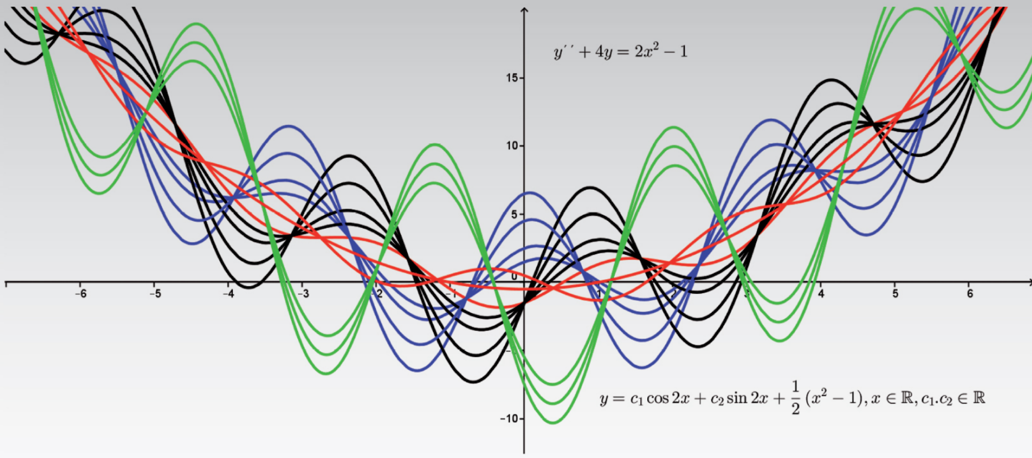
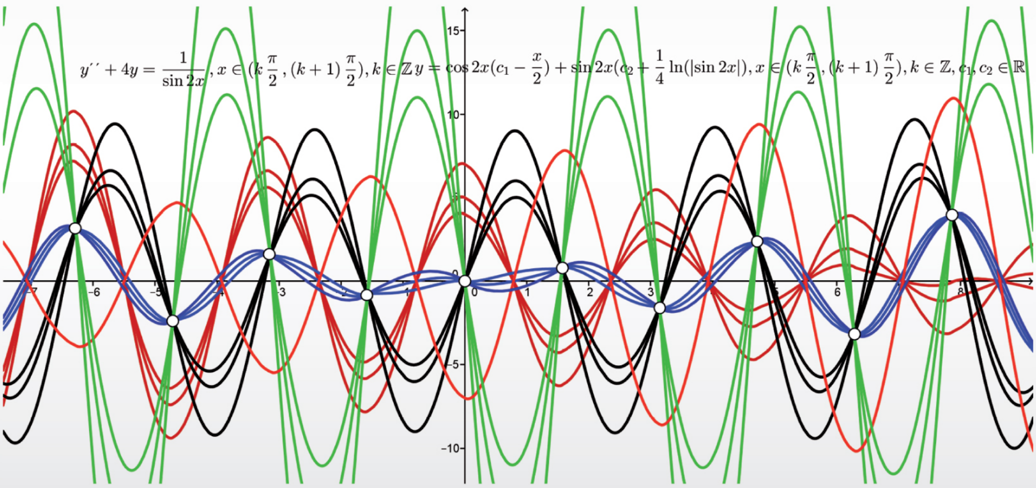
MATHEMATICS I





Daniela Velichová

STU

SLOVENSKÁ TECHNICKÁ
UNIVERZITA V BRATISLAVE

MATHEMATICS I

MATHEMATICS I

Daniela Velichová

SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
2014

Publikácia Mathematics I je základnou učebnicou rovnomenného povinného predmetu všetkých bakalárskych programov na Strojníckej fakulte STU pre štúdium v angličtine. Učebný text je spracovaný s podrobným výkladom, príkladmi, aplikáciami a grafickými ilustráciami pre všetky základné pojmy podľa akreditovaných sylabov predmetu, zoradené do kapitol: Basics of linear algebra, Defferential calculus of real functions of one real variable, Integral calculus of real functions of one real variable, Basics of ordinary defferential equations. Učebnica obsahuje okrem vyriešených príkladov aj úlohy na riešenie vo všetkých kapitolách s uvedenými správnymi výsledkami.

Recenzenti: Dr. Margareta Rebelos, Ph.D., BA (Hons), Dipl. Lang.
doc. RNDr. Ľubomír Marko, PhD.

Schválila Vedecká rada Strojníckej fakulty STU v Bratislave.

ISBN 978-80-227-4130-9

I hear and I forget.

I see and I understand.

I do and I remember.

Confucius

1 Introduction

1.1 Basic concepts from mathematical logic

Mathematical logic deals with statements and their truth values. A proposition is any statement expressed in written or oral form, whose truth value can be evaluated. Any proposition is either true or false, and we can relate a certain truth value to it, which can be symbolically denoted as 1 or 0. More complex compound statements can be formed as compositions of two or more sentences using logical connectives. Corresponding semantics of logical connectives are truth functions, whose values are expressed in the form of truth tables. The most common logical connectives are binary connectives (also called dyadic connectives) that join two sentences which can be thought of as the function's operands. Negation is considered to be a unary connective.

Examples

1. Any positive number is greater than zero.
2. Any parallelogram is a square.
3. There exist at least three divisors of number 21.
4. No equilateral triangle is a right triangle.
5. Bratislava is the largest city in the world.
6. It is Sunday today.
7. $3 + 7 = 10$

Propositions 1., 4. and 7. are true, while propositions 2., 3. and 5. are false. Truth value of the proposition 6. is sometimes 0 and sometimes it is 1, with respect to what day it is today.
Propositions will be denoted by small letters p, q, ... In case of a true proposition p we say that p truth value is 1. Truth value of a false proposition p is zero.
Any proposition p can be changed to a proposition negating p called negation of proposition p, which can be read as "it is not the case that p", "not that p", or more simply (though not fully grammatically correct) as "not p". Truth value of proposition negation is opposite to the truth value of the original proposition. Negation is denoted by logical operator $\neg$ with the meaning non, if p is a proposition, then its negation is $\neg p$.

Examples

1. p: Any parallelogram is a square.
   $\neg p$: ”It is not the case that any parallelogram is a square.", or “There exists a parallelogram, which is not a square."
2. p: It is raining in Bratislava now. $\neg p$: It is not raining in Bratislava now.
3. p: Number 25 is not a prime number. $\neg p$: Number 25 is a prime number.
4. p: All people are polite. $\neg p$: Not all people are polite.

If proposition q is negation of proposition p, then negation of proposition q is proposition p. The following two logical rules are true for negations of propositions.

Principle of contradiction consists of a logical incompatibility between propositions p and $\neg p$, which cannot have the same truth value.

Principle of bivalence states that every declarative sentence expressing a proposition has exactly one truth value, either true or false. There exists no third possibility. A logic satisfying this principle is called a two-valued or bivalent logic.
Simple propositions can be combined to more complex ones by means of several logical connectives. Logical connectives are: $\vee$ – or (disjunction, alternation), $\wedge$ – and (conjunction), $\Rightarrow$ – if, then (implication, conditional), $\Leftrightarrow$ – if and only if (equivalence, bi-conditional).
Disjunction of propositions (alternation) is a logical sum of propositions p and q, with the meaning that at least one of the two propositions is true. Disjunction is denoted by sign $\vee$, disjunction of propositions p and q is $p \vee q$ (p or q). Connective or – $\vee$ does not have the meaning of exclusion, therefore the possibility that both propositions p and q are true is not excluded. Disjunction is a true proposition, provided at least one of the propositions p and q is true, and it is false only in the case that both p and q are false propositions.

Examples
1. Number 7 is odd, or any multiple of number 7 is odd.
2. Each parallelogram has an even number of vertices or even number of sides.
3. This triangle has a right angle or it is equilateral.
4. Any integer is either positive or negative.

Conjunction of propositions is a logical product of propositions p and q, with the meaning that both propositions p and q are true. Conjunction is denoted by the sign $\wedge$, conjunction of propositions p and q is $p \wedge q$ (p and q). Conjunction is a true proposition only for both true p and q, and it is false in all other instances, when one or both from propositions p and q are false.

Examples
1. Number 2 is even and any multiple of number 2 is even.
2. Right-angled triangle has two equal angles and there exists an isosceles right-angled triangle.
3. Any rectangular has equal diagonals and there exists a rectangular with perpendicular diagonals.
4. Any positive number is greater than zero and any negative number.

Implication of propositions is a proposition composed of propositions p and q, with the meaning that proposition p is true only if proposition q is true. Implication is denoted by the sign $\Rightarrow$, implication of propositions p and q is $p \Rightarrow q$ (p implies q, or if p then q).

Examples

1. If at least two axes of symmetry of a polygon exist, then its centre of symmetry in the common point of the two axes of symmetry also exists.
2. If an even number is divisible by five, then it is divisible by ten.
3. If a triangle has a right angle, then it is not equilateral.
4. Any natural number is an integer.
5. If it rains today, then it is Wednesday tomorrow.

Equivalence of propositions is a proposition composed of propositions p and q, with the meaning that one of them is true only if the other one is true.
Equivalence is denoted by the sign $\Leftrightarrow$, equivalence of propositions p and q is $p \Leftrightarrow q$ (p is equivalent to q, or p if and only if q, or p iff q).

Examples

1. A number is divisible by six if and only if it is divisible by two and three.
2. Two triangles are congruent if they have two equal sides and angle formed by them.
3. Quadrilateral is a square if and only it has four equal sides and its diagonals are perpendicular to each other.
4. Sum of two non-zero integers is zero, if the two numbers are opposite numbers.

p	q	$\neg p$	$p \vee q$	$p \wedge q$	$p \Rightarrow q$	$p \Leftrightarrow q$
1	1	0	1	1	1	1
1	0	0	1	0	0	0
0	1	1	1	0	1	0
0	0	1	0	0	1	1

Table 1.1. Table of truth values

Composed propositions, which are true not depending on the truth values of the elementary propositions of which they are composed, are called tautologies.

Examples

1. $p \vee \neg p$ (principle of bivalence)
2. $\neg (p \wedge \neg p)$ (principle of contradiction)
3. $(p \Rightarrow q) \Leftrightarrow (\neg p \vee q)$, $\neg (p \Rightarrow q) \Leftrightarrow (p \wedge \neg q)$
4. $[(p \vee q) \vee r] \Leftrightarrow [p \vee (q \vee r)]$ (associativity of disjunction)

Propositional function is a function $V(x)$ that may be true or false depending on the values of its variable. A set of all x, for which $V(x)$ is a proposition, is called domain of definition of the propositional function. A set of all x, for which $V(x)$ is a true proposition, is called the truth domain of the propositional function. Here, $V(x)$ is referred to as the predicate, and x the subject of the proposition, as each choice of x produces a proposition.

Examples

1. $V(n)$: Any regular n-gon has a centre of symmetry.
2. If it rains on the Medards's name day, then it will rain for the following 40 days. $V(x)$: If it rains on the 8^{th} of June x, then it will also rain for the following 39 days in a given year x, therefore between 9^{th} of June x – 17^{th} of July x, $x \in \boldsymbol{N}$.
3. $V(x)$: $\{x$: x is a positive integer less than $4\}$ is the set $\{1, 2, 3\}$.
4. $V(n)$: Any regular n-gon has a centre of symmetry.

Quantifiers

Proposition ,,For any x from the set M statement $p(x)$ is true." can be symbolically written as $\forall\, x \in M: p(x)$. The symbol $\forall$ (any) is called general quantifier.

Examples

1. For any non-negative number a holds $|a| = a$.
2. $\forall\, x \in \boldsymbol{R}: x^2 \geq 0$
3. All even numbers are divisible by 2.

Proposition ,,There exists an x from the set M such that $p(x)$ is true." can be symbolically written as $\exists\, x \in M: p(x)$. The symbol $\exists$ (exists) is called existential quantifier.

Examples

1. There exists a triangle with one acute angle.
2. $\exists\, x \in \boldsymbol{R}: x^2 - 1 = 0$
3. There exists at least one equilateral triangle.

The negation of proposition ,, $V(x)$ is true for any x." is a proposition ,,There exists such x, for which $V(x)$ is false." Symbolically this can be written as

$$\neg\,(\forall\, x: V(x)) \Leftrightarrow \exists\, x: \neg\, V(x).$$

1.2 Elements of set theory

A set is a well defined collection of distinct objects which are called elements of the set. The elements or members of a set can be numbers, letters of the alphabet, other sets, points, geometric figures or transformations, functions, and so on. Sets are conventionally denoted with capital letters.

If x is an element of the set M, this can be symbolically denoted as $x \in M$, if x is not a member of M, we use the denotation $x \notin M$.

The set can be described in two ways:

- by a list of all elements written in curly brackets,
- by determining the characteristic property satisfied by all set elements.

Examples

1. $M = \{3, 6, 11, 107{,}5\}$
2. $e = \{X \in \boldsymbol{E}^2\colon |F_1X| + |F_2X| = 2a,\ a > 0,\ |F_1F_2| < 2a\}$

A set with a finite number of elements is called a finite set. A set with infinitely many elements is called an infinite set. A set with no elements is called an empty set and it is denoted $\varnothing$. Sets A and B are equal, $A = B$, if and only if they have precisely the same elements. Any element of set A is also an element of set B, and any member of set B is also a member of set A.
Set A is a part (a subset) of set B, $A \subset B$, if any element of set A is also element of set B.

$$A \subset B \Leftrightarrow \forall\, x\colon x \in A \Rightarrow x \in B$$

Alternatively the presented relation can be described using the concept of superset. Set B is superset of set A, $B \supset A$, if any member of set A is simultaneously a member of set B.

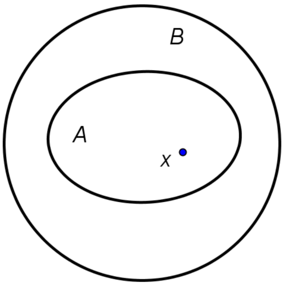


Fig. 1.1. Subset and superset

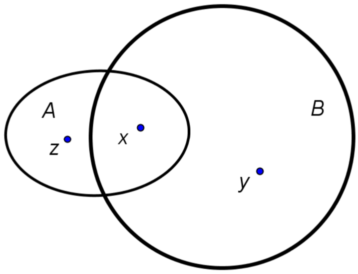


Fig. 1.2. Union of sets

For any two sets A, B holds:

$$A \subset A, \qquad \varnothing \subset A$$
$$A = B \Leftrightarrow [(A \subset B) \wedge (B \subset A)].$$

The union of sets A and B, denoted by $A \cup B$, is the set of all objects which are members of either A or B.

$$A \cup B = \{x\colon (x \in A) \vee (x \in B)\}$$

Some basic properties of union

$$A \cup B = B \cup A, \qquad A \cup (B \cup C) = (A \cup B) \cup C$$

$$A \subseteq (A \cup B)$$

$A \subseteq B$ if and only if $A \cup B = B$

$$A \cup A = A, \qquad A \cup \varnothing = A$$

Intersection of sets A, B denoted by $A \cap B$ is set of all objects which are members of both A and B.

$$A \cap B = \{x: (x \in A) \wedge (x \in B)\}$$

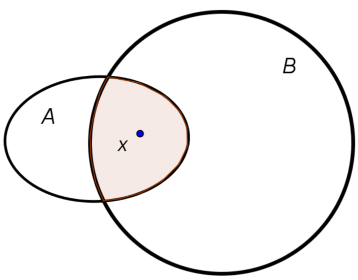


Fig. 1.3. Intersection of sets

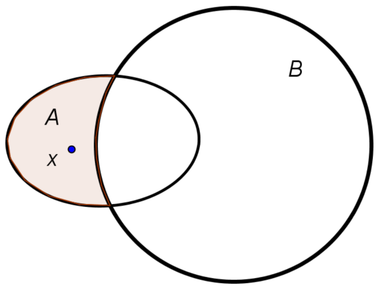


Fig. 1.4. Subtraction of sets

If $A \cap B = \varnothing$, then A and B are said to be disjoint.
Some basic properties of intersections:

$$A \cap B = B \cap A, \qquad A \cap (B \cap C) = (A \cap B) \cap C$$

$$A \cap B \subseteq A$$

$$A \cap A = A, \qquad A \cap \varnothing = \varnothing$$

$$A \subseteq B \text{ if and only if } A \cap B = A$$

Subtraction of sets A, B is a set denoted $A - B$ that consists of all objects which are members of set A but they are not members of set B.

$$A - B = \{x: (x \in A) \wedge (x \notin B)\}$$

If $B \subset A$, then subtraction $A - B$ is said to be the complement of set B in set A. Complement of set A in some basic set C is denoted A^C, whereas $A^C = C - A$.

$$A^C = \{x \in C: x \notin A\}$$

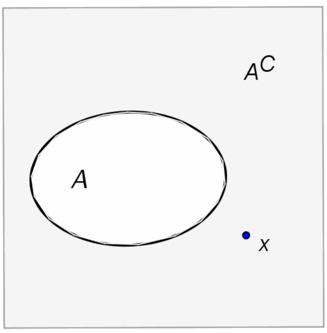


Fig. 1.5. Complement of set

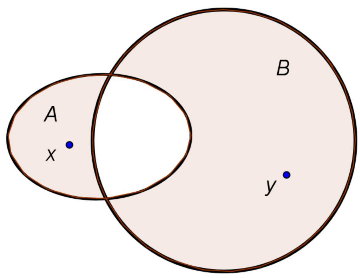


Fig. 1.6. Symmetric difference of sets

Symmetric difference of sets A, B denoted $A \Delta B$ is a set of all objects that are members of set A or set B, but they are not members of their intersection $A \cap B$.

$$A \Delta B = (\mathrm{A} - \mathrm{B}) \cup (\mathrm{B} - \mathrm{A}) = (\mathrm{A} \cup \mathrm{B}) - (\mathrm{A} \cap \mathrm{B})$$
$$A \Delta B = \{x: (x \in A - B) \vee (x \in B - A) = \{x: (x \in A \cup B) \wedge (x \notin A \cap B)\}$$

System of sets is a set consisting of sets as its members.

Examples

1. Set of all subsets of set P is denoted 2^P. If $P = \{0, 1\}$, then $2^P = \{\emptyset, \{0\}, \{1\}, \{0,1\}\}$.
2. System of sets $M = \{p = AB, q = CD\}$ is a set containing line $p = AB$ and line $q = CD$ that are infinite sets of points.

de Morgan rules

$$\neg (p \vee q) \Leftrightarrow \neg p \wedge \neg q \qquad (A \cup B)^C = A^C \cap B^C$$
$$\neg (p \wedge q) \Leftrightarrow \neg p \vee \neg q \qquad (A \cap B)^C = A^C \cup B^C$$

1.3 Relations and mappings

Relation is an association between two or more objects. It can be represented by a formula, a table or a diagram, a graph or a mapping. Symbols $=$, $\neq$ represent relations of equality or inequality between mathematical expressions, order of the number set is represented by relations $\leq$, $<$, $\geq$, $>$, while symbols $\subset$, $\cup$, $\cap$ define relations between sets.
A binary relation can be written as ordered pair (a, b), in which the objects occur in a particular order.

Examples

1. Parity is a relation between a pair of integers: if both integers are odd, or both are even, they have the same parity; if one is odd and the other is even they have different parity.
2. Transitivity is a relation between three elements such that if it holds between the first and the second one and it holds also between the second and the third one, then it must necessarily hold between the first and the third one. For any three real numbers it holds: If $a \leq b$ and $b \leq c$, then $a \leq c$.
3. Reflexiveness, reflexivity is such a relation that holds between an element and itself.
4. Cartesian product of sets A, B denoted $A \times B$ is set of all ordered pairs (a, b) such that a is a member of A and b is a member of B.

$$A \times B = \{[a, \mathrm{b}]: a \in A \wedge b \in B\}$$

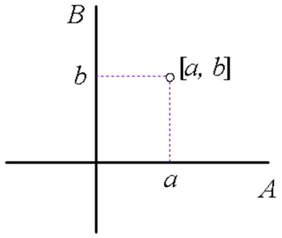


Fig. 1.7. Ordered pair – Cartesian product of sets

Ordered pairs [a, b] and [c, d] are considered equal, if the following holds: $a = c$ and simultaneously $b = d$.

$$[a, b] = [c, d] \Leftrightarrow a = c \wedge b = d$$

Mapping is an association between two sets A and B such that each element of A is associated with a unique element of B.

$$\varphi : A \to B$$
$$\forall a \in A, \exists b \in B : b = \varphi(a)$$

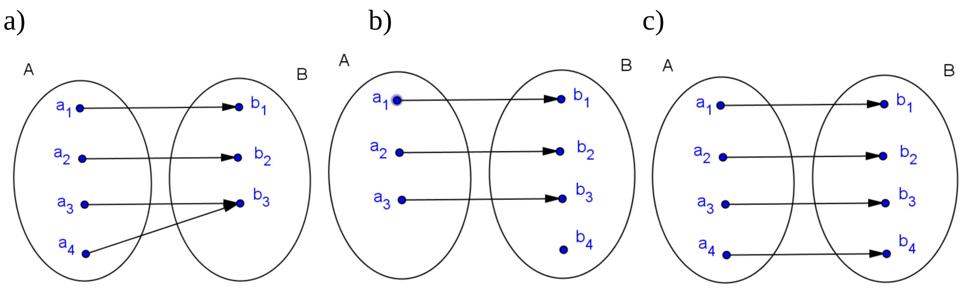


Fig. 1.8. Mappings between sets A and B: a) surjective, b) injective, c) bijective

Mapping φ is said to be surjective, if the image of set A under φ equals to B, $\varphi(A) = B$, therefore

$$\forall b \in B, \exists a \in A : b = \varphi(a) .$$

Mapping φ is said to be injective, if it maps distinct arguments to distinct images, therefore

$$\forall a_1, a_2 \in A : \varphi(a_1) = \varphi(a_2) \Rightarrow a_1 = a_2 .$$

Mapping φ is said to be bijective, if it is both surjective and injective, therefore

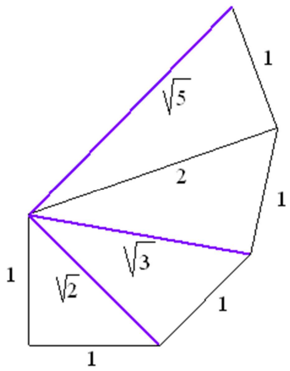
$$\forall a_1, a_2 \in A : a_1 \neq a_2 \Rightarrow \varphi(a_1) \neq \varphi(a_2) .$$

For a bijective mapping ϕ from A to B holds

$$\forall a_1, a_2 \in A : \phi(a_1) = \phi(a_2) \Leftrightarrow a_1 = a_2.$$

1.4 Notes on number sets

The concept of a number is one of the basic concepts in mathematics, whereas several different types of numbers are recognized.
Natural numbers are used to define a number of some objects in a group, or a number of elements of a finite set. Sums and products of natural numbers are again natural numbers, and these two operations are defined in the set of all natural numbers. The difference of two natural numbers is not necessarily a natural number; the operation of subtraction is not always possible in the set of all natural numbers.
Set of all natural numbers is closed with respect to operations of summation and multiplication and it is denoted $\boldsymbol{N} = \{1, 2, 3, ..., n, \dots\}$.
Extension of the set of all natural numbers by zero and negative numbers is set of numbers closed with respect to operation of subtraction, while summing up, multiplying and subtracting two integers we receive again an integer. Ratio of two integers is generally not an integer.
Set of all integers is closed with respect to operations of summation, multiplication and subtraction, and it is denoted $\boldsymbol{Z} = \{..., -2, -1, 0, 1, 2, ...\}$.
By extension of this set by all fractions we can obtain a set of numbers with defined operation of division of an integer by integer different from zero, the set of all rational numbers $\boldsymbol{Q}$. Rational number is any such number that can be represented in the form of a fraction p/q, where p and q have no common divisors and q is a natural number. The set of all rational numbers is $\boldsymbol{Q} = \{p/q: p \in \mathbf{Z}, q \in \boldsymbol{N}\}$.
Set of all natural numbers is a subset of a set of all rational numbers, as any integer can be represented in the form p/q, while $q = 1$.
All four basic operations – summation, subtraction, multiplication, division of numbers are possible in the set of all rational numbers. Nevertheless, an easy equation in the form e.g. $x^2 = 2$ does not have a solution in this set. No rational number x exists such that it satisfies this equation. One of the solutions of this equation we denote as the number $\sqrt{2}$, which is not a rational number, as it cannot be represented in the form of a fraction, and we call it an irrational number. There are infinitely many irrational numbers that appear when finding square roots, for instance $\sqrt{3}$, $\sqrt{5}$, in calculation of logarithms, finding values of goniometric functions or solving algebraic equations.
The discovery of irrational numbers dates back to ancient Greek geometers, who were able to represent irrational numbers geometrically, but they could not manage their mathematical symbolical denotation. They strived to denote all numbers as ratios of lengths of some line segments, i. e. in the form of a fraction. These numbers were called commensurable. Nevertheless, some line segments, whose length could be measured, were not able to be represented in this way, which was not well understood and such line segments were denoted as incommensurable.



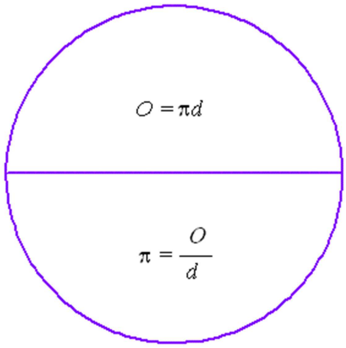


Fig. 1.9. Irrational numbers

Fig. 1.10. Ratio of circumference and radius of circle - π

Transcendent numbers are such numbers that are not solutions of an algebraic equation.
The best known transcendent number is Ludolf's number π, representing the ratio of a circle circumference and its radius. This number cannot be written as a fraction, but for practical calculations it is often approximated by 22/7.
Rational and irrational numbers together form a set of real numbers denoted as ***R***.
The following basic rules are valid for operations with real numbers.

Rules for relation of equality

1. $\forall a \in \boldsymbol{R}$: $a = a$ (reflexiveness)
2. $\forall a, b \in \boldsymbol{R}$: $a = b \Rightarrow b = a$ (symmetry)

Rules for summation

1. $\forall a, b \in \boldsymbol{R}$: $a + b = b + a$ (commutativity)
2. $\forall a, b, c \in \boldsymbol{R}$: $(a + b) + c = a + (b + c)$ (associativity)
3. $\forall a, b \in \boldsymbol{R}\ \exists\, x \in \boldsymbol{R}$: $a + x = b$, difference of numbers $x = b - a$

Rules for multiplication

1. $\forall a, b \in \boldsymbol{R}$: $ab = ba$ (commutativity)
2. $\forall a, b, c \in \boldsymbol{R}$: $(ab)c = a(bc)$ (associativity)
3. $\forall a, b \in \boldsymbol{R}$, $a \neq 0\ \exists\, x \in \boldsymbol{R}$: $ax = b$, ratio of numbers $x = b/a$
4. $\forall a, b, c \in \boldsymbol{R}$: $(a + b)\, c = ac + bc$ (distributivity)

Real numbers are ordered with respect to their value, symbols

$<$ − lesser than

$>$ − greater than

are used for this relation with the following properties.

1. Any given two real numbers a, b satisfy precisely one from the three relations

$a < b,\ a = b,\ a > b$ (trichotomy)

2. $\forall a, b, c \in \boldsymbol{R}: a < b \wedge b < c \Rightarrow a < c$ (transitivity)

3. $\forall a, b, c \in \boldsymbol{R}: a < b \Rightarrow a + c < b + c$ (monotonicity with respect to summation)

4. $\forall a, b \in \boldsymbol{R}: 0 < a \wedge 0 < b \Rightarrow 0 < ab$ (monotonicity of multiplication)

Any real number a corresponds exactly to one non-negative real number usually denoted $|a|$ - absolute value of number a.

$|a| = a$ for $a \geq 0$, $|a| = -a$ for $a < 0$

Absolute value of a real number represents its position on the real axis, a distance from the origin O.

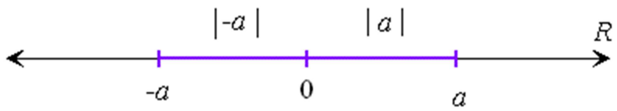


Fig. 1.11. Absolute value of a real number

Rules for calculation with absolute values

1. $|a| = |-a|$
2. $\pm a \leq |a|$
3. $|ab| = |a||b|$
4. $|a/b| = |a|/|b|$, for $b \neq 0$
5. $|a + b| \leq |a| + |b|$, $|a + b| \geq |a| - |b|$
6. $|a| \leq K, K > 0 \Rightarrow -K \leq a \leq K$

Let a, b be arbitrary real numbers, such that $a < b$. Then, the following subsets can be generated of the set of all real numbers

$(a, b) = \{x \in \boldsymbol{R}: a < x < b\}$ – open interval

$\langle a, b) = \{x \in \boldsymbol{R}: a \leq x < b\}$ – left-closed, right-open interval

$(a, b\rangle = \{x \in \boldsymbol{R}: a < x \leq b\}$ – left-open, right-closed interval

$\langle a, b\rangle = \{x \in \boldsymbol{R}: a \leq x \leq b\}$ – closed interval

Numbers a, b are boundary points of the interval, number $b - a$ is the interval length. All above intervals have a finite length or in other words they are bounded. Set of all real numbers can be denoted also as $(-\infty, \infty)$.

Unbounded subsets of the set of real numbers are infinite unbounded intervals.

$(a, \infty) = \{x \in \boldsymbol{R}: x > a\}$ – left open, right unbounded interval

$\langle a, \infty) = \{x \in \boldsymbol{R}: x \geq a\}$ – left closed, right unbounded interval

$(-\infty, a) = \{x \in \boldsymbol{R}: x < a\}$ – left unbounded, right open interval

$(-\infty, a\rangle = \{x \in \boldsymbol{R}: x \leq a\}$ – left unbounded, right closed interval

Neighbourhood of a real number a is an open interval $O_\varepsilon(a) = (a - \varepsilon, a + \varepsilon)$.
Left neighbourhood of a real number a is an open interval $O_{\varepsilon-}(a) = (a - \varepsilon, a)$.
Right neighbourhood of a real number a is an open interval $O_{\varepsilon+}(a) = (a, a + \varepsilon)$.
Any real number can be written as a rational number, using its decimal representation, which can be finite, infinite and periodic (in the case of rational numbers), or infinite (for irrational and transcendent numbers).

$$r = a_m 10^m + a_{m-1} 10^{m-1} + \ldots + a_0 + b_1 10^{-1} + b_2 10^{-2} + \ldots,\ n \in \boldsymbol{N},$$

while numbers a_i, b_i are integers from the set $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

The maximum (minimum) of a non-empty set M is such number $a \in M$, that for any $x \in M$ it holds that $a \geq x$ ($a \leq x$); it is the greatest (least) number from the set M.
Any non-empty finite set has its maximum and minimum. Infinite sets of numbers may not have a maximum or a minimum.

Non-empty set of numbers M is said to be

1. bounded above, if such real number h exists (upper bound of set M), that for all $x \in M$ is $x \leq h$
2. bounded below, if such real number d exists (lower bound of set M), that for all $x \in M$ is $x \geq d$
3. bounded if it is both bounded above and below, otherwise it is said to be unbounded.

The number set M is bounded if and only if such number $K > 0$ exists that for all $x \in M$ holds $|x| \leq K$.
Any bounded set has an infinite quantity of upper and lower bounds.

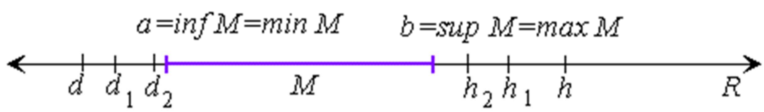


Fig. 1.12. Bounded set

Let M be a non-empty number set.

1. The least upper bound of set M (if it exists) is called supremum of set M and is denoted sup M.
2. The greatest bottom bound of set M (if it exists) is called infimum of set M and it is denoted inf M.

Properties of supremum S of a non-empty set M

1. for any $x \in M$ it holds that $x \leqq S$
2. at least one $x \in M$ exists in arbitrary left neighbourhood of S - $(S - \varepsilon, S\rangle$.

Properties of infimum s of a non-empty set M

1. for any $x \in M$ it holds that $x \geqq s$
2. at least one $x \in M$ exists in arbitrary right neighbourhood of S - $\langle s, s + \varepsilon)$.

The minimum (maximum) of set M (if it exists), equals to the infimum (supremum) of M.
Many mathematical problems have no solution in the set of real numbers. Simple quadratic equation $x^2 + 1 = 0$ has no solution in real numbers, as its discriminant is negative, $D = -4$. Set of all real numbers must be therefore extended to such set of numbers, in which all quadratic equations have a solution.
Complex numbers are all the numbers in the form $z = a + b\mathrm{i}$, where $a, b \in \boldsymbol{R}$ and i is an imaginary unit, a number, for which equality $\mathrm{i}^2 = -1$ is true.
Powers of imaginary unit i are

$$\mathrm{i}^2 = -1,\ \mathrm{i}^3 = -\mathrm{i},\ \mathrm{i}^4 = 1,\ \mathrm{i}^5 = \mathrm{i},\ \mathrm{i}^6 = -1,\ \ldots$$

Number a is the real part of the complex number denoted also Re(z), number b is the imaginary part of the complex number also denoted Im(z).
Any real number is a member of a set of all complex numbers, while its imaginary part is $b = 0$. Complex numbers with real part $a = 0$ are said to be pure imaginary numbers. The set of all complex numbers is usually denoted as $\boldsymbol{C}$.
Any complex number can have its complex conjugate number attached, $\bar{z} = a - b\mathrm{i}$.

The rules for summation and subtraction, multiplication and division of complex numbers are

1. $(a + b\mathrm{i}) \pm (c + d\mathrm{i}) = (a \pm c) + (b \pm d)\mathrm{i}$
2. $(a + b\mathrm{i})(c + d\mathrm{i}) = (ac - bd) + (ad + bc)\mathrm{i}$
3. $\dfrac{a + b\mathrm{i}}{c + d\mathrm{i}} = \dfrac{ac + bd}{a^2 + b^2} + \dfrac{bc - ad}{a^2 + b^2}\mathrm{i}$

For complex conjugate numbers $z = a + bi, \bar{z} = a - bi$ it holds that

$$z + \bar{z} = 2a,\ z\bar{z} = a^2 + b^2.$$

For arbitrary complex numbers z_1, z_2 it holds that

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2},\ \overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}.$$

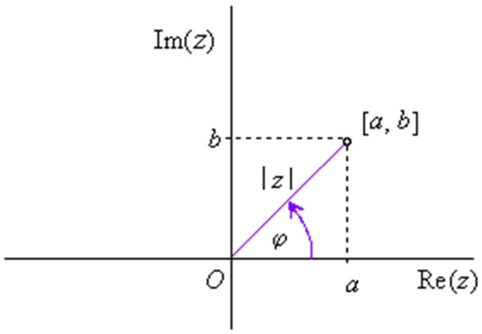


Fig. 1.13. Geometric interpretation of complex number

The geometric interpretation of a set of complex numbers is a plane. Any complex number $z = a + b$i can be represented as a point in the plane, whose Cartesian coordinates in the determined orthogonal coordinate system with origin $O = [0, 0]$ are given as ordered pair $[a, b]$ also called the algebraic form of a complex number. Coordinate axis x is a real axis; and coordinate axis y is an imaginary axis.

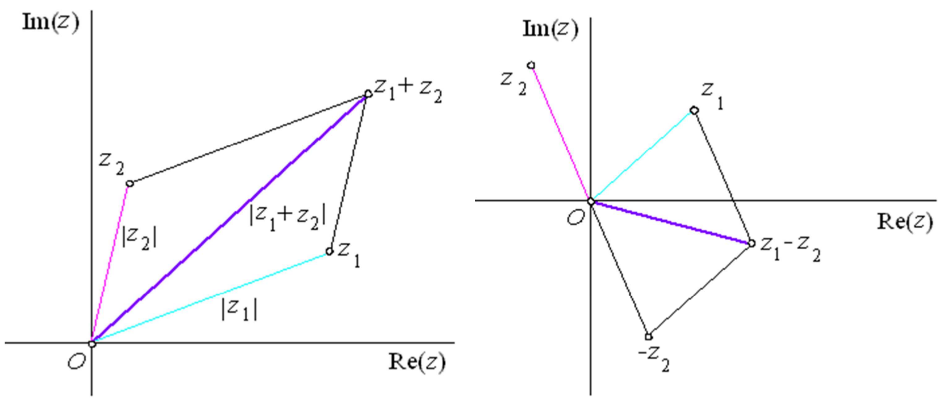


Fig. 1.14. Geometric interpretation of sum and difference of two complex numbers

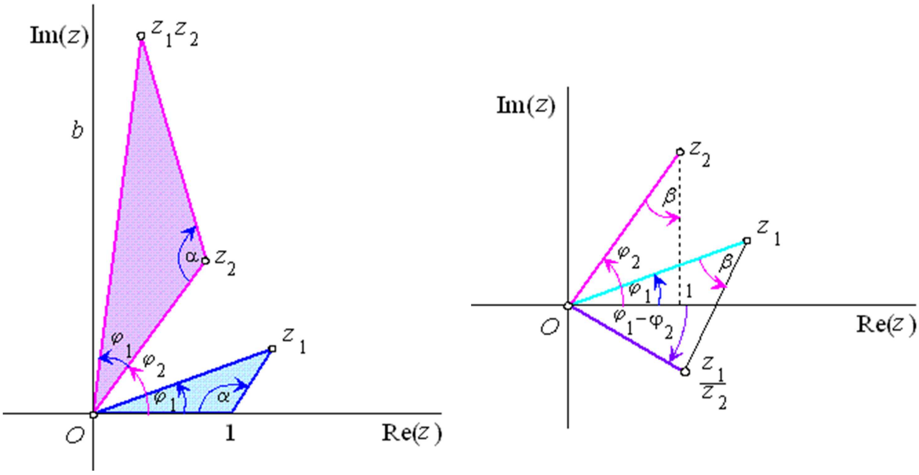


Fig. 1.15. Geometric interpretation of product and ratio of two complex numbers

Absolute value of complex number $z = a + b\mathrm{i}$ is a non-zero real number

$$|z| = |a + b\mathrm{i}| = \sqrt{a^2 + b^2}.$$

Geometric interpretation of the absolute value of a complex number $z = a + b\mathrm{i}$ is the distance of point $z = [a, b]$ from the origin O.
Real number φ, for which

$$\cos\varphi = \frac{a}{|z|},\ \sin\varphi = \frac{b}{|z|}$$

is the argument (amplitude) of the complex number $z = a + b\mathrm{i}$, denoted also as arg z, and it is the size of an angle formed by the line segment with end points in origin O and point $[a, b]$, and axis x.

$$\arg z = \varphi + 2k\pi,\ k = 0, \pm1, \pm2, \ldots$$

Only one of the arguments of complex number $z \neq 0$ satisfies condition $0 \leq \varphi < 2\pi$ and this is called the principle argument (amplitude) of complex number z and is denoted Arg z.
If $z = 0$, than $|z| = 0$, $\varphi = 0$.

Complex number $z = a + b\mathrm{i} \neq 0$ can be written in the form

$$z = |z|(\cos\varphi + \mathrm{i}\sin\varphi)$$

called the goniometric form of the complex number.

Complex number, whose absolute value equals one, is denoted as complex unit

$$z_j = (\cos\varphi + \mathrm{i}\sin\varphi).$$

Let two complex numbers be given in goniometric form

$$z_1 = |z_1|(\cos\varphi_1 + \mathrm{i}\sin\varphi_1),\ z_2 = |z_2|(\cos\varphi_2 + \mathrm{i}\sin\varphi_2).$$

Then their product and ratio are

$$z_1 z_2 = |z_1||z_2|(\cos(\varphi_1 + \varphi_2) + \mathrm{i}\sin(\varphi_1 + \varphi_2))$$

$$\frac{z_1}{z_2} = \frac{|z_1|}{|z_2|}(\cos(\varphi_1 - \varphi_2) + \mathrm{i}\sin(\varphi_1 - \varphi_2)).$$

The Moivre formula for powers of complex unit is

$$(\cos\varphi + \mathrm{i}\sin\varphi)^n = \cos n\varphi + i\sin n\varphi.$$

For a complex number $z = |z|(\cos\varphi + \mathrm{i}\sin\varphi)$, $n \in \boldsymbol{N}$ it holds that

$$z^n = |z|^n(\cos\varphi + \mathrm{i}\sin\varphi)^n = |z|^n(\cos n\varphi + i\sin n\varphi).$$

Exponential form of complex number can be written as

$$z = |z|(\cos\varphi + \mathrm{i}\sin\varphi) = |z|e^{\mathrm{i}\varphi},$$

where number e = 2,71828 is a transcendent number called Euler number.

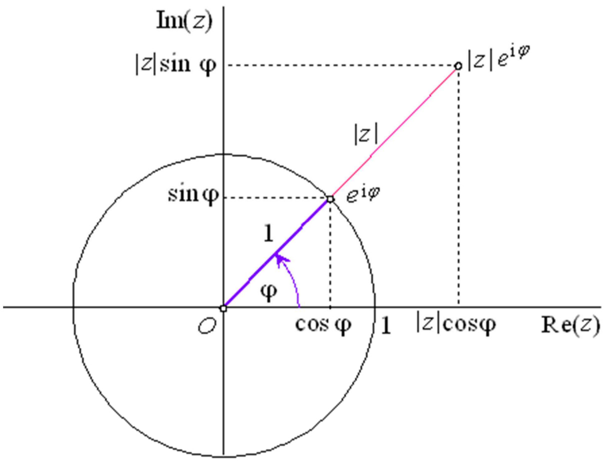


Fig. 1.16. Exponential form of complex number

The Euler formula is the following relation true for complex units

$$e^{\mathrm{i}\varphi} = \cos\varphi + \mathrm{i}\sin\varphi.$$

One of the most beautiful mathematical equations, in which all special mathematical numbers appear, is the equation

$$e^{\mathrm{i}\pi} + 1 = 0.$$

2 Chapters from linear algebra

2.1 Matrices and determinants

Let m and n be natural numbers. A matrix of $m \times n$ type is a table – rectangular array consisting of elements a_{ij}, $i = 1, 2, ..., m$, $j = 1, 2, ..., n$ ordered into m rows and n columns. Matrix is usually written in the form

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ . & . & . & \dots & . \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}$$

and denoted $\mathbf{A} = (a_{ij}) = \mathbf{A}_{m \times n}$, for $i = 1, 2,..., m$, $j = 1, 2, ..., n$.
The elements a_{ij}, which are most often specific numbers (real, complex), are called the entries of the matrix. A matrix of the type $m \times n$, where $m \neq n$, is said to be rectangular matrix, matrix of the type $n \times n$ is a square matrix of degree n. Matrix of the type $1 \times n$ is a row vector, while matrix of the type $m \times 1$ is a column vector. A vector $(a_{11}, a_{22}, ..., a_{nn})$ in a square matrix of degree n is called major (principal, leading) diagonal, a vector $(a_{1n}, a_{2\,n-1}, ..., a_{n1})$ is called minor diagonal.
Let $\mathbf{A}$ be a matrix of the type $m \times n$, then the matrix $\mathbf{A}^{\mathrm{T}}$ of the $n \times m$ type generated from matrix $\mathbf{A}$ exchanging rows and columns (in the given order), is said to be a transpose of the matrix $\mathbf{A}$.
A square matrix with all zero elements but entries on the major diagonal that are equal to number 1 is a unit (identity) matrix denoted $\mathbf{E}$.

Examples

1. Transpose to the matrix $\mathbf{A} = \begin{pmatrix} 1 & 5 & 8 & 0 \\ 7 & 2 & 6 & 9 \\ 9 & 6 & 3 & 7 \\ 0 & 8 & 5 & 4 \end{pmatrix}$

   is the matrix

$$\mathbf{A}^{\mathrm{T}} = \begin{pmatrix} 1 & 7 & 9 & 0 \\ 5 & 2 & 6 & 8 \\ 8 & 6 & 3 & 5 \\ 0 & 9 & 7 & 4 \end{pmatrix}$$

2. Unit matrix of order three is the matrix $\mathbf{E} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Operations on matrices

Let $\mathbf{A} = (a_{ij})$, $\mathbf{B} = (b_{ij})$, $\mathbf{C} = (c_{ij})$ and $\mathbf{D} = (d_{ij})$ be matrices of the same type $m \times n$. Then the following relations hold

1. Matrix equality.
   Matrices $\mathbf{A}$ and $\mathbf{B}$ are equal, $\mathbf{A} = \mathbf{B}$, if $a_{ij} = b_{ij}$ for all $i = 1, 2, ..., m$, $j = 1, 2, ..., n$.
2. Matrix addition.
   Matrix $\mathbf{C} = (c_{ij})$, where $c_{ij} = a_{ij} + b_{ij}$ for all $i = 1, 2, ..., m$, $j = 1, 2, ..., n$ is called the sum of matrices $\mathbf{A}$ and $\mathbf{B}$, written as $\mathbf{C} = \mathbf{A} + \mathbf{B}$.
3. Multiplying matrices by numbers.
   Let k be a number. Matrix $\mathbf{D} = (d_{ij})$, where $d_{ij} = ka_{ij}$ for all $i = 1, 2, ..., m$, $j = 1, 2, ..., n$ is called a multiple of matrix $\mathbf{A}$ by number k, written as $\mathbf{D} = k.\mathbf{A}$.
4. Multiplication of matrices.
   Let $\mathbf{A} = (a_{ij})$ be a matrix of type $m \times n$ and $\mathbf{B} = (b_{ij})$ a matrix of type $n \times p$. Matrix $\mathbf{C} = (c_{ij})$ of the $m \times p$ type, such that $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + ... + a_{in}b_{mj}$ for $i = 1, 2, ..., m$, $j = 1, 2, ..., p$ is called the product of matrices $\mathbf{A}$ and $\mathbf{B}$, written as $\mathbf{C} = \mathbf{A}.\mathbf{B}$.

Entries c_{ij} of matrix $\mathbf{C}$ are obtained as summation of all products of corresponding entries in the i-th row of matrix $\mathbf{A}$ and the j-th column of matrix $\mathbf{B}$. Two matrices can be multiplied only when the number of columns in the first matrix equals to the number of rows in the second one.
Multiplication of matrices, in general, is not commutative, $\mathbf{A}.\mathbf{B} \neq \mathbf{B}.\mathbf{A}$.
Let $\mathbf{A}$ be an arbitrary square matrix. For a unit matrix of the same type as $\mathbf{A}$ it holds that

$$\mathbf{A}.\mathbf{E} = \mathbf{E}.\mathbf{A} = \mathbf{A}.$$

Square matrix $\mathbf{A}^{-1}$, for which $\mathbf{A}^{-1}.\mathbf{A} = \mathbf{E}$, $\mathbf{A}.\mathbf{A}^{-1} = \mathbf{E}$ is called the inverse matrix to matrix $\mathbf{A}$.
For the matrix transpose $\mathbf{A}^{\mathrm{T}}$ it holds that $(\mathbf{A}^{\mathrm{T}})^{-1} = (\mathbf{A}^{-1})^{\mathrm{T}}$.

Examples

1. Matrix $\mathbf{B} = \begin{pmatrix} 2 & 5 & 8 & 0 \\ 7 & 3 & 6 & 9 \\ 9 & 6 & 4 & 7 \\ 0 & 8 & 5 & 5 \end{pmatrix}$ is the sum and matrix $\mathbf{C} = \begin{pmatrix} 0 & 5 & 8 & 0 \\ 7 & 1 & 6 & 9 \\ 9 & 6 & 2 & 7 \\ 0 & 8 & 5 & 3 \end{pmatrix}$ is the difference of matrices $\mathbf{A}$ and $\mathbf{E}$ from previous examples.

2. The product of matrices **A** and $\mathbf{A}^{\mathrm{T}}$ from previous examples

$$\mathbf{P} = \mathbf{A}.\mathbf{A}^{\mathrm{T}} = \begin{pmatrix} 1 & 5 & 8 & 0 \\ 7 & 2 & 6 & 9 \\ 9 & 6 & 3 & 7 \\ 0 & 8 & 5 & 4 \end{pmatrix}$$ is different from their product

$$\mathbf{R} = \begin{pmatrix} 1 & 5 & 8 & 0 \\ 7 & 2 & 6 & 9 \\ 9 & 6 & 3 & 7 \\ 0 & 8 & 5 & 4 \end{pmatrix}$$ in different order, $\mathbf{R} = \mathbf{A}^{\mathrm{T}}.\mathbf{A}$.

Determinant of a matrix

A determinant of a matrix **A** is a number denoted det**A** = |**A**|, which can be calculated as follows:

1. For matrix **A** of degree $n = 1$, i.e. $\mathbf{A} = (a_{11})$, it holds that $\det\mathbf{A} = |\mathbf{A}| = a_{11}$.

2. For all $n \geqq 2$ the determinant of matrix **A** equals

$$\det\mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ . & . & . & \dots & . \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix} =$$
$$= a_{11}|\mathbf{A}_{11}| - a_{12}|\mathbf{A}_{12}| + a_{13}|\mathbf{A}_{13}| - \dots + (-1)^{1+n}|\mathbf{A}_{1n}|$$

where $|\mathbf{A}_{1j}|$, $j = 1, 2, \dots, n$ are the subdeterminants of matrices obtained from the matrix **A** excluding (deleting!) its first row and j-th column.

The following calculation holds for the determinant of a matrix of degree three

$$D = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}|\mathbf{A}_{11}| - a_{12}|\mathbf{A}_{12}| + a_{13}|\mathbf{A}_{13}| =$$
$$= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} =$$
$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) =$$
$$= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$$

The determinant of a matrix of degree three can also be evaluated using the following easy rule.

Sarus rule

Add the first and second rows of matrix **A** under its third row and sum up the products of three entries of this new determinant in the direction of major and minor diagonal, while multiplying the products in the direction of minor diagonal by –1. The result is the value of the matrix **A** determinant.

$$\begin{array}{c} + \\ + \\ + \\ - \\ - \end{array}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{vmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$$

The determinant of a matrix with at least one zero row or column equals zero.

Inverse matrix

There exists an inverse matrix $\mathbf{A}^{-1}$ to any square matrix **A**, with a non-zero determinant det**A** $\neq 0$, which can be calculated as follows

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\begin{pmatrix} D_{11} & D_{12} & D_{13} & \dots & D_{1n} \\ D_{21} & D_{22} & D_{23} & \dots & D_{2n} \\ . & . & . & \dots & . \\ D_{n1} & D_{n2} & D_{n3} & \dots & D_{nn} \end{pmatrix}^T \text{, where}$$

$$D_{ij} = (-1)^{i+j}\left|\mathbf{A}_{ij}\right|, \; i, j = 1, 2, \dots, n$$

Element D_{ij} is called the algebraic complement of matrix **A** entry a_{ij}, which is derived from the determinant det**A** excluding its i-th row and j-th column.

