}x + 3 \cdot \int 1 \,\mathrm{d}x - \int x^{\frac{1}{2}} \,\mathrm{d}x - \int x^{-\frac{3}{4}} \,\mathrm{d}x =$$
$$= 4 \cdot \frac{x^{2+1}}{2+1} - 2 \cdot \frac{x^{1+1}}{1+1} + 3 \cdot x - \frac{x^{\frac{1}{2}+1}}{\frac{1}{2}+1} - \frac{x^{-\frac{3}{4}+1}}{-\frac{3}{4}+1} + c =$$
$$= 4 \cdot \frac{x^3}{3} - 2 \cdot \frac{x^2}{2} + 3 \cdot x - \frac{x^{\frac{3}{2}}}{\frac{3}{2}} - \frac{x^{\frac{1}{4}}}{\frac{1}{4}} + c =$$
$$= \frac{4}{3}x^3 - x^2 + 3x - \frac{2}{3}\sqrt{x^3} - 4\sqrt[4]{x} + c.$$

(b) We modify the integrand to such form that we can use the rules for integration and integration formulas.

$$\int \frac{(2-x)^2}{x^3} \,\mathrm{d}x = \int \left(\frac{4 - 2x + x^2}{x^3}\right) \mathrm{d}x = \int \left(\frac{4}{x^3} - \frac{2x}{x^3} + \frac{x^2}{x^3}\right) \mathrm{d}x =$$
$$= \int \frac{4}{x^3} \,\mathrm{d}x - \int \frac{2x}{x^3} \,\mathrm{d}x + \int \frac{x^2}{x^3} \,\mathrm{d}x = 4 \cdot \int x^{-3} \,\mathrm{d}x - 2 \cdot \int x^{-2} \,\mathrm{d}x + \int \frac{1}{x} \,\mathrm{d}x =$$
$$= 4\frac{x^{-3+1}}{-3+1} - 2\frac{x^{-2+1}}{-2+1} + \ln|x| + c = 4\frac{x^{-2}}{-2} - 2\frac{x^{-1}}{-1} + \ln|x| + c =$$
$$= \frac{-2}{x^2} + \frac{2}{x} + \ln|x| + c.$$

(c) Integrand is the product of a polynomial and trigonometric function $\cos x$. In such cases, we integrate this function using the integration by parts.

$$\int x^2 \cdot \cos x \,\mathrm{d}x \overset{\text{PP}}{=} \left| \begin{array}{ll} u = x^2 & v' = \cos x \\ u' = 2x & v = \sin x \end{array} \right| = x^2 \cdot \sin x - \int 2x \cdot \sin x \,\mathrm{d}x =$$
$$= x^2 \cdot \sin x - 2 \cdot \int x \cdot \sin x \,\mathrm{d}x \overset{\text{PP}}{=} \left| \begin{array}{ll} u = x & v' = \sin x \\ u' = 1 & v = -\cos x \end{array} \right| =$$
$$= x^2 \cdot \sin x - 2 \cdot \left[x \cdot (-\cos x) - \int 1 \cdot (-\cos x) \,\mathrm{d}x\right] =$$
$$= x^2 \cdot \sin x + 2 \cdot x \cdot \cos x - 2 \cdot \int \cos x \,\mathrm{d}x =$$
$$= x^2 \cdot \sin x + 2x \cdot \cos x - 2 \cdot \sin x + c.$$

(d) If we look at integrand and we realize what it is equal to the derivative of the function $\arctan x$, so for the calculation of indefinite integral is convenient to use the substitution method.

$$\int \frac{(\operatorname{arctg} x)^5}{x^2+1}\,\mathrm{d}x \stackrel{\text{Sub}}{=} \left| \begin{array}{lcl} \operatorname{arctg} x & = & t \\ \dfrac{1}{1+x^2}\,\mathrm{d}x & = & 1\,\mathrm{d}t \end{array} \right| = \int t^5\,\mathrm{d}t = \frac{t^{5+1}}{5+1} + c =$$
$$= \frac{t^6}{6} + c = \frac{1}{6}\cdot t^6 + c = \frac{1}{6}\cdot(\operatorname{arctg} x)^6 + c. \qquad \checkmark$$

**Example 4.2** Calculate the given definite integral by the trapezoidal method with the accuracy $\varepsilon = 10^{-3}$:

$$\int_1^2 \ln(x^2+4)\,\mathrm{d}x.$$

*Solution:*
Denote integrand as the function $f\colon\ y = \ln(x^2+4)$. To determine the number of subintervals $n$, we need to calculate the first and the second derivative of the function $f$, and then to use the estimate of the error (4.22) in calculating the definite integral by the trapezoidal method.

$$y' = \Big(\ln(x^2+4)\Big)' = \frac{2x}{x^2+4} \qquad y'' = \left(\frac{2x}{x^2+4}\right)' = \frac{8-2x^2}{(x^2+4)^2}.$$

We calculate the value:

$$M_2 \geqq \max_{x\in\langle a,b\rangle} |f''(x)| = \max_{x\in\langle 1,2\rangle} \left|\frac{8-2x^2}{(x^2+4)^2}\right| = 0{,}24.$$

To determine the number of dividing $n$ we use the relation (4.22) on page 109, from which we express the variable $n$. We get the inequality:

$$n \geqq \sqrt{\frac{(b-a)^3\cdot M_2}{12\cdot\varepsilon}} = \sqrt{\frac{1^3\cdot 0{,}24}{12\cdot 0{,}001}} = \sqrt{20} \doteq 4{,}47.$$

We know that $n\in\mathbb{N}$ therefore, from the inequality $n \geqq 4{,}47$ implies that it is sufficient to choose $n = 5$. For $h$ holds: $h = \dfrac{b-a}{n} = 0{,}2$. We create the table for the calculation of function values (see Table 4.1).

Table 4.1: Trapezoidal method.

| $k$ | $x_k$ | $y_k = f(x_k)$ |
|---|---|---|
| 0 | 1,0 | 1,609438 |
| 1 | 1,2 | 1,693779 |
| 2 | 1,4 | 1,785071 |
| 3 | 1,6 | 1,880991 |
| 4 | 1,8 | 1,979621 |
| 5 | 2,0 | 2,079442 |

We can calculate the approximate value of the specified definite integral based on the values from the Table 4.1:

$$\int\limits_1^2 \ln\left(x^2+4\right) \mathrm{d}x \approx \frac{h}{2} \cdot \left[(y_0 + y_5) + 2 \cdot \sum_{i=1}^{4} y_i\right] =$$

$$= \frac{0{,}2}{2} \cdot \Big[(1{,}609438 + 2{,}079442) + 2 \cdot (1{,}693779 + 1{,}785071 + 1{,}880991 + 1{,}979621)\Big] =$$

$$= 0{,}1 \cdot 18{,}36780216 = 1{,}83678.$$

The approximate value of the definite integral is:

$$\int\limits_1^2 \ln\left(x^2+4\right) \mathrm{d}x \approx 1{,}83678.$$

We estimate the upper bound of the error of our calculation using the relation (4.22):

$$\left| s(5) - \int_a^b f(x)\,\mathrm{d}x \right| \leqq \frac{(b-a)^3}{12n^2} \cdot M_2 = \frac{(2-1)^3}{12 \cdot 5^2} \cdot 0{,}24 = 0{,}0008 < 0{,}001.$$

√

# 4.6 Unsolved Tasks

**4.1** Calculate the given indefinite integrals.

a) $\displaystyle\int \left(x^6 - 4 \cdot x^3 + 3\right) \mathrm{d}x,$

b) $\displaystyle\int \left(1 - 2 \cdot x^2\right)^2 \mathrm{d}x,$

c) $\displaystyle\int \frac{(x-2)^3}{x^2} \mathrm{d}x,$

d) $\displaystyle\int \frac{\sqrt[5]{x} - 6 \cdot \sqrt[4]{x^3}}{\sqrt{x}} \mathrm{d}x,$

e) $\displaystyle\int (5 \cdot \sin^4 x \cdot \cos x) \mathrm{d}x,$

f) $\displaystyle\int \operatorname{tg} x \ \mathrm{d}x,$

g) $\displaystyle\int \operatorname{tg}^2 x \, \mathrm{d}x,$

h) $\displaystyle\int \mathrm{e}^x \cdot (x^2 + 3 \cdot x - 4) \mathrm{d}x,$

i) $\displaystyle\int \sin(3x - 7) \mathrm{d}x,$

j) $\displaystyle\int \mathrm{e}^{8-3x} \mathrm{d}x,$

k) $\displaystyle\int \mathrm{e}^{1-4x^2} \cdot 2x \, \mathrm{d}x,$

l) $\displaystyle\int (1 - 6x) \cdot \ln(2x - 6x^2) \mathrm{d}x,$

m) $\displaystyle\int \frac{3x^2 + 2}{x^3 + 2x - 6} \mathrm{d}x,$

n) $\displaystyle\int \frac{\ln x}{2x} \mathrm{d}x,$

o) $\displaystyle\int \frac{\ln^2 x}{6x} \mathrm{d}x,$

p) $\int \ln x \, \mathrm{d}x,$

r) $\int (x^3 + 3x) \cdot \ln x \, \mathrm{d}x,$

s) $\int \frac{\operatorname{tg}^3 x}{\cos^2 x} \, \mathrm{d}x,$

t) $\int (4x + 2)^{15} \, \mathrm{d}x,$

u) $\int \frac{1}{x \cdot \ln^4 x} \, \mathrm{d}x,.$

**4.2** Calculate the given definite integrals.

a) $\int_{-1}^{1} \left(x^3 + 4x^2 - 2x + 1\right) \mathrm{d}x,$

b) $\int_{0}^{1} \left(x^3 \cdot \mathrm{e}^{2x}\right) \mathrm{d}x,$

c) $\int_{\frac{\pi}{4}}^{\frac{\pi}{2}} 2 \cdot \sin x \cdot \cos x \, \mathrm{d}x,$

d) $\int_{1}^{\mathrm{e}} x \cdot \ln x \, \mathrm{d}x,$

e) $\int_{3}^{9} \left(\frac{1}{3}x - 2\right)^8 \mathrm{d}x,$

f) $\int_{0}^{\pi} \operatorname{tg} x \, \mathrm{d}x,$

g) $\int_{0}^{1} 2x \cdot \mathrm{e}^{x^2} \, \mathrm{d}x.$

**4.3** Calculate definite integrals using the trapezoidal method for the given number of subintervals, $n$ on the interval $I = \langle a, b \rangle$ and extimate the upper bound of the error.

a) $\int_{-1}^{1} \left(2x^3 - 5x + 1\right) \mathrm{d}x, \qquad n = 16,$

b) $\int_{0}^{1,2} \left(\mathrm{e}^{2x^2}\right) \mathrm{d}x, \qquad n = 10,$

c) $\displaystyle\int_{\frac{\pi}{4}}^{\frac{\pi}{2}} 3 \cdot \sin^3 2x \cdot \cos^2 x \,\mathrm{d}x, \qquad n = 6,$

d) $\displaystyle\int_1^{\mathrm{e}} x^3 \cdot \ln \sqrt{x^2 + 1}\,\mathrm{d}x, \qquad n = 10,$

e) $\displaystyle\int_3^4 \left(\frac{1}{3}x^2 - 1\right)^5 \mathrm{d}x, \qquad n = 8,$

f) $\displaystyle\int_0^{\frac{\pi}{4}} \operatorname{tg} x \ \mathrm{d}x, \qquad n = 12,$

g) $\displaystyle\int_0^1 \sin x \cdot \mathrm{e}^{x^2} \mathrm{d}x, \qquad n = 10,$

h) $\displaystyle\int_1^2 \frac{\sin x}{x} \,\mathrm{d}x, \qquad n = 8.$

**4.4** Calculate the definite integral using the Simpson's method with the accuracy $\varepsilon = 10^{-5}$ and estimate the error of calculation:

$$\int_{-1}^{0} \frac{1 - x}{(x + 2)^2} \,\mathrm{d}x.$$

**4.5** Calculate the integral from task 4.3 letter f) by the rectangular method and estimate the error of calculation, if accuracy is given $\varepsilon = 10^{-3}$.

