1. The basic theorems of differential calculus

Local extrema of functions. Fermat's theorem

Local extrema of functions

18A/37:04 (12:31)

DEFINITION 1.

Let f act from E to $\mathbb{R}$, let the point $x \in E$ be the limit point of the set E. It is said that the point x_0 is the *local minimum point* of the function f if there exists a neighborhood U_{x_0} of the point x_0 such that for any $x \in U_{x_0} \cap E$ the inequality $f(x) \geq f(x_0)$ holds:

$$\exists\, U_{x_0} \quad \forall\, x \in U_{x_0} \cap E \quad f(x) \geq f(x_0).$$

The value of the function f at the point x_0 is called the *local minimum*.

The local maximum point is defined similarly. The point x_0 is called the *local maximum point* of the function f if

$$\exists\, U_{x_0} \quad \forall\, x \in U_{x_0} \cap E \quad f(x) \leq f(x_0).$$

The value of the function f at the local maximum point is called the *local maximum*.

We also introduce the concepts of strict local maximum and minimum.

DEFINITION 2.

Let f act from E to $\mathbb{R}$, let the point $x_0 \in E$ be the limit point of the set E. The point x_0 is called the *point of a strict local minimum* of the function f if there exists a punctured neighborhood $\overset{\circ}{U}_{x_0}$ of the point x_0 such that for any $x \in \overset{\circ}{U}_{x_0} \cap E$ the inequality $f(x) > f(x_0)$ holds:

$$\exists\, \overset{\circ}{U}_{x_0} \quad \forall\, x \in \overset{\circ}{U}_{x_0} \cap E \quad f(x) > f(x_0).$$

The point of a strict local maximum is defined similarly. The point x_0 is called the *point of a strict local maximum* of the function f if

$$\exists\, U_{x_0} \quad \forall\, x \in \overset{\circ}{U}_{x_0} \cap E \quad f(x) < f(x_0).$$

It should be noted that when we give definitions of the points of strictly local minimum and maximum, punctured neighborhoods of these points are considered, since strict inequalities of the form $f(x) > f(x_0)$ or $f(x) < f(x_0)$ cannot be fulfilled for $x = x_0$.

The points of local minima and maxima are called the *points of local extrema*, and the values at these points are called the *local extrema* of this function. As for minima and maxima, strict and non-strict local extrema can be defined.

Recall the definition of the interior point of a set. A point x_0 is called an *interior point* of the set E if it is simultaneously the limit point of the sets $E^-_{x_0} = \{x \in E : x < x_0\}$ and $E^+_{x_0} = \{x \in E : x > x_0\}$.

DEFINITION 3.

Let f act from E to $\mathbb{R}$, $x_0 \in E$. It is said that the point x_0 is a *point of an interior local minimum, maximum*, or *extremum* if it is a point of a local minimum, maximum, or extremum and at the same time it is an interior point of the set E.

Fermat's theorem

`18A/49:35 (02:40), 18B/00:00 (13:02)`

Theorem (Fermat's theorem on interior local extrema).

Let the function f act from E to $\mathbb{R}$, let $x_0 \in E$ be the point of interior local extremum of the function f. Let's suppose, moreover, that the function f is differentiable at the point x_0. Then its derivative vanishes at this point: $f'(x_0) = 0$.

Remark.

Under the conditions of the theorem, it is not assumed that the point x_0 is a point of strict local extremum, but it is required that it be a point of interior local extremum.

Rolle's theorem, Lagrange theorem, and Cauchy's mean value theorem

Rolle's theorem

`18B/13:02 (11:12)`

Theorem (Rolle's theorem).

Let the function f be defined and continuous on the segment $[a, b]$ and differentiable on the interval (a, b). Let the function take the same values at the endpoints of the segment: $f(a) = f(b)$. Then there exists a point $\xi \in (a, b)$ for which $f'(\xi) = 0$.

Lagrange theorem

`18B/24:14 (09:45)`

Theorem (Lagrange theorem).

Let the function f be defined and continuous on the segment $[a, b]$ and differentiable on the interval (a, b). Then there exists a point $\xi \in (a, b)$ for which the following relation holds:

$$f(b) - f(a) = f'(\xi)(b - a). \tag{1}$$

Remark.

The Lagrange theorem is a generalization of Rolle's theorem, since in the case $f(a) = f(b)$ the left-hand side of equality (1) turns to 0, which immediately implies that $f'(\xi) = 0$.

Corollaries of the Lagrange theorem

`19A/00:00 (18:01)`

Corollary 1.

If the function f is defined and continuous on the segment $[a, b]$, differentiable on the interval (a, b), and for any $x \in (a, b)$ its derivative vanishes, $f'(x) = 0$, then the function f is a constant on the segment $[a, b]$:

$$\forall\, x \in [a, b] \quad f(x) = c.$$

Remark.