Example

1. Inverse matrix to the matrix $\mathbf{A} = \begin{pmatrix} 2 & 4 & 0 \\ 3 & 7 & -2 \\ 3 & 0 & 1 \end{pmatrix}$ can be determined as follows

$$|\mathbf{A}| = \begin{vmatrix} 2 & 4 & 0 \\ 0 & 7 & -2 \\ 3 & 0 & 1 \end{vmatrix} = -10,$$

$$D_{11} = (-1)^{1+1}\begin{vmatrix} 7 & -2 \\ 0 & 1 \end{vmatrix} = 7,\ D_{12} = (-1)^{1+2}\begin{vmatrix} 0 & -2 \\ 3 & 1 \end{vmatrix} = -6,\ D_{13} = (-1)^{1+3}\begin{vmatrix} 0 & 7 \\ 3 & 0 \end{vmatrix} = -21,$$

$$D_{21} = (-1)^{2+1}\begin{vmatrix} 4 & 0 \\ 0 & 1 \end{vmatrix} = -4,\ D_{22} = (-1)^{2+2}\begin{vmatrix} 2 & 0 \\ 3 & 1 \end{vmatrix} = 2,\ D_{23} = (-1)^{2+3}\begin{vmatrix} 2 & 4 \\ 3 & 0 \end{vmatrix} = 12,$$

$$D_{31} = (-1)^{3+1}\begin{vmatrix} 4 & 0 \\ 7 & -2 \end{vmatrix} = -8,\ D_{32} = (-1)^{3+2}\begin{vmatrix} 2 & 0 \\ 0 & -2 \end{vmatrix} = 4,\ D_{33} = (-1)^{3+3}\begin{vmatrix} 2 & 4 \\ 0 & 7 \end{vmatrix} = 14,$$

$$\mathbf{A}^{-1} = -\frac{1}{10}\begin{pmatrix} 7 & -6 & -21 \\ -4 & 2 & 12 \\ -8 & 4 & 14 \end{pmatrix}^{\mathrm{T}} = \frac{1}{10}\begin{pmatrix} -7 & 4 & 8 \\ 6 & -2 & -4 \\ 21 & -12 & -14 \end{pmatrix}$$

Rank of matrix

A rank of a matrix is a natural number determining the number of non-zero, linearly independent rows (columns) of this matrix. A matrix with no zero rows (columns), i.e. none of its rows (columns) is a multiple of other rows (columns), with a value equal to the number of its rows (columns) is called a regular matrix. The determinant of a regular matrix is non-zero. The rank of matrix **A** is denoted $h(\mathbf{A})$.

Matrix $\mathbf{A} = (a_{ij})$ is said to be upper triangular (lower triangular), if for all $i > j$ ($i < j$) it holds that $a_{ij} = 0$. All entries of a square upper triangular (lower triangular) matrix below (under) major (minor) diagonal are equal to zero.

Triangular matrix of type 3×5 has the form

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ 0 & a_{22} & a_{23} & a_{24} & a_{25} \\ 0 & 0 & a_{33} & a_{34} & a_{35} \end{pmatrix}.$$

Equivalent manipulations with matrices are such operations, which do not change the rank of the matrix, but are just the following:

1. interchanging of any two rows (columns)
2. multiplying any row (column) by an arbitrary number $c \neq 0$
3. multiplying any row (column) by a number and adding the result to any other row (column)
4. deleting zero rows (columns) and those, which are multiples of another row (column).

Example

1. Inverse matrix to the matrix **A** can also be calculated by means of equivalent manipulations with the matrix composed of a given matrix **A** and a unit matrix **E** of the same rank. Manipulations with rows (columns) must be carried out so that the matrix **A** will be transformed to the unit matrix **E**, while unit matrix **E** will be consequently transformed to the inverse matrix $\mathbf{A}^{-1}$. We can calculate the inverse matrix for the matrix from the previous example in the following way

$$\begin{pmatrix} 2 & 4 & 0 & 1 & 0 & 0 \\ 0 & 7 & -2 & 0 & 1 & 0 \\ 3 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \cong \begin{pmatrix} 2 & 4 & 0 & 1 & 0 & 0 \\ 0 & 7 & -2 & 0 & 1 & 0 \\ 0 & -12 & 2 & -3 & 0 & 2 \end{pmatrix} \cong$$

$$\cong \begin{pmatrix} 2 & 4 & 0 & 1 & 0 & 0 \\ 0 & -5 & 0 & -3 & 1 & 2 \\ 0 & -12 & 2 & -3 & 0 & 2 \end{pmatrix} \cong \begin{pmatrix} 2 & 4 & 0 & 1 & 0 & 0 \\ 0 & -5 & 0 & -3 & 1 & 2 \\ 0 & 0 & 10 & 21 & -12 & -14 \end{pmatrix} \cong$$

$$\cong \begin{pmatrix} 10 & 0 & 0 & -7 & 4 & 8 \\ 0 & 10 & 0 & 6 & -2 & -4 \\ 0 & 0 & 10 & 21 & -12 & -14 \end{pmatrix} \cong$$

$$\cong \begin{pmatrix} 1 & 0 & 0 & -7/10 & 4/10 & 8/10 \\ 0 & 1 & 0 & 6/10 & -2/10 & -4/10 \\ 0 & 0 & 1 & 21/10 & -12/10 & -14/10 \end{pmatrix}$$

2.2 Linear systems of equations

A system of m equations with n unknowns

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n = b_1 \\
&a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n = b_2 \\
&a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots + a_{3n}x_n = b_3 \\
&\ldots \\
&a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \ldots + a_{mn}x_n = b_m
\end{aligned}$$

where coefficients a_{ij} and absolute coefficients b_i are real numbers for i = 1, 2,... m, j = 1, 2, ... n can be written in the matrix form as

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ . & . & . & \dots & . \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ . \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ . \\ b_m \end{pmatrix}$$

$$\mathbf{A}_{m\times n}.\mathbf{X}_{n\times 1} = \mathbf{B}_{m\times 1}$$

Matrix **A** = (a_{ij}), i = 1, 2, ... m, j = 1, 2, ..., n is called the matrix of the system, matrix **C** in the form

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} & b_2 \\ . & . & . & \dots & . & . \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} & b_m \end{pmatrix}$$

is called the extended matrix of the system.

The system of equations is said to be

a) homogeneous (without the right side), if $b_i = 0$ for all i = 1, 2, ..., m

b) non-homogeneous (with the right side), if $b_i \neq 0$ for at least one i.

The solution of the system of m equations with n unknowns is any such ordered n-tuple of real numbers (r_1, r_2, r_3, ..., r_n), i. e. such n-dimensional column vector, that satisfies the system of equations. Insertion of a vector (r_1, r_2, r_3, ..., r_n) to the equations from the given system instead of a vector (x_1, x_2, x_3, ..., x_n) leads to the true statements.
Two systems of linear equations with the same number of unknowns are said to be equivalent, if any solution of one of them is also a solution of the other one.

To solve linear systems we use the following equivalent transformations:

1. interchanging of any two equations

2. multiplying any equation by an arbitrary number $c \neq 0$

3. multiplying any equation by a number and adding the result to any other equation

4. deleting the equation which is the multiple of another equation in the system, or 0.

A system, obtained from the original system by applying a finite number of the above elementary transformations, is equivalent to the primary system. Equivalent transformations of a system result in a system equivalent to the original one.
Matrices of two equivalent linear systems of equations can be transformed by equivalent manipulations with rows (columns) from one to the other. Therefore, the solution of a linear system of equations can be obtained by performing equivalent transformations with the extended matrix of the system.

This method is an universal method for solving systems of any number of equations with any number of unknowns, called the Gauss‘ elimination method. It consists of the reduction of a given system to a triangular system by means of a finite number of elementary transformations applied on the extended matrix of the system.
Then, after the triangulation process, we can consider the following possibilities.

A) The new equivalent system has the same number of non-zero equations (extended matrix rows) as is the number of unknowns – the system has a unique solution.

B) There appears an equation in the form $0 = c$, $c \neq 0$ – the system has no solution.

C) The resulting triangular system consists of less non-zero equations (extended matrix rows) than is the number of unknowns and it does not contain any equation in the form $0 = c$, $c \neq 0$ – the system has infinitely many solutions.

Since homogeneous systems (or systems equivalent to them) never contain equations in the form $0 = c$, $c \neq 0$, it follows that they always have a solution, at least a zero solution called trivial solution (type A), or infinitely many solutions (type C).

Examples

1. Solution of a linear system with 4 unknowns

$$\begin{aligned}
&x_1 + 3x_2 - x_4 = 4\\
&2x_1 - 2x_2 + x_3 = 2\\
&x_1 + 2x_3 - 2x_4 = 0\\
&x_1 + x_2 + x_3 + x_4 = 4
\end{aligned}$$

can be found by finding an equivalent triangular system by means of equivalent transformations of the extended matrix of the original system

$$\begin{pmatrix} 1 & 3 & 0 & -1 & 4\\ 2 & -2 & 1 & 0 & 2\\ 1 & 0 & 2 & -2 & 0\\ 1 & 1 & 1 & 1 & 4 \end{pmatrix} \cong \begin{pmatrix} 1 & 1 & 1 & 1 & 4\\ 0 & 2 & -1 & -2 & 0\\ 0 & 4 & 1 & 2 & 6\\ 0 & 1 & -1 & 3 & 4 \end{pmatrix} \cong \begin{pmatrix} 1 & 1 & 1 & 1 & 4\\ 0 & 1 & -1 & 3 & 4\\ 0 & 1 & 0 & -5 & -4\\ 0 & 0 & 5 & -10 & -10 \end{pmatrix} \cong$$

$$\cong \begin{pmatrix} 1 & 1 & 1 & 1 & 4\\ 0 & 1 & -1 & 3 & 4\\ 0 & 0 & 1 & -8 & -8\\ 0 & 0 & 1 & -2 & -2 \end{pmatrix} \cong \begin{pmatrix} 1 & 1 & 1 & 1 & 4\\ 0 & 1 & -1 & 3 & 4\\ 0 & 0 & 1 & -8 & -8\\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

The resulting triangular system has a unique solution (2, 1, 0, 1) that is also the solution of the original linear system.

2. Homogeneous system of linear equations

$$x_1 + x_2 + x_3 = 0$$
$$2x_1 - x_2 + 2x_3 = 0$$
$$x_1 + 2x_2 - x_3 = 0$$

can be transformed to the triangular system $x_1 = 0, x_2 = 0, x_3 = 0$ with a unique trivial solution (0, 0, 0).

3. Linear system with 3 unknowns

$$2x_1 - x_2 + x_3 = 2$$
$$x_1 + 2x_2 - 2x_3 = 1$$
$$3x_1 + x_2 - x_3 = 3$$

can be transformed

$$\begin{pmatrix} 2 & -1 & 1 & 2 \\ 1 & 2 & -2 & 1 \\ 3 & 1 & -1 & 3 \end{pmatrix} \cong \begin{pmatrix} 1 & 2 & -2 & 1 \\ 0 & -5 & 5 & 0 \\ 0 & -5 & 5 & 0 \end{pmatrix} \cong \begin{pmatrix} 1 & 2 & -2 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

to the triangular system consisting of 2 equations

$$x_1 + 2x_2 - 2x_3 = 1$$
$$x_2 - x_3 = 0$$

while it contains no equation in the form 0 = c, $c \neq 0$, therefore both systems have infinitely many solutions represented parametrically as (1, p, p), for an arbitrary number p.

Geometric interpretation of a system of linear equations

The solution of a system of two linear equations with two unknowns

$$\begin{matrix} a_{11}x + a_{12}y = b_1 \\ a_{21}x + a_{22}y = b_2 \end{matrix} \Leftrightarrow \mathbf{A}.\mathbf{X} = \mathbf{B}, \mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \mathbf{X} = \begin{pmatrix} x \\ y \end{pmatrix}, \mathbf{B} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

can be geometrically interpreted as a determination of a set of all common points of two lines in a plane. Each line is represented by one of the two equations that satisfy the Cartesian coordinates of its points.

Two lines in a plane can have only one of the three possible superpositions:

a) intersecting lines, the system of equations has a unique solution

b) coinciding lines, the system of equations has infinitely many solutions

c) parallel lines, the system of equations has no solution.

If $a_{11}a_{22} - a_{12}a_{21} \neq 0$, then the system has a unique solution

$$x = \frac{b_1 a_{22} - b_2 a_{12}}{a_{11}a_{22} - a_{12}a_{21}},\ y = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}}.$$

Denoting

$$D = |\mathbf{A}| = a_{11}a_{22} - a_{12}a_{21}, D_1 = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} = b_1 a_{22} - b_2 a_{12}, D_2 = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix} = a_{11}b_2 - a_{21}b_1$$

then, if $D \neq 0$, $x = \dfrac{D_1}{D}, y = \dfrac{D_2}{D}$.

Analogously we can solve a system of 3 linear equations in 3 unknowns:

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= b_1 \\ a_{21}x + a_{22}y + a_{23}z &= b_2 \\ a_{31}x + a_{32}y + a_{33}z &= b_3 \end{aligned}$$

Any equation with three unknowns can be represented as an equation of a plane in the three-dimensional space that is satisfied by Cartesian coordinates of its points. Solutions of the system of three linear equations represent the coordinates of all common points of the three determined planes. According to a possible superposition of three planes in the space we obtain:

a) a unique solution, all planes meet in one common point

b) no solution, at least two of the planes are parallel

c) infinitely many solutions, the planes have one common line.

Generally, the system of *m* linear equations with *n* unknowns can have:

1. a unique solution
2. infinitely many solutions
3. no solution.

Cramer rule

A linear system of *n* equations $\mathbf{A}_{n\times n}.\mathbf{X}_{n\times 1} = \mathbf{B}_{n\times 1}$ with *n* unknowns has a unique solution if the determinant of the matrix of the system is nonzero, det**A** = $D \neq 0$, and the solution is

$$(x_1, x_2, ..., x_n) = \left(\frac{D_1}{D}, \frac{D_2}{D}, ..., \frac{D_n}{D}\right),$$

where D_1, D_2, ..., D_n are the determinants of matrices derived from the matrix **A** by exchanging entries in the *i*-th column with absolute coefficients b_i, for $i = 1, 2, ..., n$.

The linear system can also be solved using an inverse matrix. Multiplying the equation **A**.**X** = **C** by inverse matrix $\mathbf{A}^{-1}$ from left, we obtain

$$\mathbf{A}^{-1}.\mathbf{A}.\mathbf{X} = \mathbf{A}^{-1}.\mathbf{C} \Rightarrow \mathbf{E}.X = \mathbf{A}^{-1}.\mathbf{C} \Rightarrow \mathbf{X} = \mathbf{A}^{-1}.\mathbf{C}$$

Examples

1. Solution of the linear system of equations

$$x - 3y + z = -8$$
$$2x + y - z = 8$$
$$x - y - z = 0$$

can be obtained using the Cramer rule, because the determinant of the matrix of this system is nonzero

$$D = \begin{vmatrix} 1 & -3 & 1 \\ 2 & 1 & -1 \\ 1 & -1 & -1 \end{vmatrix} = -8 \neq 0 .$$

For determinants D_1, D_2, D_3 it holds that

$$D_1 = \begin{vmatrix} -8 & -3 & 1 \\ 8 & 1 & -1 \\ 0 & -1 & -1 \end{vmatrix} = -16,\ D_2 = \begin{vmatrix} 1 & -8 & 1 \\ 2 & 8 & -1 \\ 1 & 0 & -1 \end{vmatrix} = -24,\ D_3 = \begin{vmatrix} 1 & -3 & -8 \\ 2 & 1 & 8 \\ 1 & -1 & 0 \end{vmatrix} = 8 ,$$

therefore the unique solution for the system is

$$(x_1, x_2, x_3) = \left(\frac{D_1}{D}, \frac{D_2}{D}, \frac{D_3}{D}\right) = \left(\frac{-16}{-8}, \frac{-24}{-8}, \frac{8}{-8}\right) = (2,3,-1) .$$

2. The above linear system can also be solved using the inverse matrix to the matrix **A**, which is matrix $\mathbf{A}^{-1}$

$$\mathbf{A}^{-1} = -\frac{1}{8}\begin{pmatrix} -2 & -4 & 2 \\ 1 & -2 & 3 \\ -3 & -2 & 7 \end{pmatrix},$$

$$\mathbf{X} = \mathbf{A}^{-1}.\mathbf{C} = -\frac{1}{8}\begin{pmatrix} -2 & -4 & 2 \\ 1 & -2 & 3 \\ -3 & -2 & 7 \end{pmatrix}.\begin{pmatrix} -8 \\ 8 \\ 0 \end{pmatrix} = -\frac{1}{8}\begin{pmatrix} -16 \\ -24 \\ 8 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$$

3 Differential calculus of functions with one real variable

3.1 Definition of function

The concept of function is a basic concept in mathematics used for the description of dependence of various quantities. In the majority of simple dependencies certain value of the investigated variable depends on the value of one or more independent variables. Developments in the understanding of our real world are a direct consequence of the discoveries made in the understanding of dependencies that describe interrelations of phenomena and processes in the nature and in the society.

Examples

1. The tone height of a guitar string depends directly on the string tension.
2. Newton's law of universal gravitation states that every point mass in the universe attracts every other point mass with a force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between them.
3. The pressure exerted on a container's sides by an ideal gas is proportional to its temperature.
4. Market price of a product is primarily determined by the interaction of supply and demand.
5. Distance covered by a moving object within a time interval is proportional to the object velocity.
6. Circumference of a circle is directly related to its radius.
7. Area of a parallelogram is determined by the lengths of its sides.
8. Volume of a solid changes according to its dimensions.
9. Age of a human person influences her/his appearance.
10. Mood of a human being is influenced by the level of endorphins in the blood.

Mathematical models of the above situations describing proportionality between changing quantities in various contexts are the functions with one or more variables. Italian scientist Galileo Galilei (1564-1642) was the first who used quantitative methods of investigation of dependencies between quantities in his study of processes in dynamics. The idea to describe these interrelations among investigated quantities in an effective and precise way led to the development of many new mathematical concepts, such as mutual correspondence of number sets, mappings of sets, dependent and independent variables and function. They became the primary subject of study in mathematics for more than two centuries, and culminated in the establishment of a separate domain of mathematics known as mathematical analysis, differential and integral calculus.

The mapping of a number set to another number set is called function, usually denoted by letters f, g etc.
Let M be a non-empty set of real numbers: $M \neq \emptyset$, $M \subset \boldsymbol{R}$. A rule f, that assigns exactly one element (real number) $y \in R$ to each element (real number) $x \in M$

$$f: M \to \boldsymbol{R},\ x \to y = f(x)$$

is called a real function of a real variable, briefly a function $f(x)$. The set M is called the domain of definition of function f and it is usually denoted by $D(f)$ and we say that function f is defined on $D(f)$. The number $y = f(x)$ is the value of the function f at the point $x \in M = D(f)$. The set of real numbers which are the values of function f,

$$R(f) = \{y: \exists\, x \in D(f): y = f(x)\}$$

is called the range of function f, y is said to be the dependent and x the independent variable (or argument).
A function may be written in a roster form (set of ordered pairs), in a table form, as an arrow diagram, in a graph or in an equation form (formula). If a function is given by an analytic formula, without specifying its domain $D(f)$, then we are interested in those real x, for which the formula makes a sense. The set of all those x is then accepted as the domain $D(f)$ of the given function, and it is said to be the natural (maximal) domain of definition.
The set

$$G(f) = \{[x\ , y]: x \in D(f), y = f(x)\} \subset \boldsymbol{R} \times \boldsymbol{R}$$

is called the graph of function f.
The graph of a function with one real variable can be sketched in the Cartesian plane with the orthogonal coordinate system Oxy as a set of all points with coordinates $[x, f(x)]$. Coordinate axis x represents the independent variable and coordinate axis y the dependent variable. Domain $D(f)$ is the orthogonal projection of the function graph $G(f)$ onto the coordinate axis x, while range $R(f)$ is the orthogonal projection of $G(f)$ onto the coordinate axis y.
A set of points in the Cartesian plane is the graph of a function, if each straight line parallel to the coordinate axis y has at most one common point with it.
Geometrically the graph of a function interprets a lot of important information on the function behaviour, such as continuity, zero points, increasing or decreasing character of functional dependence, stationary points, points of extreme values and points of inflexion.

Examples

1. Function f is defined by formula $f: y = \sqrt{4 - x^2}$, therefore its domain of definition is set $D(f) = \langle -2, 2\rangle$ and range is set $R(f) = \langle 0, 2\rangle$. Function graph is semicircle with centre in origin and radius 2, in fig. 3.1, left.

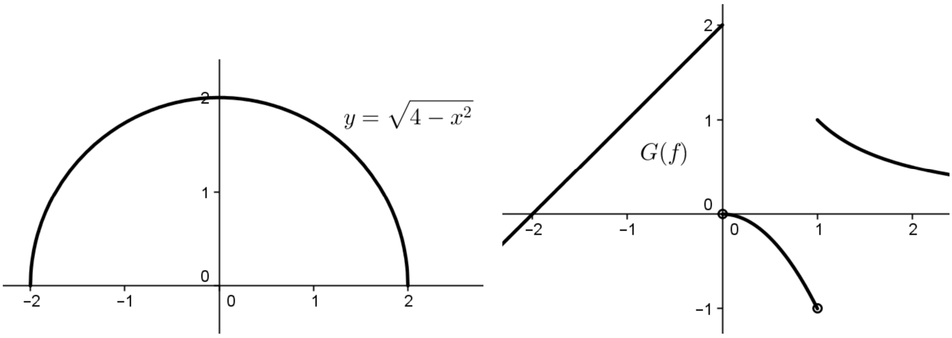


Fig. 3.1. Graphs of functions

2. Function g is defined on $\boldsymbol{R} = (-\infty, \infty)$ by the following formulae

$$g(x) = x + 2, \quad x \in (-\infty, 0\rangle$$

$$g(x) = -x^2, \quad x \in (0,1)$$

$$g(x) = \frac{1}{x}, \quad x \in \langle 1, \infty)$$

while its range is interval $R(f) = (-\infty,\ 2\rangle$. The function graph consists of 3 separate parts, see fig. 3.1, right.

3. Graph of function $f(x) = |x|$ defined on all real numbers $\boldsymbol{R}$ is a pair of semi-lines with a common point in the origin, in the fig. 3.2, left.

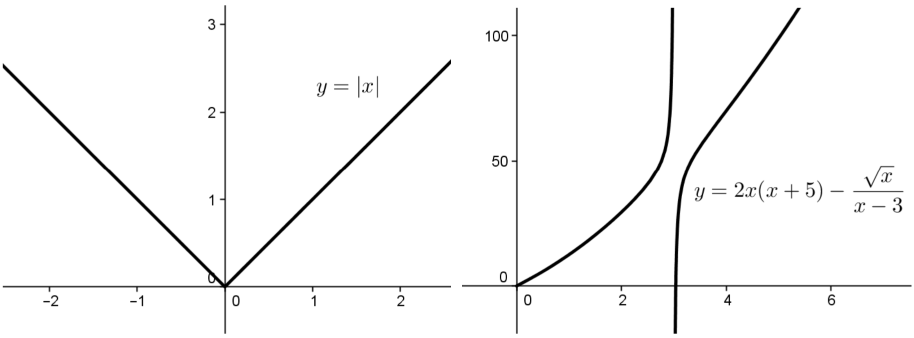


Fig. 3.2. Graphs of functions

4. Graf of function $f(x)$ defined on $\boldsymbol{R} = (-\infty,\infty)$ in the following way
   $f(x) = 1$ for all rational numbers x
   $f(x) = -1$ for all irrational numbers x
   can be described into words, but cannot be sketched in the Cartesian plane.

5. Function defined on set $D = \{-3, -2, -1, 0, 1, 2, 3\}$ with values corresponding to arguments in the given order $R = \{-1, 0, 5, -8, 4, 2, 1\}$ can be presented also in the form of a table, where a series of arguments in one row corresponds to the series of respective function values in the second row.

x	−3	−2	−1	0	1	2	3
y	−1	0	5	−8	4	2	1

Table 3.1. Function values

In addition to the case of functions defined on finite discrete sets, it is sometimes useful to also use the table form for functions whose domain of definition is an infinite number set. Then only important values, those that are of interest for specific reasons, are given in special selected points from function domain. This function definition is mostly used in natural and technical sciences, where the dependence of one variable on others is determined experimentally or by investigation. Logarithmic tables or tables of values of trigonometric functions were also prepared in this way, as these were the most frequently used functional dependencies in different sciences such as physics, astronomy, chemistry, or technical engineering sciences.

Sometimes, a function can be described only with words, e.g. Euler function ϕ is defined for any natural number n so that its value $\Phi(n)$ is the number of natural numbers that are less than n and have no common divisors with n, for instance

$$\Phi(2) = 1,\ \Phi(4) = 2,\ \Phi(5) = 4,\ \Phi(6) = 2$$

and its graph is a set of separate points in the Cartesian plane.

The advantage of an analytic function formula dwells in the fact that strong analytic methods were developed to analyse function behaviour by means of the concept of function derivatives. Moreover, function values can be expressed in arbitrary points from its domain and extremal values can be easily detected. Certain disadvantage can be seen with regard to the unsatisfactory exemplification and insight into the defined dependence, but this can be avoided using the function graph in the Cartesian coordinate system.

3.2 Operations on functions

Let function $f(x)$ be defined on set M.

Absolute value of function

Function $h(x)$, whose domain of definition is M and for all $x \in M$ it holds that $h(x) = |f(x)|$ is called the absolute value of function $f(x)$,

$$h\colon M \to \boldsymbol{R}_0^+,\ x \to y = |f(x)|$$

where $\boldsymbol{R}_0^+$ is the set of all positive real numbers.

Product of number and function

Let k be a real number. Function $h(x)$, whose domain of definition is M and for all $x \in M$ it holds that $h(x) = k \cdot f(x)$ is called the product of number k and function $f(x)$.

Sum, difference, product and quotient of functions

Let f and g be two functions with domains $D(f)$ and $D(g)$.

- Functions f and g are equal to one another, if $D(f) = D(g)$ and if for each x from the domain their values are equal, $f(x) = g(x)$.
- Function F defined on $D(F) = D(f) \cap D(g)$ is called the sum (difference, product, quotient) of functions f and g and denoted $f + g$ ($f - g$, $f \cdot g$, f / g), if for each $x \in D(F)$

  $$F(\mathrm{x}) = f(x) + g(x),\ (F(\mathrm{x}) = f(x) - g(x),\ F(\mathrm{x}) = f(x) \cdot g(x),\ F(x) = f(x) / g(x)).$$

  Apparently, points at which $g(x) = 0$ must be excluded from $D(f) \cap D(g)$, to obtain the domain of the quotient f / g.

Composite function

Suppose that the values of a function g with the domain $D(g)$ can be used as arguments of a function $f(x)$ with the domain $D(f)$. It is then possible to blend f and g together to form a new function $F(x)$, whose inputs are arguments of function f and whose values are numbers $f(g(x))$.
Function $F(x)$

$$F\colon M \to \boldsymbol{R},\ x \to (f \circ g)(x) = f(g(x))$$

is said to be composite function composed from functions $f(x)$ and $g(x)$, if its domain of definition $D(F)$ is the set of all such numbers from the domain of definition of function $g(x)$, in which the function $g(x)$ value is a number from the domain of definition of function $f(x)$, and for all $x \in D(F)$ it holds that $F(x) = f(g(x))$.
Value of function F at the point x equals to the value of function f at the point u that is the value of function g at the point x, $u = g(x)$. Function $f(u)$ is the major (outside) part (component) and function $u = g(x)$ is the minor (inside) part (component) of the composite function $F(x)$.

Examples

1. Function $h(x) = 2x(x+5) - \dfrac{\sqrt{x}}{x-3}$ is the difference of functions f and g. The domain of definition of function $f(\mathrm{x}) = 2x(x + 5)$ is $D(f) = \boldsymbol{R}$, as it is the product of functions $f_1(x) = 2x$ (product of the real number 2 and function $y = x$) with domain $D(f_1) = \boldsymbol{R}$ and function $f_2(x) = x + 5$ that is also defined on $D(f_2) = \boldsymbol{R}$, therefore $D(f) = D(f_1) \cap D(f_2) = \boldsymbol{R}$. Function $g(x) = \dfrac{\sqrt{x}}{x-3}$ is quotient of functions $g_1(x) = \sqrt{x}$ with the domain $D(g_1) = \{x \in \boldsymbol{R}\colon x \geq 0\}$ and $g_2(x) = x - 3$ with the domain $D(g_2) = \boldsymbol{R}$, while the domain of definition of function g_1 / g_2 is a set of all those points from intersection $D(g_1) \cap D(g_2)$, for which $g_2(x) \neq 0$,

therefore $x \neq 3$, which is the set $D(g) = \langle 0, 3) \cup (3, \infty)$. The definition domain of function $h(x)$ is finally set $D(h) = D(f) \cap D(g)$ that equals to $\langle 0, 3) \cup (3, \infty)$, see the graph in fig. 3.2, right.

2. Function $F(x) = \sin(x - \pi)$ is a composite function with a major part f: $y = \sin x$ and a minor part g: $y = x - \pi$. Its value at the point $x = 0$ is $F(0) = 0$, as $g(0) = -\pi$, $\sin(-\pi) = 0$.

3.3 Some special classes of functions

Bounded functions

Function $f(x)$ defined on set $D(f)$ is called bounded (bounded above, bounded below), if such a real number K exists that for all $x \in D(f)$ it holds that

$$\left| f(x) \right| \leq K \quad (f(x) \leq K,\ f(x) \geq K)$$

It means that a function is bounded (bounded above, bounded below), if its range $R(f)$ is a bounded (bounded above, bounded below) set of real numbers. Function that is not bounded is called unbounded.

The property of function to be bounded can be geometrically interpreted as follows:

- graph of function $f(x)$ bounded from above is under or in the line $y = K$, (see fig. 3.1, on the right)
- graph of function $f(x)$ bounded from below is over or in the line $y = K$, (see fig. 3.2, on the left)
- graph of bounded function $f(x)$ is in-between parallel lines $y_1 = -K$, $y_2 = K$, (see fig. 3.1, left).

Function f is said to be bounded on set $M \subset D(f)$, if and only if number $K > 0$ exists and such that for all $x \in M$ it holds that $|f(x)| \leq K$.

Examples

1. Function $f(x) = \dfrac{1}{x^2 + 4}$ defined on $\boldsymbol{R}$ is bounded. For any real number x it holds that

$$x^2 > 0 \Rightarrow x^2 + 4 > 4 \Rightarrow \frac{1}{x^2+4} < \frac{1}{4}$$

$$x^2 > 0 \Rightarrow x^2 + 4 > 0 \Rightarrow \frac{1}{x^2+4} > 0 \Rightarrow \left| \frac{1}{x^2+4} \right| = \frac{1}{x^2+4}$$

and if $K = f(0) = ¼$, $|f(x)| \leq K$. The graph of function f is sketched in fig. 3.3.

2. Function $g(x) = 2 - x^2$ is bounded from above, as for all $x \in D(f) = \boldsymbol{R}$ it holds that $g(x) \leq 2$. Its graph is a parabola with vertex in the point $V = [0, 2]$, fig. 3.4.

3. Function $h(x) = |x|$ is bounded from below, $h(x) > 0$, (fig. 3.2, on the left).

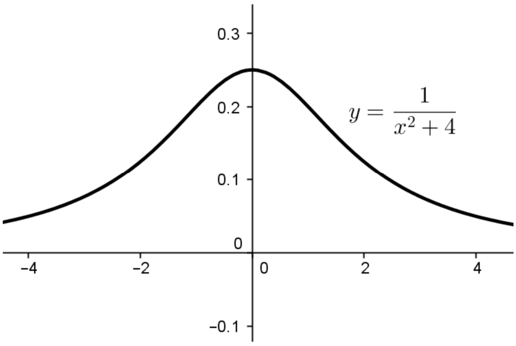


Fig. 3.3. Function graph

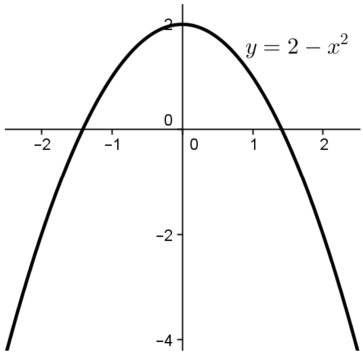


Fig. 3.4. Function graph

Monotone functions

There can be distinguished 4 types of monotone functions: increasing, decreasing, non-decreasing and non-increasing functions.

A function f is called increasing (decreasing, non-decreasing, non-increasing), if for any two points x_1, x_2 from its domain of definition the following is valid:

$$x_1 < x_2 \Rightarrow f(x_1) < f(x_2)\ (f(x_1) > f(x_2),\ f(x_1) \leq f(x_2),\ f(x_1) \geq f(x_2)).$$

It is clear that any increasing function is non-decreasing, and any decreasing function is non-increasing, but the opposite is not true. Constant functions represent the only possible type of functions, which are non-decreasing and non-increasing simultaneously. Increasing and decreasing functions are said to be strictly monotone.

Examples

1. Function $y = x^2$ is not monotone in the above sense, as none of the required conditions are fulfilled for each pair of points from its domain of definition $D(f) = \boldsymbol{R}$. Nevertheless, restricting its domain to the set of non-negative numbers, it is easy to see that f is increasing there, and similarly f is decreasing on the set of non-positive numbers.
2. Function $f : y = 2 - \frac{1}{\sqrt{x}}$ is strictly monotone, as for all x from its domain of definition $D(f) = \boldsymbol{R}^+$ (all non-negative real numbers) it holds that

$$0 < x_1 < x_2 \Rightarrow \sqrt{x_1} < \sqrt{x_2} \Rightarrow \frac{1}{\sqrt{x_1}} > \frac{1}{\sqrt{x_2}} \Rightarrow$$

$$\Rightarrow -\frac{1}{\sqrt{x_1}} < -\frac{1}{\sqrt{x_2}} \Rightarrow 2 - \frac{1}{\sqrt{x_1}} < 2 - \frac{1}{\sqrt{x_2}}$$

3. Function $h(x) = |x|$ is decreasing on $(-\infty, 0\rangle$ and increasing on $\langle 0, \infty)$, (fig. 3.2, on the left).

4. Function $g(x) = \frac{|x|}{x}$, $x \neq 0$ is non-decreasing (non-increasing) on its domain of definition, see graph on fig. 3.5.

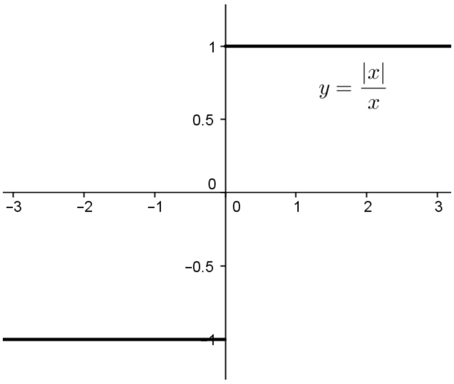


Fig. 3.5. Function graph

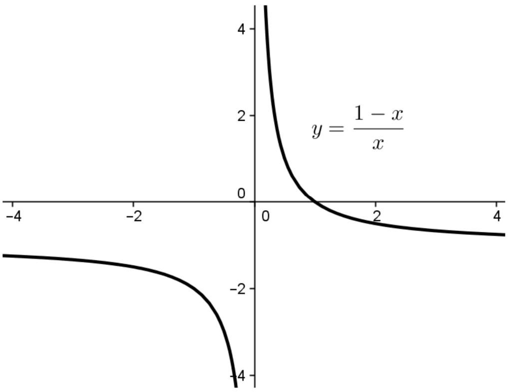


Fig. 3.6. Function graph

Periodicity of function

A function *f* is called periodic, if a positive number *p* exists such that if $x \in D(f)$, then also $x \pm p \in D(f)$, and $f(x + p) = f(x)$ for each $x \in D(f)$. Number *p* is called the period of the function *f*.

All functional values of a periodic function repeat themselves infinitely many times, which means that the part of graph on any interval of the length *p* is also repeated infinitely many times and the whole graph of function consists of copies of it.

Example

1. Trigonometric functions are the most frequently used periodic functions. Period of the functions sine and cosine is $p = 2\pi$, while minimal period of the functions tangent and cotangent is $p = \pi$.

Parity of functions

Let function *f* be defined on set *M* such that for each $x \in M$ is also $-x \in M$. Function *f* is said to be even, if for any $x \in M$ it holds that $f(-x) = f(x)$, and it is said to be odd, if for any $x \in M$ it holds that $f(-x) = -f(x)$.

Function $f: M \to \boldsymbol{R}$, $x \to f(x)$ is on the set *M*:

a) even, if $\forall x \in M: -x \in M \wedge f(-x) = f(x)$

b) odd, if $\forall x \in M: -x \in M \wedge f(-x) = -f(x)$.

The graphs of even functions are symmetrical with respect to the coordinate axis *y*, while the graphs of odd functions are symmetrical with respect to the origin of the coordinate system.

Examples

1. Function $y = \cos x$ is an even function, $\cos(-x) = \cos x$ and the function graph is symmetric with respect to the coordinate axis y.
2. Function $y = \sin x$ is an odd function, $\sin(-x) = -\sin x$ and the function graph is symmetric with respect to the origin of the coordinate system.
3. Function $f(x) = \dfrac{1-x}{x}$ is neither odd, nor even, see the function graph presented in fig. 3.6, as $f(-x) = -\dfrac{1+x}{x}$, $\quad -f(x) = \dfrac{x-1}{x}$.

One-to-one function

Let f be a function defined on $D(f)$. If for any two x_1, $x_2 \in D(f)$, $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$, then the function is said to be one-to-one. A function is one-to-one if each straight line parallel to the coordinate axis x has at most one common point with the function graph $G(f)$. Any strictly monotone function is on-to-one.

Inverse function

Let f be a one-to-one function with the domain $D(f)$ and the range $R(f)$ and let function $f^{-1}(x)$ be defined on $R(f)$ as follows: for each $y_0 \in R(f)$: $f^{-1}(y_0) = x_0 \in D(f)$, if $f(x_0) = y_0$. Then function f^{-1} is called the inverse function of function f.
Obviously: $D(f^{-1}) = R(f)$, $R(f^{-1}) = D(f)$, $(f^{-1})^{-1} = f$.
Any inverse function is again one-to-one. The graphs of two mutually inverse functions f and f^{-1}, therefore $G(f)$ and $G(f^{-1})$ are symmetric with respect to the straight line $y = x$.

Examples

1. Function $f(x) = 2x + 3$, whose domain and range is a set of all real numbers $\boldsymbol{R}$ is strictly monotone (increasing) function, as

$$x_1 < x_2 \Rightarrow 2x_1 < 2x_2 \Rightarrow 2x_1 + 3 < 2x_2 + 3 \Rightarrow f(x_1) < f(x_2)$$

therefore it is one-to-one, and its inverse function exists determined by exchanging depended and independent variables as follows

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

and the inverse function $f^{-1}(x) = \dfrac{x-3}{2}$ is defined on $\boldsymbol{R}$. The graphs of f and f^{-1} are in fig. 3.7.

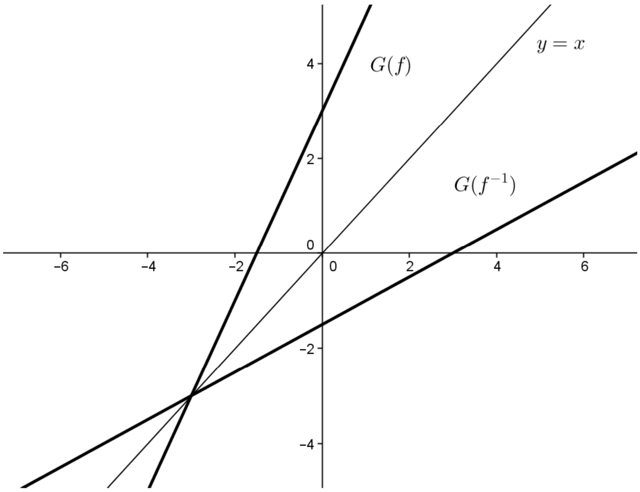


Fig. 3.7. Graphs of inverse functions

2. Trigonometric functions $y = \sin x$, $y = \cos x$, $y = \tan x$, $y = \cot x$ are periodic, it means that they acquire each value from their ranges infinitely many times, so it follows that they are not one-to-one and they have no inverse functions.

3.4 Elementary functions

A constant or power function with real exponent, exponential and logarithmic functions, trigonometric and their inverse cyclometric functions or hyperbolic and hyperbolometric functions are collectively called elementary functions.
Any function represented by means of a finite number of operations such as sum, difference, product or quotient on these functions, or any function composed from a finite number of presented elementary functions is also considered an elementary function.
An elementary function is considered to be defined for all those values of argument x, at which the given analytic formula makes sense, and it reaches a real value.
Elementary functions are frequently used in mathematics and its applications. Many functions (often with rather complicated analytic formulas) describing for example the dependence of various physical variables, properties and behaviour of complex systems that are mathematical models of certain problems from technical practise, or determining influence of specific parameters on functionality of some technical devices, are elementary functions. Their properties can be investigated by means of differential calculus.

Rational functions
A polynomial function is P in x defined on $\boldsymbol{R}$ by the formula

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$$

where n is a non-negative integer and real numbers $a_0, a_1, \ldots, a_n$ are its coefficients.

Polynomial is of degree n if $a_n \neq 0$. The graph of the polynomial function is a curve of degree n, which means that it has at most n intersection points with a straight line passing not parallel to the coordinate axis y.
For $a_0 = a_1 = ... = a_n = 0$ is $P(x) = 0$ for all $x \in \boldsymbol{R}$. This polynomial is called zero polynomial. For $n = 0$, $a_0 \neq 0$ it holds that $P(x) = a_0$ for all $x \in \boldsymbol{R}$, which is a non-zero constant function with a graph in a straight line $y = a_0$ parallel to the coordinate axis x.
For $n = 1$, $P(x) = a_1 x + a_0$ for all $x \in \boldsymbol{R}$ is a polynomial of degree 1, which is a linear function with a graph in a line, increasing or decreasing on $\boldsymbol{R}$.
For $n = 2$, $P(x) = a_2 x^2 + a_1 x + a_0$ for all $x \in \boldsymbol{R}$ is a polynomial of degree 2, which is a quadratic function with a graph in a parabola with axis parallel to the coordinate axis y, fig. 3.4.
For $n = 3$, $P(x) = a_3 x^3 + a_2 x^2 + a_1 x + a_0$ for all $x \in \boldsymbol{R}$ is a polynomial of degree 3, which is a cubic function. The graph of a cubic function is a curve of degree three called a cubic curve.
Let $P(x)$ be an arbitrary polynomial, $Q(x)$ be a non-zero polynomial, and let M be a set of all those real numbers x, for which $Q(x) \neq 0$. Function $R(x)$ defined on M by the formula

$$R(x) = \frac{P(x)}{Q(x)}$$

is called a rational function. If the degree of polynomial in the enumerator is less than the degree of polynomial in the denominator, then the function is called purely rational, otherwise it is called not-purely rational.

Power functions
Function f: $y = x^r$ for $x > 0$, where r is a real number, is a power function.
For natural number $r = n$, function $y = x^n$ is defined on $\boldsymbol{R}$ and it is called a power function with natural exponent. In the case of an even number n, the range is $H = \langle 0, \infty)$ and the function is even, increasing on interval $\langle 0, \infty)$ and decreasing on interval $(-\infty, 0\rangle$. For all odd numbers n the range of power function with natural exponent is $H = \boldsymbol{R}$, while the function is odd and increasing on $\boldsymbol{R}$.
If exponent r is a negative integer number, then $y = x^r$ defines a rational function on the set $D = (-\infty,0) \cup (0,\infty)$.
If $r = \frac{1}{q}, q \in \boldsymbol{N}$, then function $y = r = x^{\frac{1}{q}} = \sqrt[q]{x}$ is defined on $\boldsymbol{R}$ for an odd q, while it is defined on interval $\langle 0, \infty)$ for an even q.
The graph of a power function $y = x^r$ for $r = 0$, $r = 1$ is a straight line $y = 1$, $y = x$, for $r = ½$ it is a part of parabola with axis in the coordinate axis x and vertex in the origin, fig. 3.8, left.
Hyperbola with equal semi-axes, axes in the coordinate axes x and y and the centre in the origin is the graph of a power function with exponent $r = -1$, $y = x^{-1}$.

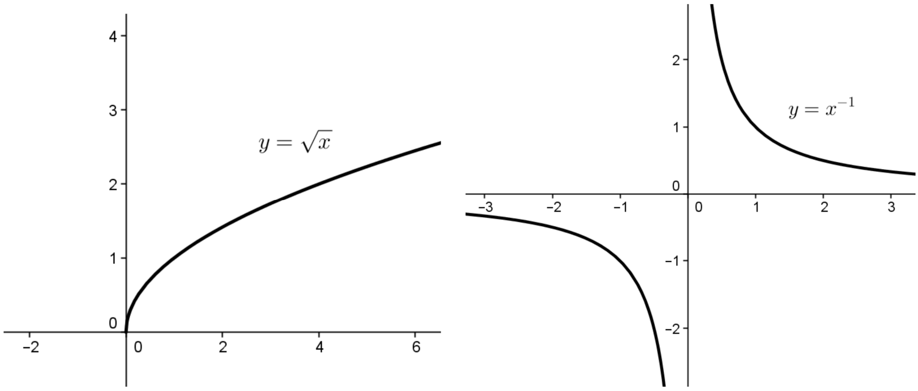


Fig. 3.8. Graphs of power functions

Exponential function

Function $f(x) = a^x$, $a > 0$, $a \neq 1$ defined on $\boldsymbol{R}$ is an exponential function increasing for $a > 1$ and decreasing for $0 < a < 1$, with the range $H(f) = (0, \infty)$.

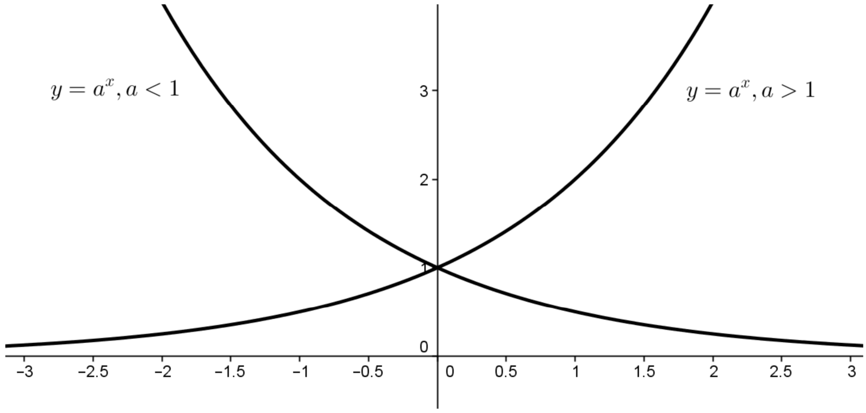


Fig. 3.9. Graphs of exponential functions

Exponential function is a constant function for $a = 1$ with the value 1, its graph is a straight line parallel to coordinate axis x. This function is not one-to-one, therefore it is usually assumed that $a \neq 1$, if not stated differently.

Especially important function is the natural exponential function $y = e^x$ with the base equal to the transcendent Euler number $e = 2{,}718281...$.

Let x_1, x_2 be two real numbers, then the following holds:

1. $a^{x_1} \cdot a^{x_2} = a^{x_1 + x_2}$

2. $\dfrac{a^{x_1}}{a^{x_2}} = a^{x_1 - x_2}$

3. $\left(a^{x_1}\right)^{x_2} = a^{x_1 \cdot x_2}$

Logarithmic function

Function $f(x) = \log_a x, a > 0, a \neq 1$ defined on interval $D(f) = (0, \infty)$ with range $H(f) = \boldsymbol{R}$ is a logarithmic function, increasing for $a > 1$ and decreasing for $0 < a < 1$. Logarithmic function is an inverse to exponential function. Its graph is symmetric to the graph of exponential function with respect to line $y = x$.

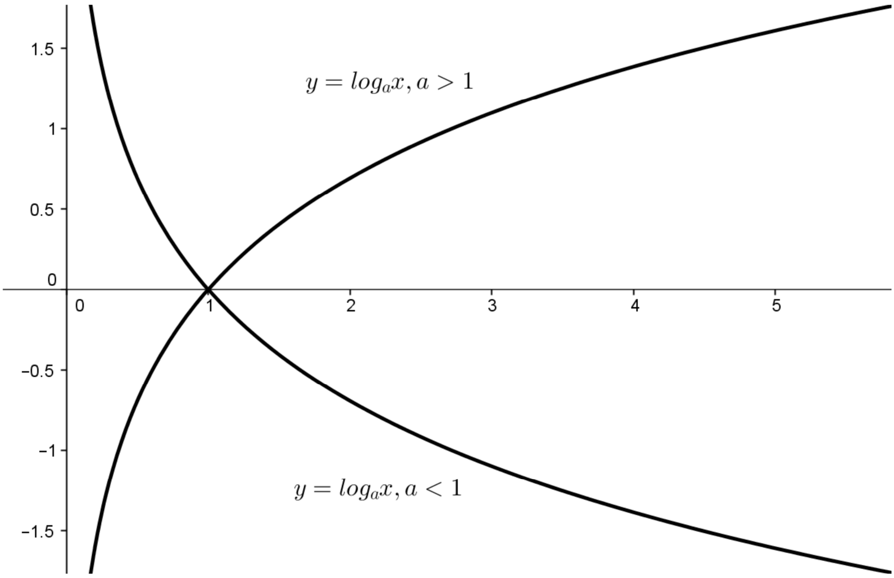


Fig. 3.10. Graphs of logarithmic functions

The value of logarithmic function at the point x is logarithm of number x for the base a, $\log_a x$, i.e. it is the value of the exponent to which base a should be raised to obtain x.

The most frequently used logarithms in practical calculations are decimal logarithms with the base $a = 10$, the decimal logarithmic function is usually denoted $y = \lg x = \log_{10} x$, and natural logarithms with the base in Euler number e, while the natural logarithmic function is denoted $y = \ln x = \log_e x$.

For any $x > 0, y > 0, a > 0, a \neq 1$ the following holds:

1. $\log_a xy = \log_a x + \log_a y$
2. $\log_a \frac{x}{y} = \log_a x - \log_a y$
3. $\log_a x^r = r \log_a x, r \in \boldsymbol{R}$
4. $x = a^{\log_a x}$

Trigonometric functions

Functions sine, cosine, tangent and cotangent are collectively called trigonometric functions.

The function sine defined on $\boldsymbol{R}$, with range $H = \langle -1, 1\rangle$,

$$\sin x: \boldsymbol{R} \to \langle -1, 1\rangle: y = \sin x$$

is an odd function, periodic with period 2π, bounded, and its graph is a sinusoidal curve in fig. 3.11.