## 4.7 Results of Unsolved Tasks

**4.1** **a)** $\frac{1}{7}\cdot x^7 - x^4 + 3x + c$ **b)** $\frac{4}{5}\cdot x^5 - \frac{4}{3}\cdot x^3 + x + c$ **c)** $\frac{1}{2}\cdot x^2 - 6x + 6\cdot \ln x + \frac{8}{x} + c$ **d)** $\frac{10}{7}\cdot \sqrt[10]{x^7} - \frac{24}{5}\cdot x\sqrt[4]{x} + c$ **e)** $\sin^5 x + c$ **f)** $-\ln|\cos x| + c$ **g)** $\operatorname{tg} x - x + c$ **h)** $\mathrm{e}^x\cdot(x^2 + x - 5) + c$ **i)** $-\frac{1}{3}\cdot\cos(3x-7) + c$ **j)** $-\frac{1}{3}\cdot \mathrm{e}^{8-3x} + c$ **k)** $-\frac{1}{4}\cdot \mathrm{e}^{1-4x^2} + c$ **l)** $(3x^2-1)\cdot(2-\ln(2x-6x^2)) + c$ **m)** $\ln|x^3+2x-6| + c$ **n)** $\left(\frac{1}{2}\cdot\ln x\right)^2 + c$ **o)** $\frac{1}{18}\cdot\ln^3 x + c$ **p)** $x\cdot(\ln x - 1) + c$ **r)** $\left(\frac{1}{4}x^4 + \frac{3}{2}x^2\right)\cdot\left(\ln x - \frac{1}{4}\right) + c$ **s)** $\frac{1}{4}\cdot\operatorname{tg}^4 x + c$ **t)** $\frac{1}{64}\cdot(4x+2)^{16} + c$ **u)** $\frac{-1}{3\cdot\ln^3 x} + c$

**4.2** **a)** $\frac{14}{3}$ **b)** $\frac{1}{8}\cdot\mathrm{e}^2 + \frac{3}{8}$ **c)** $\frac{1}{2}$ **d)** $\frac{1}{4}\cdot(\mathrm{e}^2+1)$ **e)** $\frac{2}{3}$ **f)** 0 **g)** $\mathrm{e}-1$

**4.3** **a)** 2,0000000000 **b)** 4,8046382541 **c)** 0,3168332310 **d)** 11,7688920106 **e)** 442,9627421238 **f)** 0,3469302083 **g)** 0,7829408000 **h)** 0,6591551087

**4.4** 1

**4.5** 0

**4.6** $-1$

# Chapter 5

# Linear Algebra

## 5.1 Vector Space

**Definition 5.1** A *vector space over a field $F$* is a set $V$, on which are defined the operations addition of elements from $V$ and scalar multiplication of elements from $V$ by elements of $F$, such that the following applies:

| | |
|---|---|
| (1) $\forall \vec{a}, \vec{b} \in V$: | $\vec{a} + \vec{b} \in V$, |
| (2) $\forall \vec{a} \in V$, $\forall \alpha \in F$: | $\alpha \cdot \vec{a} \in V$, |
| (3) $\forall \vec{a}, \vec{b} \in V$: | $\vec{a} + \vec{b} = \vec{b} + \vec{a}$, |
| (4) $\forall \vec{a}, \vec{b}, \vec{c} \in V$: | $\vec{a} + (\vec{b} + \vec{c}) = (\vec{a} + \vec{b}) + \vec{c}$, |
| (5) $\exists \vec{0} \in V$, $\forall \vec{a} \in V$: | $\vec{a} + \vec{0} = \vec{a}$, |
| (6) $\forall \vec{a} \in V$, $\exists \vec{b} \in V$: | $\vec{a} + \vec{b} = \vec{0}$, |
| (7) $\forall \vec{a}, \vec{b} \in V$, $\forall \alpha \in F$: | $\alpha \cdot (\vec{a} + \vec{b}) = \alpha \cdot \vec{a} + \alpha \cdot \vec{b}$, |
| (8) $\forall \alpha, \beta \in F$, $\forall \vec{a} \in V$: | $(\alpha + \beta) \cdot \vec{a} = \alpha \cdot \vec{a} + \beta \cdot \vec{a}$, |
| (9) $\forall \alpha, \beta \in F$, $\forall \vec{a} \in V$: | $(\alpha \cdot \beta) \cdot \vec{a} = \alpha \cdot (\beta \cdot \vec{a})$, |
| (10) $\exists 1 \in F$, $\forall \vec{a} \in V$: | $1 \cdot \vec{a} = \vec{a}$. |

If a field $F$ is a set of real numbers $\mathbb{R}$ and elements of the set $V$ are $n$-tuples of real numbers from the set $\mathbb{R}^n$, then the vector space $V = \mathbb{R}^n$ is called *$n$-dimensional arithmetic vector space*. Denote the *zero vector* $\vec{0} = (0, 0, 0, \ldots, 0) \in \mathbb{R}^n$. *Unit vector* with the unit on the $k$-th position is denoted by letter $\vec{e}_k = (0, 0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^n$.

**Remark 5.1** From Definition 5.1 it is easy to see the following consequences:

(a) uniqueness of zero element (origin),
(b) *cancellation law for vector addition* (i. e. $\vec{a}+\vec{u}=\vec{b}+\vec{u} \Longrightarrow \vec{a}=\vec{b}$),
(c) uniqueness of the opposite vector (i. e. $(\forall\vec{a}\in V)(\exists!(-\vec{a})\in V)$: $\vec{a}+(-\vec{a})=\vec{0}$),
(d) $\alpha\cdot\vec{a}=\vec{0} \iff (\alpha=0 \vee \vec{a}=\vec{0}) \quad$ for $\alpha\in F \wedge \vec{a}\in V$,
(e) $(-1)\cdot\vec{a}=-\vec{a} \quad$ for $\vec{a}\in V$.

**Definition 5.2** We call *subspace* of the vector space $V$ a non-empty set $W$, $\emptyset\neq W\subseteq V$, such that:

(1) $\forall a,b\in W: \qquad a+b\in W$,
(2) $\forall a\in W, \forall\alpha\in F: \qquad \alpha\cdot a\in W$.

**Definition 5.3** Let $\vec{a}_1,\vec{a}_2,\ldots,\vec{a}_n\in V$ and $\alpha_1,\ldots,\alpha_n\in F$. A *linear combination* of vectors $\vec{a}_1,\vec{a}_2,\ldots,\vec{a}_n$ is any vector $\vec{u}$ of the form: $\vec{u}=\alpha_1\cdot\vec{a}_1+\alpha_2\cdot\vec{a}_2+\cdots+\alpha_n\cdot\vec{a}_n$.

**Definition 5.4** The span of S is the set of all finite linear combinations of elements of $S$: $\operatorname{span}(S)=\{\vec{x}\in V : \vec{x}=\sum_{i=1}^{n}\alpha_i\cdot v_i$; where $v_i\in S$; and $n\geqq 1\}$. Denote it $[S]$.

**Remark 5.2** The span of $S$ is also called the *Linear cover* or *Linear cover of vector system* $S$ from the vector space $V$, i. e. the span$(S)$ is the set of all linear combinations of all finite subsystems of the system $S$.[1]

**Definition 5.5** Let $\vec{a}_i\in V$ and $\alpha_i\in F$ for $i=1,2,\ldots,k$. A *nontrivial linear combination* of vectors $\vec{a}_1,\ldots,\vec{a}_k$ is such linear combination $\alpha_1\cdot\vec{a}_1+\alpha_2\cdot\vec{a}_2+\cdots+\alpha_k\cdot\vec{a}_k$, that at least one of the coefficients $\alpha_1,\ldots,\alpha_k$ is different from zero. A *trivial linear combination* of vectors $\vec{a}_1,\ldots,\vec{a}_k$ is the linear combination $0\cdot\vec{a}_1+0\cdot\vec{a}_2+\cdots+0\cdot\vec{a}_k$ (i. e. $\forall i\in\{1,2,\ldots,k\}$: $\alpha_i=0$).

**Definition 5.6** *System of vectors* (*group of vectors*) $\{\vec{a}_1,\ldots,\vec{a}_k\}\subseteq V$ is called *linearly dependent* if there is the nontrivial linear combination of vectors $\vec{a}_1,\ldots,\vec{a}_k$ for which it holds: $\alpha_1\cdot\vec{a}_1+\alpha_2\cdot\vec{a}_2+\cdots+\alpha_k\cdot\vec{a}_k=\vec{0}$. The system of vectors $\{\vec{a}_1,\ldots,\vec{a}_k\}\subseteq V$ is called *linearly independent* if it is not linearly dependent.

[1] Let $v_1,\ldots,v_n\in V$. The span of $(v_1,\ldots,v_n)$ is a subspace of $V$. Moreover, $(v_1,\ldots,v_n)$ is a spanning set in this subspace. We denote the span of $v_1,\ldots,v_n$ by $[v_1,\ldots,v_n]$.

**Definition 5.7** A system $B$, $\emptyset \neq B \subseteq V$, which satisfies the following conditions:

(1) $B$ is a linearly independent system of $V$,

(2) system $B$ generates the vector space $V$ i. e. $[B] = V$,

is called a *basis* $B$ of the vector space $V$.

**Definition 5.8** Let $V$ be the vector space over the field $F$. We say that the vector space $V$ is *finite-dimensional*, if there exists finite system of vectors $S \subseteq V$, which generates vector space $V$. An *infinite-dimensional* vector space is one that is not finite-dimensional.

**Definition 5.9** A *dimension* of finite-dimensional vector space $V$ is the number of vectors of a basis of $V$. We will denote it by $\dim(V)$, $\dim(\vec{0}) = 0$.[2]

**Definition 5.10** A *rank* of a system $S$ of finite-dimensional vector space $V$ is $\dim([S])$.[3]

**Definition 5.11** The groups of vectors $S$ and $T$ over the vector space $V$ are called *equivalent*, if it holds $[S] = [T]$.[4]

**Elementary Operations on Vector Space:**[5]

(1) interchange the order of determining (generating) vectors,
(2) multiplication of vector by any non-zero scalar $\alpha \in F$, $\alpha \neq 0$,
(3) adding a linear combination of other vectors to some vector,
(4) omitting a vector that is a linear combination of the other vectors,
(5) adding a vector which is a linear combination of the other vectors.

---

[2]If vector space $V$ is finite-dimensional, then $V$ has a finite basis. Every basis $B$ for $V$ contains the same number of vectors. The unique number of vectors in each basis $B$ is the dimension of the vector space $V$.

[3]The rank of a system $S$ is given uniquely.

[4]One system of vectors has infinitely many bases. If two bases define the same vector system then one of the bases was created from the second base using the elementary operations.

[5]These are changes respectively, equivalent modifications, which do not change the vector space.

**Theorem 5.1** Let $\vec{a} \in V$. System $\{\vec{a}\}$ is linearly dependent if and only if $\vec{a} = \vec{0}$.

**Theorem 5.2** System of vectors $S$, $\emptyset \neq S \subseteq V$ is linearly dependent, if there exists a finite subsystem of vectors $\{\vec{a}_1, \ldots, \vec{a}_k\} \subseteq S$, where the system of vectors $\{\vec{a}_1, \ldots, \vec{a}_k\}$ is linearly dependent.

**Theorem 5.3** System of vectors $S$, $\emptyset \neq S \subseteq V$ is linearly independent if every finite subsystem of $S$ is linearly independent.

**Theorem 5.4** The system of vectors $\{\vec{a}_1, \ldots, \vec{a}_k\} \subseteq V$, where $k > 1$ is linearly dependent if and only if at least one of the vectors can be expressed as a linear combination of the other vectors.

**Theorem 5.5** The system of vectors $\{\vec{a}_1, \ldots, \vec{a}_k\} \subseteq V$, where $k > 1$ is linearly independent if and only if neither of the vectors can be expressed as a linear combination of other vectors.

**Theorem 5.6** The system of vectors $\{\vec{a}_1, \ldots, \vec{a}_k\} \subseteq V$ is linearly independent, if holds:

(1) Every nontrivial linear combination of vectors $\vec{a}_1, \ldots, \vec{a}_k$ is different from zero.

(2) Only trivial linear combination of vectors $\vec{a}_1, \ldots, \vec{a}_k$ is equal to zero.

(3) $\alpha_1 \cdot \vec{a}_1 + \alpha_2 \cdot \vec{a}_2 + \cdots + \alpha_k \cdot \vec{a}_k = \vec{0} \Longrightarrow \alpha_1 = 0 \wedge \alpha_2 = 0 \wedge \cdots \wedge \alpha_k = 0$, $\alpha_i \in F$, $i = 1, 2, \ldots, k$ .

**Theorem 5.7** In vector space $V = \mathbb{R}^n$ is each system of vectors containing more than $n$ vectors linearly dependent.

**Theorem 5.8** Each finite-dimensional vector space different from $\{\vec{0}\}$ has a finite basis.

**Theorem 5.9** If the vector space $V$ has a finite basis consisting of $n$ vectors, then each basis of vector space $V$ has exactly $n$ elements.

**Theorem 5.10** Suppose the vector space $V$ with $\dim(V) = n$ is given. Each group of vectors from $V$ which is consisting of a greater number than $n$ vectors is linearly dependent.

**Theorem 5.11** Each linearly independent system of vectors can be extended to a basis in finite-dimensional vector spaces.

**Theorem 5.12** Suppose $V$ is a finite-dimensional vector space with $\dim(V) = n > 0$. Let $\{\vec{a}_1, \dots, \vec{a}_n\}$ be an arbitrary group of vectors from the vector space $V$. Then the following conditions are equivalent:

(1) $\{\vec{a}_1, \dots, \vec{a}_n\}$ is linearly independent group of vectors,

(2) $[\{\vec{a}_1, \dots, \vec{a}_n\}] = V$,

(3) $\{\vec{a}_1, \dots, \vec{a}_n\}$ is basis of the vector space $V$.

**Theorem 5.13** The groups of vectors $S$ and $T$ are equivalent if and only if $S \subseteq [T] \wedge T \subseteq [S]$.

**Theorem 5.14** Each of the following systems of vectors (1) – (4) is equivalent to the system of vectors $\{\vec{a}_1, \dots, \vec{a}_i, \dots, \vec{a}_j, \dots, \vec{a}_k\}$:

(1) $\{\vec{a}_1, \dots, \vec{a}_i, \dots, \vec{a}_j, \dots, \vec{a}_k\}$,

(2) $\{\vec{a}_1, \dots, \alpha \cdot \vec{a}_i, \dots, \vec{a}_j, \dots, \vec{a}_k\}$, where $\alpha \neq 0$, $\alpha \in F$,

(3) $\{\vec{a}_1, \dots, \vec{a}_i, \dots, \vec{a}_j + \alpha \cdot \vec{a}_i, \dots, \vec{a}_k\}$, $\alpha \in F$,

(4) $\{\vec{a}_1, \dots, \vec{a}_i, \dots, \vec{a}_j, \dots, \vec{a}_k, \alpha_1 \cdot \vec{a}_1 + \dots + \alpha_k \cdot \vec{a}_k\}$, $\alpha_i \in F$.

## 5.2 Matrix

**Definition 5.12** A matrix of type $m \times n$ is system of the elements which are written to a table with $m$ rows and $n$ columns, where $m$ and $n$ are positive integers. For writing matrices we use parentheses

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

where $a_{ij}$ are elements (entries) of the matrix $A$ for $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$. We usually denote the matrix by capital letters and also we use the shorthand notation of the matrix: $A = \{a_{ij}\}$, $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.[6]

**Definition 5.13** (Special types of matrices:)

**Rectangular Matrix:** A general matrix of type $m \times n$, matrix with $m$ rows and $n$ columns.

**Square Matrix:** A matrix, which has the same number of rows and columns, matrix of type $n \times n$.

**Column Matrix:** A matrix of type $m \times 1$. It is also called the column vector.

**Row Matrix:** A matrix of type $1 \times n$. It is also called the row vector.

**Zero Matrix:** A matrix of type $m \times n$, which has all elements equal zero, i. e. $\forall i, j : a_{ij} = 0$. We will denote it $\mathbf{0}$.