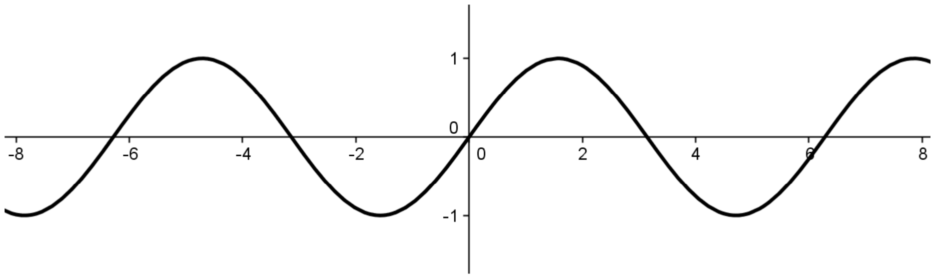


Fig. 3.11. Graf of function sine – sinusoidal curve

Function cosine defined on $\boldsymbol{R}$, with range $H = \langle -1, 1\rangle$,

$$\cos x: \boldsymbol{R} \to \langle -1, 1\rangle: y = \cos x$$

is an even function, periodic with period 2π, bounded, and its graph is a shifted sinusoidal curve.

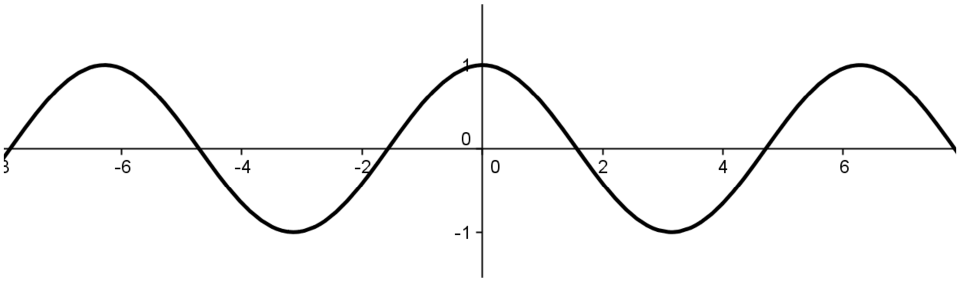


Fig. 3.12. Graf of function cosine – shifted sinusoidal curve

Function tangent is defined as the quotient of functions sine and cosine for those $x \in \boldsymbol{R}$, for which $\cos x \neq 0$, its domain is $D = \left\{ x \in \boldsymbol{R} : x \neq \frac{(2k+1)\pi}{2}, k \in \boldsymbol{Z} \right\}$, while its range is $\boldsymbol{R}$.

$$\tan x\ (\mathrm{tg}\, x): D \to \boldsymbol{R}:\ y = \frac{\sin x}{\cos x} = \tan x\ (= \mathrm{tg}\, x)$$

It is an odd function, periodic with period π, unbounded, and increasing on all intervals $\left(-\frac{\pi}{2}+k\pi, \frac{\pi}{2}+k\pi\right), k \in \boldsymbol{Z}$.

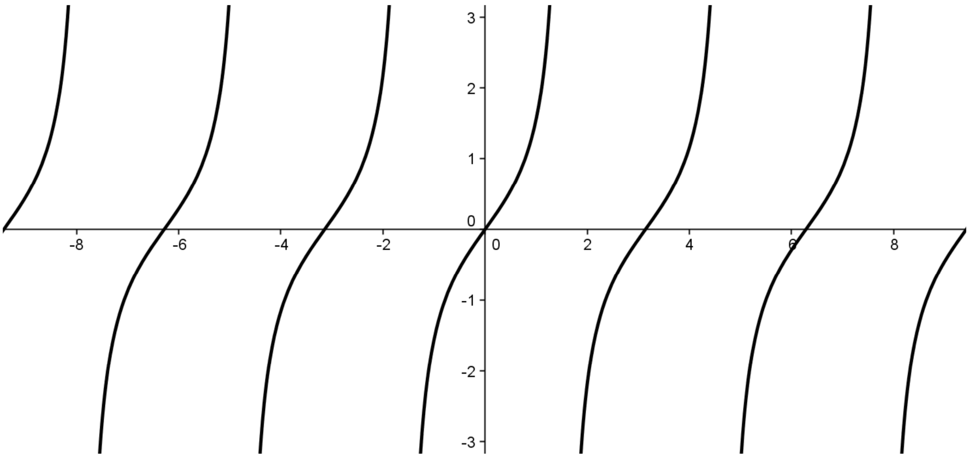


Fig. 3.13. Graf of function tangent

Function cotangent is defined as the quotient of functions cosine and sine for those $x \in \boldsymbol{R}$, for which sin $x \neq 0$, its domain is $D = \{x \in \boldsymbol{R} : x \neq k\pi, k \in \boldsymbol{Z}\}$, while its range is $\boldsymbol{R}$.

$$\cot x\colon D \to \boldsymbol{R}\colon\; y = \frac{\cos x}{\sin x} = \cot x\ (= \operatorname{cotg} x)$$

It is an odd function, periodic with period 2π, unbounded, and decreasing on all intervals $(k\pi, (k+1)\pi), k \in \boldsymbol{Z}$.

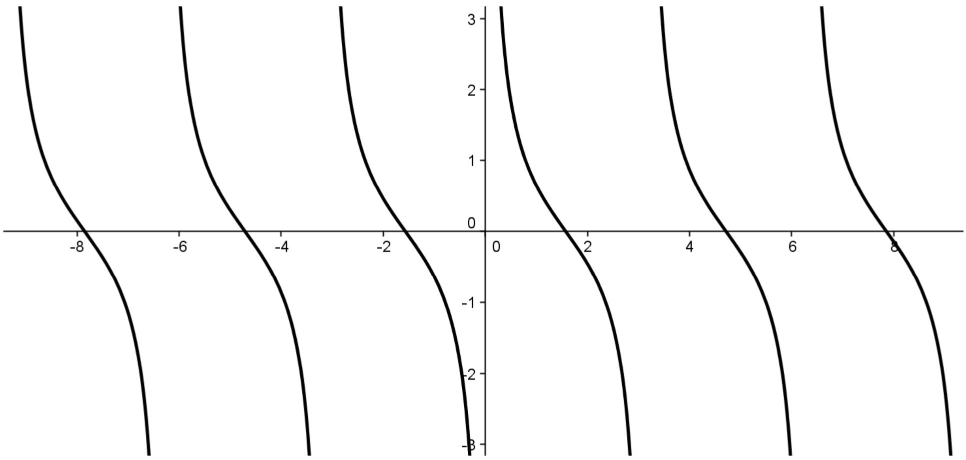


Fig. 3.14. Graf of function cotangent

Cyclometric functions

Trigonometric functions are periodic, it means that they assume each value from their ranges infinitely many times, therefore they are not one-to-one and they have no inverse functions. But if these functions are considered to be defined on those relevant parts of their natural domains on which they are one-to-one, their inverse functions exist, they are called cyclometric functions and they are defined as below.

The function arcsine is inverse to function sine on interval $\langle -\pi/2, \pi/2\rangle$, its domain is $D = \langle -1, 1\rangle$ and its range is $H = \langle -\pi/2, \pi/2\rangle$.

$$\arcsin\colon \langle -1, 1\rangle \to \langle -\pi/2, \pi/2\rangle\colon y = \arcsin x$$

It is increasing, bounded and odd, and for all $u \in \langle -1, 1\rangle$ it holds that
$v = \arcsin u \Leftrightarrow u = \sin v$.

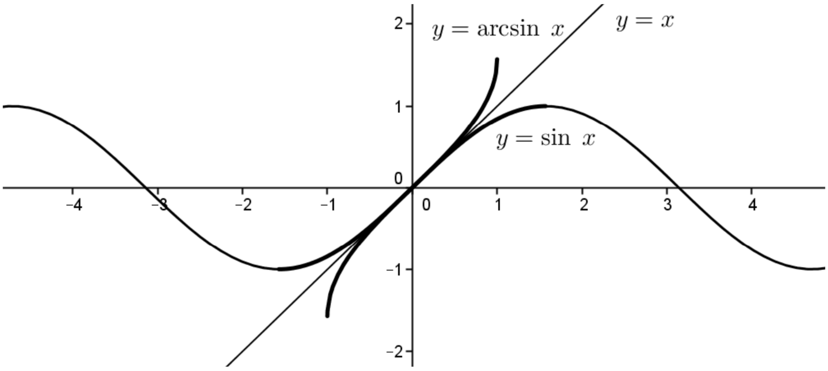


Fig. 3.15. Graph of function arcsine

The function arccosine is inverse to function cosine on interval $\langle 0, \pi\rangle$, its domain is $D = \langle -1, 1\rangle$ and its range is $H = \langle 0, \pi\rangle$.

$$\arccos\colon \langle -1, 1\rangle \to \langle 0, \pi\rangle\colon y = \arccos x$$

It is decreasing and bounded, while for all $u \in \langle -1, 1\rangle$ it holds that
$v = \arccos u \Leftrightarrow u = \cos v$.

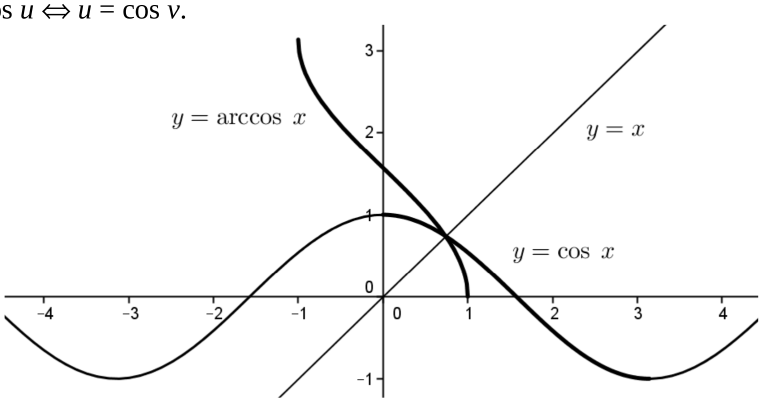


Fig. 3.16. Graph of function arccosine

The function arctangent is inverse to function tangent on interval (−π/2, π/2), its domain is $\boldsymbol{R}$ and its range is $H = (-\pi/2, \pi/2)$.

$$\text{arctan: } \boldsymbol{R} \to (-\pi/2, \pi/2)\text{: } y = \arctan x$$

It is increasing, bounded and odd, while for all $u \in \boldsymbol{R}$ it holds that
$v = \arctan u \Leftrightarrow u = \tan v$.

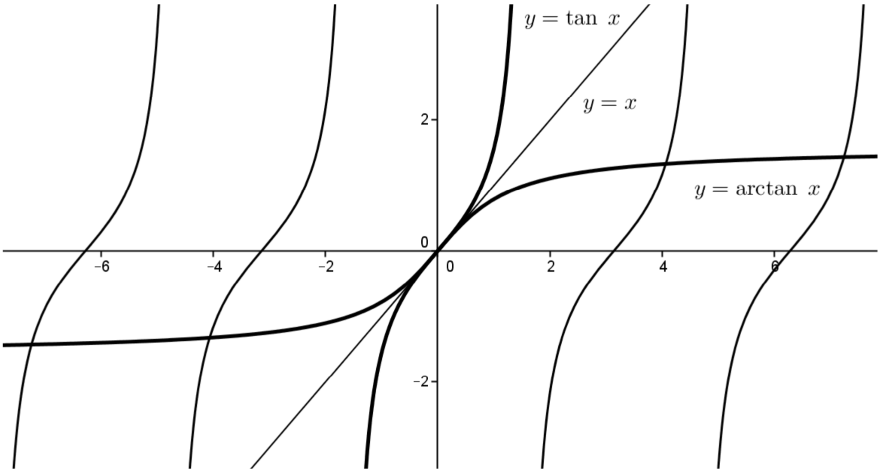


Fig. 3.17. Graph of function arctangent

The function arccotangent is inverse to function cotangent on interval (0, π), its domain is $\boldsymbol{R}$ and its range is $H = (0, \pi)$.

$$\text{arccot: } \boldsymbol{R} \to (0, \pi)\text{: } y = \operatorname{arccot} x$$

It is decreasing and bounded, while for all $u \in \boldsymbol{R}$ it holds that
$v = \operatorname{arccot} u \Leftrightarrow u = \cot v$.

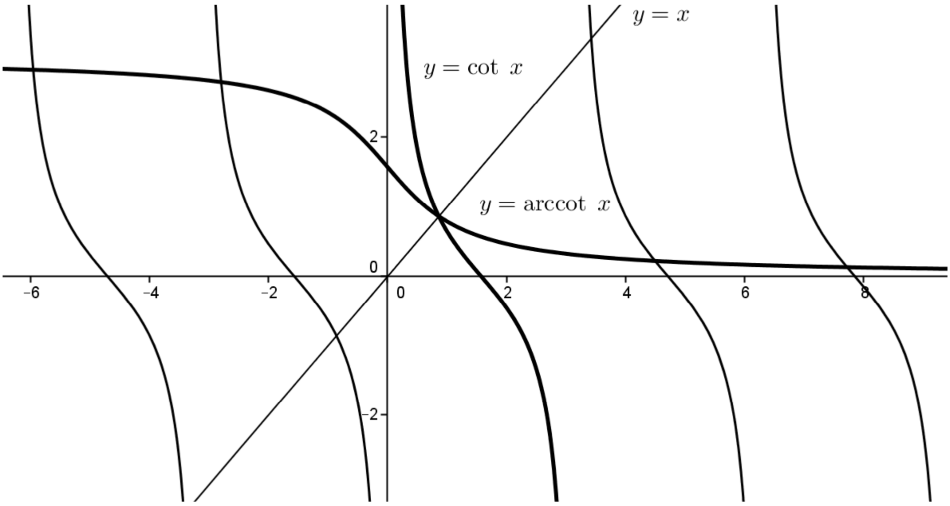


Fig. 3.18. Graph of function arccotangent

Hyperbolic functions
Function sine hyperbolic is defined by the formula

$$\sinh x\colon \boldsymbol{R} \to \boldsymbol{R}\colon\ y = \sinh x = \frac{e^x - e^{-x}}{2}$$

and it is an odd function, increasing.
Function cosine hyperbolic is defined by the formula

$$\cosh x\colon \boldsymbol{R} \to \langle 1, \infty)\colon\ y = \cosh x = \frac{e^x + e^{-x}}{2}$$

and it is an even function, decreasing on interval $(-\infty, 0)$ and increasing on interval $(0, \infty)$.

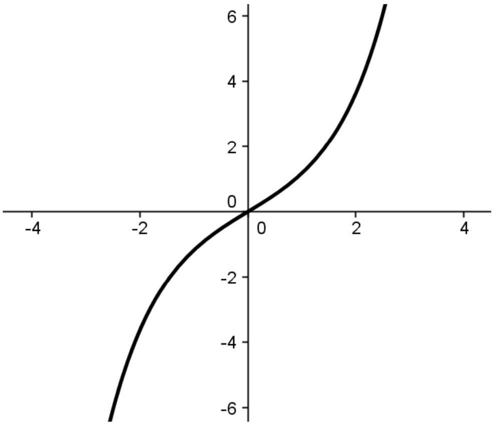


Fig. 3.19. Sine hyperbolic

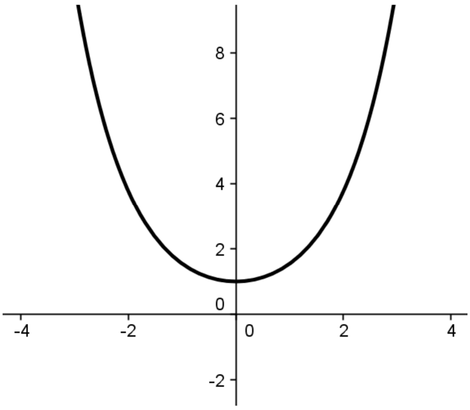


Fig. 3.20. Cosine hyperbolic

Function tangent hyperbolic is defined by the formula

$$\tanh x\colon \boldsymbol{R} \to (-1,1)\colon\ y = \tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

and it is an odd function, increasing.

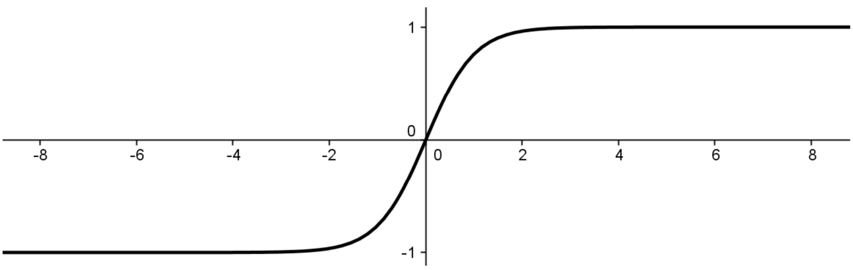


Fig. 3.21. Tangent hyperbolic

Function cotangent hyperbolic is defined by the formula

$$\coth x\colon \boldsymbol{R}-\{0\}\to \boldsymbol{R}-\langle -1,1\rangle\colon\ y=\coth\ x=\frac{e^{x}+e^{-x}}{e^{x}-e^{-x}}$$

and it is an odd function, decreasing.

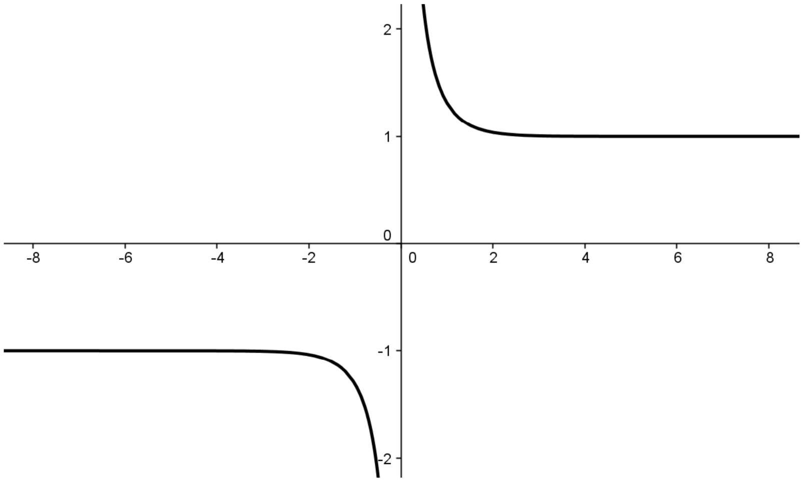


Fig. 3.22. Cotangent hyperbolic

Inverse hyperbolic functions (area hyperbolic, hyperbolometric functions)
Hyperbolometric, or area hyperbolic functions, are inverse functions to hyperbolic functions on intervals, on which these functions are strictly monotone.
Inverse hyperbolic sine

$$x=\sinh y,\ y\in \boldsymbol{R},\ y=\operatorname{arcsinh} x,\ x\in \boldsymbol{R}.$$

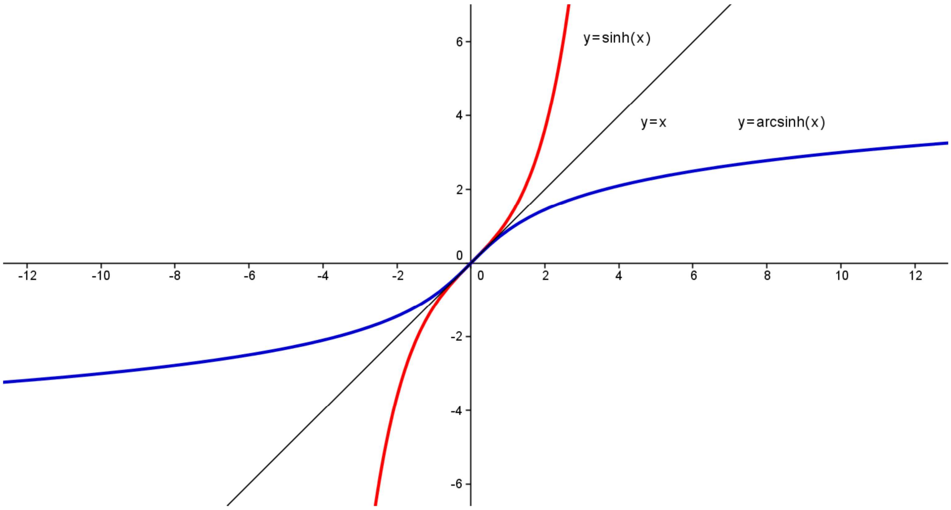


Fig. 3.23. Graph of function inverse hyperbolic sine

Inverse hyperbolic cosine

$$x = \cosh y,\ y \in \langle 0, \infty),\ y = \operatorname{arccosh} x,\ x \in \langle 1, \infty).$$

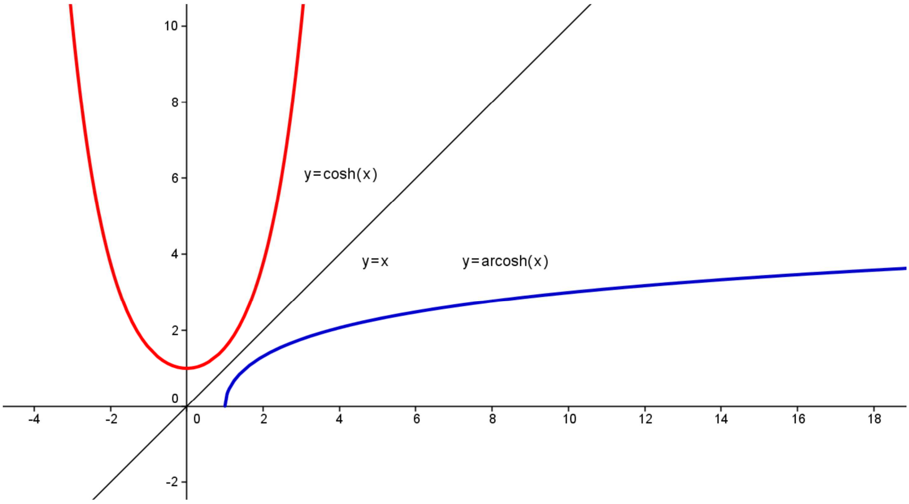


Fig. 3.24. Graph of function inverse hyperbolic cosine

Inverse hyperbolic tangent

$$x = \tanh y,\ y \in \boldsymbol{R},\ y = \operatorname{arctanh} x,\ x \in \langle -1, 1\rangle.$$

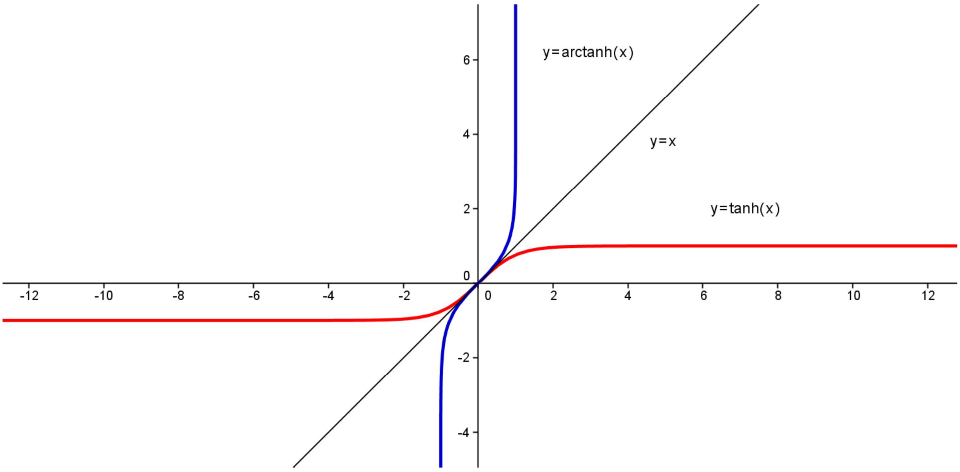


Fig. 3.25. Graph of function inverse hyperbolic tangent

Inverse hyperbolic cotangent

$$x = \coth y,\ y \neq 0,\ y = \operatorname{arccoth} x,\ |x| > 1.$$

3.5 Sequences

Every function f defined on the set of all natural numbers $D(f) = \boldsymbol{N}$, is called a sequence. If its range of values is a set of real numbers, $R(f) \subset \boldsymbol{R}$, function f is called a numerical or number sequence. The value $f(n)$, for $n \in \boldsymbol{N}$ is called the n-th term (member) of the sequence, while its usual notation is a_n instead. All terms of the sequence are referred as $\{a_n\}$, or $\{a_n\}_{n=1}^{\infty}$.

Examples

1. The following notations have the same meaning

$$f(n) = \frac{n-1}{n},\ n \in \boldsymbol{N}, \quad \{a_n\} = \left\{\frac{n-1}{n}\right\} = 0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots$$

2. Arithmetic sequence is defined as $\{a + (n-1)d\}$, where a is the first term and d is the difference, both are real numbers.

3. Geometric sequence with the first term a and quotient q, both real numbers, is defined by formula $a_n = aq^{n-1}$.

4. Sequence $\left\{\frac{1}{n}\right\} = 1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots$ is called harmonic sequence.

5. Sequence 2, 3, 5, 7, ..., p_n, ... is the sequence of prime numbers, where n-th term p_n is the n-th largest prime.

Sequences are real functions, so they can possess some of the general properties of real functions.

The graph of a sequence is a set of isolated points $\{A_n = [n, a_n],\ n \in \boldsymbol{N}$, fig. 3.26.

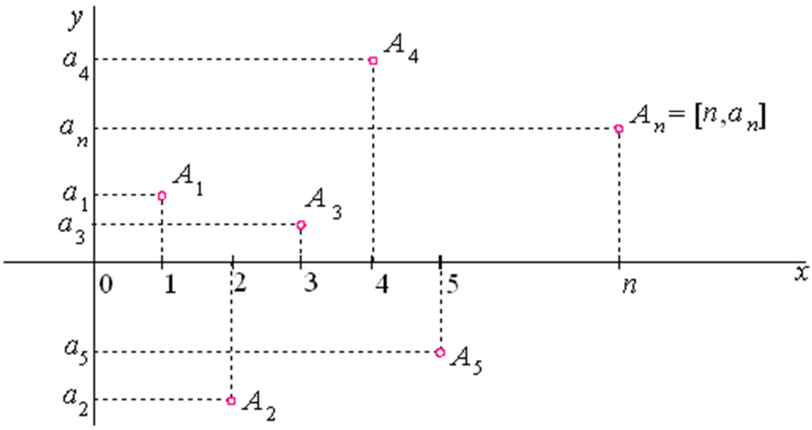


Fig. 3. 26. Graph of a sequence

The domains of definition disable sequences to be even, odd or periodic, but they can be monotone or bounded. Monotonicity and boundedness of sequences is defined in the same way as for real functions of a real variable. The definition of monotonicity for sequences can be simplified in the following way:
A sequence $\{a_n\}$ is increasing (decreasing, non-decreasing, non-increasing), if for each $n \in \boldsymbol{N}$

$$a_n < a_{n+1} \ (a_n > a_{n+1},\ a_n \le a_{n+1},\ a_n \ge a_{n+1}).$$

Increasing, non-decreasing, decreasing and non-increasing sequences are called monotone sequences, increasing and decreasing sequences are called strictly monotone sequences.

Examples

1. Arithmetic sequence is increasing for positive difference, $d > 0$, and decreasing for negative difference, $d < 0$, while it is stationary for $d = 0$.
2. Geometric sequence is decreasing for $q < 1$, and it is increasing for $q > 1$, while it is stationary for $q = 1$.
3. Harmonic sequence is decreasing for all natural n.
4. Sequence of primes is increasing.

Sequence $\{a_n\}$ is bounded (bounded below, bounded above) if such real numbers K, L exist that for all natural numbers n it holds that

$$K < a_n < L \ (K < a_n,\ a_n < L).$$

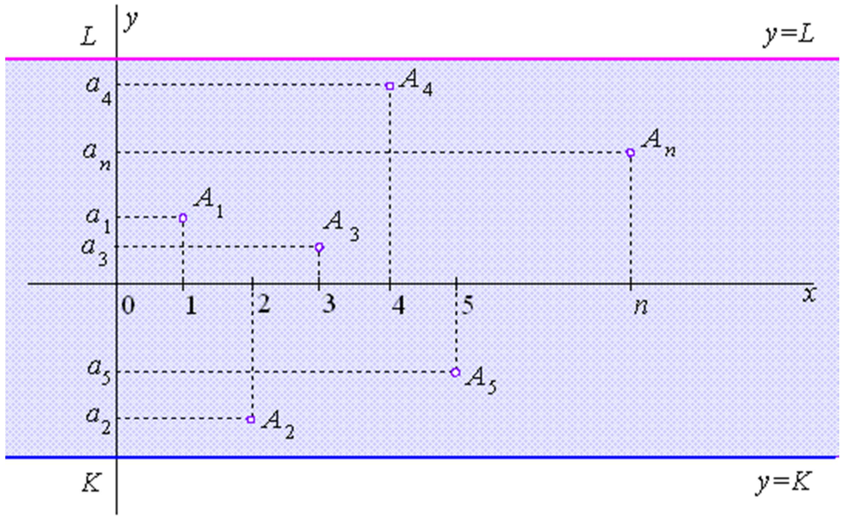


Fig. 3.27. Graph of bounded sequence

Number K or L is called the lower, or upper bound of sequence. All points on the graph of bounded sequences are in the layer between parallel lines with equations $y = L$, $y = K$. All points on the graph of a sequence bounded from above or from below are in one half-plane determined by the line with equation $y = L$.

Examples

1. Arithmetic sequence with positive difference is bounded from below; sequence with negative difference is bounded from above.
2. Harmonic sequence is bounded, as for all natural n it holds that $0 < \frac{1}{n} \le 1$.

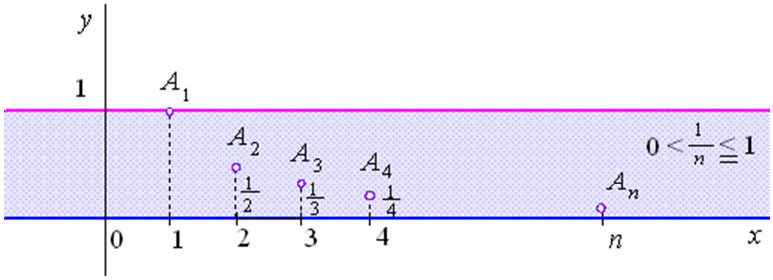


Fig. 3.28. Graph of harmonic sequence

3. Oscillating sequence $\{(-1)^n\}_{n=1}^{\infty} = -1,1,-1,1,\ldots$ is bounded.

Limit of a sequence

The concept of a limit of a sequence is one of the most important concepts in mathematics. It describes a special property of some sequences, which can be represented as the following tendency: with an increasing n the corresponding sequence terms assume a value close to a certain number called the limit of a sequence.

Let us consider a sequence

$$\left\{\frac{n}{n+1}\right\} = \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots$$

that is increasing, bounded, and it has a special property that if n tends to infinity, its n-th term tends to one. For large enough n its terms are very close to number 1. In geometric interpretation we obtain points on the sequence graph appearing close to the line $y = 1$, and their distance is diminishing with an increasing n.

Choosing an arbitrary small positive number $\varepsilon > 0$, such term always exists in the sequence, from which all consequent terms are in a distance from 1 that is lesser then the chosen ε. This leads to the concept of a limit of a sequence.

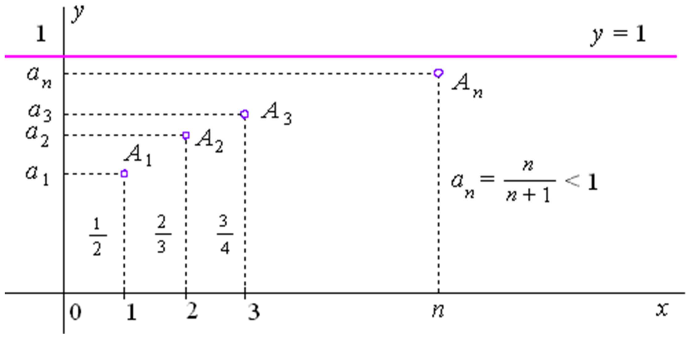


Fig. 3.29. Graph of sequence

Let $\{a_n\}$ be a sequence and a be a real number. If for any $\varepsilon > 0$ such a number $n_0(\varepsilon)$ exists that for each $n \in \boldsymbol{N}$, $n > n(\varepsilon)$ it holds that $|a_n - a| < \varepsilon$, then number a is called a (proper) limit of the sequence $\{a_n\}$, the sequence is said to be convergent to a, and it is written as

$$\lim_{n\to\infty} a_n = a$$

Briefly: $\lim\limits_{n\to\infty} a_n = a \Leftrightarrow \forall\varepsilon > 0, \exists n_0(\varepsilon) : \forall n > n_0(\varepsilon) > |a_n - a| < \varepsilon$

A sequence that is not convergent is called divergent.

Examples

1. $\lim\limits_{n\to\infty} \frac{n}{n+1} = 1$
2. $\lim\limits_{n\to\infty} \frac{1}{n} = 0$
3. Sequence $\{(-1)^n\}_{n=1}^{\infty}$ has no limit.

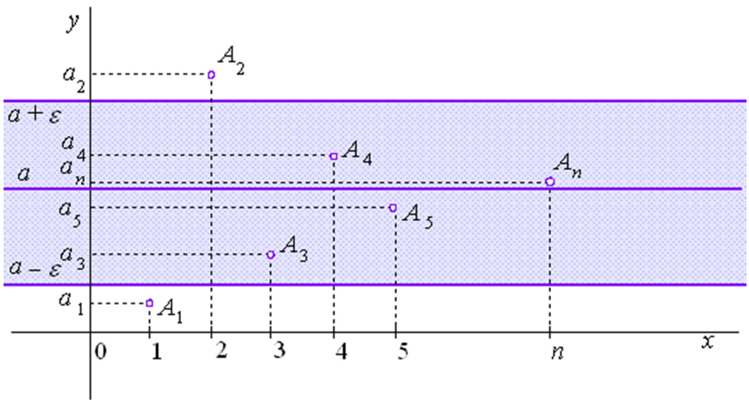


Fig. 3.30. Graph of sequence limit

Alternatively, the limit of a sequence can also be determined as a property of almost all sequence terms to be in an ε-neighbourhood of the number a that is the sequence limit, see fig. 3.30.
Number a is called the limit of a sequence $\{a_n\}$, if for any number $\varepsilon > 0$ and for almost all sequence terms a_n it holds that $|a_n - a| < \varepsilon$.
In other words:
Number a is called the limit of a sequence $\{a_n\}$, if any neighbourhood $O_\varepsilon(a)$ contains almost all sequence terms.
For relations between convergence and boundedness or monotonicity of a sequence, the following is valid.

1. Any sequence has at most one limit (i.e. none or just one).
2. Sequence $\{a_n\}$ has a limit if and only if the sequence $\{a_n - a\}$ has a limit equal to 0.
3. If $\lim\limits_{n\to\infty}|a_n| = 0$, then also $\lim\limits_{n\to\infty} a_n = 0$ and vice versa, if $\lim\limits_{n\to\infty} a_n = 0$, then also $\lim\limits_{n\to\infty}|a_n| = 0$.
4. Let it hold for all terms of sequence $\{a_n\}$ that $a_n \le A$ ($a_n \ge$ A) and let $\lim\limits_{n\to\infty} a_n = a$. Then it holds that $a \le A$ ($a \ge A$).
5. Sandwich theorem: Let $\lim\limits_{n\to\infty} a_n = a$, $\lim\limits_{n\to\infty} c_n = a$, and let it hold for all natural n that $a_n \le b_n \le c_n$, then $\lim\limits_{n\to\infty} b_n = a$.
6. Any convergent sequence is bounded. Unbounded sequence is divergent.
7. Any sequence that is both monotone and bounded is convergent.
8. Let sequences $\{a_n\}$, $\{b_n\}$ be convergent. Then the following sequences are convergent: $\{k \cdot a_n\}$, $k \in \boldsymbol{R}$, $\{a_n \pm b_n\}$, $\{a_n \cdot b_n\}$, $\left\{\dfrac{a_n}{b_n}\right\}, b_n \ne 0$ for all n.

   If $\lim\limits_{n\to\infty} a_n = a$, $\lim\limits_{n\to\infty} b_n = b$, then $\lim\limits_{n\to\infty} k \cdot a_n = k \cdot a$, $\lim\limits_{n\to\infty}(a_n \pm b_n) = a \pm b$, $\lim\limits_{n\to\infty} a_n \cdot b_n = a \cdot b$, $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = \dfrac{a}{b}, b \ne 0$.

Examples

1. Let there be two different limits $a \ne b$ of a sequence $\{a_n\}$. Then almost all sequence terms are in any neighbourhood $O_\varepsilon(a)$ and also $O_\varepsilon(b)$, it means, in their intersection. Then choosing $\varepsilon = (b - a)/3 > 0$, we obtain $O_\varepsilon(a) \cap O_\varepsilon(b) = \emptyset$, which means that almost all sequence terms can be only in one of the two neighbourhoods, therefore the sequence does not have two different limits. Property 1 has been proved.

2. For all terms of sequence $\{a_n\}=1,\frac{1}{2},\frac{1}{3},...,\frac{1}{n},...$ it holds that $a_n \leq 1$, and $\lim_{n\to\infty} a_n = 0 < 1$.
3. Sequence $\{(-1)^n\}$ is bounded, but not convergent. It is neither monotone.
4. Sequence $\left\{\left(1+\frac{1}{n}\right)^n\right\} = 2, \frac{9}{4}, \frac{64}{9},...$ is increasing and bounded, as $1 < a_n < 3$, therefore it is convergent and its limit equals to Euler number e, $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e$.

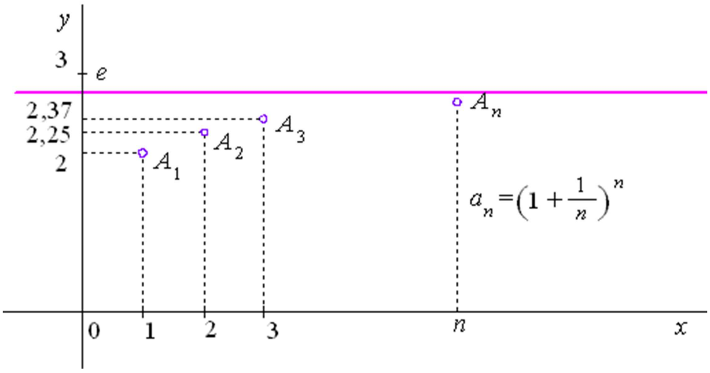


Fig 3.31. Graph of sequence

Improper limit of a sequence

Suppose sequence $\{a_n\}$ is increasing, unbounded, and therefore not convergent. For all terms of such sequence it holds that with an increasing index n the value of the sequence terms is also increasing without any bounds. It means that for any number A such a number n_0 must exist that for all $n > 0$ it is $a_n > A$, i.e. almost all the sequence terms are in any neighbourhood $O_A(\infty)$. This leads to the concept of an improper limit ∞ of a sequence.

The sequence $\{a_n\}$ has an improper limit ∞, if almost all its terms are in any neighbourhood $O_A(\infty)$, that is to any number A such number n_0 exists that for all $n > 0$ it is $a_n > A$, which means $\lim_{n\to\infty} a_n = \infty$. The sequence $\{a_n\}$ has an improper limit $-\infty$, if almost all its terms are in any neighbourhood $O_A(-\infty)$, that is to any number A such number n_0 exists, that for all $n > 0$ it is $a_n < A$, which means $\lim_{n\to\infty} a_n = -\infty$.

In brief:

$$\lim_{n\to\infty} a_n = \infty \Leftrightarrow \forall A, \exists n_o : \forall n > n_0 : a_n > A\text{, fig. 3.32}$$

$$\lim_{n\to\infty} a_n = -\infty \Leftrightarrow \forall A, \exists n_o : \forall n > n_0 : a_n < A\text{, fig. 3.33.}$$

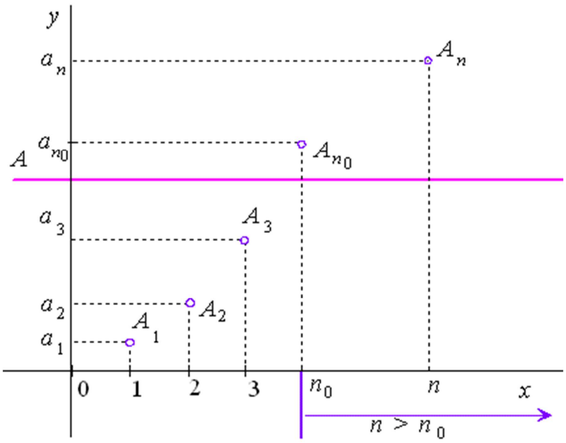


Fig. 3.32. Graph of a sequence with improper limit ∞

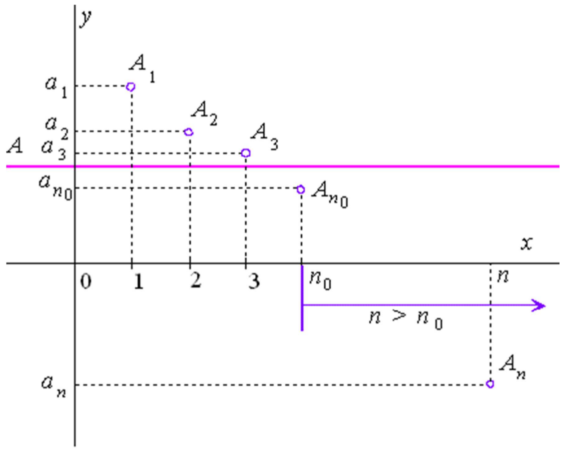


Fig. 3.33. Graph of a sequence with improper limit -∞

For a sequence $\{a_n\}$, only one of the following is true:

1. There exists a proper limit $\lim\limits_{n\to\infty} a_n = a$.

2. There exists an improper limit $\lim\limits_{n\to\infty} a_n = \infty$.

3. There exists an improper limit $\lim\limits_{n\to\infty} a_n = -\infty$.

4. There exists no proper or improper limit, the sequence is oscillating.

Examples

1. Sequence $\{n^2\}$ = 1, 4, 9, ... is increasing and unbounded, so it is divergent and $\lim\limits_{n\to\infty} n^2 = \infty$.

2. Sequence $\{1-2^n\} = -1, -3, -26, \ldots$ is decreasing and unbounded, so it is divergent and $\lim\limits_{n\to\infty}(1-2^n) = -\infty$.

3. Sequence $\{(1-n)^n\} = 0, 1, -8, 81, \ldots$ is unbounded and oscillating, so it has no limit.

Some properties of the limit of monotone sequences:

1. Let sequence $\{a_n\}$ be non-decreasing. If it is not bounded from above, then $\lim\limits_{n\to\infty} a_n = \infty$. If it is bounded from above, then it has a proper limit a and $a_n \le a$ for all n.
2. Let sequence $\{b_n\}$ be non-increasing. If it is not bounded from below, then $\lim\limits_{n\to\infty} a_n = -\infty$. If it is bounded, it has a proper limit a and $a_n \ge a$ for all n.
3. Monotone sequence is convergent if and only if it is bounded.
4. Let sequence $\{a_n\}$ be bounded and let $\lim\limits_{n\to\infty} b_n = \infty$. Then

$$\lim_{n\to\infty}(a_n + b_n) = \infty \qquad \lim_{n\to\infty}(a_n - b_n) = -\infty \qquad \lim_{n\to\infty}\frac{a_n}{b_n} = 0, b_n \ne 0\ \forall n.$$

5. If $\lim\limits_{n\to\infty} a_n = \infty$ and $\lim\limits_{n\to\infty} b_n = \infty$, then $\lim\limits_{n\to\infty}(a_n + b_n) = \infty$, $\lim\limits_{n\to\infty}(a_n \cdot b_n) = \infty$.
6. If $\lim\limits_{n\to\infty} a_n = \infty$ and $\lim\limits_{n\to\infty} b_n = -\infty$, then
$\lim\limits_{n\to\infty}(a_n - b_n) = \infty$, $\lim\limits_{n\to\infty}(a_n \cdot b_n) = -\infty$.
7. If $\lim\limits_{n\to\infty} a_n = a \ne 0$ and $\lim\limits_{n\to\infty} b_n = \infty$, then $\begin{array}{l}\lim\limits_{n\to\infty}(a_n \cdot b_n) = \infty, a > 0 \\ \lim\limits_{n\to\infty}(a_n \cdot b_n) = -\infty, a < 0\end{array}$.
8. If $\lim\limits_{n\to\infty} a_n = a \ne 0$ and $\lim\limits_{n\to\infty} b_n = 0, b_n > 0, \forall n$, then $\begin{array}{l}\lim\limits_{n\to\infty}\dfrac{a_n}{b_n} = \infty, a > 0 \\ \lim\limits_{n\to\infty}\dfrac{a_n}{b_n} = -\infty, a < 0\end{array}$.

Examples

1. $\lim\limits_{n\to\infty} 2^n = \infty$, therefore $\lim\limits_{n\to\infty}\dfrac{1}{2^n} = 0$.
2. $\lim\limits_{n\to\infty}(n^3 - 5n - 10) = \lim\limits_{n\to\infty}\left(1 - \dfrac{5}{n^2} - \dfrac{10}{n^3}\right)n^3 = \infty$ because

$$\lim_{n\to\infty}\left(1 - \frac{5}{n^2} - \frac{10}{n^3}\right) = 1 > 0, \lim_{n\to\infty} n^3 = \infty.$$

3. $\lim\limits_{n\to\infty} \dfrac{n^2+5n+2}{n^4+4n^3+3n^2+1} = \lim\limits_{n\to\infty} \dfrac{\dfrac{1}{n^2}+\dfrac{5}{n^3}+\dfrac{2}{n^4}}{1+\dfrac{4}{n}+\dfrac{3}{n^2}+\dfrac{1}{n^4}} = 0$, because

$$\lim_{n\to\infty}\left(1+\frac{4}{n}+\frac{3}{n^2}+\frac{1}{n^4}\right)=1\neq 0, \lim_{n\to\infty}\left(\frac{1}{n^2}+\frac{5}{n^3}+\frac{2}{n^4}\right)=0.$$

3.6 Limit and continuity of function

The concept of a limit of a function is one of the basic concepts in calculus leading to other important concepts as continuity, derivative, or anti-derivative of a function. The limit of a function determines the behaviour of the function in the neighbourhood of a certain point, while the function itself is defined on a certain neighbourhood of this point but not necessarily in the point itself.

Let us investigate the function $f(x)=\dfrac{x^2-1}{x-1}$ defined for all real numbers but 1, $D(f)$ = $\boldsymbol{R}$ − {1}, in the neighbourhood of point 1. The function values are $f(0.8) = 1.8$, $f(0.9) = 1.9$, $f(1.1) = 2.1$, $f(1.2) = 2.2$, therefore the function values are approaching number 2, and it is said that the limit of function f at the point 1 equals to 2, fig. 3.34.

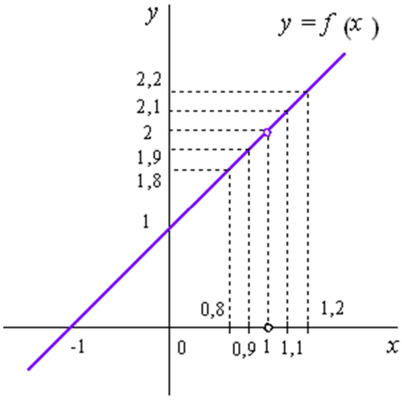


Fig. 3.34. Limit of function

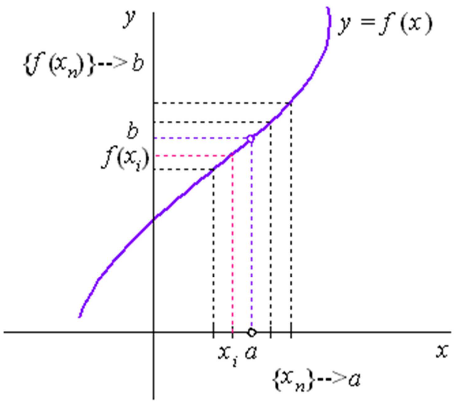


Fig. 3.35. Limit of function – Heine definition

This property can be precisely formulated in the following Heine definition of the limit of function f at the point a, where f is defined on some neighbourhood of point a, see fig. 3.35.

Let function f be defined for all $x \neq a$ from some neighbourhood of point a. Function f is said to have limit b at the point a, if for any sequence $\{x_n\}$ of points from the domain of definition of function f such that $x_n \neq a$ convergent to a the corresponding sequence of function values $\{f(x_n)\}$ has the limit b.

$$\lim_{x\to a} f(x)=b \Leftrightarrow \left[\lim_{n\to\infty} x_n = a, x_n \in D(f), x_n \neq a \Rightarrow \lim_{n\to\infty} f(x_n)=b\right]$$

The definition of the limit of function describes the fact that in points *x* slightly different from the point *a* (but different from *a*) the values *f*(*x*) differ only slightly from *b*.

Cauchy definition of the limit of function at the point *a*

Let function *f* be defined for all $x \neq a$ from some neighbourhood of point *a*. Function *f* is said to have limit *b* at the point *a*, if to any neighbourhood $O_\varepsilon(b)$ such neighbourhood $O_\delta(a)$ exists that for all $x \in O_\delta(a)$, $x \neq a$ is $f(x) \in O_\varepsilon(b)$.

$$\forall \varepsilon > 0\, \exists \delta > 0 : 0 < |x-a| < \delta \Rightarrow |f(x)-b| < \varepsilon$$

The Cauchy definition of $\lim_{x\to a} f(x)=b$ can be interpreted geometrically as the property of the function graph in the neighbourhood $O_\delta(a)$ of point *a* to be located in the layer between the parallel lines $y = b - \varepsilon$, $y = b + \varepsilon$, see fig. 3.36.

The function has no limit at the point *a*, if for some (at least one) ε-neighbourhood of number *b*, $O_\varepsilon(b) = (b - \varepsilon,\ b + \varepsilon)$, there exists no δ-neighbourhood of point *a*, $O_\delta(a) = (a - \delta,\ a + \delta)$, such that for all points from this neighbourhood different from *a*, $x \neq a$, is $f(x) \in O_\varepsilon(b)$, see fig. 3.37.

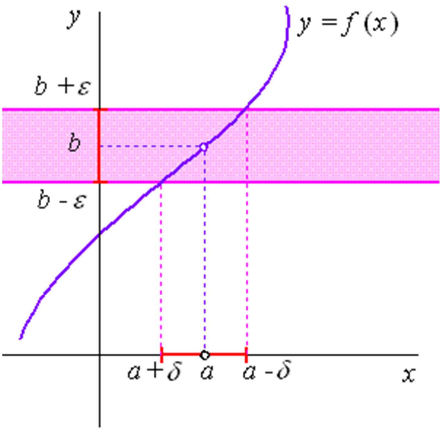


Fig. 3.36. Limit of function – Cauchy definition

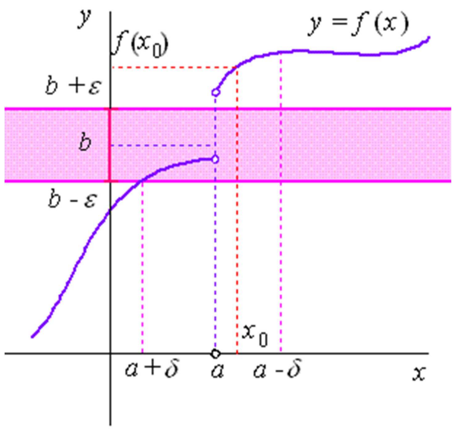


Fig. 3.37. Non-existing limit of function

Basic properties of limit of a function

P1. Any function *f* can have at most one limit.

P2. $\lim_{x\to a} f(x)=b \Leftrightarrow \lim_{x\to a}(f(x)-b)=0$

P3. Let $\lim_{x\to a} f(x)=b$ and let there exists such neighbourhood *O*(*a*) that for all $x \in O(a), x \neq a$ is $K \leq f(x) \leq L$ (the function is bounded on *O*(*a*)), then $K \leq b \leq L$.

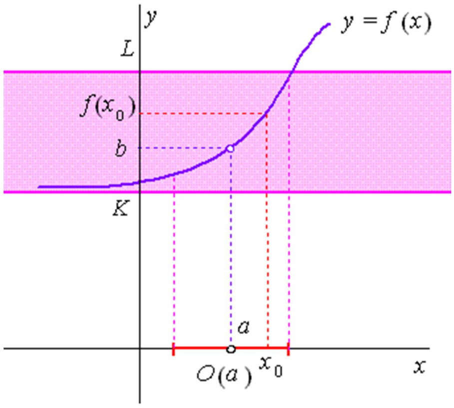


Fig 3.38. Limit of a function bounded on $O(a)$

P4. If $f(x) = c$, $c \in \boldsymbol{R}$, $D(f) = \boldsymbol{R}$, then function f has limit for all $a \in \boldsymbol{R}$ and $\lim\limits_{x\to a} f(x) = c$.

P5. Limit of three functions: Let $\lim\limits_{x\to a} f(x) = b$ and $\lim\limits_{x\to a} h(x) = b$ and let there be such neighbourhood $O(a)$ of point a that for all $x \in O(a), x \neq a$ is $f(x) \leq g(x) \leq h(x)$. Then it also holds that $\lim\limits_{x\to a} g(x) = b$.

P6. If $\lim\limits_{x\to a} f(x) = A$ and $\lim\limits_{x\to a} g(x) = B$ and k, l are real numbers, then:

a) $\lim\limits_{x\to a}\left[k \cdot f(x) \pm l \cdot g(x)\right] = k \lim\limits_{x\to a} f(x) \pm l \lim\limits_{x\to a} g(x) = kA \pm lB$

b) $\lim\limits_{x\to a} f(x) \cdot g(x) = \lim\limits_{x\to a} f(x) \cdot \lim\limits_{x\to a} g(x) = A \cdot B$

c) If $b \neq 0, \lim\limits_{x\to a} \dfrac{f(x)}{g(x)} = \dfrac{\lim\limits_{x\to a} f(x)}{\lim\limits_{x\to a} g(x)} = \dfrac{A}{B}$

d) $\lim\limits_{x\to a}[f(x)]^k = \left[\lim\limits_{x\to a} f(x)\right]^k = A^k, k \in \boldsymbol{N}$

e) $\lim\limits_{x\to a} c^{f(x)} = c^{\lim\limits_{x\to a} f(x)} = c^A, c \in \boldsymbol{R}, c \neq 0$

P7. Limit of a composite function: Let $\lim\limits_{x\to a} g(x) = b, \lim\limits_{u\to b} f(u) = B$ and let there be such neighbourhood $O(a)$ of point a that for all $x \in O(a), x \neq a$ it holds that $g(x) = b$. Then the composite function $f(g(x))$ has the limit at point a and $\lim\limits_{x\to a} f(g(x)) = \lim\limits_{u\to b} f(u) = B$.

Examples

1. Let function $f(x) = x\sin\frac{1}{x},\ x \in \boldsymbol{R}$, then $\lim\limits_{x\to 0} f(x) = 0$.

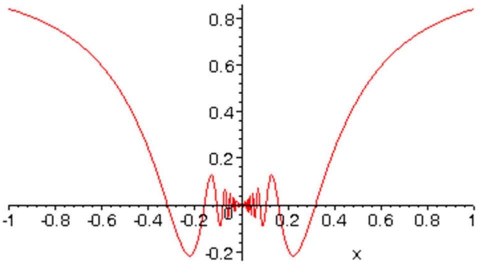


Fig. 3.39. Graph of $f(x) = x\sin\frac{1}{x},\ x \in \boldsymbol{R}$

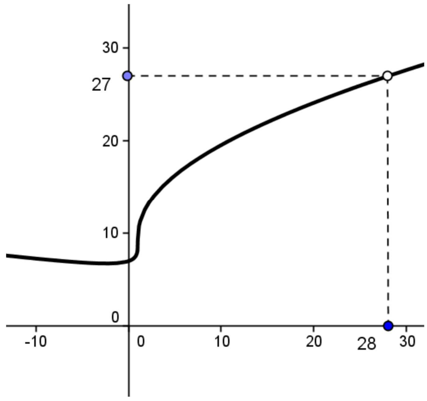


Fig. 3.40. Graph of $f(x) = \dfrac{x-28}{\sqrt[3]{x-1}-3}$

2. $\lim\limits_{x\to 3}(x^2+4) = \lim\limits_{x\to 3} x^2 + \lim\limits_{x\to 3} 4 = 9+4 = 13$

3. $\lim\limits_{x\to -1} \dfrac{2x^5}{x^2+1} = \dfrac{\lim\limits_{x\to -1} 2x^5}{\lim\limits_{x\to -1} x^2+1} = \dfrac{-2}{2} = -1$

4. $\lim\limits_{x\to 3} 2^{\frac{x}{3}-1} = 2^{\lim\limits_{x\to 3}\left(\frac{x}{3}-1\right)} = 2^{\left(\lim\limits_{x\to 3}\frac{x}{3}-\lim\limits_{x\to 3}1\right)} = 2^{1-1} = 2^0 = 1$

5. $\lim\limits_{x\to 28} \dfrac{x-28}{\sqrt[3]{x-1}-3} = 27$, because if $u = g(x) = \sqrt[3]{x-1}-3$,

then $u+3 = \sqrt[3]{x-1} \Rightarrow x-1 = (u+3)^3 \Rightarrow x-28 = (u+3)^3 - 27$

and substitution to $f(x) = \dfrac{x-28}{\sqrt[3]{x-1}-3}$ results in

$$f(g(x)) = f(u) = \frac{(u+3)^3-27}{u} = u^2+9u+27$$

that yields

$$\lim_{x\to 28} \frac{x-28}{\sqrt[3]{x-1}-3} = \lim_{u\to 0}\frac{(u+3)^3-27}{u} = \lim_{x\to 28}\left(u^2+9u+27\right) = 27$$

6. $\lim\limits_{x\to 1} x^{2x-1} = 1$, because if $u = g(x) = 2x-1$, then $x = \dfrac{u+1}{2}$, $f(u) = \left(\dfrac{u+1}{2}\right)^{u}$

and $\lim\limits_{x\to 1} x^{2x-1} = \lim\limits_{u\to 1}\left(\dfrac{u+1}{2}\right)^{u} = \lim\limits_{u\to 1}\left(\dfrac{u+1}{2}\right)^{\lim\limits_{u\to 1} u} = 1^1 = 1$

Improper limit of a function

In the definition of a limit of function the letters *a*, *b* denote numbers. Exchanging letter *b* by symbol ∞ or −∞, the definition of improper limit of function at the point *a* is defined.

Let function *f* be defined for all $x \neq a$ from some neighbourhood of point *a*. Function *f* has an improper limit ∞, or −∞, at the point *a*, if from

$$\lim_{n\to\infty} x_n = a, x_n \in D(f), x_n \neq a \text{ implies}$$

$$\lim_{n\to\infty} f(x_n) = \lim_{x\to a} f(x) = \infty, \text{ or } \lim_{n\to\infty} f(x_n) = \lim_{x\to a} f(x) = -\infty .$$

Briefly:

$$\lim_{x\to a} f(x) = \infty \Leftrightarrow \left[\lim_{n\to\infty} x_n = a, x_n \in D(f), x_n \neq a \Rightarrow \lim_{n\to\infty} f(x_n) = \infty\right]$$

$$\lim_{x\to a} f(x) = -\infty \Leftrightarrow \left[\lim_{n\to\infty} x_n = a, x_n \in D(f), x_n \neq a \Rightarrow \lim_{n\to\infty} f(x_n) = -\infty\right]$$

The geometric interpretation of the fact that function *f* has an improper limit ∞ or −∞ at the point *a* means that if *x* is approaching *a*, then the values of *f*(*x*) are increasing above any bounds.

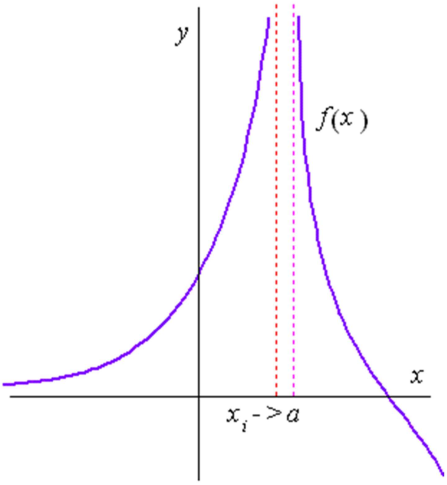


Fig. 3.41. Improper limit ∞ of function *f*

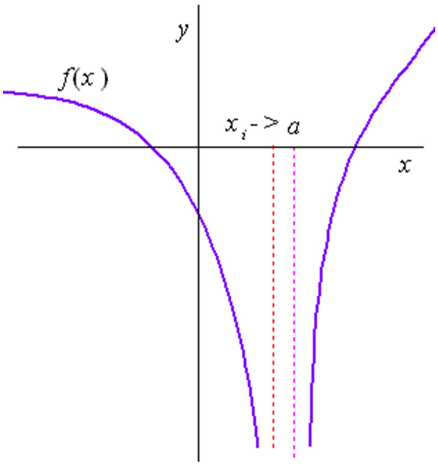


Fig. 3.42. Improper limit −∞ of function *f*

Limit of function at improper point

Exchanging letter a in definition of the limit of function by symbol ∞ or $-\infty$, the definition of the limit of function at the improper point ∞ or $-\infty$ is defined.

Let function f be defined on interval $(-\infty, a)$ or (a, ∞). Function f has limit b at the improper point $-\infty$ or ∞, if for any neighbourhood $O(b)$ such number $A > 0$ exists that for all $x < -A$ or all $x > A$ it holds that $f(x) \in O(b)$, which is denoted as

$$\lim_{x\to-\infty} f(x) = b, \text{ or } \lim_{x\to\infty} f(x) = b .$$

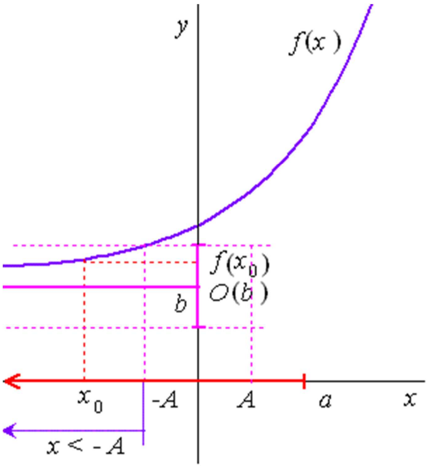


Fig. 3.43. Limit of function f at $-\infty$

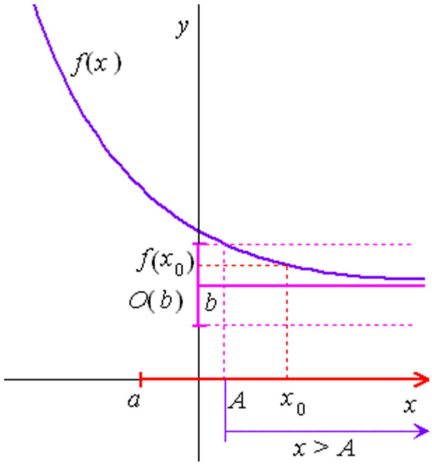


Fig. 3.44. Limit of function f at ∞

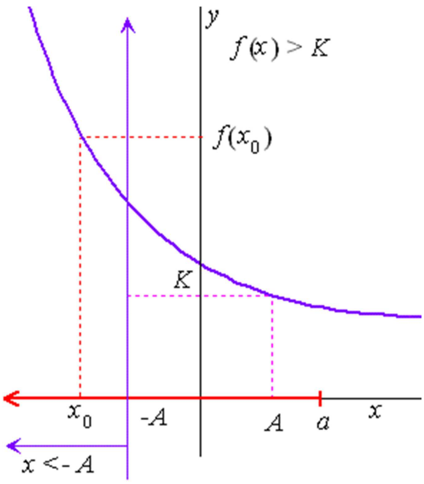


Fig. 3.45. Improper limit ∞ of f at $-\infty$

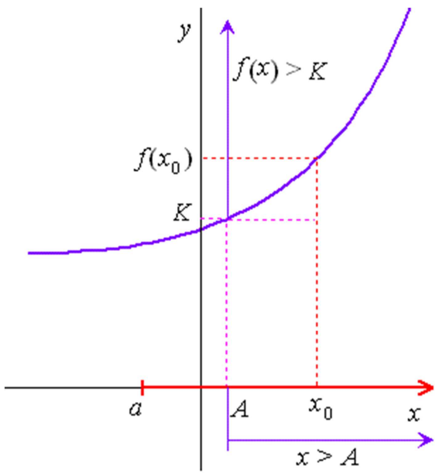


Fig. 3.46. Improper limit ∞ of f at ∞

Improper limit of function at improper point

Let function f be defined on interval $(-\infty, a)$ or (a, ∞). The function f has improper limit ∞ in the improper point $-\infty$ or ∞, if for any $K > 0$ such number $A > 0$ exists that for all $x < -A$ or $x > A$ it holds that $f(x) > K$.

$$\lim_{x\to-\infty} f(x) = \infty, \text{ or } \lim_{x\to\infty} f(x) = \infty$$

Similarly, the improper limit $-\infty$ at improper point $-\infty$ or ∞ can be defined.

Properties of improper limit of a function

P8. If $\lim\limits_{x\to a} f(x) = b \neq 0, \lim\limits_{x\to a} g(x) = 0$ and for all $x \neq a$ from some neighbourhood $O(a)$ it holds that

$$\frac{f(x)}{g(x)} < 0, \text{ then } \lim_{x\to a}\frac{f(x)}{g(x)} = -\infty,$$

$$\frac{f(x)}{g(x)} > 0, \text{ then } \lim_{x\to a}\frac{f(x)}{g(x)} = \infty.$$

P9. Let function $f(x)$ be bounded on some neighbourhood of point a and let $\lim\limits_{x\to a} g(x) = \infty$, then

$$\lim_{x\to a}\left[f(x) + g(x)\right] = \infty, \quad \lim_{x\to a}\frac{f(x)}{g(x)} = 0.$$

P10. If $\lim\limits_{x\to a} f(x) = b \neq 0$ and such number $p > 0$ exists that for all $x \neq a$ from some neighbourhood of point a is $g(x) > p$, then $\lim\limits_{x\to a} f(x)\cdot g(x) = \infty$.

Examples

1. $\lim\limits_{x\to 2}\dfrac{1+x}{(x-2)^2} = \infty$, because $\lim\limits_{x\to 2}(1+x) = 3, \lim\limits_{x\to 2}(x-2)^2 = 0 \Rightarrow \lim\limits_{x\to 2}\dfrac{1+x}{(x-2)^2} = \infty$

2. $$\lim_{x\to\infty}\frac{2x^3 + x^2 - 1}{5x^3 - x} = \lim_{x\to\infty}\frac{2 + \frac{1}{x} - \frac{1}{x^3}}{5 - \frac{1}{x^2}} = \frac{2+0-0}{5-0} = \frac{2}{5}$$

3. $\lim\limits_{x\to\infty}(x + \sin x) = \infty$, because $|\sin x| \le 1$, and $\lim\limits_{x\to\infty} x = \infty$

4. $\lim\limits_{x\to\infty}\dfrac{\sin x}{x} = 0$, because $\lim\limits_{x\to\infty} x = \infty$ and $\sin x$ is the bounded function.

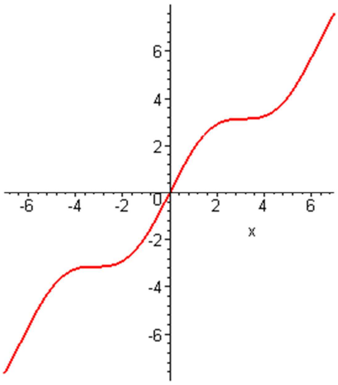


Fig. 3.47. Graph of function $x + \sin x$

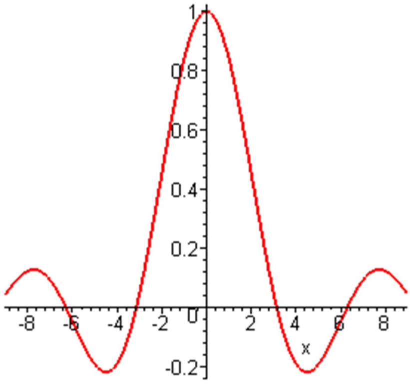


Fig. 3.48. Graph of function $\dfrac{\sin x}{x}$

5. $\lim\limits_{x\to\infty}\dfrac{1}{x} = \lim\limits_{x\to-\infty}\dfrac{1}{x} = 0 \Rightarrow \lim\limits_{x\to\infty}\dfrac{1}{x^k} = \left(\lim\limits_{x\to\infty}\dfrac{1}{x}\right)^k = 0^k = 0,$

$$\lim_{x\to-\infty}\frac{1}{x^k} = \left(\lim_{x\to-\infty}\frac{1}{x}\right)^k = 0^k = 0$$

One sided limits

If we replace neighbourhood $O_\delta(a)$ of point a in the definitions of limits by the left $O^-_\delta(a)$ or the right $O^+_\delta(a)$ neighbourhood, we obtain definitions of one-sided limits.

Limit on the right of $f(x)$ at a

$$\lim_{x\to a^+} f(x) = b \Leftrightarrow \forall\varepsilon > 0, \exists\delta > 0 : x \in O^+_\delta(a) \Rightarrow f(x) \in O_\varepsilon(b)$$

Limit on the left of $f(x)$ at a

$$\lim_{x\to a^-} f(x) = b \Leftrightarrow \forall\varepsilon > 0, \exists\delta > 0 : x \in O^-_\delta(a) \Rightarrow f(x) \in O_\varepsilon(b)$$

If $b = \infty$ or $-\infty$, function f has improper limit on the right or on the left at the point a.

Theorem. The limit of function f at the point a exists if and only if both one-sided limits exist at the point a and they are equal

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

If some of the one-sided function limits does not exist at the point a, then the limit of function f does not exist at the point a.

All properties P1 – P10 are equally valid for one-sided limits.

Examples

1. $\lim\limits_{x\to -2^+} \sqrt{x+2} = 0$

2. $\lim\limits_{x\to 0^+} \frac{1}{x} = \infty$, $\lim\limits_{x\to 0^-} \frac{1}{x} = -\infty$, therefore $\lim\limits_{x\to 0} \frac{1}{x}$ does not exist.

3. $\lim\limits_{x\to 0^-} \frac{|x|}{x} = \lim\limits_{x\to 0^-} \frac{-x}{x} = -1$, $\lim\limits_{x\to 0^+} \frac{|x|}{x} = \lim\limits_{x\to 0^+} \frac{x}{x} = 1$, therefore $\lim\limits_{x\to 0} \frac{|x|}{x}$ does not exist.

4. $\lim\limits_{x\to \infty} \frac{1}{x} = \lim\limits_{x\to -\infty} \frac{1}{x} = 0$

Continuity of function

Function $f(x)$ is said to be continuous at a point a if $\lim\limits_{x\to a} f(x) = f(a)$, which means:

1. $f(x)$ is defined at a ($a \in D(f)$),
2. there exists $\lim\limits_{x\to a} f(x)$,
3. this limit is equal to $f(a)$.

It is said that function $f(x)$ is continuous at a on the right (on the left) if

$$\lim_{x\to a^+} f(x) = f(a) \text{ (or } \lim_{x\to a^-} f(x) = f(a)\text{) }.$$

A function $f(x)$ is said to be continuous on an interval $\langle a, b\rangle$ if it is continuous at each $x \in \langle a, b\rangle$ and moreover, if it is continuous at a on the right and at b on the left.
If functions $f(x)$ and $g(x)$ are continuous at a point a, then the following functions are also continuous at this point:

1. $f(x) \pm g(x)$
2. $c \cdot f(x), c \in \boldsymbol{R}$
3. $f(x) \cdot g(x)$
4. $\frac{f(x)}{g(x)}$, if $g(a) \neq 0$
5. $[f(x)]^k, k \in \boldsymbol{N}$

A function continuous at each point of its domain of definition is said to be continuous. The graph of a continuous function is a non-interrupted curve. All elementary functions are continuous at each point of their domains of definition.

The points at which function f is not continuous are called the points of discontinuity of f. Point a is a point of discontinuity of function f, if the function:

a) has no limit at a

b) is not defined at a

c) has a limit at a that is not equal to the function value $f(a)$.

Properties of functions continuous on a closed interval

T1. Function continuous on a closed interval is bounded on this interval.

T2. Function continuous on a closed interval $\langle a, b\rangle$ acquires the greatest value (maximum) and the least value (minimum) in this interval, i.e. such points $c_1, c_2 \in \langle a, b\rangle$ exist that $f(c_1) \leq f(x) \leq f(c_2), \forall x \in \langle a,b \rangle$.

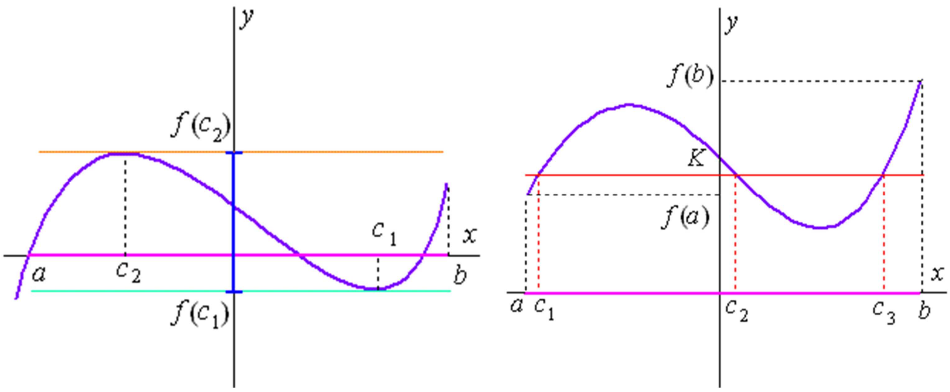


Fig. 3.49. Properties T2 and T3 of functions continuous on a closed interval

T3. If function f is continuous on a closed interval $\langle a, b\rangle$ and $f(a) \neq f(b)$, then for arbitrary K such that $f(a) \leq K \leq f(b)$ at least one $c \in (a, b)$ exists at which $f(c) = K$.

T4. If function f is continuous on a closed interval $\langle a, b\rangle$ and $f(a) \cdot f(b) < 0$, then such $c \in \langle a, b\rangle$ exists that $f(c) = 0$.

T5. If function f is continuous on a closed interval $\langle a, b\rangle$, then the image of this interval, set $f(\langle a,b \rangle) = \{f(x) : x \in \langle a,b \rangle\}$, is again a closed interval, or a one point set (in the case of a constant function f).

T6. If function f is increasing (decreasing) and continuous on a closed interval $J \subset \boldsymbol{R}$, then its inverse function f^{-1} is also increasing (decreasing) and continuous on the range of function f, $R(f) \subset \boldsymbol{R}$.

T7. If function $g(x)$ is continuous at the point x_0 and function $f(u)$ is continuous at the point $u_0 = g(x_0)$, then the composite function $f(g(x))$ is also continuous at the point x_0.

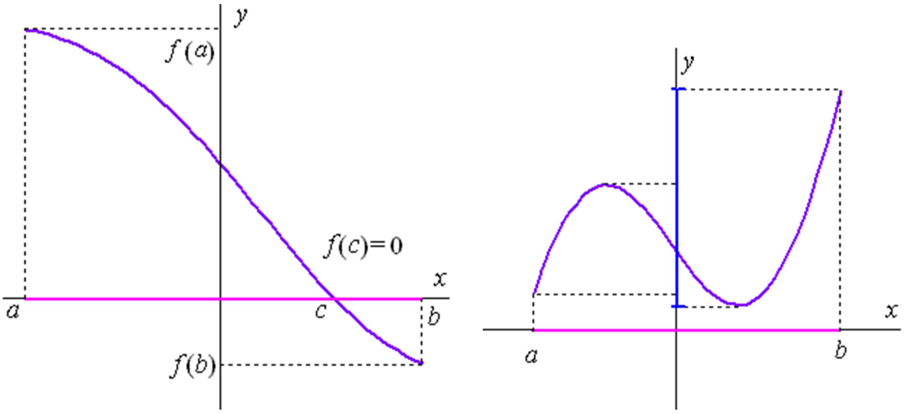


Fig. 3.50. Properties T4 and T5 of functions continuous on a closed interval

Asymptotes to graph of function

The behaviour of functions on neighbourhoods of points which are not in their domains of definition can be analysed by means of lines known as asymptotes to graphs of functions.

Asymptotes are lines coming close to the graph of function. A straight line is an asymptote to the graph of a function, if the distance from the variable point *M* of the graph to this line approaches zero, as the point *M* tends to infinity (asymptotes are tangents at infinity).

Two forms of asymptotes must be distinguished: vertical asymptotes (without the slope) and inclined asymptotes (with the slope).

Asymptotes that are not perpendicular to coordinate axis *x* are called asymptotes with a slope, asymptotes without slope are perpendicular to coordinate axis *x*.

Line $y = kx + b$ is called asymptote with slope (inclined asymptote) to the graph of function *f*, if

$$\lim_{x\to\infty}\left[f(x)-(kx+b)\right]=0, \text{ or } \quad \lim_{x\to-\infty}\left[f(x)-(kx+b)\right]=0 .$$

Line $y = kx + b$ is asymptote with slope (inclined asymptote) to the graph of function f as x approaches infinity, $x \to \infty$, or as x approaches minus infinity, $x \to -\infty$, if and only if

$$\lim_{x\to\infty}\frac{f(x)}{x}=k, \quad \lim_{x\to\infty}\left(f(x)-kx\right)=b \text{ , or}$$

$$\lim_{x\to-\infty}\frac{f(x)}{x}=k, \quad \lim_{x\to-\infty}\left(f(x)-kx\right)=b .$$

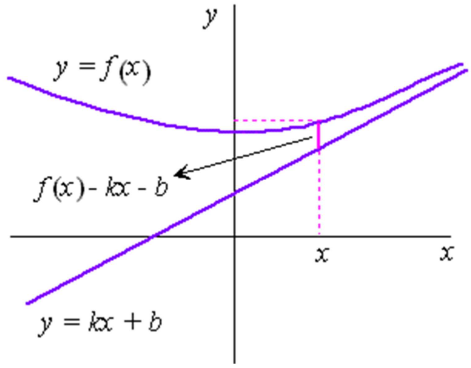


Fig. 3.51. Asymptote to graph of function

In particular, if the function f tends to a finite limit as x approaches infinity, that is

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

then obviously $k = 0$ and $G(f)$ has a horizontal asymptote (regarded as a special case of the inclined asymptote) parallel to the coordinate axis x, namely $y = b$. Similar relations hold for $x \to -\infty$.

The asymptotic behaviour of a function may be of a different character when x becomes positively or negatively infinite, and therefore the cases $x \to +\infty$ and $x \to -\infty$ should be treated separately. If for $x \to +\infty$ and $x \to -\infty$ numbers k and b coincide, then both asymptotes form a common straight line.

Line $x = a$ is called asymptote without slope to the graph of function f, if at least one of above relations is true

$$\lim_{x\to a^+} f(x) = \infty, \quad \lim_{x\to a^+} f(x) = -\infty, \quad \lim_{x\to a^-} f(x) = \infty, \quad \lim_{x\to a^-} f(x) = -\infty .$$

Examples

1. Line $x = 0$ is asymptote without slope and line $y = 0$ is horizontal asymptote to the graph of function $y = \frac{1}{x}$, as $\lim_{x\to 0^+} \frac{1}{x} = \infty$ and also $\lim_{x\to 0^-} \frac{1}{x} = -\infty$, and

$$k = \lim_{x\to\infty} \frac{f(x)}{x} = \lim_{x\to\infty} \frac{1}{x^2} = \lim_{x\to-\infty} \frac{1}{x^2} = 0,$$

$$q = \lim_{x\to\infty} \big(f(x) - k.x\big) = \lim_{x\to\infty} f(x) = \lim_{x\to-\infty} f(x) = 0 .$$

2. Lines $y = x$ and $y = -x$ are inclined asymptotes to the graph of function $f(x) = \sqrt{x^2 - 9}$. The function is defined on set $M = (-\infty, -3\rangle \cup \langle 3, \infty)$ in real numbers, and the following holds

$$k=\lim_{x\to\infty}\frac{f(x)}{x}=\lim_{x\to\infty}\frac{\sqrt{x^2-9}}{x}=\lim_{x\to\infty}\sqrt{\frac{x^2-9}{x^2}}=\sqrt{\lim_{x\to\infty}\left(1-\frac{9}{x^2}\right)}=1,$$

$$q=\lim_{x\to\infty}\left(f(x)-kx\right)=\lim_{x\to\infty}\left(\sqrt{x^2-9}-x\right)=\lim_{x\to\infty}\frac{\left(\sqrt{x^2-9}-x\right)\left(\sqrt{x^2-9}+x\right)}{\left(\sqrt{x^2-9}+x\right)}=$$

$$=\lim_{x\to\infty}\frac{x^2-9-x^2}{\sqrt{x^2-9}+x}=\lim_{x\to\infty}\frac{-9}{\sqrt{x^2-9}+x}=0$$

and similarly holds $k=\lim\limits_{x\to-\infty}\frac{f(x)}{x}=-1,\ q=\lim\limits_{x\to-\infty}\left(f(x)-kx\right)=0$.

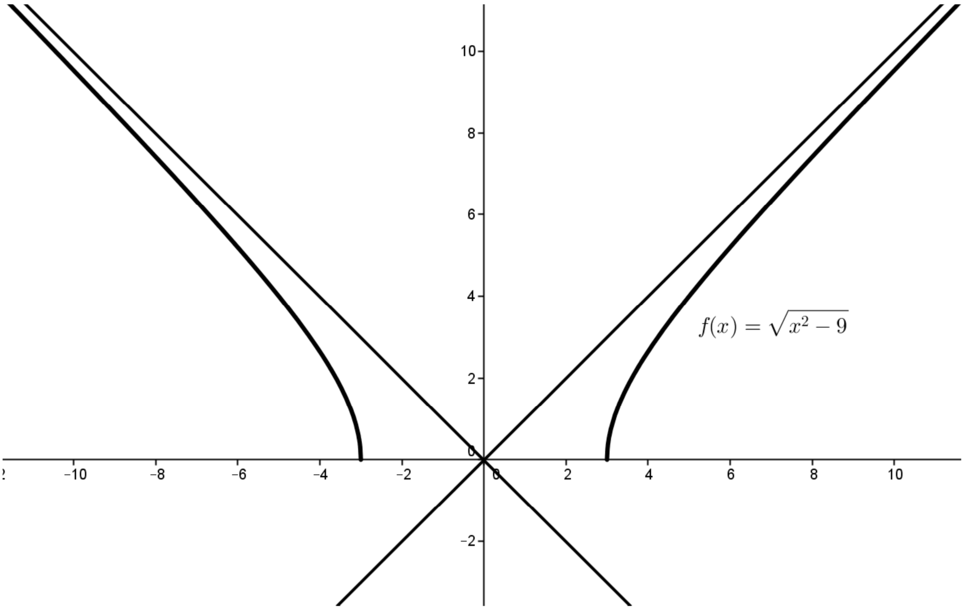


Fig. 3.52. Graph of function $f(x)=\sqrt{x^2-9}$ with asymptotes

3.7 Derivative of function

Let f be a function defined on domain $D(f)$, with a graph in a certain curve. Choosing arbitrary numbers x, $x_0 \in D(f)$, the difference $\Delta x = x - x_0$ will be called increment of the argument, therefore $x = x_0 + \Delta x$. The difference of function values at points x, x_0 denoted $\Delta f = f(x) - f(x_0)$ is called the increment of function f value at the point x_0, corresponding to the increment Δx of the argument, or the function f difference at the point x_0, while

$$f(x) = f(x_0 + \Delta x) = f(x_0) + \Delta f\,.$$

Quotient

$$\frac{\Delta f}{\Delta x} = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

is called the differential quotient of function f at the point x_0, or the relative increment of the function f value at the point x_0.
The differential quotient of function f characterises relative change of the function f values with respect to the change of its argument.

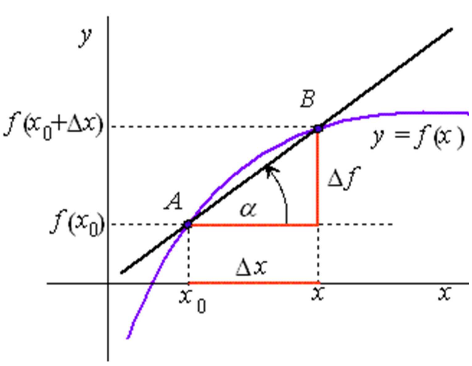


Fig. 3.53. Differential quotient of f

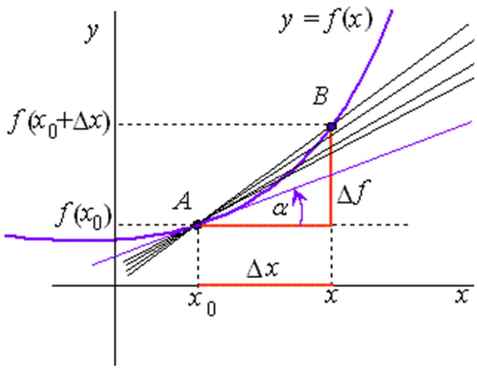


Fig. 3.54. Geometric interpretation of derivative

The differential quotient of function f at the point x_0 can be geometrically interpreted as a slope of line passing through the points $A = [x_0, f(x_0)]$, $B = [x_0 + \Delta x, f(x_0 + \Delta x)]$ on the function f graph, which intersects curve $G(f)$ in these points

$$\tan\alpha = \frac{\Delta f}{\Delta x}\,.$$

Let function f be defined at the point x_0 and on some neighbourhood of this point. If limit (the proper limit)

$$\lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

exists, then this limit is said to be the derivative of function f at the point x_0 and it is denoted $f'(x_0)$

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}, \Delta x = x - x_0 .$$

If real number $f'(x_0)$ exists, function f is called differentiable at x_0.

The derivative of a function at the point x_0 therefore determines the slope of the curve that is the graph of function f, so it is the slope of a line, which is the limit position of the line intersecting the graph and passing through points A and B − a tangent to the function graph at the point $[x_0, f(x_0)]$.

To denote derivative of function $y = f(x)$ at the point x_0 various symbols are used

$$f'(x_0) = \frac{df}{dx}(x_0) = \left[\frac{df}{dx}\right]_{x=x_0} = y'(x_0) = \left[\frac{dy}{dx}\right]_{x=x_0} .$$

Examples

1. Derivative of function $f(x) = x^2$ at the point x_0 can be determined as

$$f'(x_0) = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} =$$

$$= \lim_{\Delta x \to 0} \frac{(x_0 + \Delta x)^2 - x_0^2}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta x(2x_0 - \Delta x)}{\Delta x} = 2x_0$$

2. Function $f(x) = |x|$ has no derivative at the point $x_0 = 0$, as

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to 0} \frac{f(x) - f(0)}{x} = \lim_{x \to 0} \frac{|x|}{x}$$

   does not exist.

If the derivative of function f exists at all points of some set, then the function is said to have derivative on this set and it is called differentiable on this set. The set of points, on which derivative of function f exists, is a number set $M = \{x \in D(f) : \exists\, f'(x)\}$. Number $f'(x_0)$ can be attached to any number $x_0 \in M$, thus a new function is defined on the set M called the derivative of function f that is denoted f', or $f'(x_0)$ and the following holds

$$f' : M \to \boldsymbol{R}, x \to y = f'(x).$$

Necessary condition of differentiability: If a function $f(x)$ is differentiable at a point x_0, it is continuous at this point:

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} (x - x_0) + f(x_{0)} \right) = f(x_0) .$$

Continuity is not a sufficient condition of differentiability.

Improper derivative

If $\lim\limits_{x \to x_0} \dfrac{f(x) - f(x_0)}{x - x_0} = +\infty$, or $-\infty$, it is said that f has an improper derivative at the point x_0 (but function f is not differentiable at the point x_0).

One-sided derivatives

If limit $\lim\limits_{x \to x_0^+} \dfrac{f(x) - f(x_0)}{x - x_0}$, or $\lim\limits_{x \to x_0^-} \dfrac{f(x) - f(x_0)}{x - x_0}$ exists, it is called the derivative of function f at the point x_0 on the right, or on the left, and denoted by symbol $f'_+(x_0)$, or $f'_-(x_0)$.

If these limits are improper, we speak about improper derivative of function f at the point x_0 on the right, or on the left.

Function f is said to have derivative on a closed interval $\langle a, b \rangle$, if it has derivative at all points $x \in (a, b)$, and a derivative on the left at the point a and a derivative on the right at the point b.

Geometric meaning of derivative

If the value of the derivative of a function $f(x)$ at a point x_0 is $f'(x_0)$, then the straight line $y = f(x_0) = f'(x_0)(x - x_0)$ is the tangent to the graph $G(f)$ of function f at the point $[x_0, f(x_0)]$. Hence $f'(x_0)$ is the slope of the tangent to $G(f)$ at the point $[x_0, f(x_0)]$.

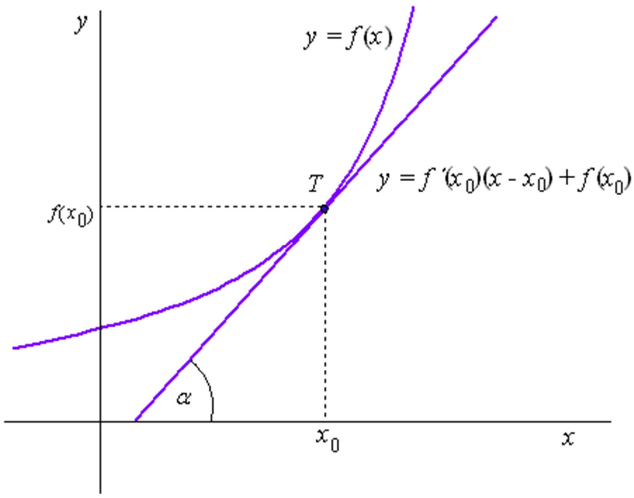


Fig. 3.55. Graph of function f and its tangent

Physical meaning of derivative

If a point moves along a straight line and its law of motion is $s = f(t)$, where variable t means time, then the ratio $\dfrac{f(t) - f(t_0)}{t - t_0}$ is an average velocity of the motion, corresponding to the time interval $\Delta t = t - t_0$. Then

$$v(t_0) = f'(t_0) = \lim_{t \to t_0} \frac{f(t) - f(t_0)}{t - t_0}$$

is called the velocity of the rectilinear motion $s = f(t)$, at the given moment $t = t_0$.

Examples

1. Function $f(x) = x^2$ has a derivative at all points of its domain of definition, $x \in (-\infty, \infty)$, and $f'(x) = 2x$. The graph of function f therefore has a tangent line at any point $\left[x_0, x_0^2\right]$, and it is determined by the equation $y - x_0^2 = 2x_0(x - x_0)$. The tangent at the point $[-1, 1]$ has the equation $2x + y + 1 = 0$, fig. 3.56.

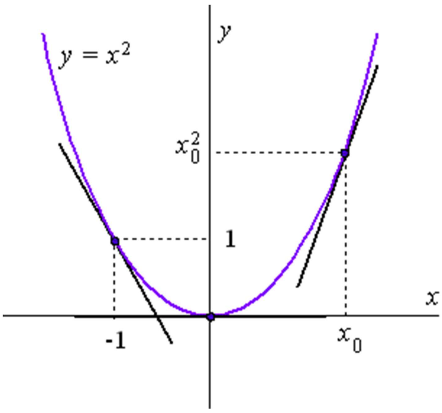


Fig. 3.56. Tangent to the graph of function

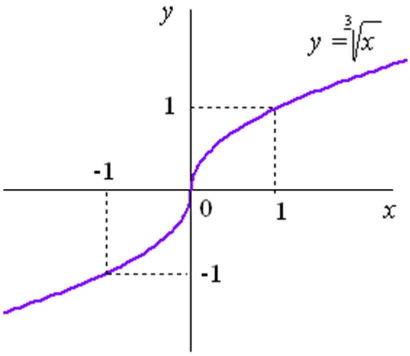


Fig. 3.57. Graph of function

2. Function $f(x) = \sqrt[3]{x}$ has an improper derivative at the point $x_0 = 0$, as

$$f'(0) = \lim_{x \to 0} \frac{\sqrt[3]{x}}{x} = \lim_{x \to 0} \sqrt[3]{\frac{x}{x^3}} = \lim_{x \to 0} \frac{1}{\sqrt[3]{x^2}} = \infty .$$

   Because it is continuous at the point 0, at the point $[0, 0]$ its graph has a tangent line with the equation $x = 0$, fig. 3.57.

3. Graph of function $f(x) = |x|$ has no tangent line at the point $[0, 0]$, as no derivative defined at the point 0, nor improper derivative at this point exist, see fig. 3.1, right.

4. Function $f(x)=\sqrt{1-x^2}$ defined for all real x such that $|x|\le 1$ has a derivative at all interior points of its domain, interval $(-1,\,1)$, and it has improper one-sided derivatives in the points 1 and -1, as the following holds

$$f'_{-}(-1)=\lim_{x\to -1^-}\frac{\sqrt{1-x^2}}{x+1}=\infty,\qquad f'_{+}(1)=\lim_{x\to 1^+}\frac{\sqrt{1-x^2}}{x-1}=-\infty\,,$$

and the tangent line equations at these points are $x=-1$, $x=1$, see fig. 3.58.

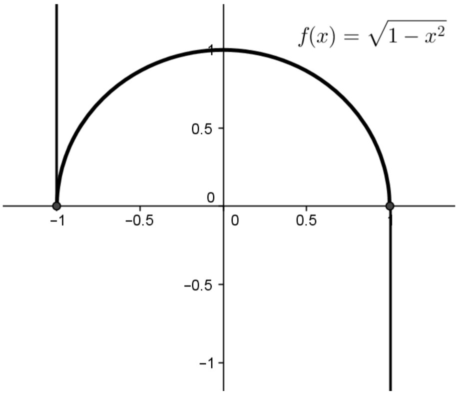


Fig. 3.58. Graph of function

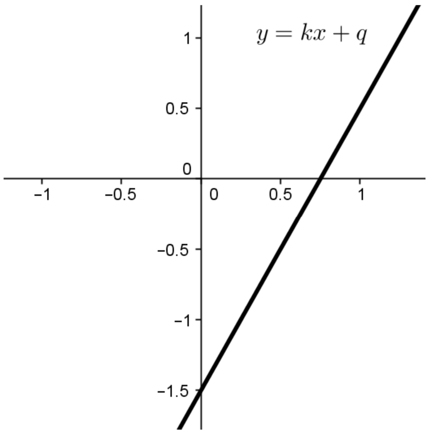


Fig. 3.59. Graph of function

5. Function $f(x) = kx + q$, for arbitrary constants k, $q\in$ $\boldsymbol{R}$ has a constant derivative equal to k at all points of its domain of definition $\boldsymbol{R}$, while the graph of function coincides with the tangent line $y = kx + q$, fig. 3.59.
6. Function has no derivative at any point of its domain of definition $\boldsymbol{R}$.

Basic properties of differentiation

If function $f(x)$ and $g(x)$ are differentiable on the set M, then also functions $c\cdot f(x)$, $f(x)\pm g(x)$, $f(x)\cdot g(x)$, $\dfrac{f(x)}{g(x)}$, if $g(x)\ne 0\ \ \forall x\in M$, are differentiable on M, while

$$\left[c\cdot f(c)\right]'=c\cdot f'(x)$$

$$\left[f(x)\pm g(x)\right]'=f'(x)\cdot g'(x)$$

$$\left[f(x)\cdot g(x)\right]'=f'(x)\cdot g(x)+f(x)\cdot g'(x)$$

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

Chain rule

If function $g(x)$ has a derivative at the point x_0 and function $f(x)$ has a derivative at the point $u_0 = g(x_0)$, then the composite function $F(x) = f(g(x))$ also has a derivative at the point x_0 and it holds that

$$F'(x_0) = f'(u_0)g'(x_0), u_0 = g(x_0),$$

which can also be written for all suitable x as

$$\frac{dF(x)}{dx} = \frac{df(g(x))}{dx} = \frac{df(u)}{du} \cdot \frac{dg(x)}{dx} = g(x_0), \text{ for } u = g(x),$$

and denoting $y = f(u)$ we receive $\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$.

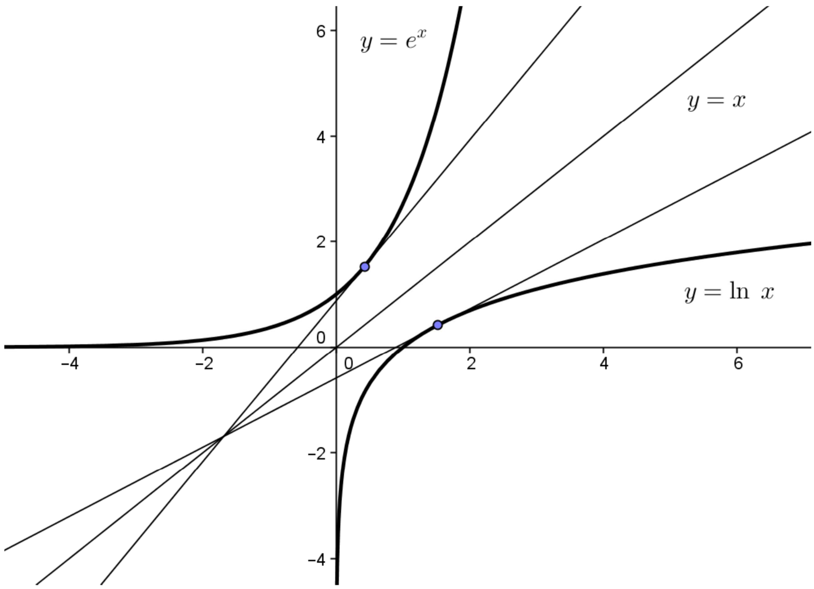


Fig. 3.60. Inverse functions

Derivative of inverse function

Let function f be one-to-one and continuous on interval (a, b), and let there exist a non-zero derivative f' of this function on the interval. Then the inverse function f^{-1} also has a derivative on its domain of definition (that is the range of function f), and for all $x \in (a, b)$ it holds that

$$\left[f^{-1}(x) \right]' = \frac{1}{f'(f^{-1}(x))}.$$

Logarithmic differentiation

If function f has a derivative f' on interval (a, b) and $f(x) > 0$ on this interval, then for all $x \in (a, b)$ it holds that

$$[\ln f(x)]' = \frac{f'(x)}{f(x)}.$$

The derivatives of elementary functions can be derived from the definition of derivative and from the rules on differentiation. Some of these are presented in the following.

Derivatives of basic elementary function

1. Power function

$$\left[x^n\right]' = nx^{n-1}, n \in \boldsymbol{N}, x \in \boldsymbol{R},$$

and generally $\left[\boldsymbol{x^a}\right]' = ax^{a-1}, a \in \boldsymbol{R}, x \in (0, \infty)$

2. Rational function

$$\left[\frac{P(x)}{Q(x)}\right]' = \frac{P'(x)\cdot Q(x) - P(x)\cdot Q'(x)}{Q^2(x)}, \forall x \in \boldsymbol{R}: Q(x) \neq 0$$

3. Exponential function

$$\left[a^x\right]' = a^x \ln a, a \in \boldsymbol{R}, a > 0, a \neq 1, x \in \boldsymbol{R},$$

while particularly $\left[e^x\right]' = e^x$

4. Logarithmic function

$$[\log_a x]' = \frac{1}{x \ln a}, a \in \boldsymbol{R}, a > 0, a \neq 1, x \in (0, \infty),$$

while particularly $[\ln x]' = \dfrac{1}{x}$

5. Goniometric functions

$$[\sin x]' = \cos x, x \in \boldsymbol{R}$$

$$[\cos x]' = \sin x, x \in \boldsymbol{R}$$

$$[\tan x]' = \frac{1}{\cos^2 x}, x \neq (2k-1)\frac{\pi}{2}, k \in \boldsymbol{Z}$$

$$[\cot x]' = -\frac{1}{\sin^2 x}, x \neq k\pi, k \in \boldsymbol{Z}$$

6. Cyclometric functions

$$[\arcsin x]' = \frac{1}{\sqrt{1-x^2}}, x \in (-1,1)$$

$$[\arccos x]' = -\frac{1}{\sqrt{1-x^2}}, x \in (-1,1)$$

$$[\arctan x]' = \frac{1}{1+x^2}, x \in \boldsymbol{R}$$

$$[\operatorname{arc}\cot x]' = -\frac{1}{1+x^2}, x \in \boldsymbol{R}$$

Examples

1. Derivative of function $y = \sinh x$ is

$$y' = \left[\frac{e^x - e^{-x}}{2}\right]' = \frac{e^x - (-1)e^{-x}}{2} = \frac{e^x + e^{-x}}{2} = \cosh x,$$

and in a similar way it can be shown that

$$[\cosh x]' = \sinh x, [\tanh x]' = \frac{1}{\cosh^2 x}, [\coth x]' = \frac{1}{\sinh^2 x}.$$

2. Derivative of function $y = x^5 \cos x$ is $y' = 5x^4 \cos x - x^5 \sin x$.

3. Derivative of function $y = \sqrt[3]{2x^2 - 5}$ is function $y' = \dfrac{4x}{3\sqrt[3]{\left(2x^2-5\right)^2}}$.

4. Derivative of composite function $y = e^{\sin x^3}$ can be calculated using denotation

$$y = e^u, u = \sin v, v = x^3 .$$

Then we can write

$$y' = \frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dv} \cdot \frac{dv}{dx}, \text{and } y' = e^{\sin x^3} \cdot \cos x^3 \cdot 3x^2 .$$

5. For function $y = \ln\left(x + \sqrt{x^2 - a^2}\right)$ it holds that

$$y' = \frac{1}{x+\sqrt{x^2-a^2}}\left(1 + \frac{2x}{2\sqrt{x^2-a^2}}\right) =$$

$$= \frac{1}{x+\sqrt{x^2-a^2}} \cdot \frac{x+\sqrt{x^2-a^2}}{\sqrt{x^2-a^2}} = \frac{1}{\sqrt{x^2-a^2}}, x^2 \neq a^2.$$

6. Looking for the derivative of function $y = x^x$ we can write $y = e^{\ln x^x} = e^{x\ln x}$, hence denoting $y = e^u$, $u = x\ln x$ we receive

$$\frac{dy}{dx} = \frac{dy}{du}\cdot\frac{du}{dx} = e^u\left(1\cdot\ln x + x\frac{1}{x}\right) = x^x\left(1+\ln x\right).$$

7. Another possibility is to use the logarithmic differentiation and write

$$\ln y = \ln x^x = x\ln x\text{, while}$$

$$[\ln y]' = [x\ln x]' = \ln x + x\frac{1}{x} = 1 + \ln x.$$

Then from $[\ln y]' = \dfrac{y'}{y} = 1 + \ln x$ it follows $y' = y(1+\ln x) = x^x(1+\ln x)$.