**Diagonal Matrix:** A square matrix of type $n \times n$, where $a_{ij} = 0$, for $i \neq j$, $i, j \in \{1, 2, \ldots, n\}$. We will denote it: $\mathrm{diag}\{a_{11}, a_{22}, \ldots, a_{nn}\} = \mathrm{diag}\{a_{ii}\}_{i=1}^{n}$.

**Identity Matrix:** Diagonal matrix $I = \mathrm{diag}\{a_{11}, a_{22}, \ldots, a_{nn}\}$, where elements $a_{ii} = 1$, $i \in \{1, 2, \ldots, n\}$. We will denote it $I$, $I_n$, or $I_{n \times n}$. Matrix $I$ is the multiplicative identity for the matrices.[7]

**Upper Triangular Matrix:** A square matrix of type $n \times n$ with zeros below the main diagonal, i. e. $a_{ij} = 0$, for $j < i$, $i, j \in \{1, 2, \ldots, n\}$.

**Lower Triangular Matrix:** A square matrix of type $n \times n$ with zeros above the main diagonal, i. e. $a_{ij} = 0$, for $i < j$, $i, j \in \{1, 2, \ldots, n\}$.

**Transpose Matrix:** Transpose matrix $A^\top$ to the matrix $A$ of type $m \times n$ is matrix of type $n \times m$, for which holds: $a_{ij}^\top = a_{ji}$, $i \in \{1, 2, \ldots, m\}$, $j \in \{1, 2, \ldots, n\}$, i. e. we exchanged the rows and the columns of the matrix $A$ with each other.

[6] The $m \times n$ is the type of matrix $A$ and it is also called *dimension of matrix A*.

[7] A square matrix in which the elements in the leading diagonal are all equal to one and all other elements are equal to zero is also called *unit matrix*, but the name unit matrix is also used in the context of a square matrices for which all elements are equal to one, we will denote it $E$.

**Symmetric Matrix:** Matrix $A$ is symmetric, if for elements of a matrix $A$ holds: $a_{ij} = a_{ji}$ for all $i, j$, i. e. $A = A^\top$.

**Definition 5.14** We say that the matrices $A = \{a_{ij}\}$ and $B = \{b_{ij}\}$ *are equal* (we write $A = B$), if matrices $A$ and $B$ are of the same type (have the same number of rows and columns) and $a_{ij} = b_{ij}$ for all $i, j$.

Basic operations with matrices.

**Definition 5.15** A *sum* (*addition*) of two matrices $A = \{a_{ij}\}$ and $B = \{b_{ij}\}$ of the type $m \times n$ is the matrix $C = \{c_{ij}\}$ (we write $C = A + B$) of the type $m \times n$ in which for all entries holds: $c_{ij} = a_{ij} + b_{ij}$, $\forall i \in \{1, 2, \dots, m\}$, $\forall j \in \{1, 2, \dots, n\}$.

**Remark 5.3** The sum of two matrices $A$ and $B$ is matrix $A + B$ which is defined by adding corresponding entries: $(A + B)_{ij} = a_{ij} + b_{ij}$.

**Definition 5.16** We say that the matrix $D$ is $\alpha$ *multiple of matrix* $A$ (we write $D = \alpha \cdot A$, $\alpha \in \mathbb{R}$), if the matrices $A$ and $D$ are of the same type and for all entries of the matrix $D$ holds: $d_{ij} = \alpha \cdot a_{ij}$ for all $i$, $j$.

**Remark 5.4** A *scalar multiplication* is defined as $(\alpha \cdot A)_{ij} = \alpha \cdot a_{ij}$.

**Definition 5.17** Let matrices $A$ of type $m \times p$ and $B$ of type $p \times n$ be given. The *product of two matrices* $A$ and $B$ is the matrix $C$ of type $m \times n$ (we write $C = A \cdot B$), for which holds:

$$c_{ij} = \sum_{l=1}^{p} a_{il} \cdot b_{lj}$$

for $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$.

**Remark 5.5** *Matrix Product*: $(i, j)$ entry of matrix $A \cdot B$ is obtained by multiplying each element in the $i^{th}$ row of matrix $A$ by the corresponding element in the $j^{th}$ column of the matrix $B$ and summing, i. e.

$$(A \cdot B)_{ij} = \sum_{l=1}^{p} a_{il} \cdot b_{lj}.$$

**Basic Properties of Matrices**

(1) $A = A$, the equality relation "=" is reflexive.

(2) If $A = B$, then $B = A$, the equality relation "=" is symmetric.

(3) If $A = B$ and $B = C$, then $A = C$, the equality relation "=" is transitive.

(4) $(-1) \cdot A = -A$.

(5) $A + (-A) = A - A = \mathbf{0}$.

(6) $A + A^\top$ is symmetric matrix.

(7) $A \cdot I = I \cdot A = A$, where $A$ is arbitrary square matrix.

(8) $A \cdot B \neq B \cdot A$ in general case.[8]

(9) $A + B = B + A$, commutativity of addition "+".

(10) $A + (B + C) = (A + B) + C$, associativity of addition "+".

(11) $A + \mathbf{0} = \mathbf{0} + A = A$, additive identity.

(12) $\alpha \cdot (A + B) = \alpha \cdot A + \alpha \cdot B$, distributivity .

(13) $(\alpha + \beta) \cdot A = \alpha \cdot A + \beta \cdot A$, distributivity.

(14) $(A \cdot B) \cdot C = A \cdot (B \cdot C)$, associativity of multiplication "$\cdot$".

(15) $(A + B) \cdot C = A \cdot C + B \cdot C$, right distributivity.

(16) $A \cdot (B + C) = A \cdot B + A \cdot C$, left distributivity.

(17) $(\alpha \cdot \beta) \cdot A = \alpha \cdot (\beta \cdot A)$, associativity of scalar multiplication.

(18) $1 \cdot A = A \cdot 1 = A$, scalar identity.

(19) $(A \cdot B)^\top = B^\top \cdot A^\top$.

(20) $(A_1 \cdot A_2 \cdot \dots \cdot A_n)^\top = A_n^\top \cdot \dots \cdot A_2^\top \cdot A_1^\top$.

[8] If there exists a product of matrices $A \cdot B$, then the product $B \cdot A$ needs not to exist.

**Remark 5.6** (*Zero Product Property*) Just because a product of two matrices is the zero matrix does not mean that one of them was the zero matrix.

**Remark 5.7** (*Multiplicative Property of Equality*) If $A = B$, then $AC = BC$. This property is still true, but the converse is not necessarily true. Just because $AC = BC$ does not mean that $A = B$. If $A = B$, then $AC = BC$ or $CA = CB$, but $AC \neq CB$ for the general matrix $C$.

**Definition 5.18** Let $A$ be a matrix $m \times n$. Let $h_1$ be dimension of the system of row vectors and $h_2$ be dimension of the system of column vectors. The number $h = h_1 = h_2$ is called the *rank* of matrix $A$ and it is denoted by the symbol $\text{rank}(A) = h$.

**Remark 5.8** Alternative definition of the rank of a matrix $A$: Let $A$ be a $m \times n$ matrix over a field $\mathbb{R}$. We say that the column rank of the matrix $A$ is the maximum number of linearly independent columns of the matrix $A$, while the row rank of the matrix $A$ is the maximum number of linearly independent rows of the matrix $A$. (We regard columns or rows as vectors in $\mathbb{R}^m$ or $\mathbb{R}^n$, respectively.)

**Definition 5.19** Let $A$ be a square matrix of the type $n \times n$. We say that the matrix $A$ is *regular*, if $\text{rank}(A) = n$ (*singular*, if $\text{rank}(A) < n$).

The *elementary operations* on matrices are such adjustments to the rows (or columns) of the matrix, which do not change the rank of the matrix. The elementary operations are usually also called row (or column) matrix transformations (modifications).

We can transform the non-upper triangular matrix to an upper triangular matrix using so called elementary row and column operations, which do not change the rank of the matrix.

The procedure of transformation of an arbitrary matrix to an upper triangular matrix (all of whose elements on the main diagonal are different from zero) by means of the elementary row and column operations is called the *Gauss algorithm.*

**Elementary Operations:**

(1) change of order of rows,

(2) multiplication of some row by a non-zero real number,

(3) addition of a linear combination of the other rows to some row,

(4) omission of a row which is a linear combination of the other rows,

(5) omission of a zero row.

All elementary operations can also be performed with columns.

**Definition 5.20** A matrix $B$ is called *equivalent* to the matrix $A$, if the matrix $B$ can be derived from the matrix $A$ using a finite number of elementary transformations. The relation "equivalence" defined above is an equivalence relation on the set of all $m \times n$ matrices; that is, it is reflexive, symmetric, and transitive.

Let be given an arbitrary matrix $A$ of the type $m \times n$. We know that for the rank of the matrix $A$ applies: $0 \leqq \operatorname{rank}(A) \leqq \min\{m, n\}$. The rank of the zero matrix is zero. For arbitrary non-zero matrix $A$ applies: $\operatorname{rank}(A) \in \mathbb{N}$. The rank of the matrix $A$ is the dimension of the linear span of column or row vectors of the matrix $A$.

**Theorem 5.15** The following four assertions hold:[9]

(a) Elementary column operations don't change the column rank of a matrix.

(b) Elementary row operations don't change the column rank of a matrix.

(c) Elementary column operations don't change the row rank of a matrix.

(d) Elementary row operations don't change the row rank of a matrix.

**Definition 5.21** Let $A$ be a square matrix of the type $n \times n$. If there exists a matrix $B$ to the matrix $A$, that following applies: $A \cdot B = B \cdot A = I$, then the matrix $B$ is called the *inverse matrix* of the matrix $A$, and it is denoted by the symbol $A^{-1}$.

**Theorem 5.16** Let $A$ be a square matrix of the type $n \times n$. A necessary and sufficient condition for the existence of the inverse matrix $A^{-1}$ to the matrix $A$ is, that the matrix $A$ is regular, i. e. $\operatorname{rank}(A) = n$ ($\det(A) \neq 0$). We say that matrix $A$ is *invertible*.

[9] The rank is independent of the row and column operations used to compute it.

**Properties of inverse matrices:**

(1) $(A^{-1})^{-1} = A$,

(2) $A \cdot A^{-1} = A^{-1} \cdot A$,

(3) $\operatorname{diag}(a_1, a_2, \ldots, a_n) \cdot \operatorname{diag}(b_1, b_2, \ldots, b_n) = I$ if and only if $a_i \cdot b_i = 1$,

(4) $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$,

(5) $(A_1 \cdot A_2 \cdot \cdots \cdot A_{k-1} \cdot A_k)^{-1} = A_k^{-1} \cdot A_{k-1}^{-1} \cdot \cdots \cdot A_2^{-1} \cdot A_1^{-1}$.

## 5.3 Determinant

Let $A = \{a_{ij}\}$ be a square matrix of the type $n \times n$. We can assign a real number to the matrix $A$, which we call the *determinant*. We will define this using the following definitions.

**Definition 5.22** Let $A = \{a_{ij}\}$ be a square matrix of the type $n \times n$. A *Minor* of the matrix $A$ which is corresponding to the element $a_{ij}$ of matrix $A$ is determinant of the $(n-1) \times (n-1)$ square sub-matrix which arises from $A$ by omission of the $i$-th row and the $j$-th column, for $i, j \in \{1, 2, \ldots, n\}$. We will denote it $\det(A_{ij})$.[10]

**Definition 5.23** A *co-factor of element* $a_{ij}$ of $n \times n$ square matrix $A = \{a_{ij}\}$ is called the product of number $(-1)^{i+j}$ and minor of matrix $A$ which is corresponding to the element $a_{ij}$ for $i, j \in \{1, 2, \ldots, n\}$. We will denote it $\overline{\det}(A_{ij}) = (-1)^{i+j} \cdot \det(A_{ij})$.

**Definition 5.24** Let $A = \{a_{ij}\}$ be a square matrix of the type $n \times n$. *Determinant* of the matrix $A$ is a real number, which is denoted by $\det(A)$ or $|A|$ and for which holds:

(1) $\det(A) = a_{11}$, if $A = \{a_{ij}\}$ is a square matrix of type $1 \times 1$,

(2) $\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$,

[10] Determinant $\det(A_{ij})$ is called the minor, which is the abbreviation for "minor determinant".

(3) $\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \cdot a_{22} \cdot a_{33} + a_{21} \cdot a_{32} \cdot a_{13} + {} + a_{31} \cdot a_{12} \cdot a_{23} - (a_{13} \cdot a_{22} \cdot a_{31} + a_{23} \cdot a_{32} \cdot a_{11} + a_{33} \cdot a_{12} \cdot a_{21}),$

(4) $\det(A) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1(n-1)} & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2(n-1)} & a_{2n} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ a_{(n-1)1} & a_{(n-1)2} & \dots & a_{(n-1)(n-1)} & a_{(n-1)n} \\ a_{n1} & a_{n2} & \dots & a_{n(n-1)} & a_{nn} \end{vmatrix} =$

$$= \sum_{k=1}^{n} a_{kj} \cdot \overline{\det}(A_{kj}) = \sum_{l=1}^{n} a_{il} \cdot \overline{\det}(A_{il})$$

for arbitrary $i \in \{1, 2, \dots, n\}$, $j \in \{1, 2, \dots, n\}$.

**Remark 5.9** If $A = \{a_{ij}\}$ is an $n \times n$ square matrix (for $n > 1$) then we choose an arbitrary $i$-th row of the matrix $A$ and we put:

$$\det(A) = \sum_{l=1}^{n} a_{il} \cdot \overline{\det}(A_{il}) =$$

$$= a_{i1} \cdot \overline{\det}(A_{i1}) + a_{i2} \cdot \overline{\det}(A_{i2}) + a_{i3} \cdot \overline{\det}(A_{i3}) + \cdots + a_{in-1} \cdot \overline{\det}(A_{in-1}) + a_{in} \cdot \overline{\det}(A_{in}).$$

The sum $a_{i1} \cdot \overline{\det}(A_{i1}) + a_{i2} \cdot \overline{\det}(A_{i2}) + \cdots + a_{in-1} \cdot \overline{\det}(A_{in-1}) + a_{in} \cdot \overline{\det}(A_{in})$ is called the *expansion of the determinant according to the $i$-th row*. The *expansion of the determinant according to the $j$-th column* reads:[11]

$$\det(A) = \sum_{k=1}^{n} a_{kj} \cdot \overline{\det}(A_{kj}) =$$

$$= a_{1j} \cdot \overline{\det}(A_{1j}) + a_{2j} \cdot \overline{\det}(A_{2j}) + a_{3j} \cdot \overline{\det}(A_{3j}) + \cdots + a_{n-1j} \cdot \overline{\det}(A_{n-1j}) + a_{nj} \cdot \overline{\det}(A_{nj}).$$

**Theorem 5.17** Let $A^\top$ be the transposed matrix to the square matrix $A$, then $\det(A) = \det(A^\top)$.

[11] It is also called co-factor expansion along the $i$-th row (the $j$-th column).