3.8 Basic theorems of calculus

Fermat theorem

If function f attains minimal or maximal value at the point ξ and a derivative of function exists at this point, then $f'(\xi) = 0$.

The geometric interpretation of this theorem is very simple:

if function f is differentiable at the point ξ, in which it attains maximal or minimal value, then the tangent line to the graph of function f at the point $T = [\xi, f(\xi)]$ is parallel to the coordinate axis x.

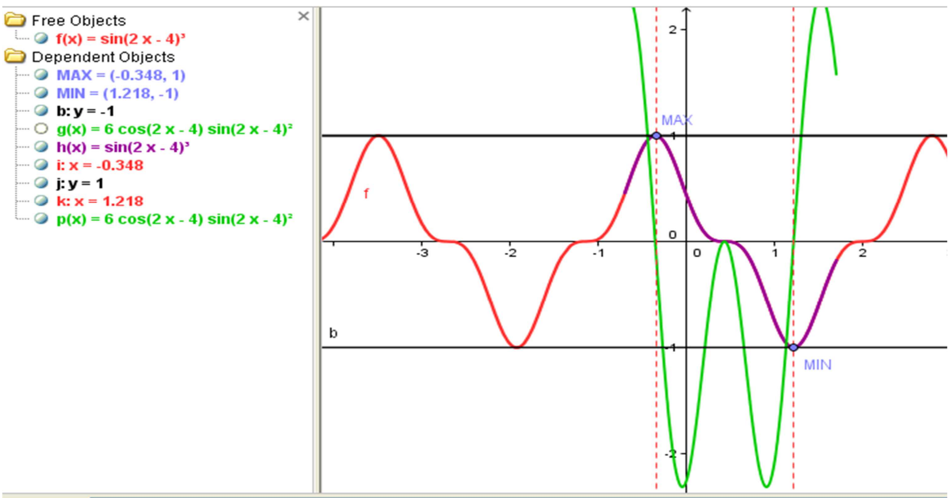


Fig. 3.61. Fermat theorem

Rolle theorem
Let the following properties hold for function f

1. it is continuous on closed interval $\langle a, b\rangle$
2. it is differentiable at each point of the open interval (a, b)
3. $f(a) = f(b)$.

Then in the interval (a, b) at least one point ξ exists such, that $f'(\xi) = 0$.

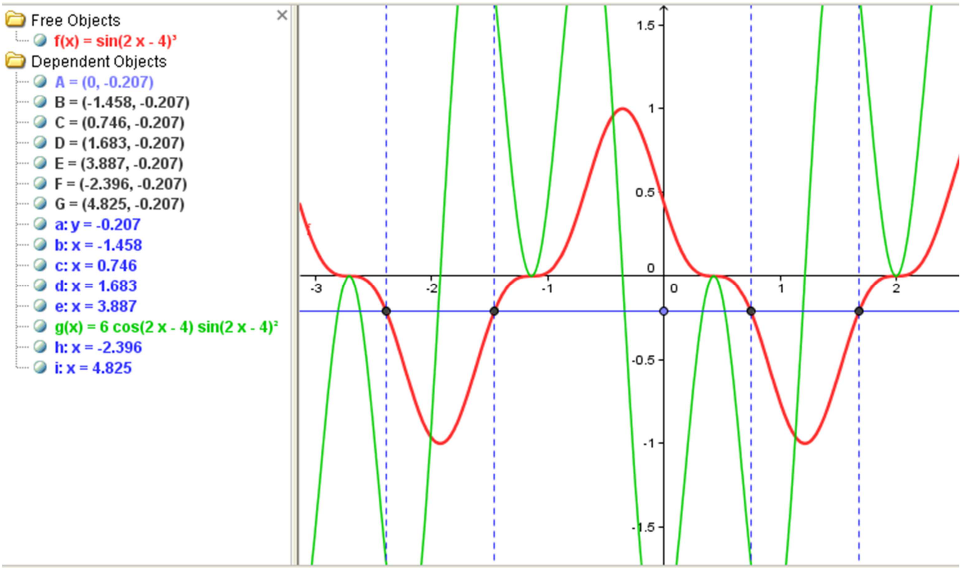


Fig. 3.62. Rolle theorem

Lagrange theorem (on function increment)
Let the following properties be true for function f

1. it is continuous on closed interval $\langle a, b\rangle$
2. it is differentiable at each point from open interval (a, b).

Then at least one point ξ exists in the interval (a, b) such, that

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

The theorem can be interpreted geometrically as follows:
Graph $G(f)$ of such function f that satisfies all of the above properties, has a tangent line at all points but the end points $A = [a, f(a)]$, $B = [b, f(b)]$. There exists at least one point $T = [\xi, f(\xi)]$ on the graph $G(f)$ such that the tangent to the graph at this point is parallel to the line segment AB, and its slope $f'(\xi)$ equals to the slope of this line segment, i.e. $\frac{f(b)-f(a)}{b-a}$.

If $f'(x) = 0$ for all $x \in (a, b)$, then the function $f(x)$ is constant on entire interval (a, b), therefore $f(x) = c$, $\forall x \in (a, b)$.
If $f'(x) - g'(x) = 0$ for all $x \in (a, b)$, then the function $f(x) - g(x)$ is constant on (a, b), therefore

$$f(x) = g(x) + c, \ \forall x \in (a, b).$$

Physical interpretation of Lagrange theorem:
Let function $s = f(t)$ represent the trajectory of a point moving on a straight line, and its derivative $f'(t)$ determine the velocity of a point in time t. The average velocity of this rectilinear motion in the time interval $\langle t_1, t_2\rangle$ is determined by the quotient

$$\frac{f(t_2) - f(t_1)}{t_2 - t_1}$$

and such moment ξ exists in this time interval at which instantaneous velocity equals to the average (mean) velocity,

$$f'(\xi) = \frac{f(t_2) - f(t_1)}{t_2 - t_1}$$

Lagrange theorem is therefore sometimes denoted the theorem about mean value.

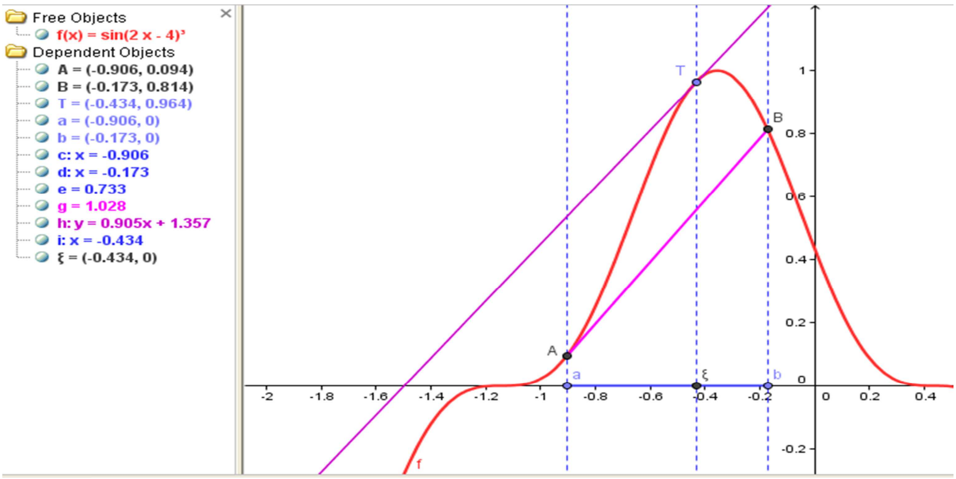


Fig. 3.63. Lagrange theorem

3.9 Higher order derivatives, differential, Taylor polynomial

Let function $f(x)$ be differentiable on a set M. If its derivative $f'(x)$ has a derivative at each point $x \in M$, then this derivative is called the second derivative of function $f(x)$ on M and it is denoted $f''(x) = \dfrac{d^2 f}{dx^2}$.

The third derivative and derivatives of higher orders can be defined analogously.
If for all points $x \in M$ the function $f^{(n-1)}(x)$ (the derivative of $(n-1)$-th order) is differentiable, then its derivative is called the n-th derivative, or derivative of the order n of the function f, it means that $f^{(n)}(x) = [f^{(n-1)}(x)]'$, for $n = 2, 3, 4, \dots$.
Another notation of the n-th derivative is $f^{(n)}(x) = \dfrac{df^{n}(x)}{dx^{n}}$.

Examples

1. Function $f(x) = \arctan\dfrac{1}{x}$ has the following first three derivatives:

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

2. The tenth derivative of function $f(x) = e^{5x}$ is $f^{(10)}(x) = 5^{10}e^{5x}$.
3. The n-th derivative of function $f(x) = \dfrac{1}{x}, x \neq 0$ is $f^{(n)}(x) = (-1)^n \dfrac{n!}{x^{n+1}}$,

   therefore $n = 1,\ f'(x) = -\dfrac{1}{x^2}$, $\quad n = 2,\ f''(x) = \dfrac{2}{x^3}$, $\quad n = 3,\ f'''(x) = -\dfrac{6}{x^4}$,

   and so on.

Differential and its meaning

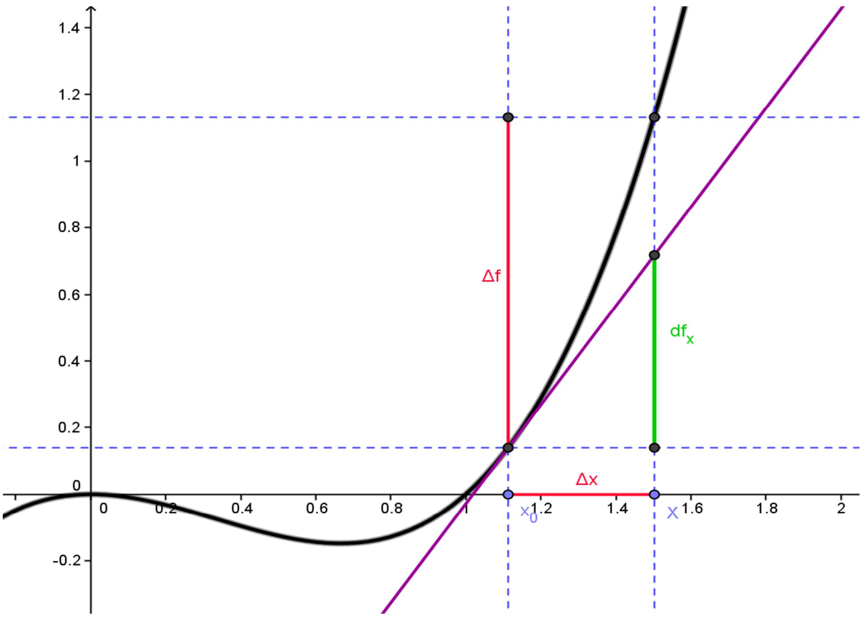


Fig. 3.64. Geometric interpretation of differential

Suppose function $f(x)$ is defined on a neighbourhood $N_\varepsilon(x_0)$ and differentiable at point x_0. The expression $f'(x_0)\cdot(x-x_0)$ is called the differential of f at the point x_0 and it is denoted

$$df_{x_o} = f'(x_0)\cdot(x-x_0).$$

The differential df_{x_o} of a function $f(x)$ at a point x_0 is equal to the increment of the y-coordinate of points on the tangent line to the function graph $G(f)$ at the point $[x_0, f(x_0)]$. If difference $x - x_0$ approaches 0, then Δf equals approximately df, thus

$$f(x)-f(x_0)=f'(x_0)\cdot(x-x_0).$$

For the function f: $y = x$, we have $df = dx = \Delta x$, that is why the differential at an arbitrary point is denoted $df = f'(x)dx$.

Examples

1. Differential of function $f(x) = \arctan(x)$ at the point $x_0 = 2$, is $df_2(x)=\dfrac{x-2}{5}$, and its value at the point $x = -3$ equals -1.
2. Approximate value of $\sin\alpha$, $\alpha = 60^\circ 18'$ can be calculated using differential of function $\sin x$, $x \in \boldsymbol{R}$.

$$60^\circ = \frac{\pi}{3} = x_0, 18' = \frac{\pi}{200} = \Delta x,\ x = x_0 + \Delta x = \frac{203\pi}{600}$$
$$f(x) = \sin x,\ f'(x) = \cos x,\ f(x_0+\Delta x) = f(x_0) + f'(x_0)\cdot\Delta x$$
$$\sin\frac{203\pi}{600} = \sin\left(\frac{\pi}{3}+\frac{\pi}{200}\right) = \sin\frac{\pi}{3} + \cos\frac{\pi}{3}\cdot\frac{\pi}{200} =$$
$$= \frac{\sqrt{3}}{2} + \frac{\pi}{400} = 0,8660254 + 0,0078539 = 0,9738793$$

3. The radius of a circle is to be increased from the initial value of $r_0 = 10$ by an amount $dr = 0.1$. An estimation of the corresponding increase in the circle area $A = \pi r^2$ can be obtained by calculating the differential dA, which can be compared with the true change ΔA.

$$r_0 = 10,\ dr = 0.1,\ A(r) = \pi r^2,\ A'(r) = 2\pi r$$
$$dA = A'(r_0)\cdot(r-r_0) = A'(r_0)\cdot dr = 2\pi\cdot 10\cdot 0.1 = 2\pi$$
$$\Delta A = A(r_0+\Delta r) - A(r_0) = A(10+0.1) - A(10) = 10{,}1^2\pi - 10^2\pi =$$
$$= (10{,}1^2 - 10^2)\pi = (10{,}1-10)(10{,}1+10)\pi = 0{,}1\cdot 20{,}1\pi = 2{,}01\pi$$

Exploring the behaviour of functions in the neighbourhood of some point is often rather difficult in the case of functions with complicated formulas. It is therefore better to consider some easier function instead, which is a good enough approximation of the original one. Functions are most frequently substituted by polynomials. These are infinitely times differentiable and their derivatives of any order are again polynomials.

In order to substitute the function f differentiable up to order n, $n \in \boldsymbol{N}$, by linear polynomials, we can at first substitute the function graph by a line, while it is natural to choose here the tangent to the function graph at point $[a, f(a)]$, the graph of function $T_1(x) = f(a) + f'(a)(x - a)$. It holds that $T_1(a) = f(a)$, $T_1'(a) = f'(a)$. A better approximation can be obtained by polynomials of higher degrees, while it is required that the values of their higher derivatives are equal to the values of the respective derivatives of function f at the point a.

There exists exactly one polynomial in the form

$$T_n(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

whose coefficients are uniquely determined by the value of function f and by the values of its first n derivatives at the point $x = a$. It is called the Taylor polynomial of function f at the point a.

An approximation of function f by the Taylor polynomial leads to a certain error. This can be estimated as the difference $f(x) - T_n(x)$, which is denoted by $R_n(x)$ and called the radical of function f by the n-th Taylor polynomial

$$R_n(x) = f(x) - T_n(x).$$

Taylor theorem

Let a, x be two different numbers, $n \geqq 0$ is the integer and J is the closed interval with end points a, x. Let f be a function differentiable on interval J with continuous derivatives on the interior of J up to order $n + 1$. Then such point ξ exists inside J, $\xi \in (a, x)$ that

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots +$$
$$+ \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$$

i.e. $f(x) = T_n(x) + R_n(x)$, where $T_n(x)$ is the n-th Taylor polynomial of f at the point a and

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$$

is the radical of function f after the n-th member of Taylor polynomial.

For $a = 0$ the form of Taylor formula is

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \ldots + \frac{f^{(n)}(0)}{n!}x^n + \frac{f^{(n+1)}(0)}{(n+1)!}x^{n+1},$$

where $0 < \xi < x$ (or $x < \xi < 0$), and it is called the MacLaurin formula.

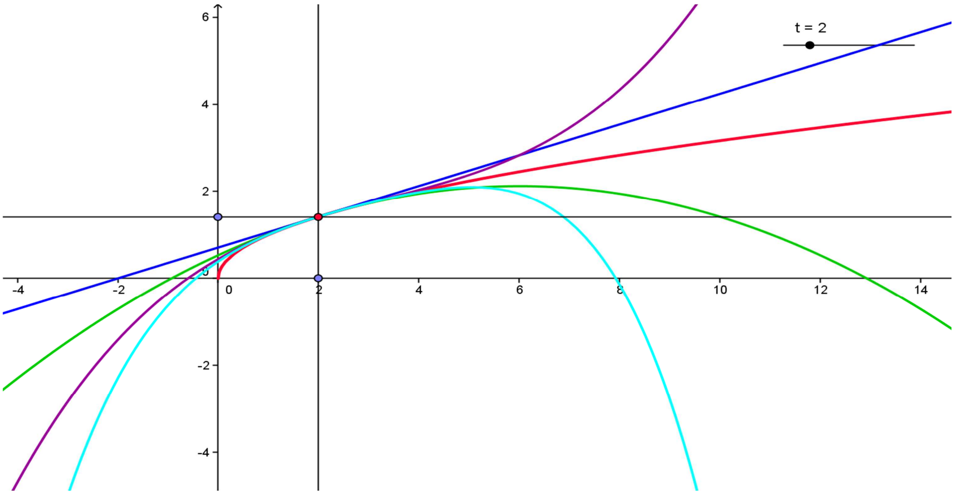


Fig. 3.65. Geometric interpretation of Taylor polynomials

The following rule can be used for the evaluation of the limits leading to one of the undetermined expressions of the type

$$\frac{0}{0} \quad \frac{\infty}{\infty} \quad 0.\infty \quad \infty - \infty \quad 0^0 \quad 1^\infty \quad \infty^0$$

L´Hospital rule

Suppose

a) $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = 0$, or $\lim\limits_{x\to a} |f(x)| = \lim\limits_{x\to a} |g(x)| = \infty$

b) there exists (proper or improper) $\lim\limits_{x\to a} \dfrac{f'(x)}{g'(x)}$.

Then $\lim\limits_{x\to a} \dfrac{f(x)}{g(x)}$ also exists, and $\lim\limits_{x\to a} \dfrac{f(x)}{g(x)} = \lim\limits_{x\to a} \dfrac{f'(x)}{g'(x)}$.

This rule is also valid for the limits at improper points and for one-sided limits.

Examples

1. Approximation of function $f(x)$ = sin x at the point a = 0 by Taylor polynomial of order 4 is

$$\sin x = \sin 0 + \frac{\cos 0}{1!}(x-0) + \frac{-\sin 0}{2!}(x-0)^2 +$$

$$+ \frac{-\cos 0}{3!}(x-0)^3 + \frac{-\sin 0}{4!}(x-0)^4 = x - \frac{x^3}{6}$$

$$\sin 1 = 1 - \frac{1}{6} = \frac{5}{6} = 0{,}83333$$

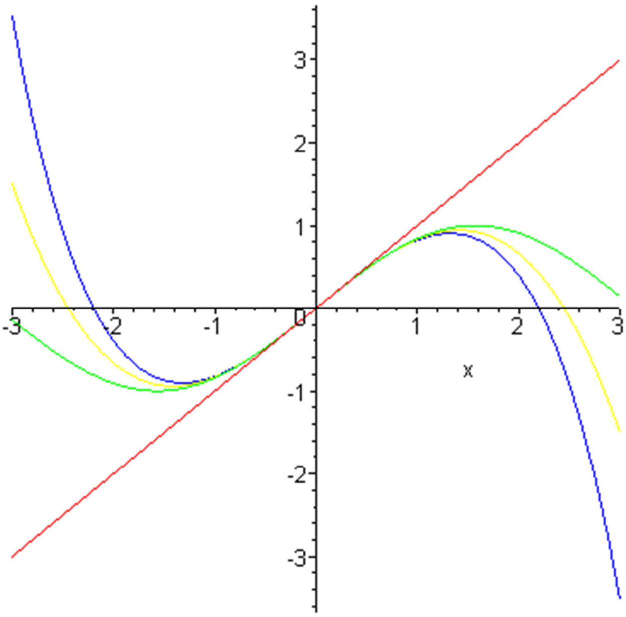


Fig. 3.66. Taylor polynomials for function sin x

2. By means of L´Hospital rule we can evaluate the following limits

a) $$\lim_{x\to 0}\frac{3^x - 1}{\sin x} = \lim_{x\to 0}\frac{3^x \ln 3}{\cos x} = \ln 3$$

b) $$\lim_{x\to 0}\frac{x - \sin x}{x^3} = \lim_{x\to 0}\frac{1-\cos x}{3x^2} = \lim_{x\to 0}\frac{\sin x}{6x} = \lim_{x\to 0}\frac{\cos x}{6} = \frac{1}{6}$$

c) $$\lim_{x\to\frac{\pi}{2}}\frac{\tan x}{1+\tan x} = \lim_{x\to\frac{\pi}{2}}\frac{\frac{1}{\cos^2 x}}{\frac{1}{\cos^2 x}} = 1$$

d) $\lim\limits_{x\to\infty} x\sin\frac{1}{x} = \lim\limits_{x\to\infty}\frac{\sin\frac{1}{x}}{\frac{1}{x}} = \lim\limits_{x\to\infty}\frac{-\frac{1}{x^2}\cos\frac{1}{x}}{-\frac{1}{x^2}} = \lim\limits_{x\to\infty}\cos\frac{1}{x} = 1$

e) $\lim\limits_{x\to 0^+} x\ln x = \lim\limits_{x\to 0^+}\frac{\ln x}{\frac{1}{x}} = \lim\limits_{x\to 0^+}\frac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim\limits_{x\to 0^+}(-x) = 0$

f) $\lim\limits_{x\to 1^-}\frac{1}{1-x}\ln x = \lim\limits_{x\to 1^-}\frac{\ln x}{1-x} = \lim\limits_{x\to 1^-}\frac{\frac{1}{x}}{-1} = \lim\limits_{x\to 1^-}\left(-\frac{1}{x}\right) = 1$

3.10 Function monotonicity and extrema

The application of the differential calculus in the investigation of functions is based on a simple relationship between the behaviour of a function and the properties of its derivatives, and particularly of the first derivative. An increase is associated with positive derivatives and a decrease with negative derivatives.
Suppose that a function $f(x)$ is differentiable at every point x of an interval J. Then

1. f is increasing on J if $f'(x) > 0, \forall x \in J$
2. f is decreasing on J if $f'(x) < 0, \forall x \in J$
3. f is non-decreasing on J if $f'(x) \geq 0, \forall x \in J$
4. f is non-increasing on J if $f'(x) \leq 0, \forall x \in J$.

In geometric terms it appears clearly evident that differentiable functions increase on intervals where their graphs have positive slopes and decrease on intervals where their graphs have negative slopes. It is evident that if the derivative of a function takes zero values at some isolated points but it retains constant sign at all other points, this function is strictly monotone (increasing or decreasing) in the given interval.

Examples

1. Function $f(x) = x^3 - 3x$ is strictly monotone on its domain of definition $\boldsymbol{R}$. Its first derivative $f'(x) = 3(x^2 - 1)$ equals zero at points -1 and 1, and as $f'(x) < 0$ for all points from interval $(-1, 1)$, function f is decreasing on this interval. It is increasing on $(-\infty, -1)$ and $(1, \infty)$, as $f'(x) > 0$ on these intervals, fig. 3.67.
2. Function $f(x) = x - \sin x$ is increasing on $\boldsymbol{R}$, whereas $f'(x) = 1 - \cos x \geq 0$ and $f'(2k\pi) = 0$, $\forall\ k = 0, \pm 1, \pm 1, ...$, fig. 3.68.

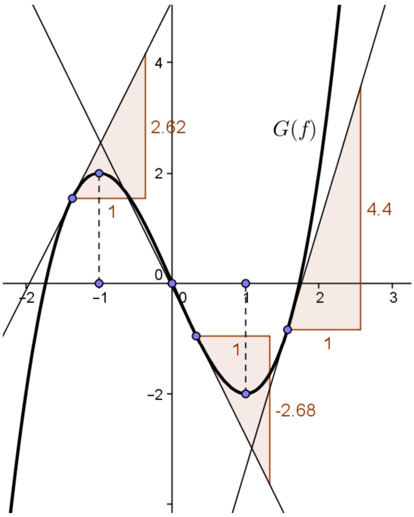


Fig. 3.67. Function monotonicity

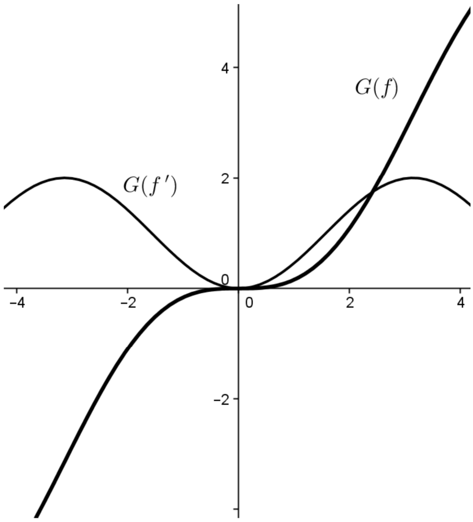


Fig. 3.68. Strictly monotone function

Local (relative) extrema

Many application problems require the determination of those arguments from the function domain of definition, in which function attains its maximal or minimal value with respect to the whole function range or on a certain interval called the function extrema, maximum or minimum.

Let f be a function defined on a neighbourhood of point x_0. The value $f(x_0)$ is said to be a local maximum of function $f(x)$ if such neighbourhood $O_\varepsilon(x_0)$ exists that

$$f(x) \le f(x_0), \forall x \in O_\varepsilon(x_0).$$

Point x_0 is called the point of local maximum.

The value $f(x_0)$ is said to be a local minimum of function $f(x)$ if such neighbourhood $O_\varepsilon(x_0)$ exists that

$$f(x) \ge f(x_0), \forall x \in O_\varepsilon(x_0).$$

Point x_0 is called the point of local minimum.

If

$$f(x) < f(x_0), \forall x \in O_\varepsilon(x_0), \text{ or } f(x) > f(x_0), \forall x \in O_\varepsilon(x_0),$$

the value $f(x_0)$ is called the strict local maximum or minimum, respectively.

When speaking of maximum or minimum at a point, we usually mean a strict extremum.

If x_0 is a point of local extremum of a function f differentiable at x_0, then $f'(x_0) = 0$.

From this assumption it follows that a function can possess local extrema at the points at which the derivative is equal to zero (these are called the stationary points) or at points at which the derivative does not exist.

Points at which the function derivative $f\,'$ is zero or it fails to exist are called the critical points of function f (for the first derivative). Existence of a local extrema at the point x_0 means that tangent line to the function graph at the point $T = [x_0, f(x_0)]$ is parallel to the coordinate axis x with equation $y = f(x_0)$, or it is parallel to the coordinate axis y with equation $x = x_0$, or no tangent line exists at this point.

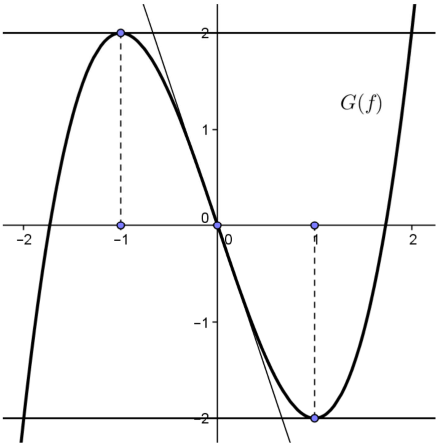


Fig. 3.69. Tangents and points of extrema

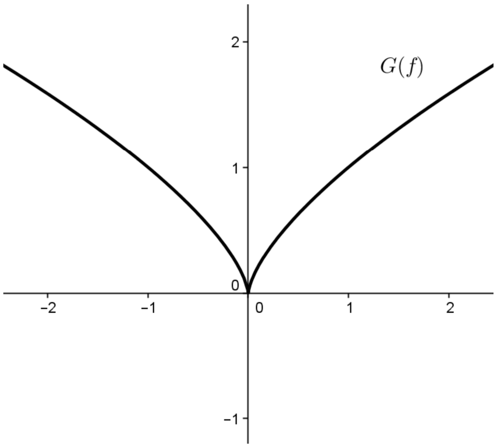


Fig. 3.70. Local minimum of function

Examples

1. Function $f(x) = x^3 - 3x$ with the first derivative $f\,'(x) = 3(x^2 - 1)$ equal to zero at points -1 and 1 attains the local maximum at -1, $f(-1) = 2$, and local minimum at 1, $f(1) = -2$. The tangent lines to function graph at points $[-1, 2]$ and $[1, -2]$ are parallel to the coordinate axis x, fig. 3.69.

2. Function $f(x) = \sqrt[3]{x^2}$ is defined on $\boldsymbol{R}$, whereas its non-zero first derivative $f\,'(x) = \dfrac{2}{3\sqrt[3]{x}}$ is not defined at the point $x = 0$. The range of the function is the interval $\langle 0, \infty)$, while the minimum 0 is the function value at the point $x = 0$. The tangent line to the function graph at the point $[0, 0]$, is the line $x = 0$, fig. 3.70.

The first derivative test

Let a function f be differentiable at each point $x \in O_\varepsilon(x_0)$, $x \neq x_0$ and $f\,'(x_0) > 0$ ($f\,'(x_0) < 0$) for all points from interval $(x_0 - \varepsilon, x_0)$ and $f\,'(x_0) < 0$ ($f\,'(x_0) > 0$) for all points from interval $(x_0, x_0 + \varepsilon)$. Then $f(x_0)$ is the strict local maximum (minimum).
It means that to attain the local maximum $f(x_0)$ the function $f(x)$ must be increasing on $O_\varepsilon^-(x_0)$ and decreasing on $O_\varepsilon^+(x_0)$, while for a local minimum $f(x_0)$ the function $f(x)$ is decreasing on $O_\varepsilon^-(x_0)$ and increasing on $O_\varepsilon^+(x_0)$.

The second derivative test

If $f'(x_0) = 0$ and $f''(x_0) \neq 0$, then x_0 is a point of local extremum. If $f''(x_0) < 0$, then $f(x_0)$ is the strict local maximum of function f and if $f''(x_0) > 0$, then $f(x_0)$ is the strict local minimum of function f.

Examples

1. Function $f(x) = x^3 - 3x$ second derivative $f'(x) = 6$ is negative at the point -1 and it is positive at the point 1, fig. 3.69.
2. Suppose we have to make a can in the shape of a right circular cylinder with a given volume $V > 0$. To find its dimensions that will use the minimum material we can consider the formula for calculation of the cylinder area, $A = 2\pi(xy + x^2)$, and its volume, $V = \pi x^2 y$, where $x > 0$ is the cylinder radius and $y > 0$ is its height. From $y = \dfrac{V}{\pi x^2}$ defined for all real $x > 0$ we obtain the function

$$A(x) = 2\pi\left(x\frac{V}{\pi x^2} + x^2\right) = 2\pi\left(\frac{V}{\pi x} + x^2\right).$$

The first derivative $A'(x) = 4\pi x - \dfrac{2V}{x^2}$ is defined for all $x > 0$, whereas $A'(x) = 0$ at the point $x_0 = \sqrt[3]{\dfrac{V}{2\pi}}$. The second derivative is $A''(x) = 4\pi + \dfrac{4V}{x^3}$, and its value at the point x_0 is $A''\left(\sqrt[3]{\dfrac{V}{2\pi}}\right) = 12\pi > 0$, therefore the function $A(x)$ attains local minimum at this point. The related height of the cylinder is $y_0 = \dfrac{V}{\pi\sqrt[3]{\dfrac{V^2}{4\pi^2}}} = \sqrt[3]{\dfrac{4V}{\pi}}$, while the minimal material consumption is the value of the area function, which is $A(x_0) = 2\pi\left(\sqrt[3]{\dfrac{4V^2}{2\pi^2}} + \sqrt[3]{\dfrac{V^2}{4\pi^2}}\right) = 3\sqrt[3]{2\pi V^2}$.

Global (absolute) extrema

Let a function $f(x)$ be defined on a set M and let x0 $\in M$. The value $f(x_0)$ is said to be a global (absolute) maximum of f on M, if for all $x \in M$ it holds that $f(x) \leq f(x_0)$. The value $f(x_0)$ is said to be a global (absolute) minimum of f on M, if for all $x \in M$ it holds that $f(x) \geq f(x_0)$. Therefore the global maximum is the greatest and the global minimum is the least value assumed by the function f on set M.

Without specifying the set M, the least or the greatest value of function f is considered to be on the function domain of definition $D(f)$.

A problem often appearing in technical applications is to find the greatest (maximum) and the least (minimum) values of a continuous function on a closed interval $\langle a, b\rangle$. Here we first specify all critical points in the open interval (a, b), then calculate the function values at these points and the values at the end points $f(a)$, $f(b)$, and finally determine the greatest and the least number among these values, which are the global extrema of the function on the closed interval $\langle a, b\rangle$.

Examples

1. Function $f(x) = 2x^3 - 3x^2 - 12x + 1$ defined on $\boldsymbol{R}$ has two critical points, the roots of quadratic equation $f'(x) = 6x^2 - 6x - 12 = 0$, $x_1 = -1$, $x_2 = 2$, while for the function second derivative $f''(x) = 12x - 6$ it holds that $f''(-1) = -18$, and $f''(2) = 6$, therefore the function attains local maximum $f(-1) = 8$ and local minimum $f(2) = -19$. To determine the function global extrema on the closed interval $\langle -2, 4\rangle$, we must consider the fact that both points of the local extrema are in the open interval $(-2, 4)$, and the function values at the endpoints are $f(-2) = -3$ and $f(4) = 33$. Therefore on interval $\langle -2, 4\rangle$ the function attains its global maximum 33 at the point 4 and global minimum -19 at the point 2. The function graph is in fig. 3.71.
2. A rectangle is to be inscribed in a semicircle with radius 2. What is the largest area the rectangle can have and what are its dimensions? The problem can be analysed for example by means of its visualisation in fig. 3.72. Considering the coordinates of one rectangle vertex located on the semicircle, $A = [x, y]$, the rectangle area P equals $2xy$, while from the circle equation in the simple form $x^2 + y^2 = 4$ we obtain $y = \sqrt{4 - x^2}$. Looking for the local maximum of function $P(x) = 2x\sqrt{4 - x^2}$, whose first derivative can be determined in the form

   $$P'(x) = 2\sqrt{4 - x^2} - \frac{2x^2}{\sqrt{4 - x^2}} = \frac{4(2 - x^2)}{\sqrt{4 - x^2}},$$

   stationary points can be calculated as the roots of the equation $P'(x) = 0$, therefore $2 - x^2 = 0$ and $x_{1,2} = \pm\sqrt{2}$, while $y = \sqrt{2}$. Because

   $$P''(x) = \frac{-8x\sqrt{4 - x^2} + 4x(2 - x^2)}{\left(4 - x^2\right)\sqrt{4 - x^2}}$$

   is negative in the point $\max_x = \sqrt{2}$, the area function has its maximum at this point, $P(\sqrt{2}) = 4$. One vertex of the obtained rectangle located on the semicircle is the point $Y = \left[\sqrt{2}, \sqrt{2}\right]$, while the rectangle dimensions are $2\sqrt{2}, \sqrt{2}$.

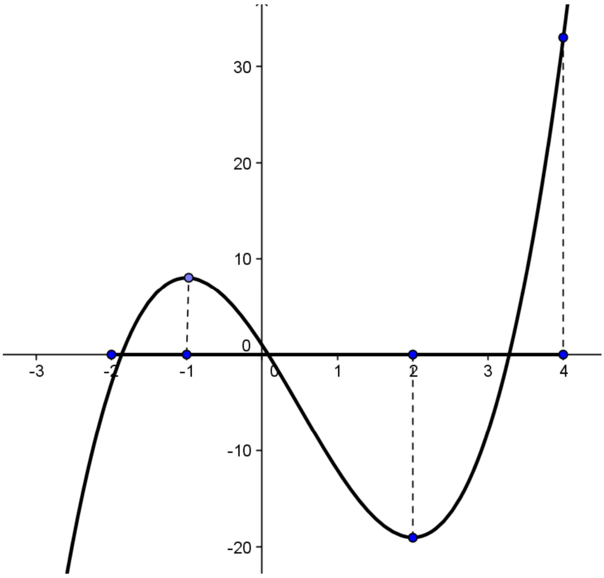


Fig. 3.71. Function global extrema on closed interval

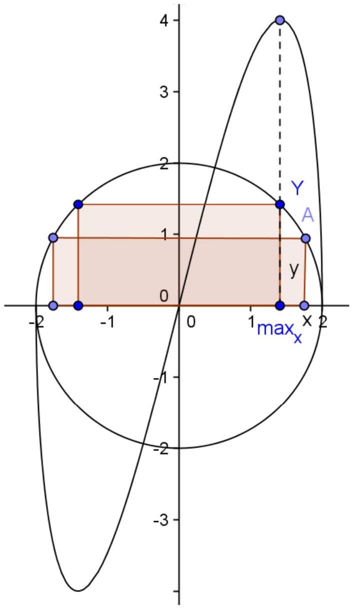


Fig. 3.72. Inscribed rectangle

In the case that function f fails to be continuous on set M, or M is not a closed interval, then the minimum and maximum of f on M can be, but need not be reached.

3.11 Convexity, concavity and points of inflexion

Let f be a function differentiable on interval J. Function f is said to be convex (concave) on interval J, if all points on its graph lie above (below) any tangent line of the function graph on this interval, with the exception of the tangent point.
Let function f be continuous on interval J and let its second derivative exist at each interior point of this interval. If for all points within the interior of the interval J it holds that

$$f''(x) > 0 \; (f''(x) < 0),$$

then function f is convex (concave) on interval J.

Suppose function f be continuous at the point x_0. If such neighbourhood $O_\varepsilon(x_0)$ of this point exists, that function f is concave (convex) on $O_\varepsilon^-(x_0)$ and it is convex (concave) on $O_\varepsilon^+(x_0)$, then the point x_0 is called the point of inflection (inflexion point) of function f.

If x_0 is the point of inflection of function f, then point $I_0 = [x_0, y_0]$ is called the point of inflection of the function graph $G(f)$. The tangent line to $G(f)$ at the point of inflection intersects the graph, which means it comes from one half-plane determined by a common boundary line in the respective tangent line to the other, see fig. 3.73.

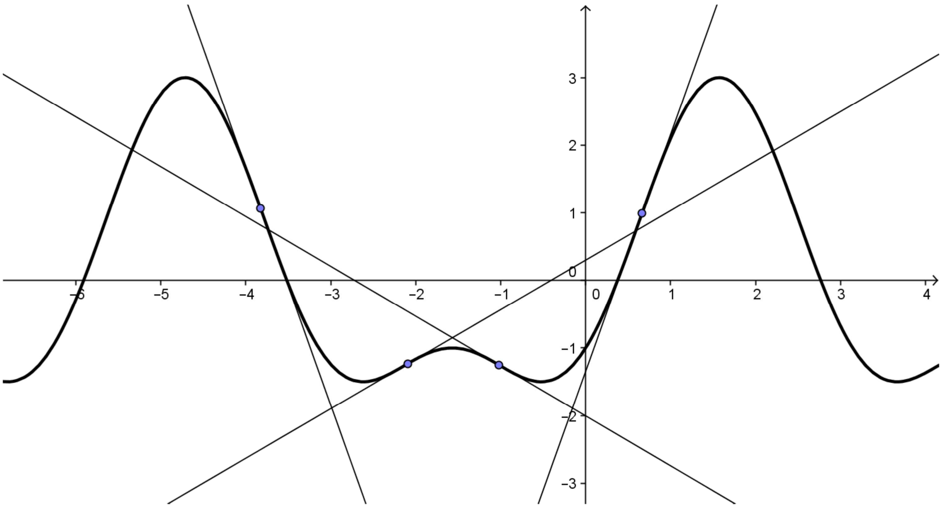


Fig. 3.73. Geometric interpretation of convexity and concavity

Let the third derivative of function f exists at the point x_0 , and let $f''(x_0) = 0$, while $f'''(x_0) \neq 0$. Then x_0 is the point of inflection of the function f. Moreover, if x_0 is the point of inflection of a function f, then either $f''(x_0) = 0$, or $f''(x_0)$ does not exist.

Let $f'(x_0) = f''(x_0) = ... = f^{(n-1)}(x_0) = 0$, but $f^{(n)}(x_0) \neq 0$. If n is an even number, then function f attains a local extrema at the point x_0 , which is

a strict local maximum if $f^{(n)}(x_0) < 0$,

a strict local minimum if $f^{(n)}(x_0) > 0$.

If n is an odd number, then the point x_0 is the point of inflection of the function f.

Examples

1. To find intervals of convex and concave behaviour of function $f(x) = x^3 - 3x$ means to find the zero points of its second derivative $f''(x) = 6x$, which is the point $x = 0$. The function is concave on interval $(-\infty, 0)$, because on this interval is $f''(x) < 0$, and it is convex on interval $(0, \infty)$, as on this interval holds $f''(x) > 0$. Point $x = 0$ is the point of inflection, as $f'''(x) = 6 \neq 0$, fig.3.69.

2. Function $y = \ln\frac{1+x}{1-x}$ is defined on interval $(-1, 1)$, its derivatives are $y' = \frac{2}{1-x^2}$, $y'' = \frac{4x}{(1-x^2)^2}$, $y''' = \frac{12x^2+4}{(1-x^2)^3}$. The second derivative equals zero at the point $x = 0$, while the third derivative is non-zero at this point, which is therefore the point of inflection of this function.

Function behaviour

To investigate the function behaviour means to determine the following:

1. domain of definition, points of discontinuity and zero points
2. parity or periodicity
3. intervals of strict monotonicity, and points of local (global) extrema
4. points of inflection and intervals of convex and concave behaviour
5. equations of asymptotes to the function graph
6. coordinates of some points on the function graph – the table of function values
7. sketch the graph of function

Example

The behaviour of function $f(x) = x^3 - 3x$ has been investigated in the previous examples.

1. $D(f) = \boldsymbol{R}$, the function is continuous on $\boldsymbol{R}$, the zero points are $-\sqrt{3}$, 0, $\sqrt{3}$.
2. Function is odd, as $f(-x) = -f(x)$.
3. Function is increasing on intervals $(-\infty, -1)$ and $(1, \infty)$, decreasing on $(-1, 1)$, and it has the local maximum at the point $x = -1$, $f(-1) = 2$, the local maximum at the point $x = 1$, $f(1) = -2$, the function does not have a global extrema.
4. Point $x = 0$ is the point of inflection, the function is concave on interval $(-\infty, 0)$, convex on interval $(0, \infty)$.
5. No asymptotes to the graph of function exist.
6. Function values at selected points

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

Table 3.2. Function values

7. Function graph is in fig. 3.69.

4 Integral calculus of functions with one real variable

4.1 Indefinite integral

An inverse problem to differentiation often appears in many scientific and technical applications, which can be formulated as follows: find function $F(x)$ to a given function $f(x)$ such, that $F'(x) = f(x)$. If such interval J exists that for each $x \in J: F'(x) = f(x)$, then the function $F(x)$ is said to be antiderivative of function $f(x)$ on interval J. The set of all antiderivatives of function $f(x)$ on interval J is called the indefinite integral of f on J and the notation is $\int f(x)dx$.

If formula $F(x) + C$ gives all antiderivatives, we indicate this with the expression

$$\int f(x)dx = F(x) + C, \forall x \in J\ .$$

To integrate a function means to find all its antiderivatives, thus its indefinite integral. Operations of differentiation and integration are inverse to each other

$$\int f'(x)dx = f(x) + C, \quad \left[\int f(x)dx\right]' = f(x).$$

If the function $f(x)$ find is continuous on an open interval J, then it possesses antiderivative on J, or it is said to be integrable on J.

Let F_i be an antiderivative of a function f_i on an open interval J for all $i = 1, 2, ..., n$. Then function $F = k_1F_1 + k_2F_2 +...+ k_nF_n$, where k_i are constants, is antiderivative of function $f = k_1f_1 + k_2f_2 +...+ k_nf_n$ on interval J, therefore

$$\int \left(k_1 f_1(x) + k_2 f_2(x) + ... + k_n f_n(x)\right) dx =$$

$$= k_1 \int f_1(x)dx + k_2 \int f_2(x)dx + ... + k_n \int f_n(x)dx .$$

In particular

$$\int [f(x) \pm g(x)]dx = \int f(x)dx \pm \int g(x)dx$$

$$\int k \cdot f(x)dx = k \cdot \int f(x)dx .$$

Examples

1. Antiderivative of function $f(x) = 1 - 3x^2$ is the function $F(x) = x - x^3$, but also any function $F(x) + C$, $C \in \boldsymbol{R}$, because for its derivative it holds that

   $[F(x) + C]' = [x - x^3 + C]' = 1 - 3x^2$.

The antiderivative whose graph is passing through the point [1, 2] is a particular function $F_p(x) = F(x) + C_p$ with a specific value of the constant C_p such that $F_p(1) = 2$, it means $C_p = 2$, $F_p(x) = x - x^3 + 2$. The graph is in fig. 4.1.

2. Function $F(x) = e^x - \sin x$ whose graph is passing through the point [0, 1] is one antiderivative of function $f(x) = e^x - \cos x$, while this integral curve is one of a system of parallel curves representing the system of functions that are antiderivatives of function $f(x)$. These can be determined as the indefinite integral of function $f(x)$, which is denoted as

$$\int \left(e^x - \cos x\right)dx = e^x - \sin x + C,\ C \in \boldsymbol{R}\ .$$

The system of curves is presented in fig. 4.2.

3. Indefinite integral of function $f(x) = \frac{1}{x} + \frac{1}{\cos^2 x}$ is the system of functions

$$f(x) = \ln|x| + \tan x + C\ ,$$

which is formally denoted as

$$\int \left(\frac{1}{x} + \frac{1}{\cos^2 x}\right)dx = \ln|x| + \tan x + C$$

and defined for all $x \neq k\frac{\pi}{2}, k \in \boldsymbol{Z}, x \neq 0$. Some integral curves are illustrated in fig. 4.3.

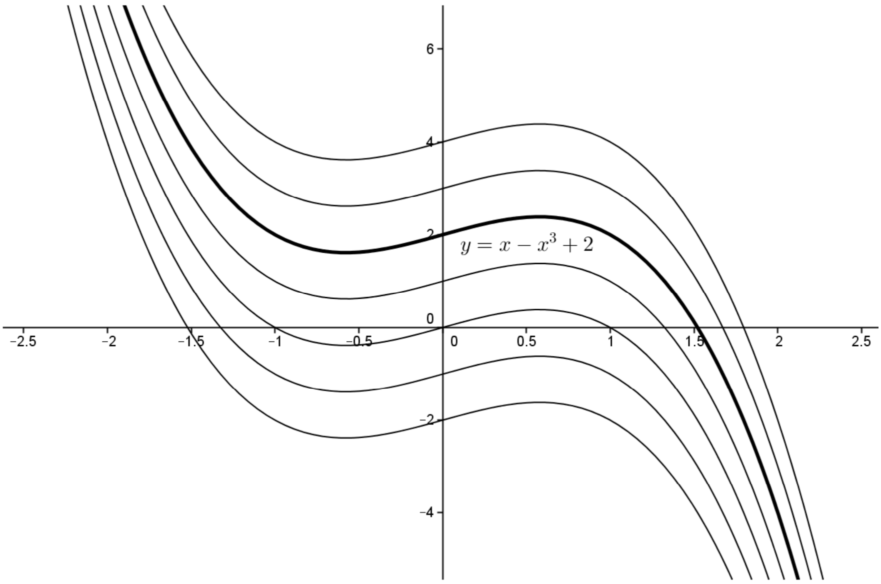


Fig. 4.1. Graphs of antiderivatives

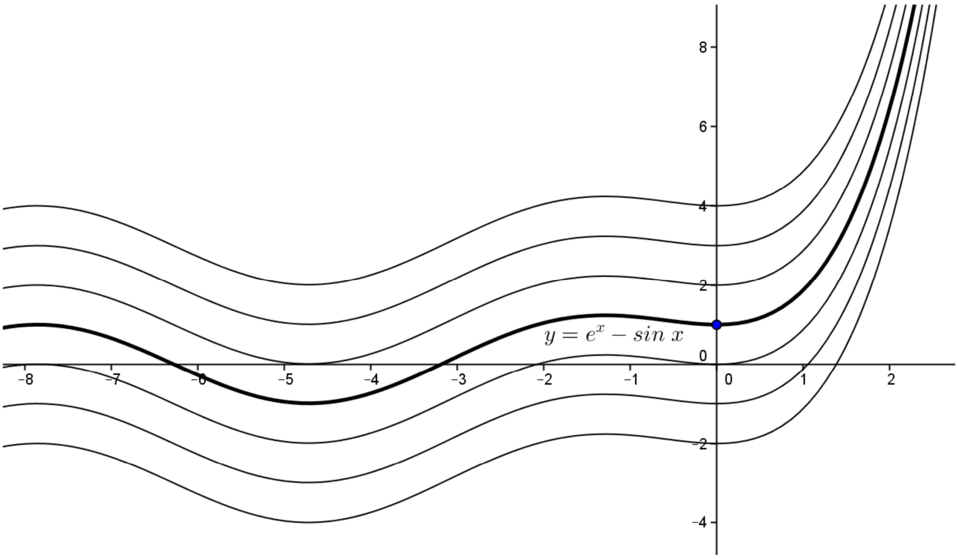


Fig. 4.2. System of antiderivative graphs

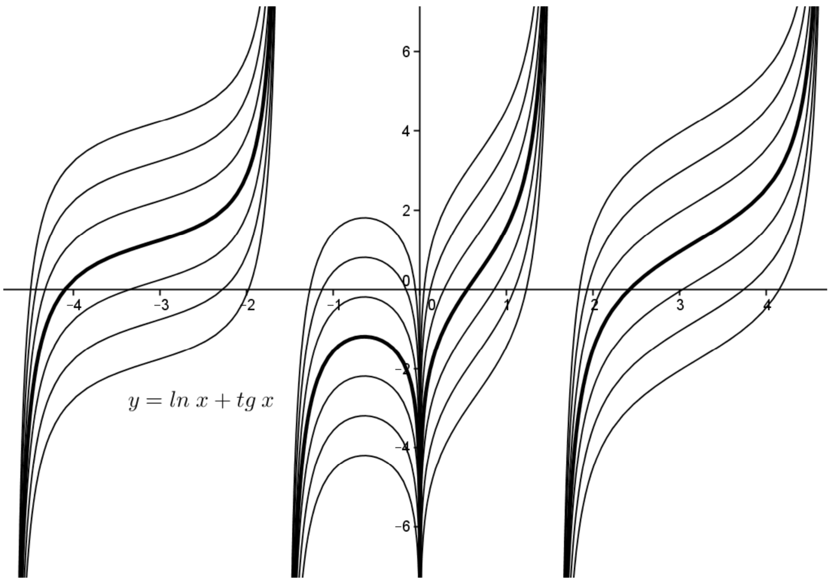


Fig. 4.3. Graphs of system of antiderivatives

4.2 Integration of elementary functions

The formulas for antiderivatives of elementary functions are valid on any open intervals that are parts of the domains of definition of corresponding antiderivatives on the right side.

Basic integration formulae

1. $\int x^n dx = \frac{x^{n+1}}{n+1} + c, n \neq -1$

2. $\int \frac{1}{x} dx = \ln|x| + c$

3. $\int a^x dx = \frac{a^x}{\ln a} + c, 0 < a \neq 1$

4. $\int e^x dx = e^x + c$

5. $\int \sin x dx = -\cos x + c$

6. $\int \cos x dx = \sin x + c$

7. $\int \frac{1}{\cos^2 x} dx = \tan x + c$

8. $\int \frac{1}{\sin^2 x} dx = -\cot x + c$

9. $\int \frac{1}{\sqrt{1-x^2}} dx = \arcsin x + c, \ \int \frac{1}{\sqrt{a^2 - x^2}} dx = \arcsin \frac{x}{a} + c, a > 0$

10. $\int \frac{1}{1+x^2} dx = \arctan x + c, \ \int \frac{1}{a^2 + x^2} dx = \frac{1}{a} \arctan \frac{x}{a} + c, a > 0$

11. $\int \frac{1}{\sqrt{x^2 \pm a^2}} dx = \ln\left|x + \sqrt{x^2 \pm a^2}\right|, a > 0$

12. $\int \frac{f'(x)}{f(x)} dx = \ln|f(x)| + c$

To integrate a function means to calculate its indefinite integral using the basic integration formulas and properties of indefinite integrals.

Examples.

1. $\int \left(2x^5 - x^3 + 5\right) dx = \frac{x^6}{3} - \frac{x^4}{4} + 5x + c$

2. $\int \frac{\sqrt[3]{x^2}-2x^3+3}{x^4}dx = \int \left(x^{\frac{2}{3}-4} - 2x^{-1} + 3x^{-4} \right) dx =$

$$= \int x^{-\frac{10}{3}} dx - 2\int \frac{1}{x} dx + 3\int x^{-4} dx = \frac{x^{-\frac{10}{3}+1}}{-\frac{10}{3}+1} - 2\ln|x| + 3\frac{x^{-4+1}}{-4+1} + c =$$

$$= \frac{-3}{7\sqrt[3]{x^7}} + \ln\frac{1}{x^2} - \frac{1}{x^3} + c$$

3. $\int \frac{x^2}{x^2+4}dx = \int \frac{x^2+4-4}{x^2+4}dx = \int dx - \int \frac{2^2}{x^2+2^2}dx = x - \arctan\frac{x}{2} + c$

4. $\int \frac{3x}{x^2+2}dx = \frac{3}{2}\int \frac{2x}{x^2+2}dx = \frac{3}{2}\ln\left|x^2+2\right| + c = \ln\sqrt{\left(x^2+2\right)^3} + c$

5. $\int \left(3^x - \tan x\right)dx = \int 3^x dx + \int \frac{-\sin x}{\cos x}dx = \frac{3^x}{\ln 3} + \ln|\cos x| + c$

6. Let the velocity of a rectilinear motion be given by the relation $v(t) = t^2$, where t is the time of motion. To find the law of the motion we are looking for the function $s(t)$ representing the trajectory of this motion under the condition $s(0) = 1$, which is a particular antiderivative of function $v(t)$. Therefore

$$s(t) = \int t^2 dt = \frac{t^3}{3} + c,$$

from which $s(0) = c = 1$, $s(t) = \frac{t^3}{3} + 1$.

4.3 Basic integration methods

Integration by parts

This method is most frequently used for integration of such expressions that may be represented in the form of a product of two functions $u(x)$ and $v(x)$ in such a way, that the finding of the function $u(x)$ and the evaluation of the integral $\int u \cdot v' dx$ is a simpler problem than the direct evaluation of the original integral $\int u' \cdot v dx$.

Let $u(x)$ and $v(x)$ be functions possessing continuous derivatives. Then

$$\int u'(x) \cdot v(x) dx = u(x) \cdot v(x) - \int u(x) \cdot v'(x) dx.$$

Examples

1. $\int xe^x dx = xe^x - \int e^x dx = e^x(x-1) + c$

2. $\int x^2 \cos x dx = x^2 \sin x - \int 2x \sin x dx =$

$$= x^2 \sin x - \left(-2x\cos x - \int -2\cos x dx\right) + c =$$

$$= x^2 \sin x + 2x\cos x - 2\sin x + c$$

3. $\int \sin^2 x dx = \int \sin x \sin x dx = -\sin x \cos x - \int -\cos^2 x dx =$

$$= -\sin x \cos x + \int \left(1-\sin^2 x\right) dx =$$

$$= -\sin x \cos x + \int 1 dx - \int \sin^2 x dx =$$

$$= -\sin x \cos x + x - \int \sin^2 x dx + C$$

$$2\int \sin^2 x dx = x - \sin x \cos x + C \Rightarrow \int \sin^2 x dx = \frac{x}{2} - \frac{1}{2}\sin x \cos x + c$$

4. $\int \ln x dx = \int 1 \ln x dx = x\ln x - \int x\frac{1}{x}dx = x\ln x - x + c = x(-1+\ln x) + c$

5. $\int \frac{\ln x}{x} dx = \int \frac{1}{x}\ln x dx = \ln^2 x - \int \frac{1}{x}\ln x dx + C$

$$2\int \frac{\ln x}{x} dx = \ln^2 x + C \Rightarrow \int \frac{\ln x}{x} dx = \frac{1}{2}\ln^2 x + c$$

6. $\int \arctan x dx = \int 1 \arctan x dx = x\arctan x - \int \frac{x}{x^2+1} dx =$

$$= x\arctan x - \frac{1}{2}\ln\left|x^2+1\right| + c = x\arctan x + \ln\frac{1}{\sqrt{x^2+1}} + c$$

7. $\int e^x \sin x dx = e^x \sin x - \int e^x \cos x dx = e^x \sin x - \left(e^x \cos x + \int e^x \sin x dx\right) + C$

$$2\int e^x \sin x dx = e^x\left(\sin x - \cos x\right) + C \Rightarrow \int e^x \sin x dx = \frac{1}{2}e^x\left(\sin x - \cos x\right) + c$$

Substitution integration method

A change of variable can often transform an unfamiliar integral into one which can be evaluated directly or by another known method. This method is called the substitution method, while it is mostly used if the integrand is a function of the form

$$f(\varphi(x)) \cdot \varphi'(x).$$

Let $\int f(u)du = F(u) + C$ on (α, β). Then

$$\int f(\varphi(x))\varphi'(x)dx = F(\varphi(x)) + C \text{ on } (a, b),$$

which is in short

$$\int f(\varphi(x))\varphi'(x)dx = \begin{vmatrix} u = \varphi(x) \\ du = \varphi'(x)dx \end{vmatrix} = \int f(u)du = F(u) + C = F(\varphi(x)) + C .$$

Examples

1. $\int 2x\cos(x^2+1)dx = \begin{vmatrix} u = x^2+1 \\ du = 2xdx \end{vmatrix} = \int \cos u du = \sin u + c = \sin(x^2+1) + c$

2. $\int \sin 5x dx = \begin{vmatrix} u = 5x \\ du = 5dx \end{vmatrix} = \frac{1}{5}\int \sin u du = -\frac{1}{5}\cos u + c = -\frac{1}{5}\cos 5x + c$

3. $\int \cos^7 x \sin x dx = \begin{vmatrix} u = \cos x \\ du = -\sin x dx \end{vmatrix} = -\int u^7 du = -\frac{1}{8}u^8 + c = -\frac{1}{8}\cos^8 x + c$

4. $\int \frac{\ln^4 x}{x} dx = \begin{vmatrix} u = \ln x \\ du = \frac{1}{x}dx \end{vmatrix} = \int u^4 du = \frac{1}{5}u^5 + c = \frac{1}{5}\ln^5 x + c$

5. $\int \frac{\sin x}{2+\cos x} dx = \begin{vmatrix} u = 2+\cos x \\ du = -\sin x dx \end{vmatrix} = -\int \frac{1}{u} du = -\ln|u| + c =$

$$= -\ln|2+\cos x| + c = \ln\frac{1}{|2+\cos x|} + c$$

6. $\int 3xe^{x^2} dx = \begin{vmatrix} u = x^2 \\ du = 2xdx \end{vmatrix} = \frac{3}{2}\int e^u du = \frac{3}{2}e^u + c = \frac{3}{2}e^{x^2} + c$

7. $\int \frac{\arctan^2 x}{1+x^2} dx = \begin{vmatrix} u = \arctan x \\ du = \frac{1}{1+x^2} dx \end{vmatrix} = \int u^2 du = \frac{1}{3}u^3 + c = \frac{1}{3}\arctan^3 x + c$

Another form of possible substitution is to take x as a function of u, i.e. $x = \varphi(u)$. In this case a suitable function φ should be chosen so that one can evaluate the obtained indefinite integral and determine the inverse function φ^{-1}.

$$\int f(x)dx = \left| \begin{matrix} x = \varphi(u) \Rightarrow u = \varphi^{-1}(x) \\ dx = \varphi'(u)du \end{matrix} \right| = \int f\left(\varphi(u)\right)\cdot\varphi'(u)du =$$
$$= F(u) + C = F(\varphi^{-1}(x)) + C$$

Examples

1. $\int \sin\sqrt{x}dx = \left| \begin{matrix} x = u^2 \Rightarrow u = \sqrt{x} \\ dx = 2udu \end{matrix} \right| = \int 2u\sin udu = -2u\cos u + 2\sin u + c =$
$$= -2\sqrt{x}\cos\sqrt{x} + 2\sin\sqrt{x} + c$$

2. $\int \sqrt{a^2 - x^2}dx = \left| \begin{matrix} x = a\sin u \Rightarrow u = \arcsin\dfrac{x}{a} \\ dx = a\cos udu \end{matrix} \right| = \int \sqrt{a^2 - a^2\sin^2 u}\,a\cos udu =$
$$= a^2\int \cos^2 udu + c = a^2\left(\frac{u}{2} + \frac{1}{2}\sin u\cos u\right) + c =$$
$$= \frac{a^2}{2}(u + \sin u\cos u) + c = \frac{a^2}{2}\left(u + \sqrt{1 - \sin^2 u}\sin u\right) + c =$$
$$= \frac{a^2}{2}\left(\arcsin\frac{x}{a} + \frac{x}{a}\sqrt{1 - \frac{x^2}{a^2}}\right) + c$$

3. $\int \dfrac{1}{x^2}\sin\dfrac{1}{x}dx = \left| \begin{matrix} u = \dfrac{1}{x} \Rightarrow x = \dfrac{1}{u} \\ du = -\dfrac{1}{x^2}dx \end{matrix} \right| = -\int \sin udu = \cos u + c = \cos\dfrac{1}{x} + c$

Generally, the integration process consists of transforming the given integral by means of algebraic transformation of the integrand to an integral already known, which can be evaluated by means of integration by parts or by change of variable.

Examples

1. $\int e^{\sqrt{x}}dx = \left| \begin{matrix} x = u^2 \Rightarrow u = \sqrt{x} \\ dx = 2udx \end{matrix} \right| = \int 2ue^u du = 2e^u(u - 1) + c = 2e^{\sqrt{x}}\left(\sqrt{x} - 1\right) + c$

2. $\int \dfrac{\ln(\ln x)}{x}dx = \left| \begin{matrix} u = \ln x \Rightarrow x = e^u \\ du = \dfrac{1}{x}dx \end{matrix} \right| = \int \ln udu = u(-1 + \ln u) + c =$
$$= \ln x\left(-1 + \ln(\ln x)\right) + c$$

3. $\int \cos(\ln x)dx = \left| \begin{array}{c} u = \ln x \Rightarrow x = e^u \\ du = \frac{1}{x}dx \end{array} \right| = \int e^{-u} \cos u du = u(-1 + \ln u) + c =$

$$= \ln x\big(-1 + \ln(\ln x)\big) + c$$

4. $\int \sin x \cos x dx = \left| \begin{array}{c} u = \sin x \\ du = \cos x dx \end{array} \right| = \int u du = \frac{u^2}{2} + c = \frac{1}{2}\sin^2 x + c$

$$\int \sin x \cos x dx = \left| \begin{array}{c} u = \cos x \\ du = -\sin x dx \end{array} \right| = -\int u du = -\frac{u^2}{2} + c = -\frac{1}{2}\cos^2 x + c$$

where according to trigonometric identity it holds that

$$\frac{1}{2}\sin^2 x = \frac{1}{2} - \frac{1}{2}\cos^2 x$$

4.4 Integration of special functions

Rational functions

One of the most important classes of elementary functions whose antiderivatives are elementary functions that can be found in a relatively simple way, are rational functions.

Consider integrals of the type

$$\int \frac{P(x)}{x^2 + px + q} dx,$$

where $P(x)$ is a polynomial, $p, q \in \boldsymbol{R}$.

If the degree of the polynomial $P(x)$ is greater than 1, then the division of $P(x)$ by $x^2 + px + q$ results in a polynomial $Q(x)$ and a polynomial $ax + b$, as the remainder. Consequently

$$\frac{P(x)}{x^2 + px + q} = Q(x) + \frac{ax + b}{x^2 + px + q}.$$

The integration of the polynomial Q(x) can be performed without any difficulties, and hence the problem is reduced to the integration of a fraction

$$\frac{ax + b}{x^2 + px + q}, \text{ if } a^2 + b^2 \neq 0.$$

Each integral of that type can be transformed to one of the following basic types:

I. $\int \frac{dx}{x^2 + a^2} = \frac{1}{a}\arctan\frac{x}{a} + C,\ a > 0$ (basic integration formula)

II. $\displaystyle\int \frac{dx}{x^2-a^2} = \frac{1}{2a}\left(\int \frac{dx}{x-a} - \int \frac{dx}{x+a}\right) + C =$

$$= \frac{1}{2a}\left(\ln|x-a| - \ln|x+a|\right) + C = \frac{1}{2a}\ln\left|\frac{x-a}{x+a}\right| + C$$

III. $\displaystyle\int \frac{x}{x^2 \pm a^2}\,dx = \frac{1}{2}\int \frac{2x}{x^2 \pm a^2}\,dx = \frac{1}{2}\ln\left|x^2 \pm a^2\right| + C$

IV. $\displaystyle\int \frac{x}{(x \pm a)^2}\,dx = \int \frac{x+a-a}{(x \pm a)^2}\,dx = \int \frac{x \pm a}{(x \pm a)^2}\,dx \mp a\int \frac{dx}{(x \pm a)^2} + C =$

$$= \ln|x \pm a| \pm \frac{a}{x \pm a} + c$$

If two different real numbers x_1, x_2 exist, such that $x^2 + px + q = (x - x_1)(x - x_2)$, then constants A and B exist such that

$$\frac{ax+b}{x^2 + px + q} = \frac{A}{x - x_1} + \frac{B}{x - x_2}$$

The two partial fractions can be easily integrated, and the unknown constants A and B can be determined by the method of the indefinite coefficients based on the comparison of the coefficients of two polynomials of the same degree.

Examples

1. $\displaystyle\int \frac{x}{x^2+x+1}\,dx = \frac{1}{2}\int \frac{2x+1-1}{x^2+x+1}\,dx = \frac{1}{2}\int \frac{2x+1}{x^2+x+1}\,dx - \frac{1}{2}\int \frac{1}{x^2+x+1}\,dx =$

$$= \frac{1}{2}\ln\left|x^2+x+1\right| - \frac{1}{2}\int \frac{1}{\left(x+\frac{1}{2}\right)^2 + \frac{3}{4}}\,dx = \ln\sqrt{\left|x^2+x+1\right|} - \frac{1}{\sqrt{3}}\arctan\frac{2x+1}{\sqrt{3}} + c$$

2. $\displaystyle\int \frac{x^4}{x^2+1}\,dx = \int \left(x^2-1\right)dx + \int \frac{1}{x^2+1}\,dx = \frac{x^3}{3} - x + \arctan x + c$

3. $\displaystyle\int \frac{dx}{3x^2-27} = \frac{1}{3}\int \frac{dx}{x^2-9} = \frac{1}{3}\cdot\frac{1}{2\cdot 3}\left(\int \frac{dx}{x-3} - \int \frac{dx}{x+3}\right) = \frac{1}{18}\ln\left|\frac{x-3}{x+3}\right| + c$

4. To calculate $\displaystyle\int \frac{12x+2}{x^2+5x+6}\,dx$ it is suitable to reduce the given fraction to partial fractions with linear denominators, such that $(x-2)(x-3) = x^2 - 5x + 6$. Basically we have to solve the following equation.

$$\frac{12x+2}{x^2-5x+6}=\frac{A}{x-2}+\frac{B}{x-3}$$

$$12x+2=A(x-3)+B(x-2)\Rightarrow 12x+2=(A+B)x-3A-2B$$

Comparing coefficients of the polynomials we receive system of 2 equations with unknown coefficients *A* and *B*.

$$A+B=12,\ -3A-2B=2$$

$$A=12-B\Rightarrow -3(12-B)-2B=2$$

$$B=38, A=-26$$

$$\int\frac{12x+2}{x^2-5x+6}dx=\int\frac{-26}{x-2}dx+\int\frac{38}{x-3}dx=-26\ln|x-2|+38\ln|x-3|+c$$

Irrational functions

Integrals of some simple irrational functions can be transformed to the integrals of rational functions with the substitution method, choosing a suitable substitution. Integrals, in which the formula $\sqrt[n]{ax+b}, a\neq 0$ appears, are among the most frequent cases.

Examples

1. $$\int\frac{x}{\sqrt[3]{x+1}}dx=\left|\begin{matrix}t=x+1\\ dt=dx\end{matrix}\right|=\int\frac{t-1}{\sqrt[3]{t}}dt=$$

$$=\int\left(t^{1-\frac{1}{3}}-t^{-\frac{1}{3}}\right)dt=\int\left(t^{\frac{2}{3}}-t^{-\frac{1}{3}}\right)dt=\frac{3}{5}t^{\frac{5}{3}}-\frac{3}{2}t^{\frac{2}{3}}+c=$$

$$=\frac{3}{5}\sqrt[3]{(x+1)^5}-\frac{3}{2}\sqrt[3]{(x+1)^2}+c=\frac{3}{10}\sqrt[3]{(x+1)^2}(2x-3)+c$$

2. $$\int\frac{\sqrt[3]{x-1}+x}{x-1}dx=\left|\begin{matrix}t=\sqrt[3]{x-1}\Rightarrow t^3=x-1\\ dt=dx\end{matrix}\right|=\int\frac{t+t^3+1}{t^3}dt=$$

$$=\int\left(1+t^{-2}+t^{-3}\right)dt=t-t^{-1}-\frac{1}{2}t^{-2}+c=$$

$$=\sqrt[3]{x-1}-\frac{1}{\sqrt[3]{x-1}}-\frac{1}{2\sqrt[3]{(x-1)^2}}+c=\frac{2(x-1)-2\sqrt[3]{(x-1)}-1}{2\sqrt[3]{(x-1)^2}}+c$$

Other frequently appearing integrals of irrational function are integrals of type

$$\int \frac{dx}{\sqrt{ax^2+bx+c}},$$

where $a \neq 0$ and ax^2+bx+c is positive on an interval, while the case of polynomials with double roots can be excluded. Using factorisation by

$$\frac{1}{\sqrt{a}}, a>0, \text{or } \frac{1}{\sqrt{-a}}, a<0$$

this integral can be transformed to one of integrals in the form

$$\int \frac{dx}{\sqrt{x^2+px+q}}, \text{ or } \int \frac{dx}{\sqrt{-x^2+px+q}},$$

i.e. integrals leading after substitution to the integrals

$$\int \frac{dx}{\sqrt{x^2 \pm k^2}} = \ln\left|x+\sqrt{x^2 \pm k^2}\right| + c, \text{ or}$$

$$\int \frac{dx}{\sqrt{k^2-x^2}} = \arcsin\frac{x}{k} + c, \text{ respectively.}$$

Examples

1. $$\int \frac{dx}{\sqrt{3x^2+5x-4}} = \frac{1}{\sqrt{3}}\int \frac{dx}{\sqrt{x^2+\frac{5}{3}x-\frac{4}{3}}} = \frac{1}{\sqrt{3}}\int \frac{dx}{\sqrt{x^2+2\cdot\frac{5}{6}x+\frac{25}{36}-\frac{73}{36}}} =$$

$$= \frac{1}{\sqrt{3}}\int \frac{dx}{\sqrt{\left(x+\frac{5}{6}\right)^2-\frac{73}{36}}} = \left|\begin{matrix} t = x+\frac{5}{6} \\ dt = dx \end{matrix}\right| = \frac{1}{\sqrt{3}}\int \frac{dt}{\sqrt{t^2-\left(\frac{\sqrt{73}}{6}\right)^2}} =$$

$$= \frac{1}{\sqrt{3}}\ln\left|t+\sqrt{t^2-\left(\frac{\sqrt{73}}{6}\right)^2}\right| + c = \frac{1}{\sqrt{3}}\ln\left|x+\frac{5}{6}+\sqrt{x^2+\frac{5}{3}x-\frac{4}{3}}\right| + c$$

2. $$\int \frac{dx}{\sqrt{x-2x^2}}dx = \frac{1}{\sqrt{2}}\int \frac{dx}{\sqrt{-\left(-\frac{x}{2}+x^2\right)}} =$$

$$=\frac{1}{\sqrt{2}}\int\frac{dx}{\sqrt{-\left(\frac{1}{4}-x\right)^2+\frac{1}{16}}}=\left|\begin{array}{l}t=\frac{1}{4}-x\\ dt=-dx\end{array}\right|=-\frac{1}{\sqrt{2}}\int\frac{dt}{\sqrt{\frac{1}{16}-t^2}}=$$

$$=-\frac{1}{\sqrt{2}}\arcsin\frac{t}{\frac{1}{4}}+c=-\frac{1}{\sqrt{2}}\arcsin(1-4x)+c$$

Trigonometric functions

Integrals of the form $\int R(\sin x,\cos x)dx$, where the integrand is a rational function in terms of trigonometric functions $\sin x$ and $\cos x$, can be transformed by substitution $t=\tan\frac{x}{2}\Rightarrow x=2\arctan t$ to integrals of rational functions. Then it holds that

$$dx=\frac{2dt}{1+t^2},\ \sin x=\frac{2t}{1+t^2},\ \cos x=\frac{1-t^2}{1+t^2}$$

This substitution is generally convenient for the computation of integrals in the form

$$\int\frac{dx}{a\cos x+b\sin x+c}.$$

Examples

1. $$\int\frac{dx}{\sin x}=\left|\begin{array}{l}\sin x=\frac{2t}{1+t^2}\\ dx=\frac{2dt}{1+t^2}\end{array}\right|=\int\frac{1}{t}dt=\ln|t|+c=\ln\left|\tan\frac{x}{2}\right|+c$$

2. $$\int\frac{dx}{3+2\cos x}=\left|\begin{array}{l}\cos x=\frac{1-t^2}{1+t^2}\\ dx=\frac{2dt}{1+t^2}\end{array}\right|=\int\frac{1}{3+2\frac{1-t^2}{1+t^2}}\cdot\frac{2dt}{1+t^2}dt=$$

$$=\int\frac{2dt}{3(1+t^2)+2(1-t^2)}=\int\frac{2dt}{3(1+t^2)+2(1-t^2)}=$$

$$=\int\frac{2dt}{t^2+5}=\frac{2}{\sqrt{5}}\arctan\frac{t}{\sqrt{5}}+c=\frac{2}{\sqrt{5}}\arctan\frac{\tan\frac{x}{2}}{\sqrt{5}}+c$$

3. $$\int \frac{dx}{1-\sin x} = \left| \begin{matrix} \sin x = \dfrac{2t}{1+t^2} \\ dx = \dfrac{2dt}{1+t^2} \end{matrix} \right| = \int \frac{1}{1-\dfrac{2t}{1+t^2}} \cdot \frac{2dt}{1+t^2} dt = \int \frac{2dt}{3(1+t^2)-2t} =$$

$$= \int \frac{2dt}{3t^2-2t+3} = \frac{2}{3}\int \frac{dt}{t^2-\dfrac{2}{3}t+1} = \frac{2}{3}\int \frac{dt}{t^2-\dfrac{2}{3}t+\dfrac{1}{9}+\dfrac{8}{9}} =$$

$$= \frac{2}{3}\int \frac{dt}{\left(t-\dfrac{1}{3}\right)^2+\dfrac{8}{9}} = \left| \begin{matrix} u = t-\dfrac{1}{3} \\ du = dt \end{matrix} \right| = \frac{2}{3}\int \frac{du}{u^2+\dfrac{8}{9}} =$$

$$= \frac{2}{3}\arctan\frac{3u}{\sqrt{8}} + c = \frac{2}{3}\arctan\frac{3t-1}{\sqrt{8}} + c = \frac{2}{3}\arctan\frac{3\tan\dfrac{x}{2}-1}{\sqrt{8}} + c$$

If the integrand can be reduced to the form $f(\sin x)\cos x$ or $f(\cos x)\sin x$, where f is an easily integrable function, then it is advantageous to use other simplier substitution $t = \sin x$, or $t = \cos x$, respectively.

Examples

1. $$\int \sin 2x dx = \int 2\sin x \cos x dx = \left| \begin{matrix} t = \sin x \\ dt = \cos x dx \end{matrix} \right| = \int 2t dt = t^2 + c = \sin^2 x + c$$

2. $$\int 3\cos^2 x \sin x dx = \left| \begin{matrix} t = \cos x \\ dt = -\sin x dx \end{matrix} \right| = \int -3t^2 dt = -t^3 + c = -\cos^3 x + c$$

3. $$\int \frac{\sin^3 x}{\cos^2 x+1} dx = \int \frac{(1-\cos^2 x)\sin x}{\cos^2 x+1} dx = \left| \begin{matrix} t = \cos x \\ dt = -\sin x dx \end{matrix} \right| = \int \frac{-(1-t^2)}{t^2+1} dt =$$

$$= \int \frac{t^2-1}{t^2+1} dt = \int \frac{t^2+1-2}{t^2+1} dt = \int dt - 2\int \frac{dt}{t^2+1} =$$

$$= t - 2\arctan t + c = \cos x - 2\arctan(\cos x) + c$$

4. $$\int \cos^3 x dx = \int (1-\sin^2 x)\cos x dx = \left| \begin{matrix} t = \sin x \\ dt = \cos x dx \end{matrix} \right| = \int (1-t^2) dt =$$

$$= t - \frac{1}{3}t^3 + c = \sin x - \frac{1}{3}\sin^3 x + c$$

5. $$\int \frac{\sin x \cos x}{\sin^2 x + 2\sin x + 3}dx = \left| \begin{array}{c} t = \sin x \\ dt = \cos x dx \end{array} \right| = \int \frac{t}{t^2 + 2t + 3}dt =$$

$$= \int \frac{t}{(t+1)^2 + 2}dt = \left| \begin{array}{c} u = t+1 \\ du = dt \end{array} \right| = \int \frac{u-1}{u^2 + 2}du =$$

$$= \int \frac{u}{u^2+2}du - \int \frac{1}{u^2+2}du = \frac{1}{2}\int \frac{2u}{u^2+2}du - \int \frac{1}{u^2+2}du =$$

$$= \frac{1}{2}\ln\left|u^2+2\right| - \frac{1}{\sqrt{2}}\arctan\frac{u}{\sqrt{2}} + c =$$

$$= \frac{1}{2}\ln\left|(t+1)^2+2\right| - \frac{1}{\sqrt{2}}\arctan\frac{t+1}{\sqrt{2}} + c =$$

$$= \frac{1}{2}\ln\left|\sin^2 x + 2\sin x + 3\right| - \frac{1}{\sqrt{2}}\arctan\frac{\sin x + 1}{\sqrt{2}} + c$$

If function $f(x)$ is continuous, then the antiderivative $F(x) = \int f(x)dx$ exists, but no general method is known to determine it. Integration of an elementary function does not always lead to an elementary function, which is not the case of differentiation. It can be proved that elementary functions exist whose integrals are inexpressible in terms of elementary functions. For instance, the following integrals

$$\int \frac{dx}{\sqrt{1+x^3}},\ \int \frac{e^x}{x}dx,\ \int \frac{\sin x}{x}dx,\ \int \frac{\cos x}{x}dx,\ \int \frac{dx}{\ln x},\ \int e^{-x^2}dx$$

cannot be represented with any elementary functions. However, it is necessary to distinguish between the question of existence of a desired antiderivative and the possibility of expressing it with the aid of elementary functions. The integrals written above exist, but the class of all elementary functions which we use is insufficient for expressing these integrals. To find and represent analytically these integrals it is necessary to extend the class of used functions. This is precisely one of the tasks solved in mathematical analysis. The non-elementary functions determined by the most important integrals inexpressible in terms of elementary functions are thoroughly investigated and tabulated (elliptic, or hyper-elliptic integrals).

Antiderivatives as functions of x defined on certain intervals can also be approximately represented, for instance using methods of numerical analysis, while some important ones are included in special integral tables.

4.5 Definite integrals

Various practical technical problems are leading to the concept of definite integral. One such problem with geometric background is the determination of an area of a specific plane figure, generally called curvilinear trapezoid, see in fig. 4.4. It is bounded by coordinate axis x, by vertical lines with equations $x = a$, $x = b$, where $a, b \in \boldsymbol{R}$, $a < b$, and by graph $G(f)$ of a continuous function $f(x)$ such that for each $x \in \langle a, b\rangle$ is $f(x) > 0$.

$$L = \{[x, y] : a \le x \le b, 0 \le y \le f(x)\}$$

Let us divide interval $\langle a, b\rangle$ into n subintervals by means of points

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$$

and denote $\Delta x_i = x_i - x_{i-1}$ for $i = 1, 2, \ldots, n$. Then let us choose an arbitrary point from each subinterval

$$\xi_i \in \langle x_{i-1}, x_i \rangle, i = 1,2,\ldots,n\,.$$

Finally, let us compute the sum

$$\sum_{i=1}^{n} f(\xi_i)\Delta x_i$$

that equals to the area of a step-like figure bounded from above by a broken line, see fig. 4.5. Depending on the choice of points x_i and ξ_i we can consider this sum to be an approximation of the area of curvilinear trapezoid, which is better and more accurate with increasing value of n, which means with decreasing length Δx_i of the division subintervals.

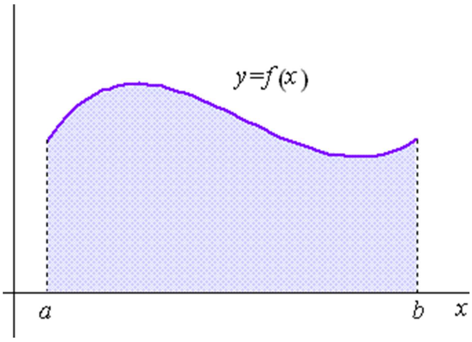


Fig. 4.4. Curvilinear trapezoid

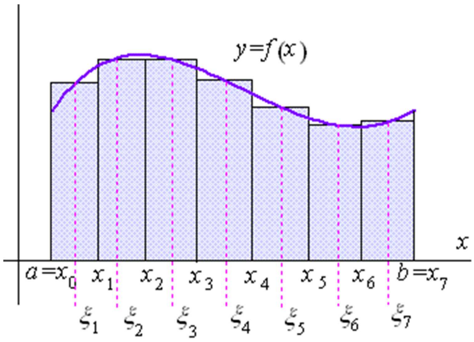


Fig. 4.5. Step-like figure

For any function f defined and bounded on a closed interval $\langle a, b\rangle$, the above sum is called the integral sum. If a limit of integral sums exists as the length of the greatest subinterval approaches zero, then it is called the definite integral of f on (or over) interval $\langle a, b\rangle$, and it is denoted as

$$\int_a^b f(x)dx = \lim_{\max \Delta x_i \to 0} \sum_{i=1}^{n} f(\xi_i)\Delta x_i ,$$

while function f is then said to be integrable on $\langle a, b\rangle$.

Sufficient condition for integrability

If a bounded function f possesses only a finite number of points of discontinuity on an interval $\langle a, b\rangle$, then it is integrable on this interval.

Every function continuous on a closed interval $\langle a, b\rangle$ is integrable on this interval.

Basic properties of definite integrals

1. **Linearity**: If functions f_1 and f_2 are integrable on an interval $\langle a, b\rangle$ and $c_1, c_2 \in \boldsymbol{R}$ are arbitrary constants, then
$$\int_a^b \left(c_1 f_1(x) + c_2 f_2(x)\right)dx = c_1 \int_a^b f_1(x)dx + c_2 \int_a^b f_2(x)dx .$$
2. **Aditivity**: If a function f is integrable on an interval $\langle a, b\rangle$ and $c \in \langle a, b\rangle$, then
$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx .$$
Defining $\int_a^a f(x)dx = 0,\quad \int_a^b f(x)dx = -\int_b^a f(x)dx$, the above equality is valid for any triplet of numbers a, b, c, provided all the included integrals exist.
3. **Monotonicity**: If functions f_1 and f_2 are integrable on an interval $\langle a, b\rangle$ and $f_1(x) \le f_2(x)$ for all $x \in \boldsymbol{R}$, then
$$\int_a^b f_1(x)dx \le \int_a^b f_2(x)dx .$$
4. **Positivity**: If function f is integrable on an interval $\langle a, b\rangle$ and $f(x) \ge 0$ for all $x \in \langle a, b\rangle$, then
$$\int_a^b f(x)dx \ge 0 .$$

Geometric interpretation

An area of a curvilinear trapezoid with sides formed by a graph of function $f(x)$ on a closed interval $\langle a, b\rangle$ and by line segments in lines with equations $x = a$, $x = b$ and coordinate axis x is

$$P = \int_a^b f(x)dx .$$

The definite integral is a measure, function attaching positive number – an area to geometric figures (trapezoids), with the following properties.

1. Number P is non-negative and it is uniquely determined by given trapezoid L.

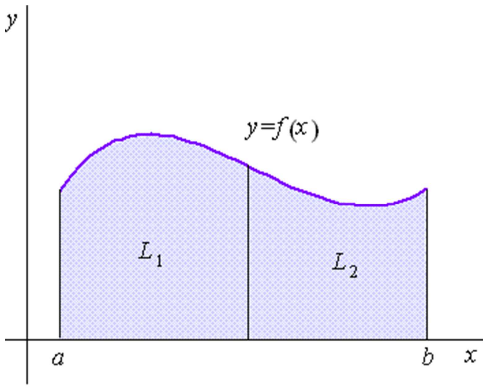


Fig. 4.6. Property 2

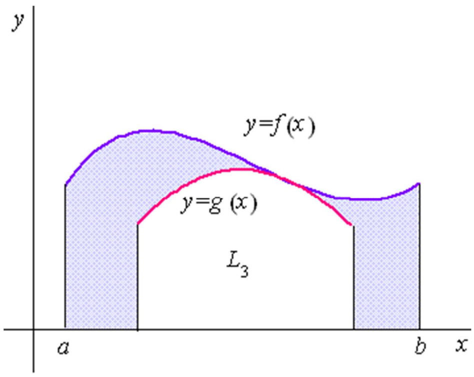


Fig. 4.7. Property 3

2. Aditivity: Dividing interval $\langle a, b\rangle$ by point c ($a < c < b$) to two intervals $\langle a, c\rangle$ and $\langle c, b\rangle$, the area P of a given curvilinear trapezoid L equals to sum of areas P_1, P_2 of curvilinear trapezoids L_1, L_2 defined over intervals $\langle a, c\rangle$ and $\langle c, b\rangle$.
3. Provided curvilinear trapezoid L_3 is subset of curvilinear trapezoid L, for its area P_3 the inequality $P_3 \leqq P$ holds.
4. If for all $x \in \langle a, b\rangle$ holds $f(x) = k > 0$, where k is constant, then

$$P = \int_a^b k dx = k(b-a)$$

is a rectangular area known from elementary geometry.

Physical interpretation

Let a force F be acting at all points of the coordinate axis x in the same direction and orientation as this axis. Let the value of the force F depend on the coordinate x of its position, and this dependence be defined by function f, therefore $F = f(x)$.

Then $\int_a^b f(x)dx$ determines the work that was done by this force, if a mass point is displaced from position $x = a$ to position $x = b$, it means it is moving on trajectory in line segment that is geometrically represented as closed interval $\langle a, b\rangle$ on the coordinates axis x.

Connection between the definite and indefinite integrals

Newton - Leibniz Formula (The evaluation theorem)

If f is a function continuous on an interval $\langle a, b\rangle$ and F is any antiderivative of f on $\langle a, b\rangle$, then

$$\int_a^b f(x)dx = \left[F(x)\right]_a^b = F(b) - F(a)\,.$$

Examples

1. $\int_0^\pi \sin x dx = [-\cos x]_0^\pi = -\cos\pi + \cos 0 = 1 + 1 = 2$

2. $\int_{-1}^1 \frac{dx}{1+x^2} = \left[\arctan x\right]_{-1}^1 = \arctan 1 - \arctan(-1) = \frac{\pi}{4} + \frac{\pi}{4} = \frac{\pi}{2}$

3. $\int_1^3 \frac{x+1}{x^2+2x+2}dx = \frac{1}{2}\int_1^3 \frac{2x+2}{x^2+2x+2}dx = \frac{1}{2}\left[\ln\left|x^2+2x+2\right|\right]_1^3 = \frac{1}{2}\left(\ln 17 - \ln 10\right)$

4.6 Integration methods for definite integrals

Integration by parts

Let $u(x)$ and $v(x)$ be functions having continuous derivatives on $\langle a, b\rangle$, then

$$\int_a^b u'(x)v(x)dx = \left[u(x)v(x)\right]_a^b - \int_a^b u(x)v'(x)dx.$$

Examples.

1. $\int_0^{\frac{\pi}{2}} x\sin x dx = [-x\cos x]_0^{\frac{\pi}{2}} + \int_0^{\frac{\pi}{2}} \cos x dx = 0 + [\sin x]_0^{\frac{\pi}{2}} = 1$

2. $\int_{-1}^1 (x+1)e^x dx = \left[(x+1)e^x\right]_{-1}^1 - \int_{-1}^1 e^x dx = 2e^2 - \left[e^x\right]_{-1}^1 = 2e^2 - e + \frac{1}{e}$

The substitution method (change of variables)

Let a function $f(x)$ be continuous on an interval $\langle a, b\rangle$, and let functions $\varphi(t)$ and $\varphi'(t)$ be continuous on an interval $\langle \alpha, \beta\rangle$, and moreover let for each $t \in \langle \alpha, \beta\rangle$ be $\varphi(t) \in \langle a, b\rangle$, while $\varphi(\alpha) = a$, $\varphi(\beta) = b$. Then the composite function $f(\varphi(t))$ is continuous on $\langle \alpha, \beta\rangle$ and

$$\int_a^b f(x)dx = \int_\alpha^\beta f(\varphi(t))\varphi'(t)dt\,.$$

Examples.

1. $$\int_0^2 \sqrt{4-x^2}\,dx = \left| \begin{array}{c} x = 2\sin t \\ x=0 \Rightarrow t=0 \\ x=2 \Rightarrow t=\frac{\pi}{2} \\ dx = 2\cos t dt \end{array} \right| = \int_0^{\frac{\pi}{2}} \sqrt{4-4\sin^2 t}\,2\cos t dt =$$

$$= 4\int_0^{\frac{\pi}{2}} \cos^2 t dt = 4\left[\frac{t}{2}+\frac{1}{2}\sin t\cos t\right]_0^{\frac{\pi}{2}} = \pi$$

2. $$\int_0^1 \frac{e^x}{e^x+1}\,dx = \left| \begin{array}{c} t = e^x+1 \\ x=0 \Rightarrow t=2 \\ x=1 \Rightarrow t=e+1 \\ dt = e^x dx \end{array} \right| = \int_2^{e+1} \frac{t+1}{t}\,dt = \int_2^{e+1}\left(1+\frac{1}{t}\right)dt = [t+\ln t]_2^{e+1} =$$

$$= e+1+\ln(e+1)-2-\ln 2 = e-1+\ln\frac{e+1}{2}$$

3. $$\int_1^e \frac{1+\ln x}{x}\,dx = \left| \begin{array}{c} t = \ln x \\ x=1 \Rightarrow t=0 \\ x=e \Rightarrow t=1 \\ dt = \frac{1}{x}dx \end{array} \right| = \int_0^1 (1+t)dt = \left[t+\frac{t^2}{2}\right]_0^1 = \frac{3}{2}$$

If $a \in \boldsymbol{R}$ and f is an even function integrable on the interval $\langle -a, a\rangle$, then

$$\int_{-a}^{a} f(x)dx = 2\int_0^a f(x)dx \ .$$

If $a \in \boldsymbol{R}$ and f is an odd function integrable on the interval $\langle -a, a\rangle$, then

$$\int_{-a}^{a} f(x)dx = 0 .$$

Examples

1. $$\int_{-5}^{5} x^5 e^{x^2}\,dx = 0$$

2. $\int_{-1}^{1} x \arctan x dx = 2\int_{0}^{1} x \arctan x dx =$

$$= \left[x^2 \arctan x\right]_0^1 - \int_0^1 \frac{x^2}{1+x^2}dx = \frac{\pi}{4} - \int_0^1 \frac{x^2+1-1}{1+x^2}dx =$$

$$= \frac{\pi}{4} - [x]_0^1 + \int_0^1 \frac{1}{1+x^2}dx = \frac{\pi}{4} - 1 + [\arctan x]_0^1 = \frac{\pi}{2} - 1$$

The mean value

Let f be a function integrable on an interval $\langle a, b\rangle$. Then the number

$$\mu = \frac{1}{b-a}\int_a^b f(x)dx$$

is called the mean value (the average value) of the function f on the interval $\langle a, b\rangle$. If the function f is continuous on $\langle a, b\rangle$, then at least one point $\xi \in (a, b)$ exists such that

$$f(\xi) = \frac{1}{b-a}\int_a^b f(x)dx.$$

In the case of a non-negative function f the last equality provides a geometric interpretation of f as the height that can be used to construct a rectangle whose base is the interval $\langle a, b\rangle$ and whose area equals the area of a given curvilinear trapezoid.

Examples

1. Mean value of function f: $y = x^2$ on interval $\langle 0, 1\rangle$ is

$$\mu = \frac{1}{1-0}\int_0^1 x^2 dx = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3},$$

   and the point, in which function attains this value is $\xi \in \langle 0, 1\rangle$

$$f(\xi) = \frac{1}{3} \Rightarrow \xi^2 = \frac{1}{3} \Rightarrow \xi = \frac{1}{\sqrt{3}}.$$

   Mean value of function f can be geometrically interpreted as the height of rectangle with the base of size 1, whose area equals to the area of the curvilinear trapezoid determined by the graph of function f, interval $\langle 0, 1\rangle$ on coordinate axis x and line segment on line $x = 1$, see fig. 4.8.

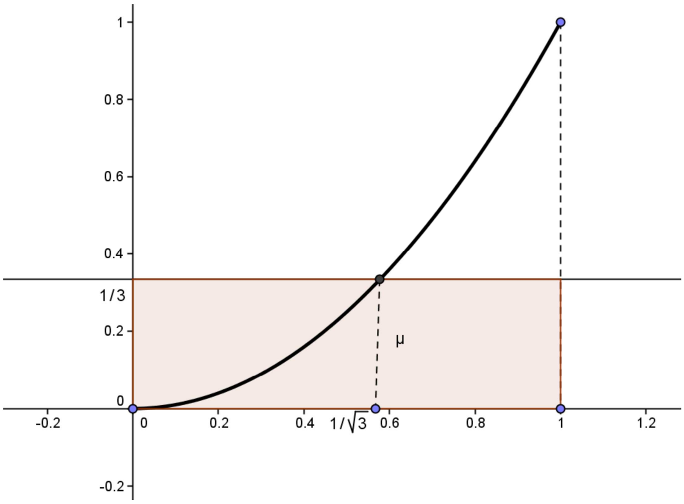


Fig. 4.8. Geometric interpretation of function mean value on interval

2. Mean value of function f: $y = 1 + 2\cos x$ on interval $\langle -\pi, \pi \rangle$ is

$$\mu = \frac{1}{\pi+\pi}\int_{-\pi}^{\pi}(1+2\cos x)dx = \left[x+2\sin x\right]_{-\pi}^{\pi} =$$
$$= \frac{\pi+2\sin\pi-(-\pi)-2\sin(-p)}{2\pi} = 1,$$

and because $f(\xi)=1 \Leftrightarrow 1+2\cos\xi=1 \Leftrightarrow \cos\xi=0 \Leftrightarrow \xi=\pm\frac{\pi}{2}$, two points exist in the interval $\langle -\pi, \pi \rangle$, in which function value equals to the average of function values on this interval, see in fig. 4.9.