**Theorem 5.18** Let a matrix $B$ be the matrix, which arose from the square matrix $A$ by exchange of two rows (columns). Then $\det(B) = -\det(A)$.

**Theorem 5.19** If a square matrix $B$ is obtained from a square matrix $A$ by multiplying the arbitrary $j$-th column by a non-zero scalar $\alpha$, then $\det(B) = \alpha \cdot \det(A)$. If the square matrix $B$ is obtained from the square matrix $A$ by multiplying the arbitrary $i$-th row by a non-zero scalar $\alpha$, then $\det(B) = \alpha \cdot \det(A)$.

**Theorem 5.20** If a square matrix $B$ is obtained from a square matrix $A$ by adding of a linear combination of row (column) of the matrix $A$ to some row (column) of the matrix $A$, then for the value of the determinant of the matrix $B$ holds: $\det(B) = \det(A)$.

**Theorem 5.21** Let $A$ be a square matrix. Determinant of the matrix $A$ is equal to zero ($\det(A) = 0$) if and only if at least one row (column) of the matrix $A$ is a linear combination of the remaining rows (columns) of the matrix $A$.

**Theorem 5.22** If a square matrix $A$ contains at least two identical rows (columns) then $\det(A) = 0$.

**Theorem 5.23** If a square matrix $A$ contains at least one zero row (column) then $\det(A) = 0$.

**Theorem 5.24** The following applies:

$$\det(A) = \begin{vmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

**Theorem 5.25** If a $n \times n$ square matrix $A$ is a triangular matrix (upper triangular, lower triangular) then $\det(A) = \det(\text{diag}\{a_{11}, a_{22}, \dots, a_{nn}\}) = a_{11} \cdot a_{22} \cdot \dots \cdot a_{nn}$.

**Theorem 5.26** The $n \times n$ square matrix $A$ is a regular (invertible) if and only if $\det(A) \neq 0$.

**Definition 5.25** Let $A$ be a square matrix of the type $n \times n$. Then matrix $\text{Adj}(A)$

$$\text{Adj}(A) = \begin{pmatrix} \overline{\det}(A_{11}) & \overline{\det}(A_{21}) & \dots & \overline{\det}(A_{n1}) \\ \overline{\det}(A_{12}) & \overline{\det}(A_{22}) & \dots & \overline{\det}(A_{n2}) \\ \vdots & \vdots & \dots & \vdots \\ \overline{\det}(A_{1n}) & \overline{\det}(A_{2n}) & \dots & \overline{\det}(A_{nn}) \end{pmatrix}$$

is called the *adjugate* (*adjoint*) of $A$ (it is a transposed matrix of co-factors to the elements of matrix $A$).

**Theorem 5.27** If the $n \times n$ matrix $A$ is invertible, then its inverse is equal to

$$A^{-1} = \frac{1}{\det(A)} \cdot \text{Adj}(A). \tag{5.1}$$

**Theorem 5.28** Let $A$ be a square matrix of the type $n \times n$. Then holds:

$$\det(\text{Adj}(A)) = (\det(A))^{n-1}. \tag{5.2}$$

## 5.4 Systems of Linear Equations

**Definition 5.26** *System* (set) of $m$ *linear equations* with $n$ unknowns variables over a set of real numbers $\mathbb{R}$ is called a finite sequence of equations $S$:

$$\begin{aligned} a_{11} \cdot x_1 + a_{12} \cdot x_2 + \dots + a_{1n} \cdot x_n &= b_1 \\ a_{21} \cdot x_1 + a_{22} \cdot x_2 + \dots + a_{2n} \cdot x_n &= b_2 \\ a_{31} \cdot x_1 + a_{32} \cdot x_2 + \dots + a_{3n} \cdot x_n &= b_3 \\ \vdots &= \vdots \\ a_{m1} \cdot x_1 + a_{m2} \cdot x_2 + \dots + a_{mn} \cdot x_n &= b_m, \end{aligned} \tag{5.3}$$

where $\vec{x} = (x_1, x_2, \ldots, x_n)^\top$ is the vector of unknown variables, $a_{ij} \in \mathbb{R}$ are the coefficients of the system $S$, and $\vec{b} = (b_1, b_2, b_3, \ldots, b_m)^\top$ is a vector, which is called the right hand side of the system of linear equations $S$.

The system of equations $S$ can be written in matrix form as $A \cdot \vec{x} = \vec{b}$, where

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \text{ and } \vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ a_m \end{pmatrix}.$$

The matrix $A$ is called the *matrix of the system* $S$ and the *augmented matrix* $A_b = (A|\vec{b})$ is $m \times (n+1)$ matrix (the entries of $\vec{b}$ are placed to the right of the matrix $A$). It can be written as:

$$A_b = (A|\vec{b}) = \left(\begin{array}{ccc|c} a_{11} & \ldots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} & b_m \end{array}\right). \tag{5.4}$$

If the system has at least one solution (one solution or infinitely many solutions), then it is said to be *consistent* system. If the system has no solution, then it is said to be *inconsistent* system (*singular*).

**Definition 5.27** The basic terms of systems of linear equations (SLE) $S$ having $m$ equations and $n$ unknowns:

**Homogeneous system of linear equations:** If vector $\vec{b} = \vec{0}$ (is zero vector), the system is homogeneous, i. e. if for all $i \in \{1, 2, \ldots, m\}$ holds: $b_i = 0$, then the system $S$ is called *homogeneous system of linear equations* (HSLE). The homogeneous system of linear equations $S$ is denoted by $S_h$.

**Non-homogeneous system of linear equations:** If vector $\vec{b} \neq \vec{0}$, the system is *non-homogeneous*, i. e. if there exists $i \in \{1, 2, \ldots, m\}$, such that: $b_i \neq 0$, then the system $S$ is called *non-homogeneous system of linear equations* (abbreviated SLE).

**Solution of the SLE:** Ordered $n$-tuple of real numbers $(c_1, c_2, c_3, \ldots, c_n)^\top \in \mathbb{R}^n$ is called *solution* of SLE $S$, if $\sum_{j=1}^{n} a_{ij} \cdot c_j = b_i$ for all $i \in \{1, 2, \ldots, m\}$.

**Set of solutions of the SLE:** The set of all vectors $\vec{c} = (c_1, c_2, c_3, \dots, c_n)^\top$, which are solution of the SLE $S$ is called the *set of solutions* of SLE $S$ and is denoted by $\Omega$ or $\Omega(S)$.

**Trivial solution of the HSLE:** Zero vector is called the *trivial solution* (zero solution) of HSLE.

**Nontrivial solution of the HSLE:** If the solution of the HSLE has at least one non-zero coordinate, then such a solution is called the *nontrivial solution* of the HSLE.

**Zero equation:** The equation $0 \cdot x_1 + 0 \cdot x_2 + \cdots + 0 \cdot x_n = 0$ is called *zero equation.*

**Contradictory equation:** Equation $0 \cdot x_1 + 0 \cdot x_2 + \cdots + 0 \cdot x_n = b_i$, where $b_i \neq 0$ for some $i \in \{1, 2, \dots, m\}$ is called *contradictory equation.*[12]

**Equivalent SLE:** The system $S_1$ is *equivalent* to the system $S_2$ if and only if both systems have the same set of solutions $\Omega$. We mark this relationship: $S_1 \sim S_2$ resp. $S_1 \approx S_2$.[13]

**Pivot of equation:** The *pivot* (*main coefficient*) of the $i$-th equation $q_i$ is the first nonzero coefficient from the left side in the $i$-th row ($i$-th equation), i. e. the first $a_{ij} \neq 0$ from the left side for some $i \in \{1, 2, \dots, m\}$ is called *pivot* of equation $q_i$: $a_{i1} \cdot x_1 + a_{i2} \cdot x_2 + \cdots + a_{ij} \cdot x_j + \cdots + a_{in} \cdot x_n = b_i$. Coefficients $a_{i1}$ up to $a_{i,j-1}$ are zero.

**Gaussian form of the SLE:** The system of linear equations $S$ without zero equations is in *Gaussian form*, if it holds: If $i < j$, then pivot $a_{ik}$ of the $i$-th equation $q_i$ has a smaller column index as pivot $a_{jl}$ of the $j$-th equation $q_j$ (i. e. $k < l$). The system of linear equations in Gaussian form will be denoted by $S_g$.

**Remark 5.10** Two SLE's using the same set of variables are equivalent if each of the equations in the second system can be derived algebraically from the equations in the first system, and vice-versa. Two systems are equivalent

[12] It is possible to derive a contradiction from the equations, that may always be rewritten using equivalent operations, such as the equality $0 = 1$.

[13] Two systems of equations $S_1$ and $S_2$ are called *equivalent* if they have identical sets of solutions (i. e. $\Omega(S_1) = \Omega(S_2)$).

if either both are inconsistent or each equation of any of them is a linear combination of the equations of the other one.
Elementary operations for systems of linear equations, which do not change the set of solutions:

(1) Omission zero equation.

(2) Multiplying any equation by a nonzero real number.

(3) Adding one equation to another equation.

(4) Exchanging order of equations.

(5) Adding to one equation linear combination of other equations.

*Gaussian elimination* is reducing a system of equations (lining up the variables, the equations are the rows), a matrix $A$, or an augmented matrix $A_b$ by using elementary row operations.

**Definition 5.28** Let a SLE $S$ be given and let the systems $S$ and $S_g$ be equivalent, where $S_g$ is system of linear equations in Gaussian form. The number of non-zero equations of the system $S_g$ is called the *rank of the system* $S$ and it is denoted by $\mathrm{rank}(S)$.

**Definition 5.29** Let $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k$; $\vec{w}_i = (c_{1i}, c_{2i}, \ldots, c_{ni})$; $i \in \{1, 2, \ldots, k\}$ be solutions of the SLE $S$ (also HSLE). Let $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}$. Then vector $\vec{w} = \alpha_1 \cdot \vec{w}_1 + \alpha_2 \cdot \vec{w}_2 + \alpha_3 \cdot \vec{w}_3 + \cdots + \alpha_k \cdot \vec{w}_k$ is called *linear combination of solutions* $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k$.

**Definition 5.30** The set of solutions $\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k\}$ is called *fundamental system of solutions* of HSLE $S_h$, if holds:

(1) Each solution $\vec{w} \in \Omega(S_h)$ is a linear combination of solutions $\vec{w}_1, \ldots, \vec{w}_k$, i. e. $(\forall w \in \Omega(S))$: $\vec{w} = \sum_{i=1}^{k} \alpha_i \cdot \vec{w}_i$.

(2) Solutions $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k$ are linearly independent, i. e. no solution $\vec{w}_i$ for $i \in \{1, 2, 3, \ldots, k\}$ is a linear combination of solutions $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_{i-1}$, $\vec{w}_{i+1}, \ldots, \vec{w}_k$.

**Theorem 5.29** Each system $S$ of $m$ linear equations of $n$ unknowns can be transformed by elementary modifications to the Gaussian form.

**Theorem 5.30** Let $S \in S_{m,n}(\mathbb{R})$[14] and suppose that $S \sim S_1 \sim S_2 \sim \cdots \sim S_g$. Then the following statements hold:

(1) $\Omega(S) = \Omega(S_g)$.

(2) If elimination produces a contradiction, i. e. $S_g$ contains a contradictory equation, then $\Omega(S) = \emptyset$ (system $S$ has no solutions).

(3) If the number of non-zero equations of the system $S_g$ equals to the number of unknowns, then $|\Omega(S)| = 1$ (system $S$ has a unique solution).

(4) If $\Omega(S) \neq \emptyset$ and the number of non-zero equations of the system $S_g$ is less than the number of unknown variables, then $\Omega(S)$ is an infinite set (system $S$ has infinitely many solutions).

**Theorem 5.31** Let $S_h$ be HSLE and let $\vec{w}_1$ and $\vec{w}_2$ be solutions of HSLE $S_h$. Then also $\vec{w} = \vec{w}_1 + \vec{w}_2$ is a solution of the HSLE $S_h$.

**Theorem 5.32** Let $S_h$ be HSLE and let $\vec{w}_1$ be a solution of HSLE $S_h$ and $\alpha \in \mathbb{R}$. Then also $\vec{w} = \alpha \cdot \vec{w}_1$ is a solution of the HSLE $S_h$.

**Theorem 5.33** Let $S_h$ be HSLE. If $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k$ are solutions of HSLE $S_h$, then also their linear combination $\vec{w} = \alpha_1 \cdot \vec{w}_1 + \alpha_2 \cdot \vec{w}_2 + \alpha_3 \cdot \vec{w}_3 + \cdots + \alpha_k \cdot \vec{w}_k$ is a solution of the HSLE $S_h$, where $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}$.

**Theorem 5.34** Let $S_h$ be HSLE, $S_h \in S_{m,n}(\mathbb{R})$. The system $S_h$ has only trivial solution if and only if $\text{rank}(S_h) = n$.

**Theorem 5.35** The fundamental system of solutions of the HSLE $S_h$ has $q = n - \text{rank}(S_h)$ elements, if $\text{rank}(S_h) < n$.

**Theorem 5.36** Let $S \in S_{m,n}(\mathbb{R})$ and let $S_h$ be HSLE corresponding to the system $S$.[15] If $\vec{w}_1 \in \Omega(S)$ and $\vec{w}_2 \in \Omega(S)$, then $\vec{w}_1 - \vec{w}_2 \in \Omega(S_h)$.