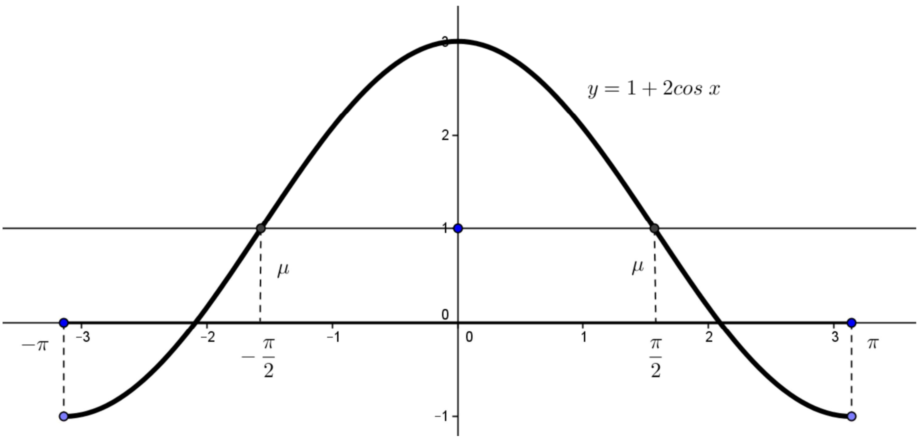


Fig. 4.9. Mean value of function

4.7 Applications of definite integrals

Geometric applications

The area of a plane figure

Let f and g be functions continuous on an interval $\langle a, b\rangle$ and such that for each $x \in \langle a, b\rangle$ it holds that $g(x) \le f(x)$. The plane region

$$R = \{[x, y] : a \le x \le b, g(x) \le y \le f(x)\}$$

is said to be a regular region and its area $A(R)$ is computed by means of the formula

$$A(R) = \int_a^b (f(x) - g(x))dx .$$

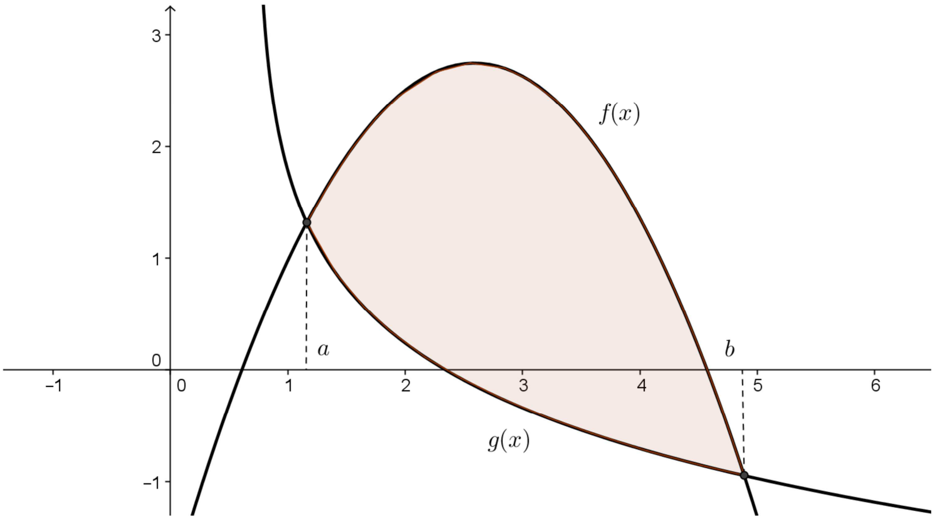


Fig. 4.10. Area of a regular region

Volume of a solid of revolution

Let us consider a solid generated by the revolution about coordinate axis x of a regular region

$$R = \{[x, y] : a \le x \le b, g(x) \le y \le f(x)\}.$$

Then the volume V of this solid can be calculated by the formula

$$V = \pi\int_a^b \left(f^2(x) - g^2(x)\right)dx.$$

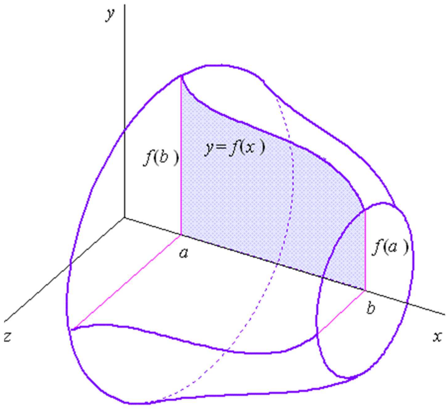


Fig. 4.11. Solid of revolution

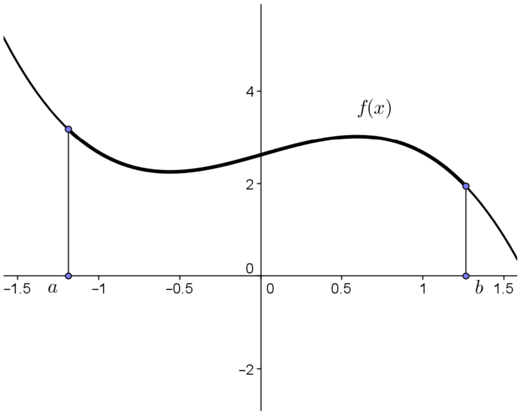


Fig. 4.12. Simple plane curve

Length of a simple plane curve

Let a simple curve be the graph of a continuous function f whose derivative is also continuous on an interval $\langle a, b\rangle$. Then it can be shown that the length L of the smooth curve given by the graph of f between lines $x = a$ and $x = b$ is determined by the formula

$$L = \int_a^b \sqrt{1 + [f'(x)]^2}\, dx\, .$$

The area of a surface of revolution

Let function f and its derivative be continuous on an interval $\langle a, b\rangle$. Let us consider a surface generated by revolving the curve that is graph of function $y = f(x)$, $x \in \langle a, b\rangle$ about coordinate axis x. It can be proved that the area of this surface can be determined by the formula

$$S = \int_a^b |f(x)| \sqrt{1 + [f'(x)]^2}\, dx\, .$$

Examples

1. Area of a region bounded by graphs of functions $f(x) = \ln x$, $g(x) = \ln^2 x$ can be calculated as definite integral with boundaries in points a, b, for which it holds that $f(a) = g(a)$, $f(b) = g(b)$, which means

$$\ln x = \ln^2 x \Leftrightarrow \ln x = 0 \vee \ln x = 1 \Leftrightarrow x = 1 \vee x = e$$

$$A(R) = \int_1^e \left(\ln(x) - \ln^2(x)\right) dx = \left[-3x + 3x\ln x - x\ln^2 x\right]_1^e = 3 - e.$$

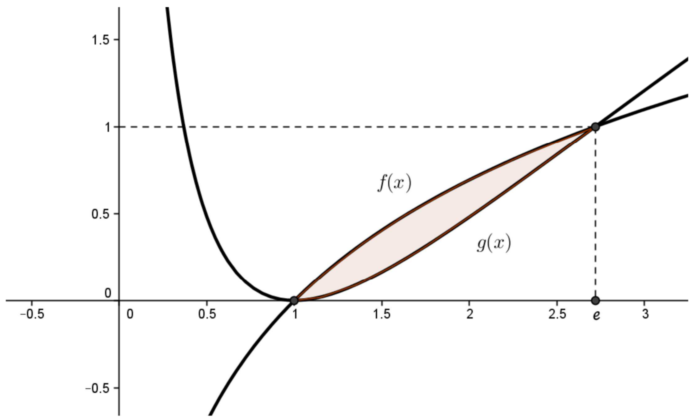
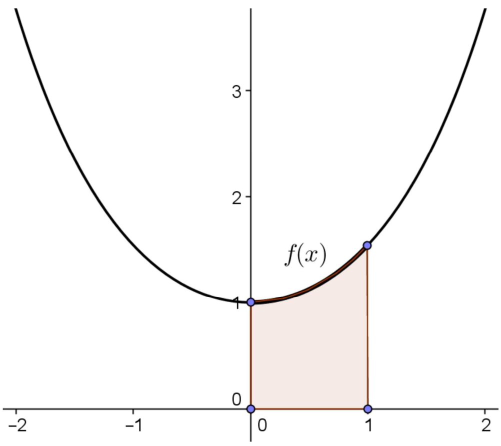


Fig. 4.13. Area of plane region

2. Area of a plane region $R = \left\{ [x, y] : 0 \le x \le 1, 0 \le y \le \dfrac{e^x + e^{-x}}{2} \right\}$ is

$$A(R) = \int_0^1 \frac{e^x + e^{-x}}{2}\, dx = \left[\frac{e^x - e^{-x}}{2} \right]_0^1 = \frac{e^2 - 1}{2e}.$$



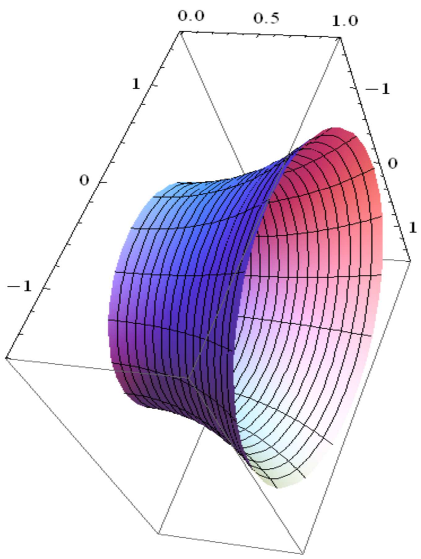


Fig. 4.14. Region *R* and surface of revolution

3. Volume of a solid of revolution generated by revolving plane region

$$R = \left\{ [x, y] : 0 \le x \le 1,\ \ 0 \le y \le \frac{e^x + e^{-x}}{2} \right\}$$

about coordinate axis x is calculated by formula

$$V = \pi\int_0^1 \frac{\left(e^x + e^{-x}\right)^2}{4}\,dx = \pi\left[\frac{1}{8}\left(e^{2x} - e^{-2x} + 4x\right)\right]_0^1 = \frac{\pi}{8}\left(e^2 + e^{-2} + 4\right).$$

4. Length of a segment of plane curve catenary, which is the graph of function $f(x) = \dfrac{e^x + e^{-x}}{2}$ on interval $\langle 0, 1\rangle$ in fig. 4.14, left, can be calculated as follows

$$L = \int_0^1 \sqrt{1 + \frac{\left(e^x - e^{-x}\right)^2}{4}}\,dx = \left[\frac{\left(e^{2x} - 1\right)\sqrt{2 + e^{2x} + e^{-2x}}}{2\left(1 + e^{2x}\right)}\right]_0^1 = \frac{e^2 - 1}{2e}.$$

5. Surface area of a surface of revolution generated by the revolving segment of a catenary about coordinate axis x, illustrated in fig. 4.14 right, is calculated as the definite integral

$$S = \int_0^1 \left|\frac{e^x + e^{-x}}{2}\right| \sqrt{1 + \frac{\left(e^x - e^{-x}\right)^2}{4}}\,dx =$$

$$= \left[\frac{\left(e^{4x} + 4xe^{2x} - 1\right)\sqrt{2 + e^{2x} + e^{-2x}}}{8e^x\left(1 + e^{2x}\right)}\right]_0^1 = \frac{e^2 + e^{-2} + 1}{8}.$$

Physical applications

Mass of a thin plane region (plate)

Let f be a positive continuous function on interval $\langle a, b\rangle$, and let the area of regular plane region

$$R = \{[x, y] : a \le x \le b, 0 \le y \le f(x)\}$$

be determined as $A(R) = \int_a^b f(x)dx$, while the specific density of the plate material is ρ. Mass m of this plate can be then calculated by means of the formula

$$m(R) = \rho \cdot A(R) = \rho\int_a^b f(x)dx.$$

Let f be a continuous and positive function on an interval $\langle a, b\rangle$, and let the volume of a solid of revolution S generated by the revolving curvilinear trapezoid

$$T = \{[x, y] : a \le x \le b, 0 \le y \le f(x)\}$$

about axis x be $V(S) = \pi\int_a^b f^2(x)dx$, and the specific density of solid material is ρ.

Let S be revolving about the coordinate axis x with the angular velocity ω. Then the following physical characteristics can be calculated for the solid of revolution S by means of definite integrals:

Mass $$m(S) = \rho \cdot V(S) = \pi\rho \int_a^b f^2(x)dx$$

Static moment $$Sx = \pi\rho \int_a^b x \cdot f^2(x)dx$$

Centre of gravity $$T = [x_T, 0, 0],\ x_T = \frac{Sx}{m} = \frac{\int_a^b x \cdot f^2(x)dx}{\int_a^b f^2(x)dx}$$

Moment of inertia $$J = \frac{\pi\rho}{2} \int_a^b f^4(x)dx$$

Kinetic energy $$J = \frac{\pi\rho\omega^2}{4} \int_a^b f^4(x)dx$$

Examples

1. Let a thin plate made from a material with specific density ρ be determined as $R = \left\{ [x, y] : 0 \le x \le a,\ 0 \le y \le \sqrt{b^2 - \frac{b^2}{a^2} x^2} \right\}$. Then the area and mass of this plate are

$$A(R) = \int_0^a \sqrt{b^2 - \frac{b^2}{a^2} x^2}\, dx = \frac{b}{a} \int_0^a \sqrt{a^2 - x^2}\, dx =$$

$$= \frac{b}{2a} \left[x\sqrt{a^2 - x^2} + a^2 \arctan \frac{x}{\sqrt{a^2 - x^2}} \right]_0^a = \frac{1}{4}\pi ab$$

$$m(R) = \rho \cdot A(R) = \frac{1}{4}\pi\rho ab\,.$$

2. Let the solid of revolution S be a part of a paraboloid of revolution determined by the revolving parabolic region bounded by parabola $y = \sqrt{2px}$, $p > 0$, coordinate axis x and line segments on lines $x = 0$, $x = h > 0$. The volume and mass of solid S made from the material with specific density ρ is

$$V(S)=\pi\int_0^h 2pxdx=\pi ph^2,\quad m(S)=\rho\cdot V=\pi\rho ph^2,$$

while the surface area of a part of the parabolic surface of revolution is

$$S=2\pi\int_0^h\left|\sqrt{2px}\right|\sqrt{1+\frac{8p}{x}}dx=2\pi\int_0^h\left|\sqrt{2px+16p^2}\right|dx=\frac{2}{3}\pi\left(\sqrt{p(2h+p)^3}-p^2\right).$$

The static moment is $Sx=\pi\rho\int_0^h 2px^2dx=2\pi\rho p\left[\frac{x^3}{3}\right]_0^h=\frac{2}{3}\pi\rho ph^3$, and the centre of gravity is in the point $T=[x_T, 0, 0]$, $x_T=\frac{Sx}{m}=\frac{\frac{2}{3}\pi\rho ph^3}{\pi\rho ph^2}=\frac{2}{3}h$.

The moment of inertia of solid S revolving about the coordinate axis x with angular velocity ω is

$$J=\frac{\pi\rho}{2}\int_0^h 4p^2x^2dx=2\pi\rho\, p^2\left[\frac{x^3}{3}\right]_0^h=\frac{2}{3}\pi\rho\, p^2h^3,$$

while its kinetic energy is $E=\frac{1}{2}\omega^2J=\frac{1}{3}\pi\rho\omega\, p^2h^3$. The solid is presented in fig. 4.15.

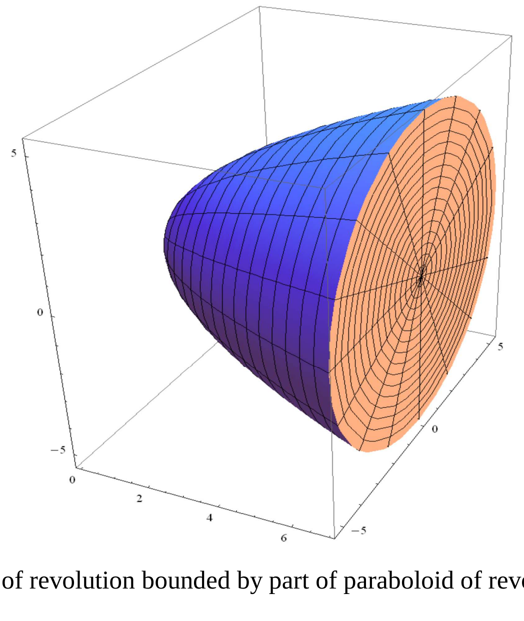


Fig. 4.15. Solid of revolution bounded by part of paraboloid of revolution and disc

4.8 Improper integrals

Integrals on unbounded intervals

Let a function f be defined on an interval $\langle a, \infty)$ and integrable on any interval $\langle a, b\rangle$, $b > a$. If a proper limit $\lim\limits_{b\to\infty}\int_a^b f(x)dx$ exists, then it is called the improper integral of f on $\langle a, \infty)$ and it is denoted by $\int_a^\infty f(x)dx$. Thus

$$\int_a^\infty f(x)dx = \lim_{b\to\infty}\int_a^b f(x)dx.$$

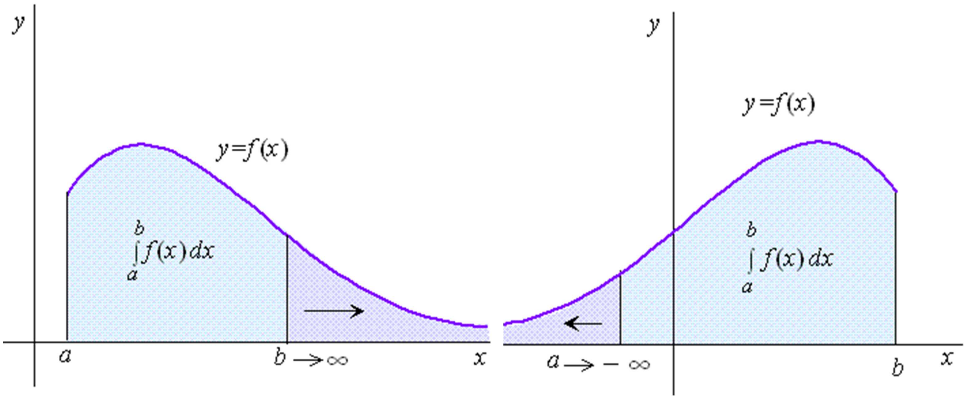


Fig. 4.16. Improper integrals

In this case it is said that this improper integral exists or converges, and in the opposite case if a proper limit does not exist this integral does not exist or diverge. Improper integrals on other types of unbounded intervals are defined similarly.

$$\int_{-\infty}^b f(x)dx = \lim_{a\to-\infty}\int_a^b f(x)dx$$

$$\int_{-\infty}^\infty f(x)dx = \int_{-\infty}^c f(x)dx + \int_c^\infty f(x)dx$$

If both improper integrals on the right hand side exist, their existence does not depend on the choice of $c \in \boldsymbol{R}$, then according to the definition above the integral on the left-hand side also exists and equals to their sum, which does not depend on the choice of c as well.

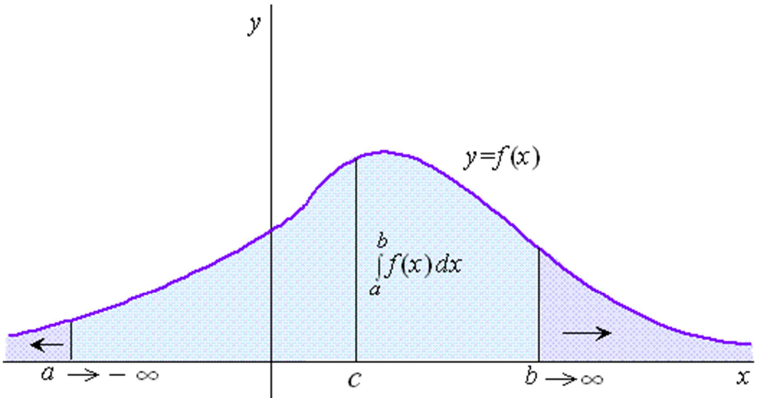


Fig. 4.17. Improper integral

The geometric meaning of improper integrals is analogical to that of the standard definite integrals. We can utilize them for example to calculate areas of unbounded curvilinear trapezoids or plane figures determined by the boundaries in graphs of continuous functions on intervals $(-\infty, b\rangle$, $\langle a, \infty)$ or $(-\infty, \infty)$.

Integrals of unbounded functions

Let function f be integrable on each interval $\langle a, \xi\rangle$, $\xi \in (a, b)$, and unbounded in some left neighbourhood $O_\varepsilon^-(b)$ of point b. If a proper limit $\lim\limits_{\xi\to b^-}\int_a^\xi f(x)dx$ (a real number) exists, then it is called the improper integral of f on $\langle a, b\rangle$. Thus

$$\int_a^b f(x)dx = \lim_{\xi\to b^-}\int_a^\xi f(x)dx\,.$$

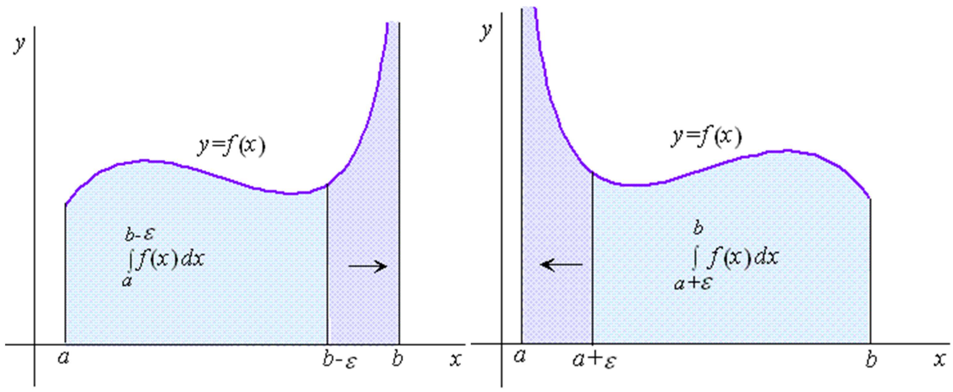


Fig. 4.18. Integrals of unbounded functions

In this case it is said that this improper integral exists or converges, in the opposite case this integral does not exist or diverges.

Analogously if f is integrable on each interval $\langle \xi, b\rangle$, $\xi \in (a, b)$ and unbounded in a right neighbourhood $O_\varepsilon^+(a)$ of point a, then

$$\int_a^b f(x)dx = \lim_{\xi\to a^+}\int_\xi^b f(x)dx\,.$$

Examples

1. Integral

$$\int_0^\infty \frac{dx}{1+x^2} = \lim_{b\to\infty}\int_0^b \frac{dx}{1+x^2} = \lim_{b\to\infty}\left[\arctan x\right]_0^b = \lim_{b\to\infty}\arctan b = \frac{\pi}{2}$$

represents geometrically the area of unbounded region, curvilinear trapezoid below the graph of function $\frac{1}{1+x^2}$ on interval $\langle 0, \infty)$, fig. 4.19.

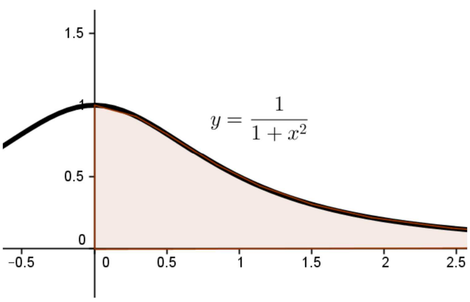


Fig. 4.19. Unbounded region with area π/2

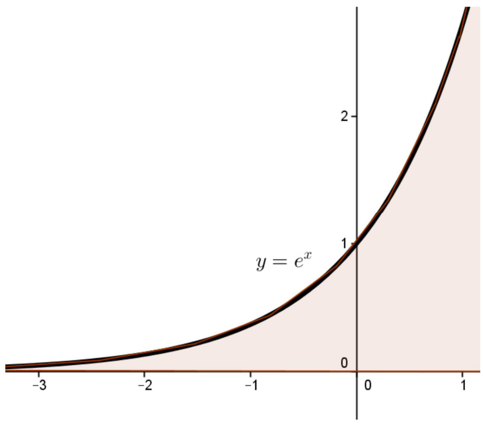


Fig. 4.20. Unbounded region without area

2. Integral

$$\int_{-\infty}^\infty e^x dx = \lim_{a\to-\infty}\int_a^c e^x dx + \lim_{b\to\infty}\int_c^b e^x dx =$$

$$= \lim_{a\to-\infty}\left[e^x\right]_a^c + \lim_{b\to\infty}\left[e^x\right]_c^b = e^c + \lim_{a\to-\infty} e^a - e^c = \infty$$

is diverging, therefore it does not determine the area of unbounded region, illustrated in fig. 4.20.

3. Integral

$$\int_{-1}^{1}\frac{1}{x^2}\,dx=\lim_{c\to 0^-}\int_{-1}^{c}\frac{1}{x^2}\,dx+\lim_{c\to 0^+}\int_{c}^{1}\frac{1}{x^2}\,dx=\lim_{c\to 0^-}\left[-\frac{1}{x}\right]_{-1}^{c}+\lim_{c\to 0^+}\left[-\frac{1}{x}\right]_{c}^{1}=\infty$$

is diverging because of the integrand function that is unbounded on interval $\langle -1, 1\rangle$, and both improper integrals are diverging, as both one-sided limits are improper. Area of unbounded region on fig. 4.21, left, does not exist.

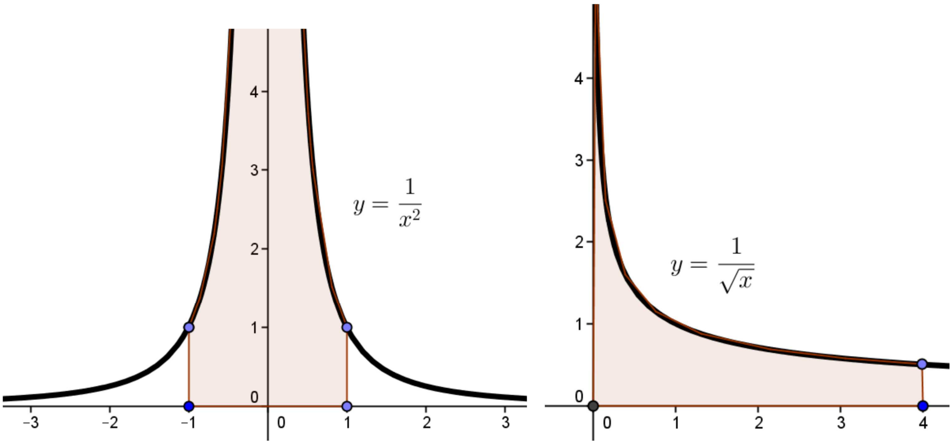


Fig. 4.21. Unbounded functions

4. Integral

$$\int_{0}^{4}\frac{dx}{\sqrt{x}}=\lim_{c\to O^+}\int_{c}^{4}\frac{dx}{\sqrt{x}}=\lim_{c\to O^+}\left[\sqrt{x}\right]_c^4=\lim_{c\to O^+}(2-\sqrt{c})=2$$

is a converging improper integral, determining the area of unbounded curvilinear trapezoid in fig. 4.21, right.

5. Integral

$$\int_{-\infty}^{\infty}e^{-x^2}dx=\lim_{a\to-\infty}\int_{a}^{c}e^{-x^2}dx+\lim_{c\to\infty}\int_{c}^{b}e^{-x^2}dx=\sqrt{\pi}\,,$$

where both improper integrals lead to elliptic integrals that can be evaluated numerically, is converging, and its value is the area of unbounded region illustrated in fig. 4.22.

6. Area of a curvilinear trapezoid in fig. 4.23 bounded by the lines $x = 0$, $x = 4$, $y = 0$ and the graph of function $y=(x-1)^{-\frac{1}{3}}$ is

$$\int_0^4 \frac{1}{\sqrt[3]{x-1}}\,dx = \lim_{c\to 1^-}\int_0^c \frac{1}{\sqrt[3]{x-1}}\,dx + \lim_{c\to 1^+}\int_c^4 \frac{1}{\sqrt[3]{x-1}}\,dx =$$

$$= \left[\frac{3}{2}\sqrt[3]{(x-1)^2}\right]_0^c + \lim_{c\to 1^+}\left[\frac{3}{2}\sqrt[3]{(x-1)^2}\right]_c^4 = \frac{3}{2}\left(1+\sqrt[3]{9}\right).$$

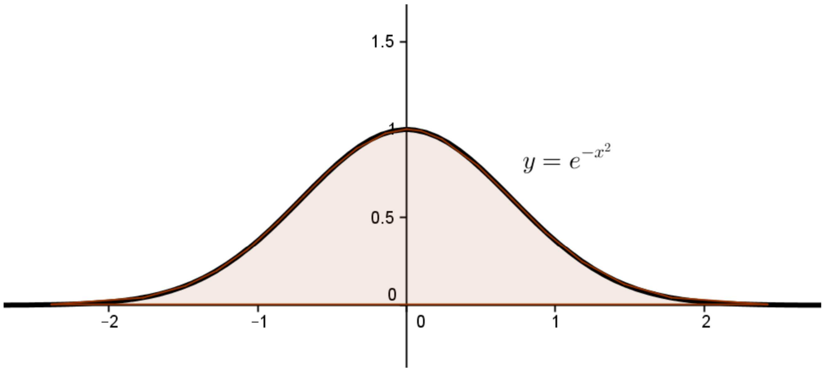


Fig. 4.22. Unbounded region with area $\sqrt{\pi}$

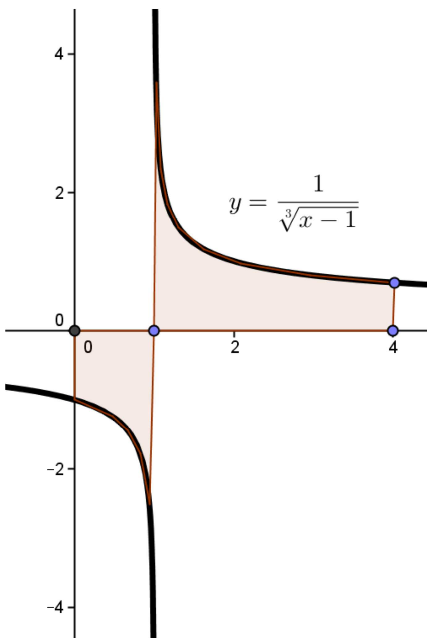


Fig.4.23. Unbounded region

5 Ordinary differential equations

5.1 Basic concepts and definitions

Many physical, chemical, biological and various technical problems lead to mathematical models dealing with solutions of differential equations.
An ordinary differential equation is an equation representing the relation between independent variable x from some set $M \subset \boldsymbol{R}$, unknown function $f(x)$ and at least one of its derivatives $f'(x), f''(x),..., f^{(n)}(x)$, $n \in N$.
Denoting $f(x) = y$, $f'(x) = y'$, ..., $f^{(n)}(x) = y^{(n)}$, the differential equation can be symbolically written in the form

$$F(x, y', ..., y^{(n)}) = 0.$$

The order of a differential equation is the order of the highest derivative which appears in the formula F. Anyhow, none of the derivatives of function $f(x)$ up to the degree $(n - 1)$, nor the function $f(x)$ or independent variable x must appear in the ordinary differential equation of order n, explicitly.
The solution of the differential equation of degree n is any function $y = f(x)$, which, when substituted together with its derivatives into the given differential equation, turns it into an identity on a set M. To solve differential equation means to find all the functions satisfying this equation and to determine set $M = (a, b) \subset \boldsymbol{R}$, which is the domain of definition of all these solutions, or to show that the respective differential equation has no solution.

The following types of solutions of differential equations can be distinguished:

1. general solution, in which the number of appearing constants equals to the order of the equation
2. particular solution, which can be derived from the general solution by an appropriate choice of constants, or a solution resulting from the given initial conditions
3. singular solution, which cannot be obtained from the general solution and is not containing any constant

Methods used to solve differential equations are called integrations of differential equation, function $y = f(x)$, which is the solution of differential equation, is also called the integral of the differential equation.
The graph of a solution is called the integral curve of the given differential equation.

Examples

1. Equation $xy' - 2y = 0$ is the differential equation of order 1, while its general solution is $y = cx^2$, $x \in \boldsymbol{R}$, where c is any real number. The graph of this general solution is the system of parabolas with axes in the coordinate axis y for $c \neq 0$, while one particular solution is the function $y = 0$, for $c = 0$, whose graph is the coordinate axis x, fig. 5.1.

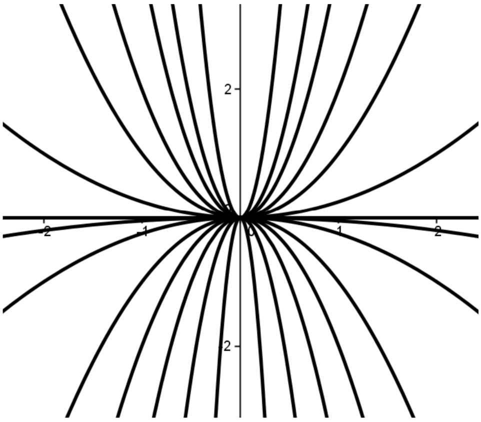


Fig. 5.1. System of integral curves

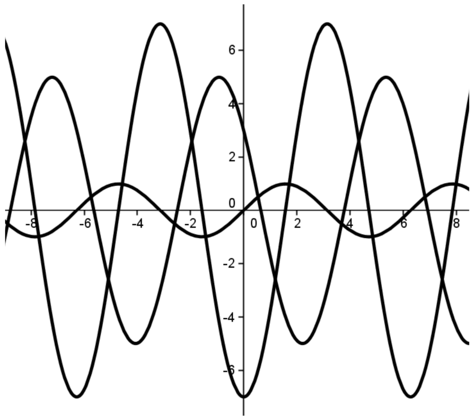


Fig. 5.2. Graphs of particular solutions

2. General solution of the second order differential equation $y'' + y = 0$ is system of functions $y = c_1\cos x + c_2\sin x$, $x \in \boldsymbol{R}$, where c_1, c_2 are arbitrary real constants. Substitution of the function y and its second derivative $y'' = -(c_1\cos x + c_2\sin x)$ in the equation yields equality $0 = 0$. Particular solutions are functions, e.g. $\sin x$, $-7\cos x$, $-4\sin x + 3\cos x$, etc. The integral curves are in fig. 5.2.

3. Second order differential equation $(y'')^2 + y^2 + x^2 = 0$, has no solution, which can be shown easily by contradiction. Let $y = f(x)$ be a solution of this differential equation. Then it holds that $(f''(x))^2 = -(y^2 + x^2)$, which cannot be true for any real number.

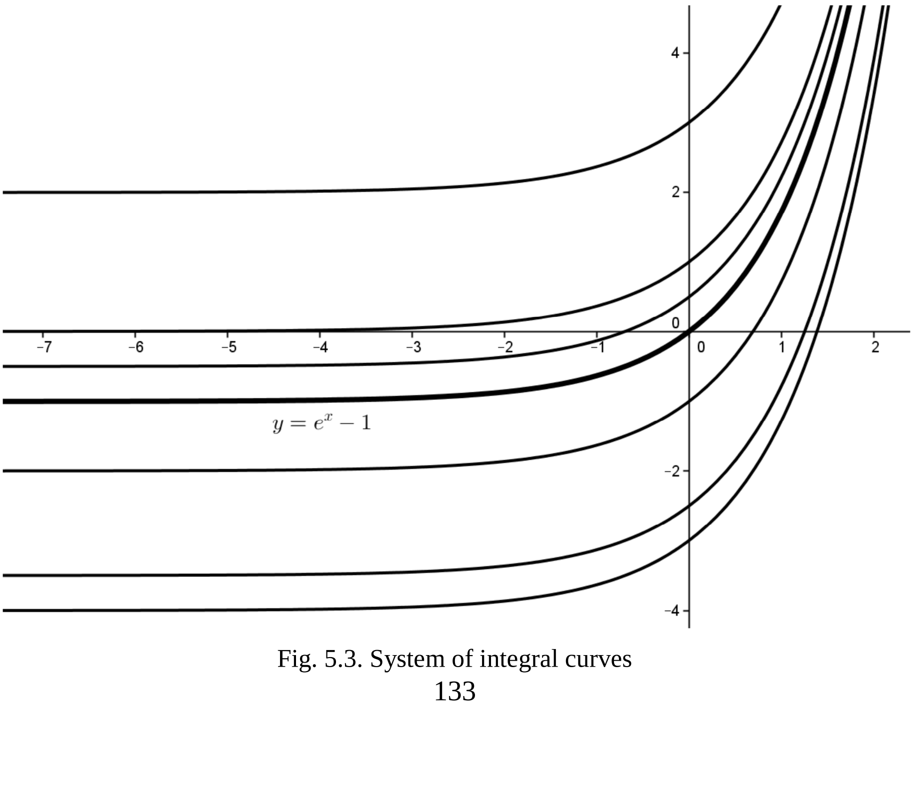


Fig. 5.3. System of integral curves

4. General solution of the first order differential equation $y'y - ye^x = 0$ is one parametric system of functions $y = e^x + c$, $x \in \boldsymbol{R}$, $c \in \boldsymbol{R}$. The solution $y = 0$ is a singular solution, which cannot be obtained from the general solution by any choice of constant c, fig. 5.3.

Cauchy initial problem for differential equation of order 1

To find the solution of a given differential equation of order 1 satisfying the initial condition $y(x_0) = y_0$, while x_0, y_0 are given numbers, is called Cauchy initial problem. This condition can be geometrically interpreted as looking for such particular solution, whose integral curve is passing through point $Q = [x_0, y_0]$ determined by initial condition.

Examples

1. Solution of the Cauchy initial problem $y'y - ye^x = 0$, $y(0) = 0$, is the function $y = e^x - 1$, $x \in \boldsymbol{R}$, whose graph is the integral curve passing through the point [0, 0], fig. 5.3.
2. Solution of Cauchy initial problem $y' = 2x$, $y(1) = 0$ is the function $y = x^2 - 1$, whose graph is a parabola with axis in the coordinate axis y passing through point [1, 0], i.e. one particular solution obtained from the general solution of a given equation in the form $y = x^2 + c$, $x \in \boldsymbol{R}$, $c \in \boldsymbol{R}$, represented geometrically by a system of coaxial parabolas in fig. 5.4.

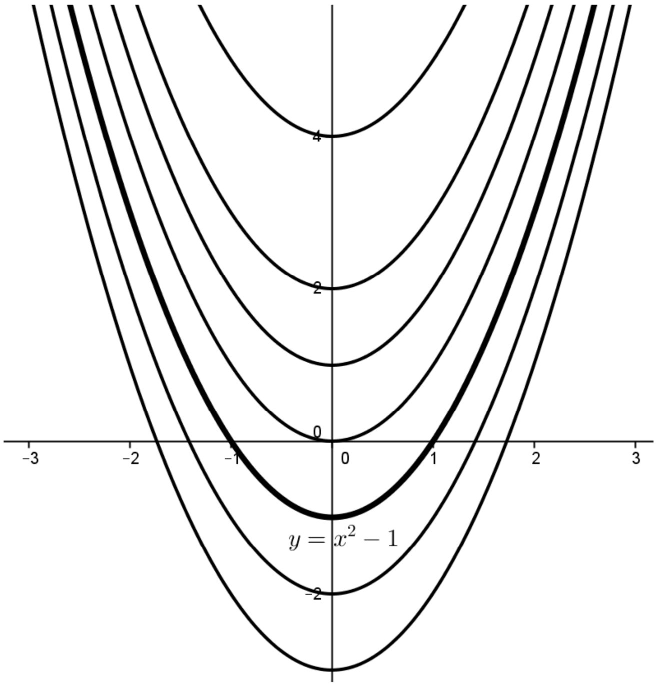


Fig. 5.4. Particular solution

3. Velocity of radioactive decay of a specific matter is proportional to the amount of matter, which is represented as the time function $M(t)$, $t \in \langle 0, \infty)$. This relation can be written as differential equation of order 1

$$\frac{dM}{dt} = -kM \Leftrightarrow M'(t) = -kM, \, t \in <0, \infty),$$

where $k > 0$ is a constant dependent on the radioactive matter and t is time as independent variable. The general solution of equation is function $M(t) = \underline{ce}^{-kt}$, $c > 0$, $t \in \langle 0, \infty)$. Different functions of decay are obtained as particular solutions for specific materials, with c characteristic for the particular matter.

Cauchy initial problem for differential equation of order 2

To find such particular solution of a given differential equation of order 2 that satisfies the initial conditions $y(x_0) = y_0$, $y'(x_0) = y_1$, while x_0, y_0 and y_1 are given numbers, is called Cauchy initial problem for differential equation of order 2. This condition can be geometrically interpreted as looking for such particular solution, whose integral curve is passing through the determined point $Q = (x_0, y_0)$, while the tangent line to the respective integral curve at this point has the slope $k = y'(x_0) = y_1$.

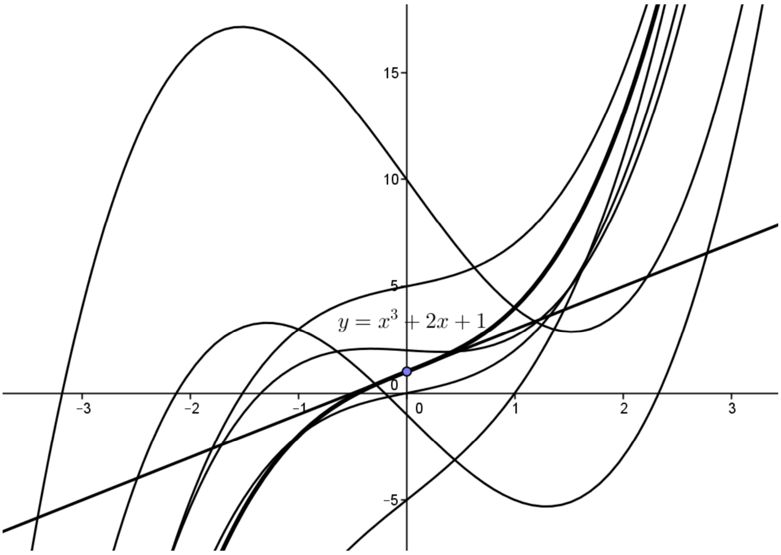


Fig. 5.5. Cauchy initial problem of order 2

Examples

1. Function $y = x^3 + c_1x + c_2$, c_1, $c_2 \in \boldsymbol{R}$ is the general solution of the second order differential equation $y'' = 6x$, where $x \in \boldsymbol{R}$. The particular solution satisfying Cauchy initial conditions $y(0) = 1$, $y'(0) = 2$ can be obtained from the general solution substituting $x = 0$, $y = 1$, $y' = 2$, with the result $c_1 = 2$, $c_2 = 1$, so it appears in the form $y = x^3 + 2x + 1$, see fig. 5.5.

2. The general solution of the second order differential equation $y'' + 4y = \frac{1}{\sin 2x}$ is $y = c_1 \cos 2x + c_2 \sin 2x - \frac{x}{2}\cos 2x + \frac{1}{4}\sin 2x \ln|\sin 2x|$, where c_1, $c_2 \in \boldsymbol{R}$ and $x \in \left(k\frac{\pi}{2}, (k+1)\frac{\pi}{2}\right), k \in \boldsymbol{Z}$. Particular solution satisfying initial conditions $y\left(\frac{\pi}{4}\right) = 1, y'\left(\frac{\pi}{4}\right) = \frac{\pi}{4}$ is function $y_p = \sin 2x - \frac{x}{2}\cos 2x + \frac{1}{4}\sin 2x \ln|\sin 2x|$ defined for $x \in \left(0, \frac{\pi}{2}\right)$ is illustrated is in fig. 5.6.

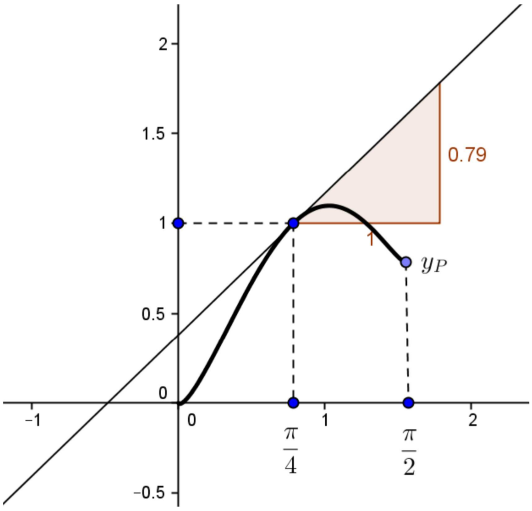


Fig. 5.6. Cauchy initial problem of order 2

While solving differential equations, certain manipulations and transformations must be performed, which transform the former equation to a new one. This new differential equation can have such solutions, that are not solutions of the former differential equation, or some solutions of the former equation are not the solutions of the new one anymore. If each solution of one of the two differential equations is also the solution of the other one on the same set, the equations are said to be equivalent. Such transformation, in which an equivalent equation is obtained from the original differential equation, is called equivalent transformation. In the case that a non-equivalent transformation is applied to solve a differential equation, the resulting differential equation can have more or fewer solutions than the original equation, or the domain of definition of the solution can be different. It is therefore necessary to analyze and specify all solutions thoroughly.

The general solution is not always expressible in an explicit form. Sometimes it can be represented as an equation, which is accepted as a solution in an implicit form

$$F(x, y) = 0.$$

5. 2 Equation with separated and separable variables

Let us start our considerations with the most common and easily solvable differential equations, in which variables x, y, y' are separated or can be easily separated.

The differential equation of the form

$$p(x) + q(y)y' = 0,$$

where $p(x)$ is continuous on (a, b), $q(y)$ is continuous on (c, d) is called differential equation of order 1 with separated variables. Any solution of this differential equation on $J \subset (a, b)$ has the form

$$\int p(x)dx + \int q(y)dy = c,\ c \in \boldsymbol{R},\ c = \text{constant}.$$

If $q(y) \neq 0$ on (c, d), then through each point in the region $D = (a, b) \times (c, d)$ is passing just one integral curve of the given differential equation.

A special case of the differential equations with separated variables are equations of the form

$$y' = f(x),$$

with the general solution

$$y = \int f(x)dx + c, \cdots c \in \boldsymbol{R}.$$

All the integral curves representing general solutions of the above differential equations are the curves defined as shifted graphs of one particular integral curve that is the graph of one particular solution.

Examples

1. Solution of differential equation $2x + \frac{y'}{y} = 0$ is the solution of equation

$$\int 2x dx + \int \frac{1}{y} dy = C,\ C \in \boldsymbol{R},\ y \neq 0,$$

$$x^2 + \ln|y| = C \Rightarrow \ln|y| = C - x^2$$

$$|y| = e^{C - x^2} \Rightarrow y = \pm e^C e^{-x^2} \Rightarrow y = ce^{-x^2}$$

which is the function $y = ce^{x^2}, c \in \boldsymbol{R}, c \neq 0$. Selected integral curves are sketched in fig. 5.7.

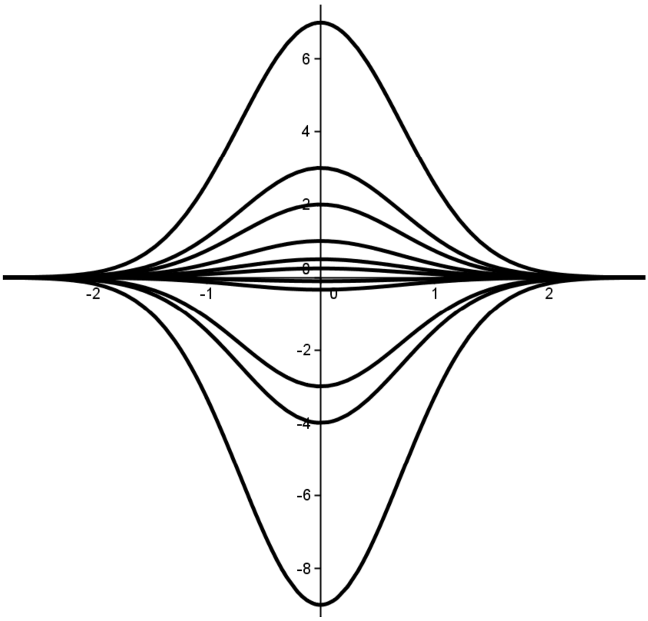


Fig. 5.7. System of integral curves

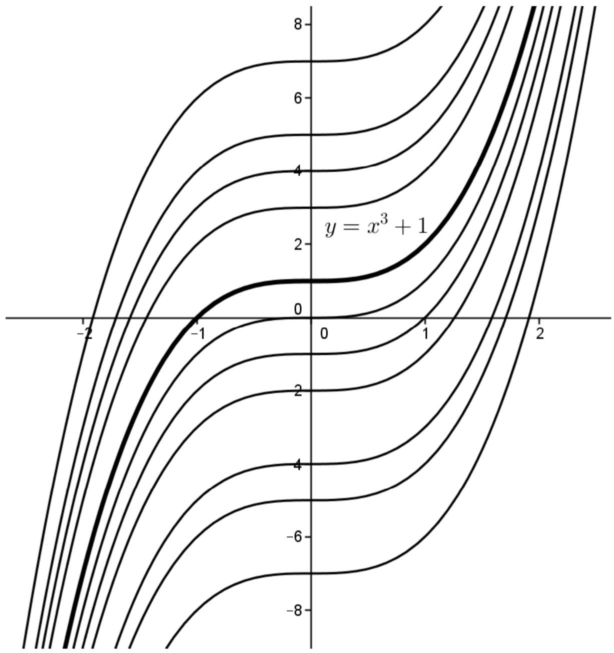


Fig. 5.8. Particular solution

2. Particular solution of the differential equation with separated variables $y' = 3x^2$, which satisfies the initial condition $y = 0$ for $x = -1$ can be derived from the general solution

$$y = \int 3x^2 dx + c = x^3 + c,\ c \in \boldsymbol{R}$$

by substituting the initial condition, $0 = (-1)^3 + c$, where the constant is $c = 1$, so it is the following function $y = x^3 + 1$, see fig. 5.8.

3. Solving differential equation $x + yy' = 0$ we obtain

$$\int x dx + \int y dy = C,\ C \in \boldsymbol{R}$$

$$\frac{x^2}{2} + \frac{y^2}{2} = C \Rightarrow x^2 + y^2 = c \Rightarrow y = \pm\sqrt{c - x^2},\ c \in \boldsymbol{R}^+,\ x \in (-c, c).$$

Integral curves are concentric semicircles with centres in the origin and radii $\sqrt{c}$.

The differential equation of the form

$$p_1(x)\, q_2(y) + q_1(x)\, p_2(y)\, y' = 0,$$

where $p_1(x)$, $q_1(x)$ are continuous on (a, b), $p_2(y)$, $q_2(y)$ are continuous on (c, d) is called the differential equation of order 1 with separable variables.
If $q_1(x).q_2(y) \neq 0$, then this differential equation can be transformed to the differential equation with separated variables

$$\frac{p_1(x)}{q_1(x)} + \frac{p_2(x)}{q_2(x)} y' = 0.$$

These two differential equations are not equivalent in general, as the assumption $q_1(x).q_2(y) \neq 0$ need not be true on the entire interval $(a, b) \times (c, d)$.
If $q_2(y) = 0$ for $y_i = b_i$, while $b_i \in (c, d)$, $i = 1, 2, \ldots, k$, $k \in N$, then functions $y_i = b_i$ are the solutions of the original differential equation, but they are not the solutions of the transformed differential equation.

The solutions of differential equation with separable variables are the functions $y_i = b_i$, $i = 1, 2, \ldots, k$, $k \in N$, where b_i, are the roots of equation $q_2(y) = 0$, and all solutions of the transformed differential equation with separated variables in the form

$$\int \frac{p_1(x)}{q_1(x)} dx + \int \frac{p_2(y)}{q_2(y)} dy = c,\ c \in \boldsymbol{R}\quad c = \text{constant.}$$

If $q(y) \neq 0$ on (c, d), then through each point in the region $D = (a, b) \times (c, d)$ just one integral curve of the given differential equation is passing.

Examples

1. Solution of differential equation $y - xy' = 0$ can be found step by step. Function $y = 0$ is one solution of the given differential equation. Transforming this equation to differential equation with separated variables we now consider $y \neq 0$ obtaining the equation

$$\frac{1}{x} - \frac{1}{y}y' = 0, x \neq 0 .$$

This equation can be solved as follows

$$\int \frac{1}{x}dx - \int \frac{1}{y}dy = C, x \neq 0, C \in \boldsymbol{R}$$

$$\ln|x| - \ln|y| = \ln e^C \Rightarrow \ln|y| = \ln\frac{|x|}{e^C} \Rightarrow |y| = \frac{|x|}{e^C} \Rightarrow y = cx, x \in \boldsymbol{R}, c \in \boldsymbol{R} .$$

Particular solution $y = 0$ of the original equation with separated variables is included in the general solution for value $c = 0$. Integral curves form a bundle of lines with a common point in the origin, fig. 5.9.

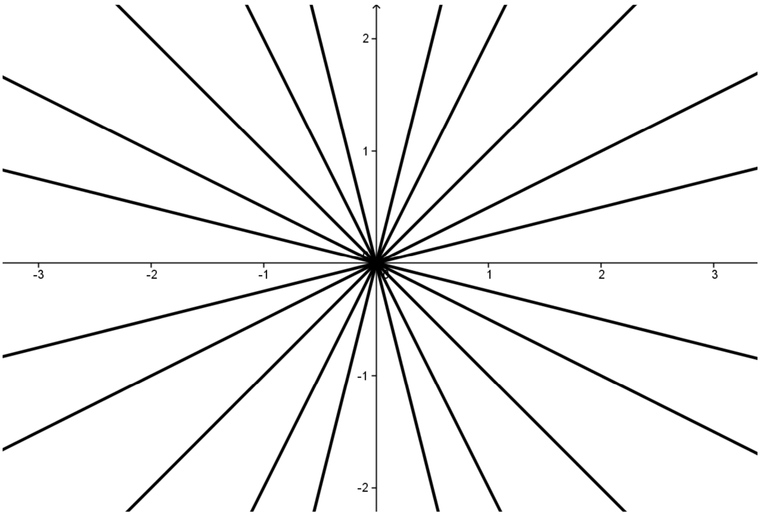


Fig. 5.9. Integral curves

2. To find the particular solution of differential equation $y' = \dfrac{2xy}{1+x^2}$ satisfying the initial condition $y(1) = 1$, let us rewrite the given equation in the form

$$\frac{2xy}{1+x^2} - y' = 0.$$

One solution of this equation is the function $y = 0$.
For $y \neq 0$ we obtain

$$\frac{2x}{1+x^2} - \frac{1}{y} y' = 0$$

$$\int \frac{2x}{1+x^2} dx - \int \frac{1}{y} dy = C$$

$$\ln\left|1+x^2\right| - \ln|y| = \ln e^C$$

$$\ln|y| = \ln\frac{1+x^2}{e^C} \Rightarrow |y| = \frac{1+x^2}{e^C} \Rightarrow y = c\left(1+x^2\right), x \in \boldsymbol{R}, c \in \boldsymbol{R}$$

For the particular solution we evaluate

$1 = c(1 + 1) \Rightarrow$ c = ½ and $y = \dfrac{1+x^2}{2}, x \in \boldsymbol{R}$.

Some of the coaxial parabolas forming the system of integral curves of the general solution, including the particular solution, are presented in fig. 5.10.