---

[14] $S_{m,n}(\mathbb{R})$ is designation of the set of all systems of $m$ linear equations on $n$ unknowns over a set of real numbers $\mathbb{R}$.

[15] HSLE corresponding to the system $S \in S_{m,n}(\mathbb{R})$ is system $S_h$, which has the same coefficients $a_{ij}$ as a system $S$, but the right hand side of the system $S_h$ is equal to the zero vector.

**Theorem 5.37** Let $S \in S_{m,n}(\mathbb{R})$ be SLE and let $S_h$ be HSLE corresponding to the system $S$. Suppose that $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_{n-r}$, where $r = \text{rank}(S) < n$ is fundamental system of solutions of the HSLE $S_h$ and $\vec{w}^\star$ is particular solution[16] of the system $S$. Then for every solution $\vec{w} \in \Omega(S)$ there exist real numbers $\alpha_1, \alpha_2, \ldots, \alpha_{n-r} \in \mathbb{R}$, such that:

$$\vec{w} = \vec{w}^\star + \alpha_1 \cdot \vec{w}_1 + \alpha_2 \cdot \vec{w}_2 + \alpha_3 \cdot \vec{w}_3 + \cdots + \alpha_{n-r} \cdot \vec{w}_{n-r}.$$

**Theorem 5.38** (*Fundamental Theorem for Homogeneous System of Linear Equations*) Suppose $A \cdot \vec{x} = \vec{0}$ is HSLE of $m$ linear equations of $n$ unknown variables and the $\text{rank}(A) = r$.[17]

(1) If $r = n$, then the trivial solution is the only solution.

(2) If $r < n$, there are infinitely many solutions and the general solution will contain $n - r$ arbitrary constants.

**Theorem 5.39** (*Fundamental Theorem*)[18] Let $S \in S_{m,n}(\mathbb{R})$ be SLE and let $S_h$ be HSLE corresponding to the system $S$. The system $S$ has a solution if and only if $\text{rank}(S) = \text{rank}(S_h)$. Furthermore,

(1) if $\text{rank}(S) = \text{rank}(S_h) = n$, then the system $S$ has exactly one solution (unique solution).

(2) If $\text{rank}(S) = \text{rank}(S_h) < n$, then the system $S$ has infinitely many solutions.

(3) If $\text{rank}(S) \neq \text{rank}(S_h)$, then the system $S$ has no solution (i. e. $\Omega(S) = \emptyset$).

**Theorem 5.40** (*Cramer's rule*) Let $S \in S_{m,n}(\mathbb{R})$ be SLE. If the determinant $D = \det(S) \neq 0$, then the system $S$ has unique solution

$$\vec{w} = \left( \frac{D_1}{D}, \frac{D_2}{D}, \frac{D_3}{D}, \ldots, \frac{D_n}{D} \right),$$

where $D$ is the determinant of the matrix of the system $S$ i. e. determinant of matrix $A = \{a_{ij}\}_{i,j=1}^{n}$ and $D_i$ is determinant $D$ with $i$-th column replaced by the right hand side $\vec{b}$, for $i \in \{1, 2, 3, \ldots, n\}$.

[16] *Particular solution* of the system $S$ is an arbitrary solution $\vec{w}^\star$ of the system $S$.

[17] This system always has at least one solution, namely the trivial solution.

[18] Fundamental theorem for the nonhomogeneous system of linear equations or also The Frobenius Theorem.

## 5.5 Numerical Methods for Solving Systems of Linear Equations

Iterative methods, unlike the exact methods, usually do not lead to the exact solution after the final number of steps of calculation. In iterative methods we choose the initial approximation of the solution and we will improve this solution by certain procedures (algorithms). This solution at each step of the iterative method will improve. Since we can not do the calculation for an infinite number of steps, we stop algorithm after finite number of steps. We obtain approximate solutions of the systems of linear equations.

### 5.5.1 Jacobi's Method

Let SLE with $n$ linear equations and $n$ unknown variables be given. Suppose main diagonal elements $a_{11}, a_{22}, \ldots, a_{nn}$ be nonzero.

$$\begin{aligned} a_{11}\cdot x_1 + a_{12}\cdot x_2 + \cdots + a_{1n}\cdot x_n &= b_1 \\ a_{21}\cdot x_1 + a_{22}\cdot x_2 + \cdots + a_{2n}\cdot x_n &= b_2 \\ a_{31}\cdot x_1 + a_{32}\cdot x_2 + \cdots + a_{3n}\cdot x_n &= b_3 \\ \vdots &= \vdots \\ a_{n1}\cdot x_1 + a_{n2}\cdot x_2 + \cdots + a_{nn}\cdot x_n &= b_n. \end{aligned}$$

From the first equation we express the variable $x_1$, from the second equation we express the variable $x_2$ and so on. From the $n$-th equation we express the variable $x_n$ and we obtain the system of equations:

$$\begin{aligned} x_1 &= \frac{1}{a_{11}}\cdot(b_1 - a_{12}\cdot x_2 - a_{13}\cdot x_3 - \cdots - a_{1n}\cdot x_n) \\ x_2 &= \frac{1}{a_{22}}\cdot(b_2 - a_{21}\cdot x_1 - a_{23}\cdot x_3 - \cdots - a_{2n}\cdot x_n) \\ x_3 &= \frac{1}{a_{33}}\cdot(b_3 - a_{31}\cdot x_1 - a_{32}\cdot x_2 - \cdots - a_{3n}\cdot x_n) \\ \vdots &= \vdots \\ x_n &= \frac{1}{a_{nn}}\cdot(b_n - a_{n1}\cdot x_1 - a_{n2}\cdot x_2 - \cdots - a_{n-1,n-1}\cdot x_{n-1}). \end{aligned}$$

At the beginning we choose any initial approximation of the solution, which will be denoted superscript zero $\vec{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, x_3^{(0)}, \dots, x_n^{(0)})^\top$. This solution will be substituted into the right-hand side of the rewritten system of equations and we obtain a new approximation of the solution $\vec{x}^{(1)} = (x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, \dots, x_n^{(1)})^\top$. We continue like this until we obtain the $k$-th approximation of the solution of the SLE in the form $\vec{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})^\top$. Then the $(k+1)$-st approximation of the solution of the SLE is calculated according to the scheme:

$$\begin{aligned}
x_1^{(k+1)} &= \frac{1}{a_{11}} \cdot \Big(b_1 - a_{12} \cdot x_2^{(k)} - a_{13} \cdot x_3^{(k)} - \cdots - a_{1n} \cdot x_n^{(k)}\Big) \\
x_2^{(k+1)} &= \frac{1}{a_{22}} \cdot \Big(b_2 - a_{21} \cdot x_1^{(k)} - a_{23} \cdot x_3^{(k)} - \cdots - a_{2n} \cdot x_n^{(k)}\Big) \\
x_3^{(k+1)} &= \frac{1}{a_{33}} \cdot \Big(b_3 - a_{31} \cdot x_1^{(k)} - a_{32} \cdot x_2^{(k)} - \cdots - a_{3n} \cdot x_n^{(k)}\Big) \\
\vdots &= \vdots \\
x_n^{(k+1)} &= \frac{1}{a_{nn}} \cdot \Big(b_n - a_{n1} \cdot x_1^{(k)} - a_{n2} \cdot x_2^{(k)} - \cdots - a_{n-1,n-1} \cdot x_{n-1}^{(k)}\Big).
\end{aligned} \tag{5.5}$$

This procedure will generate a sequence of approximations of the solution of the SLE $\{\vec{x}^{(0)}, \vec{x}^{(1)}, \vec{x}^{(2)}, \vec{x}^{(3)}, \dots, \vec{x}^{(k)}, \dots\}$, which can converge to the exact solution of the SLE. We continue in calculation until we reach required precision i. e. approximation of the solution stabilizes for the required number of decimal places or exceeds the given maximum number of iterations.

Now, this whole procedure will be described in matrix form. Given a square system of $n$ linear equations:

$$A \cdot \vec{x} = \vec{b},$$

we can write:

$$(A_c + A_d) \cdot \vec{x} = \vec{b},$$

$$A_d \cdot \vec{x} = \vec{b} - A_c \cdot \vec{x},$$

$$\vec{x} = A_d^{-1} \cdot (\vec{b} - A_c \cdot \vec{x}).$$

We obtain iterative formula:

$$\vec{x}^{(k+1)} = A_d^{-1} \cdot (\vec{b} - A_c \cdot \vec{x}^{(k)}), \tag{5.6}$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n-1} & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n-1} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n-1n-1} & a_{nn} \end{pmatrix},$$

$$A_c = \begin{pmatrix} 0 & a_{12} & \dots & a_{1n-1} & a_{1n} \\ a_{21} & 0 & \dots & a_{2n-1} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n-1n-1} & 0 \end{pmatrix},$$

$$A_d = \mathrm{diag}\{a_{11}, a_{22}, \dots, a_{n-1n-1}, a_{nn}\} = \begin{pmatrix} a_{11} & 0 & \dots & 0 & 0 \\ 0 & a_{22} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix}.$$

For each coordinate of the vector $\vec{x}^{(k+1)}$ we get the formula:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \cdot \left( b_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} \cdot x_j^{(k)} \right) \tag{5.7}$$

for $i \in \{1, 2, 3, \dots, n\}$.

Let us denote matrix $C = -A_d^{-1} \cdot A_c$ and vector $\vec{d} = A_d^{-1} \cdot \vec{b}$. Then we obtain iterative formula of the form:

$$\vec{x}^{(k+1)} = C \cdot \vec{x}^{(k)} + \vec{d}, \tag{5.8}$$

where

$$c_{ii} = 0, \qquad c_{ij} = -\frac{a_{ij}}{a_{ii}}, \qquad d_i = \frac{b_i}{a_{ii}},$$

for $i \neq j$ and $i, j \in \{1, 2, 3, \dots, n\}$.

Jacobi's method does not always converge to the exact solution of the SLE. Therefore, we have to specify the conditions under which the method converges.

**Definition 5.31** The matrix $A$ is called *strictly row-diagonally dominant* if and only if the following formula holds:

$$|a_{ii}| > \sum_{\substack{j=1 \\ i \neq j}}^{n} |a_{ij}| \qquad \text{for all} \quad i = 1, 2, \ldots, n. \tag{5.9}$$

The matrix $A$ is called *strictly column-diagonally dominant* if and only if the following formula holds:

$$|a_{jj}| > \sum_{\substack{i=1 \\ i \neq j}}^{n} |a_{ij}| \qquad \text{for all} \quad j = 1, 2, \ldots, n. \tag{5.10}$$

The matrix $A$ is said to be *diagonally dominant* if it is row or column diagonally dominant.

**Remark 5.11** Strict row diagonal dominance means, that for each row the absolute value of the diagonal element is greater than the sum of absolute values of other elements in considered row. Strict column diagonal dominance means, that for each column the absolute value of the diagonal element is greater than the sum of absolute values of other elements in considered column.

Note that this definition uses a weak inequality "$\geqq$", and is therefore sometimes called *weak diagonal dominance*. If a strict inequality "$>$" is used, this is called *strict diagonal dominance*.

The Jacobi's method sometimes converges even if the conditions (5.9) and (5.10) are not satisfied.

Now we show when the Jacobi's method converges. Using formula (5.8), we can write the SLE in the matrix form:

$$\vec{x} = C \cdot \vec{x} + \vec{d}, \tag{5.11}$$

where $C$ is the iteration matrix and $\vec{d}$ is an iterative vector of the Jacobi's method. Elements of the above matrix and vector have the form:

$$d_i = \frac{b_i}{a_{ii}}, \qquad c_{ij} = -\frac{a_{ij}}{a_{ii}}, \qquad \text{for} \quad i \neq j, \quad \text{and} \quad c_{ii} = 0.$$

The initial system of linear equations $A \cdot \vec{x} = \vec{b}$ will be modified to the system $\vec{x} = C \cdot \vec{x} + \vec{d}$. Our task is to find a solution of SLE, which corresponds to the task of finding a fixed point of mapping $F$:

$$F(\vec{x}) = C \cdot \vec{x} + \vec{d}, \tag{5.12}$$

because the solution of the initial SLE is such a vector $\vec{x}$, for which the following applies: $F(\vec{x}) = \vec{x}$. Then the general iterative step is as follows:

$$\vec{x}^{(k+1)} = F(\vec{x}^{(k)}) = C \cdot \vec{x}^{(k)} + \vec{d}. \tag{5.13}$$

We have a mapping $F$: $\mathbb{R}^n \longrightarrow \mathbb{R}^n$, where $\mathbb{R}^n$ is a vector space of all arranged $n$-tuples of real numbers. In this vector space we can define a metric by:

$$d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||,$$

where $|| \cdot ||$ is the norm. We use the norm $||\vec{x} - \vec{y}|| = \max_{i=1\ldots n}\{z_i : z_i = |x_i - y_i|\}$. Vector space $\mathbb{R}^n$ with this metric is complete. We need to check when $F$ is a contractive mapping. The following applies:

$$d(F(\vec{x}), F(\vec{y})) = ||F(\vec{x}) - F(\vec{y})|| = ||C \cdot \vec{x} + \vec{d} - (C \cdot \vec{y} + \vec{d})|| = ||C \cdot (\vec{x} - \vec{y})|| \leqq$$

$$\leqq ||C|| \cdot ||\vec{x} - \vec{y}|| = ||C|| \cdot d(\vec{x}, \vec{y}),$$

where $||C||$ is the matrix norm, $||C|| = \max_{i=1\ldots n}\{s_i : s_i = \sum_{j=1}^{n} c_{ij}\}$. We know that if $||C|| < 1$, then the $F$ is contraction mapping with a coefficient of contraction $\alpha = ||C||$. In this way it is ensured, that the sequence of successive approximations of solution using formula (5.13) converges to the fixed point of our mapping $F$. It remains to show how to check the condition $||C|| < 1$. In general case, it can be complicated, but for the SLE is easy to verify that condition, because it is related with diagonal dominance of the matrix $A$. If the matrix $A$ is row (column) diagonally dominant matrix, then the Jacobi method converges.