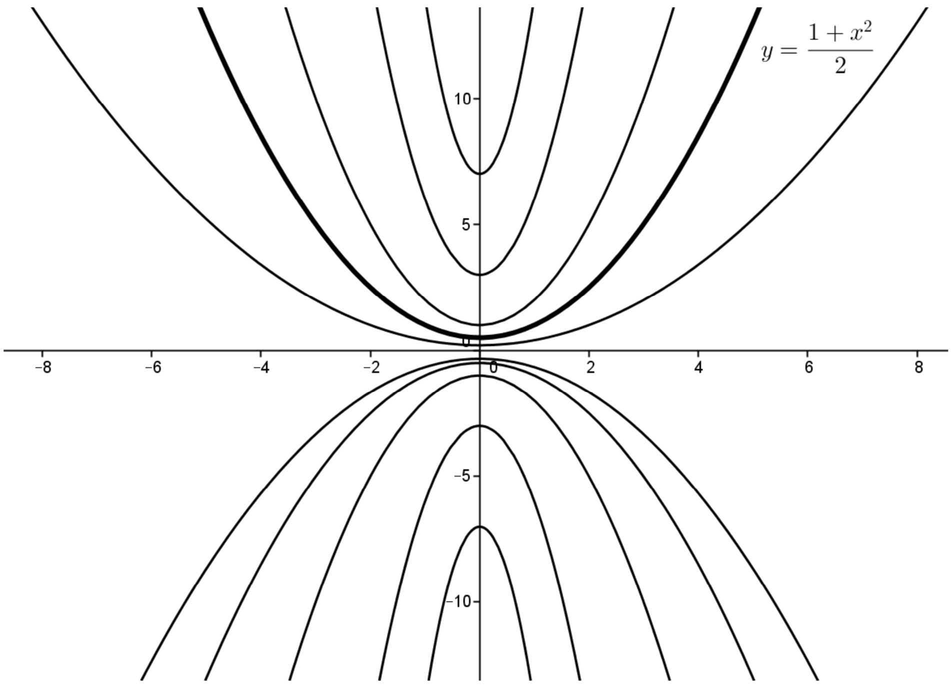


Fig. 5.10. Particular solution

5.3 Linear differential equations of the first order

Differential equation

$$y' + p(x)y = g(x)$$

where $p(x)$ and $g(x)$ are continuous on (a, b) is called a linear differential equation of the first order (or of order 1). If $g(x)$ is a nonzero function, the equation is called non-homogeneous (with a right-hand member). If $g(x) = 0$ on (a, b), it means the equation is in the form

$$y' + p(x)y = 0,$$

then it is called homogeneous (without a right-hand member).
Homogeneous linear differential equation of the first order is a differential equation with separable variables, which can be transformed to the differential equation with separated variables in the form

$$\frac{1}{y}y' + p(x) = 0, y \neq 0.$$

Solutions of the original equation on interval (a, b) are: function $y = 0$ and all solutions of the differential equation with separated variables

$$\int \frac{1}{y}dy + \int p(x)dx = C, C \in \boldsymbol{R}$$

$$\ln|y| - \ln e^C = -\int p(x)dx$$

$$\ln\frac{|y|}{e^C} = -\int p(x)dx \Rightarrow y = ce^{-\int p(x)dx}, c \in \boldsymbol{R}$$

The particular solution $y = 0$ is therefore included in the general solution for a specific value of the constant, $c = 0$.
There exists exactly one solution $y = y(x)$ of the linear homogeneous differential equation of the first order on interval (a, b) satisfying the initial condition $y(x_0) = y_0$, $x_0 \in (a, b)$ in the form

$$y = c_0 e^{-\int p(x)dx}.$$

Examples.

1. Equation $y' - 2xy = 0$ is a homogeneous differential equation of order 1, and its general solution is $y = ce^{\int 2xdx} = ce^{x^2}$, $x \in \boldsymbol{R}$, where c is any real number. The graph of this general solution is a system of exponential curves for $c \neq 0$, while one particular solution is the function $y = 0$, for $c = 0$, whose graph is the coordinate axis x, fig. 5.11.

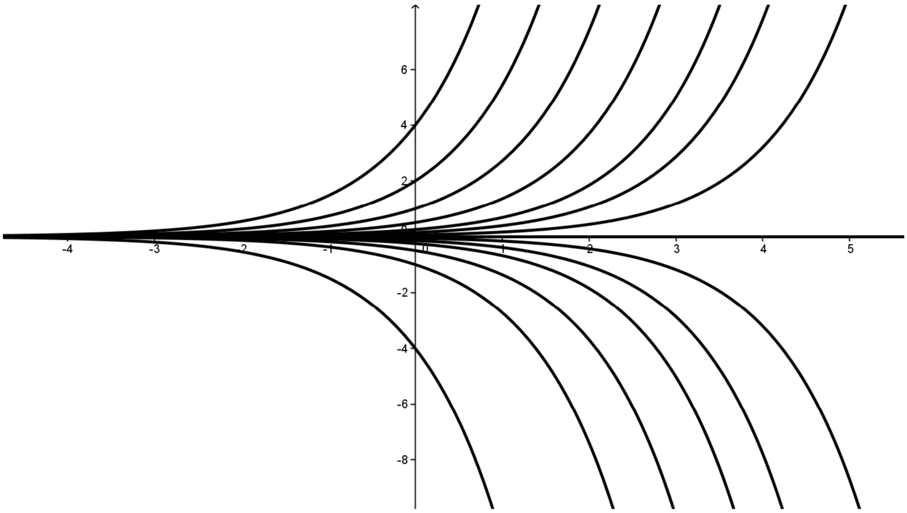


Fig. 5.11. System of integral curves

2. Solution of the first order homogeneous differential equation $y'y - ye^x = 0$ consists of a singular solution $y = 0$ and a general solution of equation $y' - e^x = 0$, which is $y' = e^x$, therefore $y = e^x + c$, $x \in \boldsymbol{R}$, $c \in \boldsymbol{R}$, see fig. 5.3.

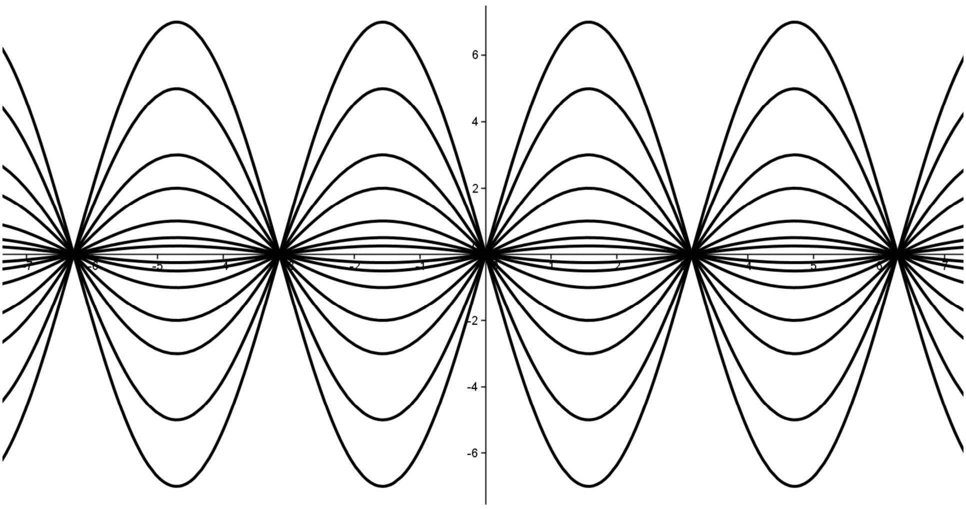


Fig. 5.12. System of integral curves

3. A particular solution of the homogeneous liner differential equation in the form $y' - y\cot x = 0$ is the function $y = \sin x$, while the general solution is determined as $y = c.\sin x$, $c \in \boldsymbol{R}$, $x \neq k\pi$, $k \in \boldsymbol{Z}$, see fig. 5.12. The function $y = 0$ is one particular solution for the constant $c = 0$, its graph is the coordinate axis x. Other particular solution $y(x) = 5\sin x$ is a solution of Cauchy problem $y(\pi/2) = 5$.

The Cauchy initial condition in the form $y(\pi) = k \neq 0$ has no solution, as cot x from the equation is not defined at $x = \pi$ (there exists no solution passing through the point $[\pi, k]$, but all solutions are passing for instance through the point $[\pi, 0]$).

A non-homogeneous linear differential equation $y' + p(x)y = g(x)$ can be solved by the method of variation of a constant. First we find the general solution of the homogeneous linear differential equation (without the right hand member)

$$y = ce^{-\int p(x)dx}, x \in (a,b).$$

Then we look for a solution of the non-homogeneous differential equation in the form

$$y(x) = c(x)e^{-\int p(x)dx}, x \in (a,b)$$

where $c(x)$ inserted into the general solution of the homogeneous differential equation instead of the constant c is such function defined on (a, b) that $y(x)$ satisfies the original equation.
Then, there must exist a derivative of the function $y(x)$ on (a, b)

$$y'(x) = c'(x)e^{-\int p(x)dx} - p(x)c(x)e^{-\int p(x)dx}.$$

Substitution to the original differential equation gives

$$c'(x)e^{-\int p(x)dx} - p(x)c(x)e^{-\int p(x)dx} + p(x)c(x)e^{-\int p(x)dx} = g(x)$$

$$c'(x)e^{-\int p(x)dx} = g(x) \Rightarrow c'(x) = g(x)e^{\int p(x)dx}$$

$$c(x) = \int g(x)e^{\int p(x)dx}dx + c, c \in \boldsymbol{R}.$$

Inserting the obtained form of function $c(x)$ to the general solution $y(x)$ yields

$$y(x) = \left(\int g(x)e^{\int p(x)dx}dx + c\right)e^{-\int p(x)dx}$$

$$y(x) = ce^{-\int p(x)dx} + e^{-\int p(x)dx}\int g(x)e^{-\int p(x)dx}dx, x \in (a,b).$$

This proves that the general solution of a non-homogeneous linear differential equation of the first order consists of a general solution for the corresponding homogeneous linear differential equation of the first order

$$y_H = ce^{-\int p(x)dx}, x \in (a,b), c \in \boldsymbol{R},$$

and one arbitrary particular solution of the original non-homogeneous linear differential equation of the first order in the form

$$y_P = e^{-\int p(x)dx}\int g(x)e^{-\int p(x)dx}dx, x \in (a,b), \; y = y_H + y_P.$$

Exactly one solution of a linear differential equation of the first order on interval (a, b) exists and such, that it satisfies the initial condition

$$y(x_0) = y_0,\ x_0 \in (a, b).$$

Examples

1. General solution for the non-homogeneous differential equation $y' + y = 2x$ can be composed from the general solution of the homogeneous differential equation $y' + y = 0$, and one particular solution of the original non-homogeneous differential equation. The homogeneous equation can be transformed to the differential equation with separated variables $y^{-1}y' + 1 = 0$, with the solution $y_H = ce^{-x}, x \in \boldsymbol{R}, c \in \boldsymbol{R}$. The particular solution of the original non-homogeneous equation can be then represented as the function $y_P = c(x)e^{-x}, x \in \boldsymbol{R}$ with the derivative $y_P' = c'(x)e^{-x} - c(x)e^{-x}$, where $c(x)$ is an unknown function. Inserting it to the original equation yields the equation $c'(x)e^{-x} = 2x$, from which follows $c'(x) = 2xe^{x} \Rightarrow c(x) = 2e^{x}(x-1)$, and $y_P = 2(x-1)$. Finally, the general solution of the original non-homogeneous differential equation is $y = y_H + y_P$, $y = ce^{-x} + 2(x-1), x \in \boldsymbol{R}, c \in \boldsymbol{R}$. Several integral curves are illustrated in fig. 5.13.

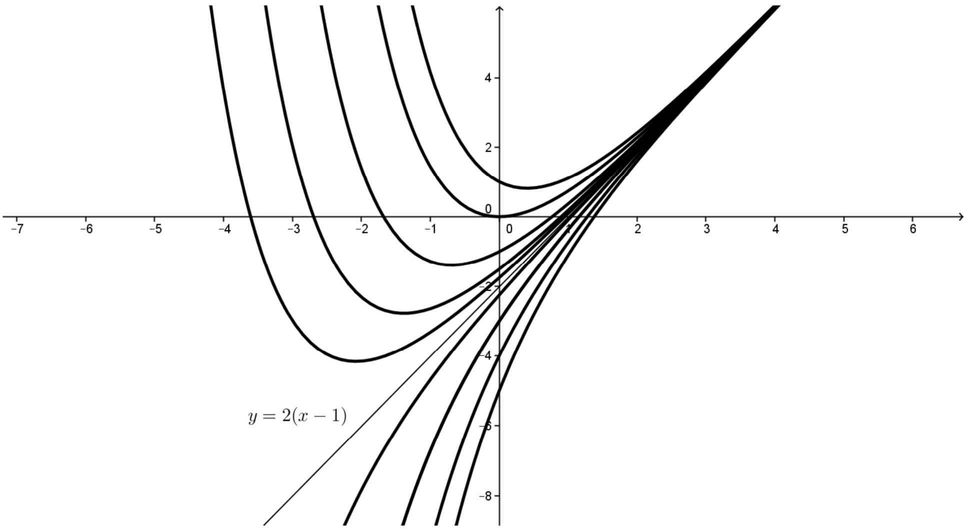


Fig. 5.13. System of integral curves

2. Solution of the initial Cauchy problem $y' + y = 2x$, $y(0) = -2$, can be obtained from the general solution of this differential equation from the previous example. The specific value of the constant c is the solution for the equation $-2 = c - 2$, therefore $c = 0$, and the requested particular solution is the linear function $y = 2(x-1), x \in \boldsymbol{R}$, while its graph is a line in fig. 5.13.

5.4 Linear differential equations of the second order

Differential equation of the form

$$y'' + p_1 y' + p_2 y = g(x),$$

where p_1 and p_2 are real numbers and $g(x)$ is a continuous function on interval (a, b), $g(x) \neq 0$, is called a non-homogeneous liner differential equation of the second order with constant coefficients. A special form of this differential equations for $g(x) \equiv 0$ on (a, b)

$$y'' + p_1 y' + p_2 y = 0,$$

is the homogeneous linear differential equation of the second order with constant coefficients.

If a_0, a_1 are arbitrary real numbers, it can be proved that there exists just one solution of the non-homogeneous (or homogeneous) liner differential equation of the second order with constant coefficients satisfying the initial conditions

$$y(x_0) = a_0,\ y'(x_0) = a_1,\ x_0 \in (a, b)\ (x_0 \in \boldsymbol{R}).$$

Let y_1, y_2 be two arbitrary solutions of the homogeneous differential equation, then any linear combination $c_1 y_1 + c_2 y_2$, $c_1, c_2 \in \boldsymbol{R}$, is also the solution of this equation.

Linear dependence and independence of solutions

Two solutions y_1, y_2 of the homogeneous differential equation are linearly dependent on $\boldsymbol{R}$, if such a number k exists, that for all $x \in \boldsymbol{R}$ it holds that

$$y_1 = k\, y_2 \Leftrightarrow y_1 - k\, y_2 = 0.$$

If the two solutions y_1, y_2 of the homogeneous differential equation are not linearly dependent on $\boldsymbol{R}$, then they are called linearly independent.

Let y_1, y_2 be two arbitrary functions differentiable on an interval J. The determinant

$$W(x) = \begin{vmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{vmatrix}$$

is called Wronskian, or Wronski determinant of functions y_1, y_2 on J.

The functions y_1, y_2 are linearly independent on J, if their Wronskian is non-zero for each $x \in J$.

Any pair of two linearly independent solutions of the homogeneous linear differential equation of the second order is called the fundamental system of solutions of this differential equation.

If y_1, y_2 form the fundamental system of solutions of the homogeneous linear differential equation of the second order, then the general solution of this differential equation is

$$c_1 y_1 + c_2 y_2,$$

where $c_1, c_2 \in \boldsymbol{R}$ are arbitrary constants.

Examples

1. Functions $y_1 = e^x$, $y_2 = 1$ form the fundamental system of solutions of the differential equation $y'' - y' = 0$, as they are both solutions of this equation, and their Wronskian is

$$W(x) = \begin{vmatrix} e^x & 1 \\ e^x & 0 \end{vmatrix} = -e^x \neq 0 \text{ for all } x \in \boldsymbol{R}.$$

   The general solution of the respective differential equation is

$$y = c_1 e^x + c_2,\ x \in \boldsymbol{R}.$$

2. Functions $y_1 = e^x$, $y_2 = xe^x$ form the fundamental system of solutions of the differential equation $y'' - y' = 0$, as they are both solutions of this equation, and their Wronskian is

$$W(x) = \begin{vmatrix} e^x & 1 \\ e^x & 0 \end{vmatrix} = -e^x \neq 0 \text{ for all } x \in \boldsymbol{R}.$$

   The general solution is

$$y = c_1 e^x + c_2,\ x \in \boldsymbol{R}.$$

Suppose that one particular solution of the homogeneous liner differential equation of the second order with constant coefficients is a function $y = e^{rx}$, r is a real constant. Then it holds that

$$y' = re^{rx},\ y'' = r^2 e^{rx}, \text{ which means } y' = ry,\ y'' = r^2 y\,.$$

Inserting function y and its derivatives to the differential equation we obtain

$$r^2 e^{rx} + p_1 re^{rx} + p_2 e^{rx} = 0$$

$$e^{rx}(r^2 + p_1 r + p_2) = 0$$

and because $e^{rx} \neq 0$ for all $x \in \boldsymbol{R}$, the above equation is true if and only if

$$r^2 + p_1 r + p_2 = 0\,.$$

Therefore, the function $y = e^{rx}$, $r \in \boldsymbol{R}$ is the solution of the homogeneous liner differential equation of the second order with constant coefficients, only if r is the root of the above quadratic equation called the characteristic equation of the respective differential equation.

There are the following three possibilities for the roots of the characteristic equation

a) two distinct real roots exist

b) a double real root exists

c) a couple of complex conjugate roots exist.

Case a)

The discriminant of the characteristic equation is positive, $D = p_1^2 - 4p_2 > 0$, and the equation has two distinct real roots $r_1 \neq r_2$. Then it can be proved that the functions $y_1 = e^{r_1 x}, y_2 = e^{r_2 x}$ are both solutions of the respective differential equation, their Wronskian is

$$W(x) = \begin{vmatrix} e^{r_1 x} & e^{r_2 x} \\ r_1 e^{r_1 x} & r_2 e^{r_2 x} \end{vmatrix} = (r_2 - r_1) e^{(r_1 + r_2)x} \neq 0 \text{ for all } x \in \boldsymbol{R},$$

hence y_1 and y_2 are linearly independent and thus they form a fundamental system of solutions, it means

$$y = c_1 e^{r_1 x} + c_2 e^{r_2 x}, \, c_1, c_2 \in \boldsymbol{R}.$$

Case b)

The discriminant of the characteristic equation equals zero, $D = p_1^2 - 4p_2 = 0$, and the equation has one double root $r = \frac{-p_1}{2}$. Then it can be proved that the functions $y_1 = e^{rx}, y_2 = xe^{rx}$ are both solutions of the respective differential equation, their Wronskian is

$$W(x) = \begin{vmatrix} e^{rx} & xe^{rx} \\ re^{rx} & e^{rx}(x+1) \end{vmatrix} = e^{2rx} \neq 0 \text{ for all } x \in \boldsymbol{R},$$

hence y_1 and y_2 are linearly independent and thus they form a fundamental system of solutions, it means

$$y = c_1 e^{rx} + c_2 xe^{rx}, \, c_1, c_2 \in \boldsymbol{R}.$$

Case c)

The discriminant of the characteristic equation is negative, $D = p_1^2 - 4p_2 < 0$, and the equation has two conjugate complex roots $r_1 = a + \mathrm{i}b$, $r_2 = a - \mathrm{i}b$. Then the complex function

$$y = e^{r_1 x} = e^{ax + \mathrm{i}bx} = e^{ax} \cos bx + ie^{ax} \sin bx$$

satisfies the respective differential equation on *R*. It can be proved that both the real and the imaginary part of this complex function *y*, it means functions

$$y_1 = e^{ax} \cos bx, y_2 = e^{ax} \sin bx$$

are real solutions of the respective differential equation.

Wronskian of the two functions is non-zero

$$W(x) = \begin{vmatrix} e^{ax}\cos bx & e^{ax}\sin bx \\ e^{ax}(a\cos bx - b\sin bx) & e^{ax}(a\sin bx + b\cos bx) \end{vmatrix} = be^{2ax} \neq 0$$

for all $x \in \boldsymbol{R}$, hence y_1 and y_2 are linearly independent and form a fundamental system of solutions, therefore the general solution is

$$y = c_1 e^{ax}\cos bx + c_2 e^{ax}\sin bx,\ c_1, c_2 \in \boldsymbol{R}.$$

Examples

1. General solution of the equation $y'' - 4y' + 3y = 0$, will have one of the forms that are described in a) - c), depending on the roots of its characteristic equation $r^2 - 4r + 3 = 0$. The discriminant is $D = (-4)^2 - 4\cdot 3 = 4 > 0$, the equation has two real roots $r_1 = 3$, $r_2 = 1$, and the general solution of the differential equation is $y = c_1 e^{3x} + c_2 e^{x},\ x \in \boldsymbol{R}, c_1, c_2 \in \boldsymbol{R}$. Some integral curves are illustrated in the fig. 5.14.

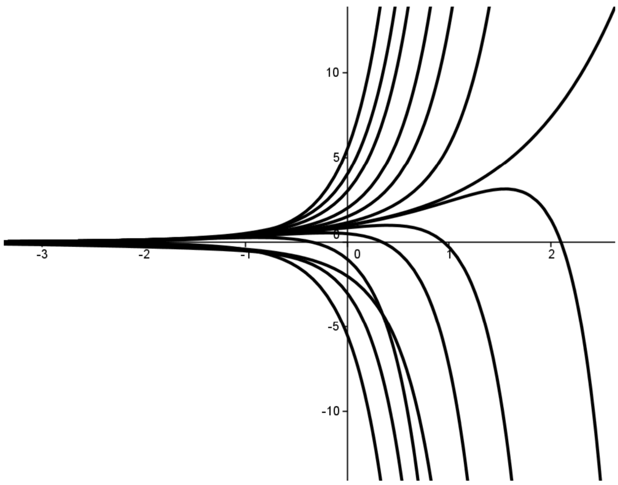


Fig. 5.14. System of integral curves

2. Equation $y'' + 6y' + 9y = 0$, with the characteristic equation $r^2 + 6r + 9 = 0$ has two solutions $y_1 = e^{-3x}$, $y_2 = xe^{-3x}$, as the discriminant is $D = (6)^2 - 4\cdot 9 = 0$, and the equation has one double real root $r = -3$, and the general solution of this differential equation is $y = c_1 e^{-3x} + c_2 xe^{-3x},\ x \in \boldsymbol{R}, c_1, c_2 \in \boldsymbol{R}$. An illustration of several integral curves is in fig. 5.15.

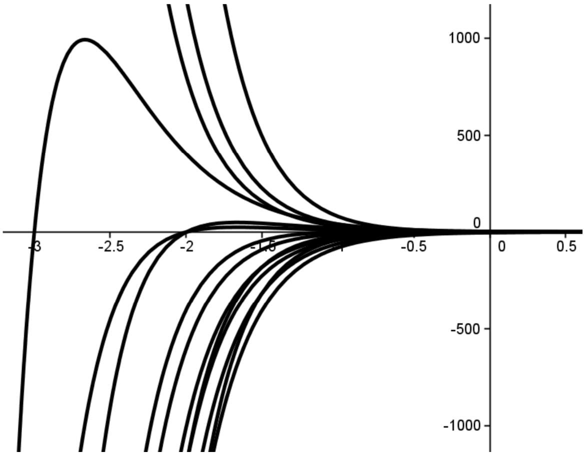


Fig. 5.15. Integral curves of general solution

3. Characteristic equation $r^2 - 6r + 13 = 0$ of the second order differential equation $y'' - 6y' + 13y = 0$ has the discriminant $D = (-6)^2 - 4 \cdot 13 = -16$, therefore its two conjugate roots in the form $r_{1,2} = 3 \pm \mathrm{i}2$ exist. The general solution of the respective differential equation is then in the form

$$y = c_1 e^{3x} \cos 2x + c_2 e^{3x} \sin 2x,\ x \in \boldsymbol{R},\ c_1, c_2 \in \boldsymbol{R},$$

see the graphs in fig. 5.16.

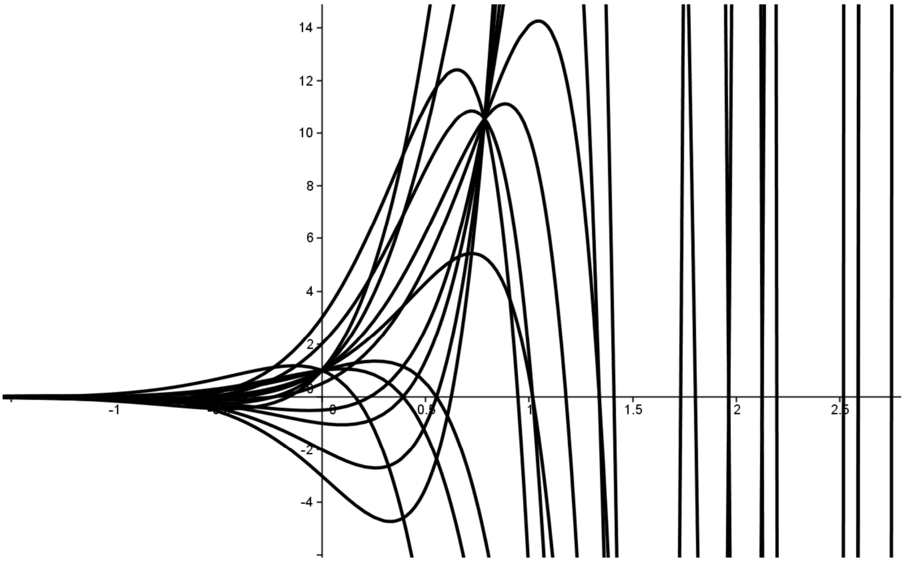


Fig. 5.16. System of integral curves

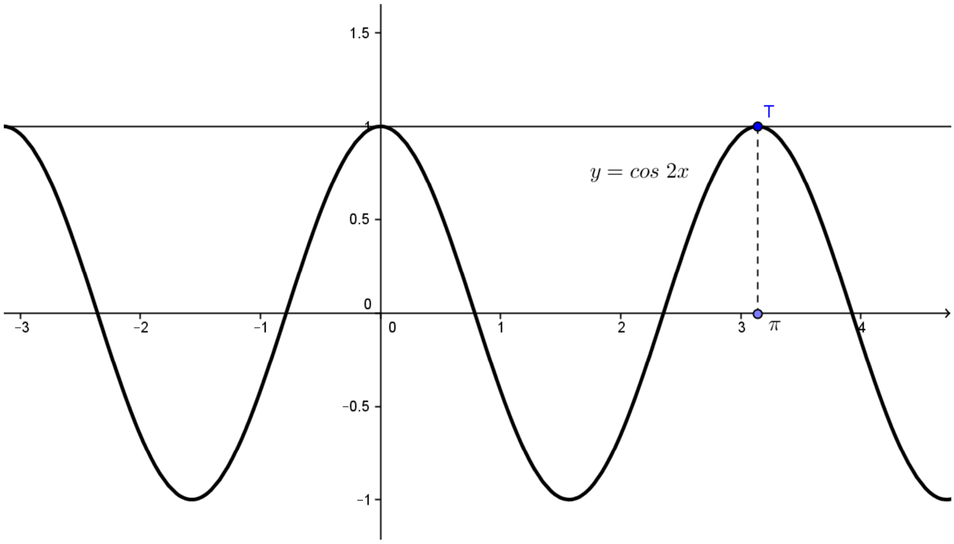


Fig. 5.17. Particular solution

4. Particular solution of the differential equation $y'' + 4y = 0$ satisfying the initial condition $y(\pi) = 1$, $y'(\pi) = 0$, is the function that can be obtained from the general solution specifying the coefficients c_1 and c_2. The characteristic quadratic equation is $r^2 + 4 = 0$, it has two complex conjugate roots $r_{1,2} = \pm 2\mathrm{i}$, therefore the general solution of the differential equation has the form

$$y = c_1 \cos 2x + c_2 \sin 2x,\ x \in \boldsymbol{R},\ c_1, c_2 \in \boldsymbol{R},$$

with the first derivative

$$y' = -2c_1 \sin 2x + 2c_2 \cos 2x\,.$$

Solving the equations

$$0 = -2c_1 \sin 2\pi + 2c_2 \cos 2\pi$$

$$1 = c_1 \cos 2\pi + c_2 \sin 2\pi$$

we obtain $0 = 2c_2$ and $1 = c_1$, from which the particular solution can be determined as function $y = \cos 2x,\ x \in \boldsymbol{R}$. The integral curve of this particular solution is presented in the fig. 5.17.

Vibrations problem plays an important role in modern engineering and physics. There are many cases when vibrations are described with linear differential equations of the second order, having constant coefficients. These equations are used as mathematical models of harmonic motions. The next example is related to the frequently appearing problem on description of a simple harmonic motion.

Example

1. Suppose that a moving body of the mass m is under the action of a force directed toward the state of equilibrium, the magnitude of the force being proportional to the deviation of the state. If we neglect the resistance of the medium, this motion is said to be a simple harmonic motion. To find its law, let us denote the distance from the body to the state of its equilibrium by s, then the force is $F = -as$, a being a positive constant. According to the Newton's 2nd law of motion it holds that

$$m\frac{d^2s}{dt^2} = -as \Rightarrow m\frac{d^2s}{dt^2} + as = 0 \quad (ms'' + as = 0) \quad .$$

   Denoting $k^2 = \frac{a}{m}$, we obtain the equation $s'' + k^2 s = 0$.

   From this it follows that $s = c_1 \cos kt + c_2 \sin kt,\ c_1, c_2 \in \boldsymbol{R}$, which means that s is a periodic function of time t with the period $T = \frac{2\pi}{k}$.

Non-homogeneous linear differential equation of the second order

$$y'' + p_1 y' + p_2 y = g(x),$$

can be solved by means of the associated homogeneous linear differential equation

$$y'' + p_1 y' + p_2 y = 0.$$

Let $Y = c_1 y_1 + c_2 y_2$, where $c_1,\ c_2 \in \boldsymbol{R}$, be the general solution of the associated homogeneous differential equation on $\boldsymbol{R}$, and y_P be the arbitrary solution of the original non-homogeneous differential equation, then

$$y = Y + y_P = c_1 y_1 + c_2 y_2 + y_P$$

is the general solution of the original non-homogeneous differential equation.

There are two methods for finding the particular solution of a non-homogeneous differential equation of the second order.

1. The method of variation of constants

Let $Y = c_1 y_1 + c_2 y_2$, where $c_1,\ c_2 \in \boldsymbol{R}$, be the general solution of the associated homogeneous differential equation on $\boldsymbol{R}$, then the particular solution of the non-homogeneous equation can be found in the form

$$y_P = c_1(x) y_1 + c_2(x) y_2 .$$

Constants c_1 and c_2 are replaced by the unknown functions $c_1(x)$ and $c_2(x)$ such, that y_P satisfies the non-homogeneous equation. Functions $c_1(x)$, $c_2(x)$ must have derivatives on some interval (a, b), as then

$$y_P' = c_1'(x) y_1 + c_1(x) y_1' + c_2'(x) y_2 + c_2(x) y_2'$$

while we will assume that $c_1'(x)y_1 + c_2'(x)y_2 = 0$, from which it follows

$$y_P' = c_1(x)y_1' + c_2(x)y_2'.$$

For the second derivative it holds that

$$y_P'' = c_1'(x)y_1' + c_1(x)y_1'' + c_2'(x)y_2' + c_2(x)y_2''.$$

Substituting y_P and its derivatives to the non-homogeneous differential equation and after some transformations we receive equation

$$c_1(x)\left(y_1'' + p_1y_1' + p_2y_1\right) + c_2(x)\left(y_2'' + p_1y_2' + p_2y_2\right) + c_1'(x)y_1' + c_2'(x)y_2' = g(x).$$

Because y_1 and y_2 are the solutions of the associated homogeneous differential equation, we obtain

$$c_1'(x)y_1' + c_2'(x)y_2' = g(x).$$

The solution of the system of two equations

$$\begin{aligned} c_1'(x)y_1 + c_2'(x)y_2 &= 0 \\ c_1'(x)y_1' + c_2'(x)y_2' &= g(x) \end{aligned}$$

with unknown functions $c_1'(x), c_2'(x)$ always exists, as the determinant of the system is the Wronskian

$$W(x) = \begin{vmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{vmatrix} \neq 0 \text{ for all } x \in \boldsymbol{R}.$$

Using the Cramer rule we obtain the unique solution in the form

$$c_1'(x) = \frac{W_1(x)}{W(x)},\ c_2'(x) = \frac{W_2(x)}{W(x)},$$

where the determinants $W_i\,(x)$ are derived from the determinant $W(x)$ exchanging its i-th column by $\begin{pmatrix} 0 \\ g(x) \end{pmatrix}$ for $i = 1, 2$.

Then the functions $c_1(x)$ and $c_2(x)$ can be determined by simple integration

$$c_1(x) = \int \frac{W_1(x)}{W(x)}\,dx,\ c_2(x) = \int \frac{W_2(x)}{W(x)}\,dx.$$

Particular solution of the non-homogeneous differential equation therefore appears in the following form

$$y_P = y_1\int \frac{W_1(x)}{W(x)}\,dx + y_2\int \frac{W_2(x)}{W(x)}\,dx.$$

Examples

1. Let us find one particular solution of the differential equation $y'' - y' = x + 1$. From the characteristic equation $r^2 - r = 0$ of the associated homogeneous differential equation we can find its general solution $Y = c_1 + c_2 e^x$. Exchanging the constants c_1, c_2 with the functions we get the particular solution of the original non-homogeneous differential equation $y_P = c_1(x) + c_2(x)e^x$, while the functions $c_1(x)$, $c_2(x)$ can be calculated as follows

$$W = \begin{vmatrix} 1 & e^x \\ 0 & e^x \end{vmatrix} = e^x, W_1 = \begin{vmatrix} 0 & e^x \\ x+1 & e^x \end{vmatrix} = -e^x(x+1), W_2 = \begin{vmatrix} 1 & 0 \\ 0 & x+1 \end{vmatrix} = x+1$$

$$c_1(x) = \int \frac{W_1}{W} dx = -\int (x+1) dx = -\frac{x^2}{2} - x,$$
$$c_2(x) = \int \frac{W_2}{W} dx = \int \frac{x+1}{e^x} dx = -(x+2)e^{-x}$$

The particular solution is then in the form $y_P = -\frac{x^2}{2} - 2x - 2$, $x \in \boldsymbol{R}$.

The integral curve is in fig. 5.18.

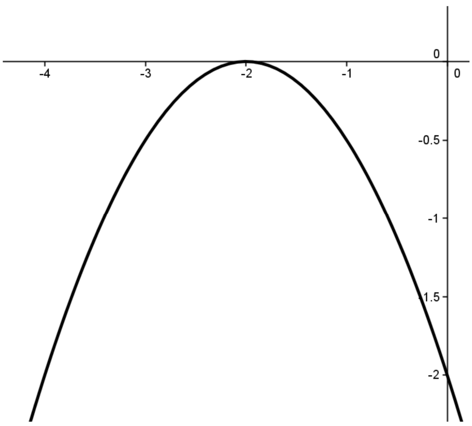


Fig. 5.18. Particular solution

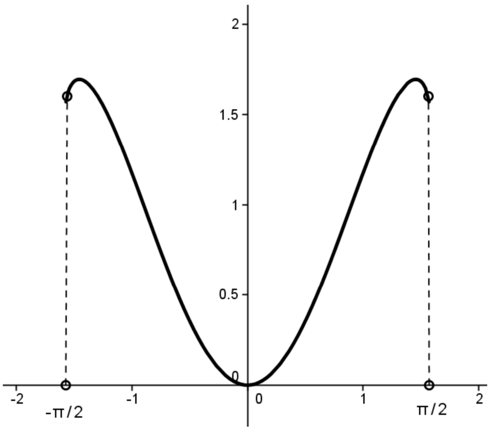


Fig. 5.19. Particular solution

2. To solve the differential equation $y'' + y = \frac{1}{\cos x}$ for all $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ let us find first the general solution $Y = c_1 y_1 + c_2 y_2$ of the associated homogeneous differential equation $y'' + y = 0$. Its characteristic equation is $r^2 + 1 = 0$ and it has the complex conjugate roots $r_{1,2} = \pm \mathrm{i}$, therefore $Y = c_1 \cos x + c_2 \sin x$.

Exchanging the constants c_1, c_2 with functions we get the form of the particular solution of the original non-homogeneous differential equation

$$y_P = c_1(x)\cos x + c_2(x)\sin x \, .$$

Then we calculate the functions $c_1(x)$, $c_2(x)$

$$W = \begin{vmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{vmatrix} = 1$$

$$W_1 = \begin{vmatrix} 0 & \sin x \\ \dfrac{1}{\cos x} & \cos x \end{vmatrix} = \tan x, W_2 = \begin{vmatrix} \cos x & 0 \\ -\sin x & \dfrac{1}{\cos x} \end{vmatrix} = 1$$

$$c_1(x) = \int \frac{W_1}{W} dx = \int \tan x dx = -\ln|\cos x|, \, c_2(x) = \int \frac{W_2}{W} dx = \int 1 dx = x$$

and determine $y_P = -\cos x \ln|\cos x| + x \sin x$, see fig. 5.19.

The general solution of the non-homogeneous differential equation is

$$y = c_1 \cos x + c_2 \sin x + x \sin x - \cos x \ln|\cos x|, \, c_1, c_2 \in \boldsymbol{R}, \, x \in \boldsymbol{R} \, .$$

2. The method of undetermined coefficients

The method is suitable for solving non-homogeneous differential equations with special form of the right-hand terms.

A) If the right-hand member is $g(x) = e^{\alpha x} \cdot P(x)$, where $\alpha \in \boldsymbol{R}$ and $P(x)$ is an polynomial of degree m, then a particular solution of the equation exists

$$y_P = x^k e^{\alpha x} \cdot P^*(x) \, ,$$

where k is the multiplicity of α considered as a root of the characteristic equation and $P^*(x) = b_0 + b_1 x + ... + b_m x^m$ is an unknown polynomial of the same degree as $P(x)$. Coefficients b_0, b_1, ..., b_m are found by the method of the undetermined coefficients.

B) If the right-hand member is $g(x) = e^{\alpha x} \cdot \left(P(x)\cos \beta x + Q(x)\sin \beta x\right)$, where α, $\beta \in \boldsymbol{R}$ and $P(x)$, $Q(x)$ are polynomials, then a particular solution of the equation exists

$$y_P = x^k e^{\alpha x} \cdot \left(P^*(x)\cos \beta x + Q^*(x)\sin \beta x\right),$$

where k is the multiplicity of $\alpha + \mathrm{i}\beta$ considered as a root of the characteristic equation and $P^*(x), Q^*(x)$ are unknown polynomials of the same degree identical with the greater degrees of the polynomials $P(x)$ and $Q(x)$. The coefficients of the polynomials $P^*(x), Q^*(x)$ are found by the method of the undetermined coefficients.

Examples

1. The right-hand term of the equation $y'' - 2y' + y = e^x$ is in the form A), while $\alpha = 1$, and the polynomial $P(x) = 1$ has just the absolute member. The characteristic equation of the associated homogeneous differential equation is $r^2 - 2r + 1 = 0$, it has one double root $r = 1$, therefore $k = 2$ and the particular solution of the equation is $y_P = x^2e^x \cdot A$, with the first two derivatives

$$y_P' = Ae^x(x^2 + 2x) \text{ and } y_P'' = Ae^x(x^2 + 4x + 2).$$

After substitution to the differential equation we obtain

$$Ae^x(x^2 + 4x + 2) - 2Ae^x(x^2 + 2x) + Ae^x x^2 = e^x$$

$$2A = 1 \Rightarrow A = \frac{1}{2}$$

and $y_P = \frac{1}{2}x^2e^x$, see fig. 5.20. The general solution of the associated homogeneous differential equation is in the form

$$Y = c_1e^x + c_2xe^x,\ c_1, c_2 \in \boldsymbol{R},$$

and the general solution of the non-homogeneous equation is

$$y = Y + y_P = c_1e^x + c_2xe^x + \frac{1}{2}x^2e^x,\ c_1, c_2 \in \boldsymbol{R},\ x \in \boldsymbol{R}.$$

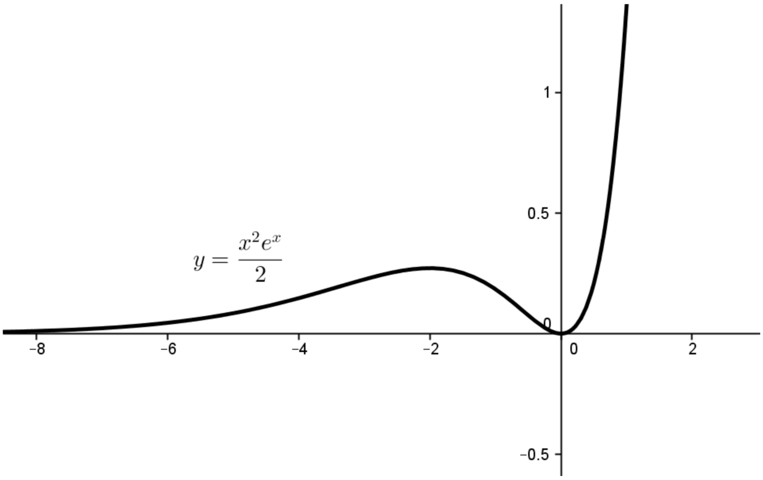


Fig. 5.20. Particular solution

2. Differential equation $y'' + 4y = \sin 2x$ has the right-hand term in the form B). By solving the characteristic equation $r^2 + 4 = 0$ of the associated homogeneous differential equation we obtain the roots $r_{1,2} = \pm 2\mathrm{i}$, which yields the general solution of this differential equation $Y = c_1 \cos 2x + c_2 \sin 2x,\ c_1, c_2 \in \boldsymbol{R}$. The special form of the particular solution of the non-homogeneous differential equation can be determined as

$$y_P = x^1 e^{0x}(A\cos 2x + B\sin 2x) = x(A\cos 2x + B\sin 2x),$$

according to $\alpha = 0$, $\beta = 2$ in the root r_1 of the characteristic equation with multiplicity $k = 1$, and both $P(x)$, $Q(x)$ are the polynomials of degree 0. Then the first two derivatives of the function y_P are

$$\begin{aligned} y'_P &= (A + 2Bx)\cos 2x + (B - 2Ax)\sin 2x, \\ y''_P &= 4(B - Ax)\cos 2x - 4(A + Bx)\sin 2x \end{aligned},$$

whereas inserting these to the equation we obtain coefficients A and B

$$\begin{aligned} &4B\cos 2x - 4A\sin 2x = \sin 2x \\ &B = 0, -4A = 1 \Rightarrow A = -\frac{1}{4} \end{aligned}.$$

Particular solution is $y_P = -\frac{1}{4} x\cos 2x$, the integral curve is in fig. 5.21.

Finally, the general solution of the non-homogeneous differential equation appears in the form

$$y = Y + y_P = c_1 \cos 2x + c_2 \sin 2x - \frac{1}{4} x\cos 2x,\ c_1, c_2 \in \boldsymbol{R}, x \in \boldsymbol{R}.$$

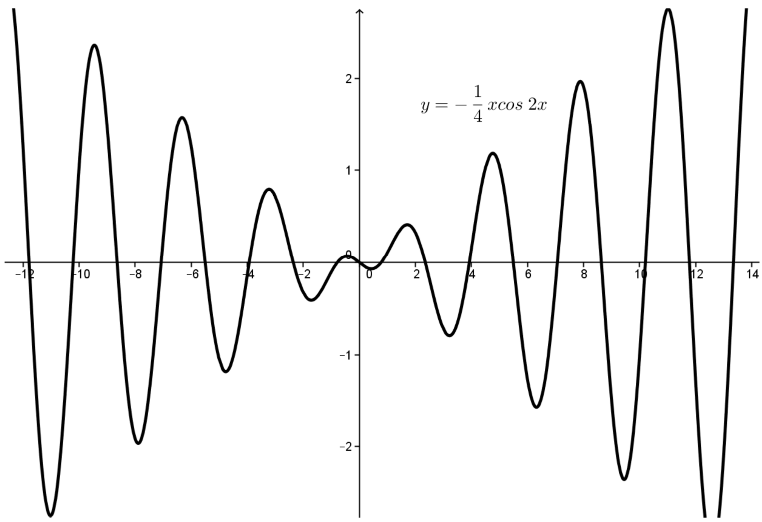


Fig. 5.21. Particular solution

The advantage of the method of the undetermined coefficients is that this method does not require integration and its application (in the case of the special form of the right-hand member) is mostly considerably simpler than the method of the variation of constants. The disadvantage is that this method is restricted to the case of the special form of the right-hand members; hence it is not always possible to use it, contrary to the method of the variation of the constants that is general, and can be used in the case of any form of the right-hand term of the respective differential equation.

Example

1. Particular solution of the differential equation $y'' - 2y' + 2y = e^x x \cos x$ satisfying the initial conditions $y(0) = 1,\ y'(0) = 1$ can be obtained with the method of the undetermined coefficients, as the right-hand term has the special form B). The characteristic equation $r^2 - 2r + 2 = 0$ of the associated homogeneous differential equation has the roots $r_{1,2} = 1 \pm i$, therefore the general solution of the homogeneous equation is $Y = c_1 e^x \cos x + c_2 e^x \sin x,\ c_1, c_2 \in \boldsymbol{R}$. One solution of the non-homogeneous equation can be found according to B) for $\alpha = 1$, $\beta = 1$ in the root r_1 of the characteristic equation with multiplicity $k = 1$, while the degree of the polynomial $P(x)$ is 1, and the degree of $Q(x)$ is 0, so

$$y_P = xe^x\left[(Ax + B)\cos x + (Cx + D)\sin x\right].$$

Then inserting the first two derivatives

$$y_P' = e^x \begin{Bmatrix} \left[(A+C)x^2 + (2A+B+D)x + B\right]\cos x + \\ + \left[(C-A)x^2 + (2C+D-B)x + D\right]\sin x \end{Bmatrix}$$

$$y_P'' = e^x \begin{Bmatrix} \left[2Cx^2 + (4A+4C+2D)x + 2(A+B+C)\right]\cos x + \\ + \left[-2Ax^2 + (-4A-2B+4C)x + 2(-B+C+D)\right]\sin x \end{Bmatrix}$$

to the non-homogeneous differential equation we obtain the equation

$$2e^x\left[(A + D + 2Cx)\cos x + (-B + C - 2Ax)\sin x\right] = e^x x \cos x$$

with solutions in coefficients A, B, C, D

$$2(A + D) = 0,\ 4C = 1,\ 2(-B + C) = 0,\ -4A = 0$$

$$A = 0,\ B = \frac{1}{4},\ C = \frac{1}{4},\ D = 0 \qquad ,$$

from which one solution of the non-homogeneous equation can be represented as

$$y_P = \frac{1}{4} e^x (x\cos x + x^2 \sin x) .$$

The general solution of the non-homogeneous equation is

$$y = c_1 e^x \cos x + c_2 e^x \sin x + \frac{1}{4} e^x (x\cos x + x^2 \sin x), c_1, c_2 \in \boldsymbol{R}$$

and the particular solution satisfying the initial conditions has the coefficients $c_1 = 1$, $c_2 = -1/4$,

$$y_B = e^x \left(\cos x - \frac{1}{4} \sin x + \frac{1}{4} x\cos x + \frac{1}{4} x^2 \sin x \right), x \in \boldsymbol{R} .$$

Applying the method of variation of constants we estimate the solution of the non-homogeneous equation according to A) as

$$y_P = c_1(x) e^x \cos x + c_2(x) e^x \sin x$$

and calculate the functions $c_1(x)$, $c_2(x)$ using Wronskian

$$W = \begin{vmatrix} e^x \cos x & e^x \sin x \\ e^x (\cos x - \sin x) & e^x (\sin x + \cos x) \end{vmatrix} = e^{2x}$$

and the determinants

$$W_1 = \begin{vmatrix} 0 & e^x \sin x \\ e^x x\cos x & e^x (\sin x + \cos x) \end{vmatrix} = -e^{2x} x \sin x \cos x,$$

$$W_2 = \begin{vmatrix} e^x \cos x & 0 \\ e^x (\cos x - \sin x) & e^x x\cos x \end{vmatrix} = e^{2x} x \cos^2 x$$

The resulting functions

$$c_1(x) = \int \frac{W_1}{W} dx = -\int x \sin x \cos x dx = \frac{1}{4} x \cos 2x - \frac{1}{8} \sin 2x,$$

$$c_2(x) = \int \frac{W_2}{W} dx = \int x \cos^2 dx = \frac{x^2}{4} + \frac{1}{8} \cos 2x + \frac{1}{4} x \sin 2x$$

inserted to the particular solution give one solution of the non-homogeneous equation

$$y_P = e^x \cos x\left(\frac{1}{4}x\cos 2x - \frac{1}{8}\sin 2x\right) +$$
$$+ e^x \sin x\left(\frac{x^2}{4} + \frac{1}{8}\cos 2x + \frac{1}{4}x\sin 2x\right)$$

and the final general solution of the non-homogeneous equation is

$$y = e^x \cos x\left(c_1 + \frac{1}{4}x\cos 2x - \frac{1}{8}\sin 2x\right) +$$
$$+ e^x \sin x\left(c_2 + \frac{x^2}{4} + \frac{1}{8}\cos 2x + \frac{1}{4}x\sin 2x\right).$$

For the particular solution determined by the initial conditions we obtain the values of both constants by substitution into y and y', $c_1 = 1$, $c_2 = -1/8$, and the particular solution illustrated in fig. 5.22 is

$$y_A = e^x \cos x\left(1 + \frac{1}{4}x\cos 2x - \frac{1}{8}\sin 2x\right) +$$
$$+ e^x \sin x\left(-\frac{1}{8} + \frac{x^2}{4} + \frac{1}{8}\cos 2x + \frac{1}{4}x\sin 2x\right).$$

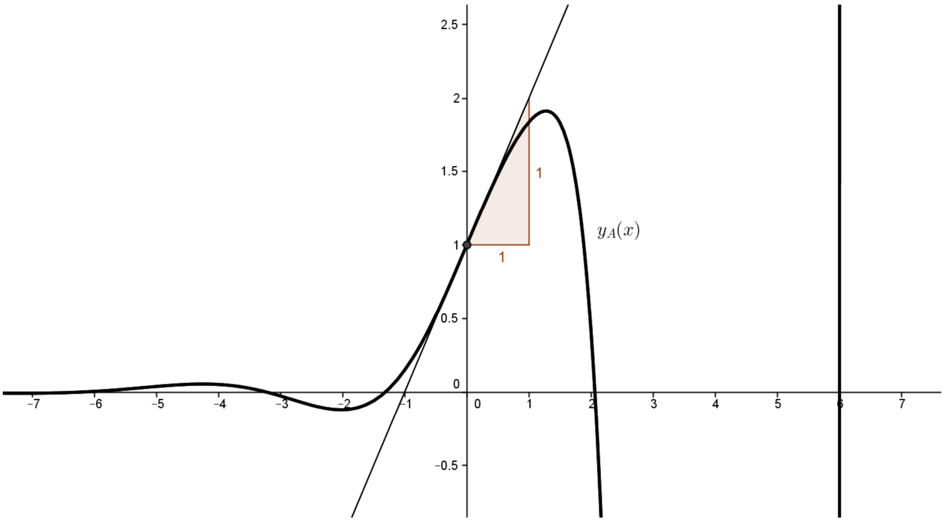


Fig. 5.22. Graph of particular solution

doc. RNDr. Daniela Velichová, CSc.

MATHEMATICS I

Vydala Slovenská technická univerzita v Bratislave v Nakladateľstve STU,
Bratislava, Vazovova 5, v roku 2014.

Edícia vysokoškolských učebníc

Rozsah 164 strán, 134 obrázkov, 3 tabuľky, 7,876 AH, 8,112 VH,
1. vydanie, edičné číslo 5758, tlač Nakladateľstvo STU v Bratislave.

85 – 206 – 2014

ISBN 978-80-227-4130-9