If the condition $||C|| < 1$ is satisfied, then for the error estimate for $k$-th iteration applies:

$$||\vec{x}^{(k)} - \vec{x}|| \leqq \frac{||C||}{1 - ||C||} \cdot ||\vec{x}^{(k)} - \vec{x}^{(k-1)}||, \tag{5.14}$$

$$||\vec{x}^{(k)} - \vec{x}|| \leqq \frac{||C||^k}{1 - ||C||} \cdot ||\vec{x}^{(0)} - \vec{x}^{(1)}||. \tag{5.15}$$

Using these estimates, we can decide when to stop the iterative process, if we are provided by $\varepsilon > 0$.

### 5.5.2 Gauss-Seidel's Method

Let SLE, which has $n$ equations and $n$ unknowns be given:

$$\begin{aligned}
a_{11}\cdot x_1 + a_{12}\cdot x_2 + \cdots + a_{1n}\cdot x_n &= b_1\\
a_{21}\cdot x_1 + a_{22}\cdot x_2 + \cdots + a_{2n}\cdot x_n &= b_2\\
a_{31}\cdot x_1 + a_{32}\cdot x_2 + \cdots + a_{3n}\cdot x_n &= b_3\\
\vdots &= \vdots\\
a_{n1}\cdot x_1 + a_{n2}\cdot x_2 + \cdots + a_{nn}\cdot x_n &= b_n.
\end{aligned}$$

The Gauss-Seidel's method is very similar to the Jacobi's method, but differs from it in the way that the calculation of further approximations of solution always uses the most recent values of the vector $\vec{x} = (x_1, x_2, x_3, \ldots, x_n)^\top$, that are available.

Main idea of the Gauss-Seidel method: With the Jacobi method, the values of $x_i^{(k)}$ obtained in the $k$-th iteration remain unchanged until the entire $(k+1)$-st iteration has been calculated. With the Gauss-Seidel's method, we use the new values $x_i^{(k+1)}$ as soon as they are known. For example, once we have computed $x_1^{(k+1)}$ from the first equation, its value is used in the second equation to obtain the new $x_2^{(k+1)}$, and so on.

Thus we obtain a modified iterative relation:

$$\begin{aligned}
x_1^{(k+1)} &= \frac{1}{a_{11}}\cdot\left(b_1 - a_{12}\cdot x_2^{(k)} - \cdots - a_{1n}\cdot x_n^{(k)}\right)\\
x_2^{(k+1)} &= \frac{1}{a_{22}}\cdot\left(b_2 - a_{21}\cdot x_1^{(k+1)} - a_{23}\cdot x_3^{(k)} - \cdots - a_{2n}\cdot x_n^{(k)}\right)\\
x_3^{(k+1)} &= \frac{1}{a_{33}}\cdot\left(b_3 - a_{31}\cdot x_1^{(k+1)} - a_{32}\cdot x_2^{(k+1)} - \cdots - a_{3n}\cdot x_n^{(k)}\right)\\
\vdots &= \vdots\\
x_n^{(k+1)} &= \frac{1}{a_{nn}}\cdot\left(b_n - a_{n1}\cdot x_1^{(k+1)} - a_{n2}\cdot x_2^{(k+1)} - \cdots - a_{n-1,n-1}\cdot x_{n-1}^{(k+1)}\right).
\end{aligned}$$

**Theorem 5.41** If a matrix $A$ is strictly diagonally dominant (row or column), then for any choice of $\vec{x}^{(0)}$, both the Jacobi's and the Gauss-Seidel's methods give sequences $\left\{\vec{x}^{(k)}\right\}_{k=0}^{\infty}$ that converge to the unique solution of the system $A\cdot\vec{x} = \vec{b}$.

## 5.6 Solved Examples

**Example 5.1** Adjust the system of linear equations by the equivalent operations to the system with diagonally dominant matrix. Perform three steps of Jacobi iterative method from the initial approximation $\vec{x}^{(0)}$ and estimate the upper bound of the error after the third step, if the SLE has the form:

$$\begin{aligned} x - 7y + 2z &= 7{,}3 \\ -x + 2y + 11z &= -8 \\ 18x - 11y - 8z &= 21{,}3. \end{aligned}$$

*Solution:*
The first step is to modify the given SLE to the diagonally dominant form. We write the third equation $q_3$ as the first and we move the first and second equations by one equation below. We get the system:

$$\begin{aligned} 18x - 11y - 8z &= 21{,}3 \\ x - 7y + 2z &= 7{,}3 \\ -x + 2y + 11z &= -8. \end{aligned}$$

We subtract the second equation from the first $(q_1 - q_2)$ and we get the SLE with diagonally dominant matrix:

$$\begin{aligned} 17x - 4y - 10z &= 14 \\ x - 7y + 2z &= 7{,}3 \\ -x + 2y + 11z &= -8. \end{aligned}$$

From this system we create a formula of Jacobi's iterative method.

$$\begin{aligned} x^{(k+1)} &= \frac{4}{17}y^{(k)} + \frac{10}{17}z^{(k)} + \frac{14}{17} \\ y^{(k+1)} &= \frac{1}{7}x^{(k)} + \frac{2}{7}z^{(k)} - \frac{7{,}3}{7} \\ z^{(k+1)} &= \frac{1}{11}x^{(k)} - \frac{2}{11}y^{(k)} - \frac{8}{11}. \end{aligned}$$

We could rewrite this system also in the form:

$$
\begin{aligned}
x^{(k+1)} &= 0{,}2353y^{(k)} + 0{,}5882z^{(k)} + 0{,}8235\\
y^{(k+1)} &= 0{,}1429x^{(k)} + 0{,}2857z^{(k)} - 1{,}0429\\
z^{(k+1)} &= 0{,}091x^{(k)} - 0{,}1818y^{(k)} - 0{,}7273.
\end{aligned}
$$

For the initial approximation we choose a vector: $x^{(0)} = 0{,}8235$, $y^{(0)} = -1{,}0429$, $z^{(0)} = -0{,}7273$ (right-hand side). We create a Table 5.1 in to which we write particular approximations.

Table 5.1: Solving of SLE by the Jacobi's iteration method.

| $k$ | $x^{(k)}$ | $y^{(k)}$ | $z^{(k)}$ |
|---|---|---|---|
| 0 | 0,8235 | −1,0429 | −0,7273 |
| 1 | 0,15031 | −1,13301 | −0,5200 |
| 2 | 0,25104 | −1,169985 | −0,61909 |
| 3 | 0,184054 | −1,1839 | −0,591905 |

We need to estimate the upper bound of the error of our solution: $x^{(3)} = 0{,}184054$, $y^{(3)} = -1{,}1839$, $z^{(3)} = -0{,}591905$. We calculate the norm of matrix $||C|| = \max\{0{,}2353 + 0{,}5882, 0{,}1429 + 0{,}2857, 0{,}091 + 0{,}1818\} =$
$= \max\{0{,}8235, 0{,}4286, 0{,}7788\} = 0{,}8235 < 1.$

Then for the error estimate holds:

$$||\vec{x}^{(k)} - \vec{x}|| \leqq \frac{||C||}{1 - ||C||} \cdot ||\vec{x}^{(k)} - \vec{x}^{(k-1)}||,$$

$$||\vec{x}^{(k)} - \vec{x}|| \leqq \frac{0{,}8235}{1 - 0{,}8235} \cdot 0{,}066986 = 0{,}3125.$$

$\surd$

## 5.7 Unsolved Tasks

**5.1** Calculate the matrix $C = 5 \cdot A + \frac{1}{2} \cdot B$, if

a) $A = \begin{pmatrix} 3 & -4 & 1 \\ -2 & 7 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -2 & 7 \\ 10 & 4 & 3 \end{pmatrix}$,

b) $A = \begin{pmatrix} 1 & 1 & -1 \\ -2 & 0 & 3 \\ 0 & 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 2 & -2 & 4 \\ 8 & 4 & 6 \\ -10 & 2 & -2 \end{pmatrix}$.

**5.2** Calculate the matrices $A^2$, $A^3$, and $A^4$, if

a) $A = \begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix}$,

b) $A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & -1 \\ 2 & -1 & 1 \end{pmatrix}$.

**5.3** Calculate the matrix $C = A \cdot B$, if

a) $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 7 \\ 1 & 3 & 6 & 10 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 & -2 & -4 \\ 0 & 1 & 0 & -1 \\ -1 & -1 & 3 & 6 \\ 2 & 1 & -6 & -10 \end{pmatrix}$,

b) $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 7 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 & -2 & -2 \\ 0 & 1 & 0 & -1 \\ 0 & -1 & 3 & 3 \\ 0 & 1 & 0 & 1 \end{pmatrix}$,

c) $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & -2 \\ 0 & 3 \\ -3 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 & -2 & -2 \\ 0 & 1 & 0 & -1 \end{pmatrix}$,

d) $A = \begin{pmatrix} 1 & 1 & 0 & -1 & -1 \\ 2 & 3 & 4 & -3 & -2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2 & -2 \\ -1 & 3 \\ 1 & 4 \end{pmatrix}$.

**5.4** Calculate the rank of the matrix $A$, if

a) $A = \begin{pmatrix} 1 & 2 & 0 & 3 \\ -2 & 0 & 1 & 1 \\ 1 & 2 & -1 & -4 \\ -1 & 2 & 0 & -3 \end{pmatrix}$,

b) $A = \begin{pmatrix} 3 & 2 & 1 \\ 6 & 1 & 8 \\ 1 & 0 & 3 \\ 4 & 3 & 2 \end{pmatrix}$,

c) $A = \begin{pmatrix} 1 & 0 & -1 & 3 \\ -6 & 0 & 6 & -18 \end{pmatrix}$,

d) $A = \begin{pmatrix} 3 & 1 & 2 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{pmatrix}$,

e) $A = \begin{pmatrix} -2 & 3 \\ 1 & \frac{3}{2} \\ 6 & -9 \\ -2 & 3 \end{pmatrix}$,

f) $A = \begin{pmatrix} 0 & 2 & 1 & 0 & 3 \\ 1 & 1 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 1 \\ 3 & 0 & 2 & 1 & 1 \end{pmatrix}$,

g) $A = \begin{pmatrix} 0 & 4 & 2 & 0 & 6 \\ 1 & 1 & 0 & 1 & 0 \\ -4 & -2 & -2 & 0 & -2 \\ 6 & 0 & 4 & 2 & 2 \\ -3 & 3 & 0 & 1 & 4 \end{pmatrix}$,

h) $A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ -4 & -1 & 3 & 0 & -2 \\ 0 & 0 & 2 & 4 & 2 \\ 0 & 0 & 0 & 4 & 5 \end{pmatrix}$.

**5.5** Verify that the matrix $A$ is a regular (invertible) or singular, if

a) $A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 2 & 1 \\ 1 & 4 & 0 \end{pmatrix}$,

b) $A = \begin{pmatrix} 1 & 1 & 1 \\ 4 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}$,

c) $A = \begin{pmatrix} 0 & 1 & 3 & 4 \\ -3 & 2 & 1 & -4 \\ -1 & 5 & 2 & -3 \\ 7 & 0 & -4 & 1 \end{pmatrix}$,

d) $A = \begin{pmatrix} 1 & 3 & 1 & 4 \\ 1 & 1 & -4 & 1 \\ 1 & 1 & 9 & 2 \\ 1 & 2 & 5 & 5 \end{pmatrix}$.

**5.6** Calculate by elementary operations the inverse matrix to the matrix $A$, if

a) $A = \begin{pmatrix} -2 & -1 & 1 \\ 2 & 0 & 0 \\ -2 & 1 & 1 \end{pmatrix}$,

b) $A = \begin{pmatrix} 2 & -1 & 2 \\ 0 & -1 & 3 \\ 3 & -1 & 1 \end{pmatrix}$,

c) $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ -5 & 2 & 1 & 0 \\ 7 & 3 & 2 & 1 \end{pmatrix}$,

d) $A = \begin{pmatrix} 3 & 0 & 1 & 0 \\ -2 & 2 & -2 & 1 \\ 0 & 2 & -3 & 2 \\ -1 & 1 & -2 & 1 \end{pmatrix}$.

**5.7** Calculate the determinant of the matrix $A$, if $\det(A) = |A|$ are given

a) $\det(A) = \begin{vmatrix} 1 & 3 & -1 \\ -1 & -2 & 2 \\ 2 & 1 & -1 \end{vmatrix}$,

b) $\det(A) = \begin{vmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 7 & 4 & -1 \end{vmatrix}$,

c) $\det(A) = \begin{vmatrix} 1 & 2 & 3 & 4 \\ -1 & 0 & -3 & -8 \\ -1 & 1 & 0 & -13 \\ 2 & 3 & 5 & 15 \end{vmatrix}$,

d) $\det(A) = \begin{vmatrix} 1 & 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & -2 & 1 \\ 1 & 1 & 2 & 1 & 1 \\ 1 & -2 & 1 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \end{vmatrix}$.

**5.8** Using Fundamental theorem (The Frobenius theorem) show the existence or absence of solutions for the following systems of linear equations $S$:

a)

$$\begin{aligned} x + y - z &= 1 \\ 2x + y - z &= 2 \\ y + z &= 4 \\ x - y - 3z &= -7, \end{aligned}$$

b)

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

**5.9** Solve the following systems of linear equations $S$:

a)

$$\begin{aligned} x_1 + 2x_2 - x_3 - 2x_4 &= -2 \\ 2x_1 + x_2 + x_3 + x_4 &= 8 \\ x_1 - x_2 - x_3 + x_4 &= 1 \\ x_1 + 2x_2 + 2x_3 - x_4 &= 4, \end{aligned}$$

b)

$$\begin{aligned} 2x_1 - x_2 - x_3 + 3x_4 &= 1 \\ 2x_1 - x_2 - 2x_4 &= 4 \\ 8x_1 - 4x_2 + x_3 - 13x_4 &= 19 \\ 6x_1 - 3x_2 - x_3 - x_4 &= 9, \end{aligned}$$

c)

$$\begin{aligned} 2x_1 + 2x_2 - 2x_3 + 5x_4 &= -6 \\ 2x_1 - x_2 + x_3 - x_4 &= 1 \\ 2x_1 - x_2 - 3x_4 &= 2 \\ 4x_1 - 2x_2 + x_3 - 4x_4 &= -3. \end{aligned}$$

**5.10** Solve the systems of linear equations by the Jacobi's iterative method with the precision $\varepsilon = 10^{-3}$, where

$$\begin{aligned} 18x_1 - 19x_2 + 2x_3 &= -12 \\ -2x_1 + x_2 - 15x_3 &= 15 \\ 2x_1 + 21x_2 - x_3 &= 23. \end{aligned}$$

## 5.8 Results of Unsolved Tasks

**5.1** **a)** $C = \begin{pmatrix} \frac{31}{2} & -21 & \frac{17}{2} \\ -5 & 37 & \frac{33}{2} \end{pmatrix}$ **b)** $C = \begin{pmatrix} 6 & 4 & -3 \\ -6 & 2 & 18 \\ -5 & 6 & 4 \end{pmatrix}$

**5.2** **a)** $A^2 = \begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix}$, $A^3 = \begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix}$, $A^4 = \begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix}$
**b)** $A^2 = \begin{pmatrix} 1 & -3 & 1 \\ -2 & 5 & -3 \\ 4 & -5 & 2 \end{pmatrix}$, $A^3 = \begin{pmatrix} 3 & -5 & 4 \\ -8 & 15 & -8 \\ 0 & -16 & 7 \end{pmatrix}$, $A^4 = \begin{pmatrix} 11 & -23 & 12 \\ -24 & 46 & -23 \\ 22 & -47 & 23 \end{pmatrix}$

**5.3** **a)** $C = \begin{pmatrix} 2 & 2 & -5 & 9 \\ 6 & 4 & -17 & -28 \\ 11 & 6 & -32 & -52 \\ 15 & 8 & -44 & -71 \end{pmatrix}$ **b)** $C = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 7 & 9 \\ 1 & 5 & 10 & 15 \end{pmatrix}$ **c)** $C = \begin{pmatrix} 1 & 1 & -2 & -3 \\ 1 & 2 & -2 & -4 \\ 1 & -2 & -2 & 0 \\ 0 & 3 & 0 & -3 \\ -3 & 0 & 6 & 6 \end{pmatrix}$ **d)** $C = \begin{pmatrix} 1 & -6 \\ 11 & -22 \end{pmatrix}$

**5.4** **a)** $h(A) = 2$ **b)** $h(A) = 3$ **c)** $h(A) = 1$ **d)** $h(A) = 3$ **e)** $h(A) = 1$ **f)** $h(A) = 4$ **g)** $h(A) = 4$ **h)** $h(A) = 5$

**5.5** **a)** regular **b)** singular **c)** regular **d)** singular

**5.6** **a)** $A^{-1} = \frac{1}{2} \cdot \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 2 & 1 \end{pmatrix}$ **b)** $A^{-1} = \begin{pmatrix} 2 & -1 & -1 \\ 9 & -4 & -6 \\ 3 & -1 & -2 \end{pmatrix}$ **c)** $A^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 11 & -2 & 1 & 0 \\ -20 & 1 & -2 & 1 \end{pmatrix}$ **d)** $A^{-1} = \begin{pmatrix} 1 & 0 & -1 & 2 \\ 1 & 1 & -1 & 1 \\ -2 & 0 & 5 & -6 \\ -4 & -1 & 6 & -10 \end{pmatrix}$

**5.7** **a)** $\det(A) = 6$ **b)** $\det(A) = -1$ **c)** $\det(A) = 24$ **d)** $\det(A) = 30$

**5.8** **a)** $h(S) = h(S_h) = 3$ **b)** $h(S) = 4 \wedge h(S_h) = 3$

**5.9** **a)** $\Omega(S) = \{\vec{x} \in \mathbb{R}^4 : \vec{x} = (1, 2, 1, 3)\}$ **b)** $\Omega(S) = \{\vec{x} \in \mathbb{R}^4 : \vec{x} = (2, 0, 3, 0) + \alpha \cdot (\frac{1}{2}, 1, 0, 0) + \beta \cdot (1, 0, 5, 1); \alpha, \beta \in \mathbb{R}\}$ **c)** $\Omega(S) = \emptyset$

**5.10** $\vec{x}_k = (0, 5; 1; -1)$

# Index

Absolute Value, 23
Adjoint Matrix
  see Adjugate Matrix, 132
Adjugate Matrix, 132
Antiderivative, 97
Aplication of Definite Integrals, 103
Approximation of Function, 85
  Interpolation, 85
  Inverse Lagrange's Interpolating Polynomial, 87
  Lagrange's Interpolating Polynomial, 85
  Least Squares Method, 87
  Normal Equations, 89

Banach's fixed-point theorem, 142
Base of Vector Space, 120
Bisection Method, 68
Bounded Function, 27

Co-factor, 129
Constant Function, 28
Cyclometric Function, 38

Definite Integral, 100
  Integral Sum, 101
  Lower Limit, 101
  Newton-Leibniz Formula, 102
  Norm of Partition, 101
  Normal Sequaence of Partitioning, 101
  Partition, 101
  Upper Limit, 101
Derivative, 49
  Bernoulli's rule, 53
  Cauchy's Theorem, 53
  L'Hospital's rule, 53
  Lagrange's Theorem, 53
  Left-hand, 49
  Mean Value Theorem, 53
  Right-hand, 49
  Rolle's Theorem, 53
Determinant of Matrix, 129
  Adjoint Matrix, 132
  Adjugate Matrix, 132
  Co-factor, 129
  Expansion of Determinant, 129
  Minor, 129
Dimension of Vector Space, 121

Elementary Functions, 28
  arccosine (arccos), 38
  arccotangent (arccotg ), 38
  arcsine (arcsin), 38
  arctangent (arctg ), 38
  Constant, 28
  cosine (cos), 36
  cotangent (cotg ), 36
  Cyclometric, 38
  Exponential, 33
  Linear, 29
  Logarithmic, 34

Polynomial, 41
Power, 32
Quadratic, 29
sine (sin), 36
tangent (tg ), 36
Trigonometric, 36
Elementary Operations, 121
Equation
Bisection Method, 68
Fixed Point Iteration Method, 71
Coefficient of Contraction, 71
Contractive Mapping, 71
Error Estimate, 72
Fixpoint of Mapping, 71
Fourier's Condition, 76
Newton's Method, 74
Non-Linear, 67
Root, 67
Separation of Roots, 67
Equivalent Matrix, 127
Equivalent SLE, 135
Equivalent Systems of Vectors, 121
Exponential Function, 33
Extreme Value, 27

Fixed Point Iteration Method, 71
Error Estimate, 72
Iteration Function, 73
Iterative Process, 73
Fourier's Condition, 76
Function, 25
Approximation of Function, 85
Bounded, 27
Bounded Above, 27
Bounded Below, 27
Composite, 26
Concave, 56
Strictly, 56
Concave Down, 56
Purely, 56
Concave Up, 56
Purely, 56
Continuous, 47
from Left, 47
from Right, 47
on Closed Interval, 47
on Open Interval, 47
on Set, 47
Convex, 56
Strictly, 56
Decreasing, 26
Dependent Variable, 25
Derivative, 49
Left-hand, 49
Right-hand, 49
Domain, 25
Elementary Functions, 28, 29, 32–34, 36, 38, 41
Even, 26
Extreme Value, 27
Global Maximum, 55
Global Minimum, 55
Graph, 25
Increasing, 26
Independent Variable, 25
Infimum, 27
Inside, 26
Integrable, 100, 101
Inverse, 26
Inverse Lagrange's Interpolating Polynomial, 87
Least Squares Method
System of Normal Equations, 90
Limit, 42
Local Maximum, 55
Local Minimum, 55

Lover Bounded, 27
Maximum, 27, 55
Minimum, 27, 55
Monotonic, 26
Non-decreasing, 26
Non-increasing, 26
Odd, 26
One-To-One, 26
Outside, 26
Period, 26
Periodic, 26
Point
Extrema, 55
Inflection, 57
Inflexion, 57
Stationary, 55
Range, 25
Strictly Local Maximum, 55
Strictly Local Minimum, 55
Strictly Monotonic, 26
Supremum, 27
Upper Bounded, 27
Function arccosine (arccos), 38
Function arccotangent (arccotg ), 38
Function arcsine (arcsin), 38
Function arctangent (arctg ), 38
Function cosine (cos), 36
Function cotangent (cotg ), 36
Function sine (sin), 36
Function tangent (tg ), 36
Fundamental system of solutions of HSLE, 135

Gauss-Seidel Methods, 143
Gaussian form of SLE, 135
Group of Vectors
Rank, 121
see also System of Vectors, 120

Indefinite Integral, 98
Integral, 97
Definite, 100
Integration by Parts, 103
Properties, 102
Substitution, 103
The Elemental Area, 103
Formulas for Integration, 98
Indefinite, 98
Integrating, 98
Integration by Parts, 99
Newton-Leibniz Formula, 102
Numerical Calculation, 105
Newton-Cotes Formulas, 105
Rectangular, 107
Simpson's, 109
Trapezoidal, 108
Per-Partes Method see Integration by Parts, 99
Riemann's Integral, 100
Rules for Integration, 98
Substitution Method, 100
Integration by Parts, 99
Interpolation, 85
Interval, 24
Bounded, 24
Closed, 24
Half-closed, 24
Half-open, 24
Open, 24
Unbounded, 24
Inverse Lagrange's Interpolating Polynomial, 87
Inverse Matrix, 127

Jacobi Iterative Method
Contraction, 142
Contraction Mapping, 142

Diagonal Dominance, 140
Diagonally Dominant Matrix, 140
Strict Column, 140
Strict Row, 140
Jacobi's Iteration Method, 138
Jacobi's Iterative Method, 138
Jacobi's Method, 138

Lagrange's Interpolating Polynomial, 85
Least Squares Method, 87
Normal Equations, 89
System of Normal Equations, 90
Limit from the Left
see Left Hand Limit, 42
Limit from the Right
see Right Hand Limit, 42
Limit of Function, 42
Bernoulli's rule, 53
Infinite, 47
L'Hospital's rule, 53
Left Hand Limit, 42
Real, 47
Right Hand Limit, 42
Linear Combinations of Solutions of SLE, 135
Linear Function, 29
Logarithmic Function, 34

Mapping
Coefficient of Contraction, 72
Contraction, 72
Fixpoint, 72
Matrix, 123
Adjugate Matrix, 132
Basic Properties, 125
Column, 124
Diagonal, 124
Element of Matrix, 123
Elementary Operations, 127
Entry of Matrix, 123
Equivalence of Matrices, 127
Identity, 124
Inverse, 127
Lower Triangular, 124
Operations with Matrices, 125
Addition, 125
Equality, 125
Product, 125
Scalar Multiplication, 125
Sum, 125
Rank, 127
Rectangular, 124
Regular, 127
Row, 124
Square, 124
Symmetric, 125
Transpose, 125
type $m \times n$, 123
Unit, 124
Upper Triangular, 124
Zero, 124
Maximum of Function, 27
Midpoint Rule, 107
Minimum of Function, 27

Newton's Method, 74
Newton-Cotes Formulas
Closed Formulas, 105
Open Formulas, 105
Newton-Leibniz Formula, 102
Number
Absolute Value, 23
Algebraic, 21
Transcendental, 21
Number Set

Number Line, 19
Real Line, 19
Number Sets
Complex, 17
Integer, 17
Irrational, 17
Natural, 17
Order, 20
Rational, 17
Real, 17

Order of Real Numbers, 20
Greater Then, 20
Greater Then or Equal, 20
Less Then, 20
Less Then or Equal, 20

Period of Function, 26
Polynomial Function, 41
Power Function, 32

Quadratic Function, 29

Real Numbers, 19
Associative Laws, 19
Commutative Laws, 19
Distributive Laws, 19
Order, 20
Real Number Line, 19
Rectangular Rule, 107

Separation of Roots, 67
Set
Bounded, 22
Bounded from Above, 22
Bounded from Below, 22
Infima, 22
Lower Bound, 22
Maximum, 22
Minimum, 22
Suprema, 22
Upper Bound, 22
Simpson's Rule, 109
Surface Area of Rotating Shape, 104
System of Linear Algebraic Equations, 132
System of Linear Equations
Contradictory Equation, 135
Cramer's Rule, 137
Elementary Operations, 135
Equivalence of SLE, 135
Frobenius Theorem, 137
Fundamental system of solutions, 135
Fundamental Theorem, 137
Gauss-Seidel Methods, 143
Gaussian form, 135
Homogeneous, 135
Jacobi's Iterative Method, 138
Linear Combinations of Solutions, 135
Non-homogeneous, 135
Pivot, 135
Pivot of equation, 135
Rank, 135
Set of Solutions, 135
Solution, 135
Nontrivial, 135
Trivial, 135
Zero Equation, 135
System of Normal Equations, 89
System of Vectors, 120
Linearly Dependent, 120
Linearly Independent, 120
Rank, 121
Systems of Linear Equations, 132
Homogeneous, 132

Non-Homogeneous, 132

Tangential Method
see also Equation, 74
see also Newton's Method, 74
The Length of the Curve, 104
Theorem
Banach's fixed-point theorem, 142
Bernoulli's rule, 53
Cauchy, 53
Cramer's Rule, 137
Fourier's Condition, 76
Frobenius, 137
Fundamental, 137
L'Hospital's rule, 53
Lagrange's, 53
Mean Value, 53
Rolle's, 53
Trapezoidal Rule, 108
Trigonometric Function, 36

Value of Elemental Surface Area, 103
Vector, 119
Linear Combination, 120
Non-trivial, 120
Trivial, 120
Span Set, 120
Unit, 119
Zero, 119
Vector Space, 119
Basis, 120
Dimension, 121
Elementary Operations, 121
Finite-Dimensional, 120
Infinite-Dimensional, 120
Rank of the System of Vectors, 121
Subspace, 119
Volume of Rotating Shape, 104

# Lexicon – Vocabulary

## English – Slovak

# A

| | |
|---|---|
| - absolute value | - absolútna hodnota |
| - absolute value of real number | - absolútna hodnota reálneho čísla |
| - algebraic number | - algebraické číslo (algebrické) |
| - addition of matrices | - súčet matíc |
| - adjoint matrix | - adjungovaná matica |
| - adjugate matrix | - adjungovaná matica |
| - antiderivative | - antiderivácia |
| - approximation of function | - aproximácia funkcie |
| - associative law | - asociatívny zákon |

# B

| | |
|---|---|
| - Banach fixed-point theorem | - Banachova veta o pevnom bode |
| - Bernoulli's rule | - Bernouliho pravidlo |
| - base of vector space | - báza vektorového priestoru |
| - bisection method | - metóda bisekcie |
| | - metóda poltenia intervalu |
| - bounded above | - ohraničená zhora |
| - bounded below | - ohraničená zdola |
| - bounded function | - ohraničená funkcia |
| - bounded interval | - ohraničený interval |
| - bounded set | - ohraničená množina |
| - bounded set from above | - ohraničená množina zhora |
| - bounded set from below | - ohraničená množina zdola |

# C

- Cauchy's theorem - Cauchyho veta
- closed interval - uzavretý interval
- coeficient of contraction - koeficient kontrakcie
- co-factor -
- column matrix - stĺpcová matica
- column vector - stĺpcový vektor
- complex number - komplexné číslo
- composite function - zložená funkcia
- commutative law - komutatívny zákon
- constant function - konštantná funkcia
- concave up function - konvexná funkcia
- concave down function - konkávna funkcia
- contimuous function - spojitá funkcia
- continuous from the left - spojitá zľava
- contimuous from the right - spojitá zprava
- continuous on the closed interval - spojitá na uzavretom intervale
- continuous on the opened interval - spojitá na otvorenom intervale
- contimuous on the set - spojitá na množine
- contraction - kontrakcia
- contraction mapping - kontraktívne zobrazenie
- contradictory equation - sporná rovnica
- Cramer's rule - Cramerovo pravislo
- cyclometric function - cyklometrické funkcie

# D

- decreasing function - klesajúca funkcia
- definite integral - určitý integrál
- dependent variable - závislá premenná
- derivative -
- determinant of matrix - determinant matice
- diagonal matrix - diagonálna matica
- diagonally dominant matrix - diagonálne dominantná matica
- dimension of vector space - dimenzia vektorového priestoru
- distributive law - distributívny zákon
- domain of function - definičný obor funkcie

# E

- element of matrix - prvok matice
- elemental area - elementárna oblasť
- elementary functions - elementárne funkcie
- elementary operations - elementárne operácie
- entry of matrix - prvok matice
- equality of matrices - rovnosť matíc
- equation - rovnica
- equivalence of matrices - ekvivalencia matíc
- equivalence of SLE - ekvivalencia SLR
- equivalent matrix - ekvivalentná matica
- equivalent SLE - ekvivalentná SLR
- equivalent system of vectors - ekvivalentný systém vektorov
- error estimate - odhad chyby
- even function - párna funkcia
- expansion of determinant - rozvoj determinantu
- exponential function - exponenciálna funkcia
- extrema point of function - extrém funkcie
- extreme value - extrémna hodnota

# F

- fixed point iteration method - metóda prostej iterácie
- fixpoint of mapping - pevný bod zobrazenia
- formulas for intergration - integračné vzorce
- Fourier's condition - Fourierova podmienka
- Frobenius theorem - Frobéniova veta
- function - funkcia
- function is lower bounded - funkcia je zdola ohraničená
- function is upper bounded - funkcia je zhora ohraničená
- fundamental system of solutions of HSLE - fundamentálny systém riešení HSLR
- fundamental theorem - základná (fundamentálna) veta

# G

- Gaussian form of SLE - Gaussov tvar SLR
- Gauss-Seidel Method - Gaussova-Seidelova metóda

- global minimum - globálne minimum
- global maximum - globálne maximum
- graph of function - graf funkcie
- greater than - väčší než
- greater then or equal - väčší alebo rovný než
- group of vectors - množina vektorov


# H


- half-closed interval - polo-uzavretý interval
- half-open interval - polo-otvorený interval
- homogeneous SLE - homogénna SLR


# I


- identity matrix - jednotková matica
- increasing function - rastúca funkcia
- indefinite integral - neurčitý integrál
- independent variable - nezávislá premenná
- infimum - infímum
- inflection point of funkction - inflexný bod funkcie
- inside function - vnútorná zložka funkcie
- integer (number) - celé číslo
- integrable function - integrovateľná funkcia
- integral - integrál
- integrating - integrovanie
- integration by parts - integrovanie metódou per-partes
- integration by substitution method - integrovanie substitučnou mtódou
- interpolation - interpolácia
- interval - interval
- inverse function - inverzná funkcia
- inverse Lagrange's interpolating polynomial - inverzný lagrangeov interpolačný polynóm
- inverse matrix - inverzná matica
- integral sum - integrálny súčet
- irrational number - irecionálne číslo
- iteration function - iteračná funkcia
- iterative process - iteračný proces

# J

| | |
|---|---|
| - Jacobi iterative method | - Jacobiho iteračná metóda |

# K

# L

| | |
|---|---|
| - Lagrange's interpolating polynomial | - Lagrangeov interpolačný polynóm |
| - Lagrange's theorem | - Lagrangeova veta |
| - least squared method | - metóda najmenších štvorcov |
| - left-hand derivative | - derivácia zľava |
| - left-hand limit | - limita zľava (ľavostranná limita) |
| - length of the curve | - dĺžka krivky |
| - less then | - menší než |
| - less then or equal | - meší alebo rovný než |
| - L'Hospital's rule | - L'Hospitalovo pravidlo |
| - limit of function | - limita funkcie |
| - linear combination of solutions of SLE | - lineárna kombinácia riešení SLR |
| - linear combination of vectors | - lineárny kombinácia vektorov |
| - linear function | - lineárna funkcia |
| - linearly dependent | - lineárne závislý |
| - linearly independent | - lineárne nezávislý |
| - local maximum | - lokálne maximum |
| - local minimum | - lokálne minimum |
| - logarithmic function | - logaritmická funkcia |
| - lower bound of function | - dolné ohraničenie funkcie |
| - lower limit | - dolná hranica |
| - lower triangular matrix | - dolná trojuholníková matica |

# M

| | |
|---|---|
| - maximum of function | - maximum funkcie |
| - minimum of function | - minimum funkcie |
| - mean value theorem | - veta o strednej hodnote |
| - minor | - |
| - monotonic function | - monotónna funkcia |

# N


- natural number - prirodzené číslo
- Newton-Cotes formulas - Newton-Cotesove vzorce
- Newton-Leibniz formula - Newtonova Leibnizova formula
- Newton's method - Newtonova metóda
- non-decreasing function - neklesajúca funkcia
- non-homogeneous SLE - nehomogánna SLR
- non-increasing function - nerastúca funkcia
- non-linear equation - nelineárna rovnica
- non-trivial linear combination - netriviálna lineárny kombinácia
- non-trivial solution of HSLE - netriviálne riešenie HSLR
- normal equations - sústava normálnych rovníc
- norm of partition - norma delenia
- normal sequaence of partitioning - postupnosť normálnych delení
- number - číslo
- number line - číselná os
- number set - číselná množina
- numerical calculation - numerický výpočet


# O


- odd function - nepárna funkcia
- one-to-one function - prostá funkcia
- opened interval - otvorený interval
- operations with matrices - operácie s maticami
- order of numbers - usporiadanie čísel
- order of real numbers - usporiadanie reálnych čísel
- outside function - vonkajšia zložka funkcie


# P


- partition - delenie
- period of funkction - perióda funkcie
- periodic function - periodická funkcia
- pivot of equation - vedúci člen rovnice
- polynomial function - polynomická funkcia
- power function - mocninová funkcia
- product of matrices - súčin matíc

# Q

| | |
|---|---|
| - quadratic function | - kvadratická funkcia |

# R

| | |
|---|---|
| - range of function | - obor hodnôt funkcie |
| - rank | - hodnosť |
| - rank of matrix | - hodnosť matice |
| - rational number | - racionálne číslo |
| - real line | - os reálnych čísel |
| - real number | - reálne číslo |
| - rectangular matrix | - obdĺžniková matica |
| - rectangular method | - obdĺžniková metóda |
| - regular matrix | - regulárna matica |
| - Rieman's integral | - Riemanov integrál |
| - right-hand derivative | - derivácia zprava |
| - right-hand limit | - limita zprava (pravostranná limita) |
| - root of equation | - koreň rovnice |
| - row matrix | - riadková matica |
| - row vector | - riadkový vektor |
| - rules for integration | - pravidlá integrovania |

# S

| | |
|---|---|
| - scalar multiplication | - násobenie skalárom |
| - Simpson's method | - Simpsonová metóda |
| - separation of roots | - separácia koreňov |
| - set of solutions | - množina riešení |
| - solution | - riešenie |
| - solution of equation | - riešenie rovnice |
| - span set of vectors | - lineárny obal množiny vektorov |
| - square matrix | - štvorcová matica |
| - stationary point of function | - stacionárny bod funkcie |
| - strictly concave up | - rýdzo konvexná |
| - strictli concave down | - rýdzo konkávna |
| - strictly monotonic function | - rýdzo monotónna funkcia |
| - subspace | - podpriestor |

- supremum - suprémum
- sum - súčet (suma)
- surface area of Rotating Shape - povrch rotačného telesa
- symetric matrix - symetrická matica
- system of linear algebraic equations - systém lineárnych algebraických rovníc
- system of linear equations - sústava lineárnych rovníc
- system of vectors - systém vektorov

# T

- transcendental number - transcendentálne číslo
- transpose matrix - transponovaná matica
- trapezoidal method - lichobežniková metóda
- trigonometric functions - trigonometrické funkcie
- trivial linear combination - triviálna lineárna kombinácia
- trivial solution of HSLE - triviálne riešenie HSLR

# U

- unit matrix - matica jednotiek
- unit vector - vektor jednotiek
- upper bound of funkction - horné ohraničenie funkcie
- upper limit - horná hranica
- upper triangular matrix - horná trojuholníková matica
- unbounded interval - neohraničený interval

# V

- value of elemental area - veľkosť elementárnej oblasti
- value of surface area - veľkosť plochy
- vector - vektor
- vector space - vektorový priestor
- volumen of rotating shape - objem rotačného telesa

# W

# Z

- zero equation - nulová rovnica

- zero matrix - nulová matica
- zero vector - nulový vektor (origin)

# Bibliography

[1] Berežný Š. – Mislivcová V.: *Prehľad základov z matematiky*, LF TU, Košice (2009) ISBN: 978-80-553-0241-6.

[2] Bučko M. – Buša J. – Schrötter Š.: *Lineárna algebra*, Elfa, Košice (2001).

[3] Buša J. – Pirč V. – Schrötter Š.: *Numerické metódy, pravdepodobnosť a matematická štatistika*, KM FEI TUKE, Košice (2006). ISBN: 80-8073-633-2

[4] Džurina J. – Grinčová A. – Pirč V.: *Matematika 1*, Elfa, Košice (2009) ISBN: 978-80-8086-115-5.

[5] Džurina J. – Pirč V.: *Calculus I*, First Edition, Elfa, Košice (2010) ISBN: 978-80-8086-165-0.

[6] Eliáš J. – Horváth J. – Kajan J.: *Zbierka úloh z vyššej matematiky 1* (6. a neskoršie vydanie), Alfa, Bratislava (1985).

[7] Flanders Harley: *Calculus*, W. H. Freeman and Company, USA (1985), ISBN: 0-7167-1643-7.

[8] Havel V. – Holenda J.: *Lineární algebra* (učebnica pre EF), SNTL, Praha (1984).

[9] Chirwin H. B. – Plumpton Ch.: *A Course of Mathematics for Engineers and Scientists*, Volume 1 – Mathematical Methods, Second Edition, Pergamon Press Ltd., Oxford (1970).

[10] Růžičková I. – Hlavička R.: *Numerické metody*, Ústav matematiky FSI VUT v Brně.

[11] Kaprálik P. – Tvarožek J.: *Zbierka riešených príkladov a úloh z lineárnej algebry a analytickej geometrie*, Alfa, Bratislava (1987).

[12] Kreyszig Erwin: *Advanced Engineering Mathematics*, John Wiley & sons Inc., $6^{th}$ edition, USA (1988), ISBN: 0-471-85824-2.

[13] Kreyszig Erwin: *Advanced Engineering Mathematics*, John Wiley & sons Inc., $8^{th}$ edition, USA (1999), ISBN: 0-471-33328-X.

[14] Kreyszig H. – Kreyszig E.: *Advanced Engineering Mathematics – Student Solutions Manual*, John Wiley & sons Inc., $8^{th}$ edition, USA (2000), ISBN: 0-471-33375-1.

[15] Daňo I. – Ostertagová E.: *Numerické metódy, pravdepodobnosť a matematická štatistika v príkladoch*, Elfa, Košice (2007).

[16] Olejník F. – Mislivcová V.: *Matematika 1*, Elfa, Košice, (2006).

[17] Rektorys K.: *Přehled užité matematiky I*, Prometheus, Praha (2000).

[18] Rektorys K.: *Přehled užité matematiky II*, Prometheus, Praha (2000).

[19] Sedláková K.: *Vektory a matice*, RVS VLA, Košice, (1999).

[20] Sedláková K.: *Lineárna algebra*, RVS, VVLŠ, Košice (1994).

[21] Sedláková K.: *Príklady z lineárnej algebry*, RVS, VVLŠ, Košice (1995).

[22] Železník V.: *Úvod do numerickej matematiky*, RVS, VLA, Košice (1999).

Názov: MATHEMATICS 1

Author: © RNDr. Štefan BEREŽNÝ, PhD., 2014

Recenzovali: Mgr. Ján BUŠA, Ph.D.
RNDr. Kristína BUDAJOVÁ, PhD.
prof. RNDr. Jozef DŽURINA, PhD.

Vydavateľ: Technická univerzita v Košiciach
Fakulta elektrotechniky a informatiky

Miesto vydania: Košice
Rok vydania: 2014
Edition: First
Rozsah: 170 pages
Náklad: 100 ks

ISBN: 978-80-553-1788-5