The converse statement (that the derivative of the constant vanishes) follows directly from the definition of the derivative.

Corollary 2.

If the function f is defined and continuous on the segment $[a, b]$, differentiable on the interval (a, b), and for any $x \in (a, b)$ its derivative is non-negative, $f'(x) \geq 0$, then the function f is non-decreasing on the segment $[a, b]$: for any $x_1, x_2 \in [a, b]$, $x_1 < x_2$, the inequality $f(x_1) \leq f(x_2)$ holds.

Remark.

The following three statements can be proved in a similar way.

If the function f is defined and continuous on the segment $[a, b]$, differentiable on the interval (a, b), and for any $x \in (a, b)$ its derivative is non-positive, $f'(x) \le 0$, then the function is non-increasing on the segment $[a, b]$.

If the function f is defined and continuous on the segment $[a, b]$, differentiable on the interval (a, b), and for any $x \in (a, b)$ its derivative is positive, $f'(x) > 0$, then the function is increasing on the segment $[a, b]$.

If the function f is defined and continuous on the segment $[a, b]$, differentiable on the interval (a, b), and for any $x \in (a, b)$ its derivative is negative, $f'(x) < 0$, then the function is decreasing on the segment $[a, b]$.

Corollary 3.

Let the function f be defined and continuous on the segment $[a, b]$, differentiable on the interval (a, b), its derivative is not equal to 0 on the interval (a, b) and is continuous on this interval. Then the function f is strictly monotonous on the segment $[a, b]$ and has the inverse function f^{-1} acting from the segment $[c, d] = f([a, b])$ into the segment $[a, b]$. The inverse function is continuous on the segment $[c, d]$, differentiable on the interval (c, d), and has the same monotonicity type as the function f.

Cauchy's mean value theorem

`19A/18:01 (13:01)`

Theorem (Cauchy's mean value theorem).

Let two functions $x(t)$ and $y(t)$ be defined and continuous on the segment $[\alpha, \beta]$, differentiable on the interval (α, β), and, in addition, $x'(t) \ne 0$ for $x \in (\alpha, \beta)$. Then there exists a point $\tau \in (\alpha, \beta)$ for which the following relation holds:

$$\frac{y(\beta) - y(\alpha)}{x(\beta) - x(\alpha)} = \frac{y'(\tau)}{x'(\tau)}. \tag{2}$$

Remark.

This theorem is a generalization of the Lagrange theorem, since a formula similar to formula (1) for the Lagrange theorem can be obtained from formula (2) if we put $y(t) = f(t)$, $x(t) = t$ in formula (2).

2. Taylor's formula

Taylor's formula for polynomials and for arbitrary differentiable functions

Taylor's formula for polynomials

19A/31:02 (09:58)

Let's consider the polynomial $P_n(x)$ of degree $n \in \mathbb{N}$:

$$P_n(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n.$$

Let's select some point $x_0 \in \mathbb{R}$. It is known from the course of algebra that any polynomial can be expanded in powers of $(x - x_0)$; the degree of the polynomial will not change:

$$P_n(x) = c_0 + c_1(x - x_0) + c_2(x - x_0)^2 + \cdots + c_n(x - x_0)^n. \tag{1}$$

We want to obtain formulas for the coefficients c_k, $k = 0, \ldots, n$, using the differentiation operation.

To find the coefficient c_0, it is enough to calculate the value of the polynomial at the point x_0:

$$c_0 = P_n(x_0).$$

To find the coefficient c_1, we firstly differentiate the polynomial:

$$P_n'(x) = c_1 + 2c_2(x - x_0) + 3c_3(x - x_0)^2 + \cdots + nc_n(x - x_0)^{n-1}.$$

The coefficient c_1 is a free term of the derivative $P_n'(x)$ and, therefore, to find it, it suffices to calculate the value of the derivative at the point x_0:

$$c_1 = P_n'(x_0).$$

Let's find the second derivative of the polynomial:

$$P_n''(x) = 2c_2 + 2 \cdot 3c_3(x - x_0) + \cdots + (n-1)nc_n(x - x_0)^{n-2}.$$

Substituting the value $x = x_0$ into this derivative, we obtain the formula for the coefficient c_2:

$$c_2 = \frac{P_n''(x_0)}{2}.$$

In this case, the formula contains not only the value of the derivative at the point x_0, but also the factor $1/2$.

A factor of $1/6$ will appear in the formula for the coefficient c_3. This factor is more convenient to represent as $1/3!$ using the factorial function $n! = 1 \cdot 2 \cdot 3 \cdots n$:

$$c_3 = \frac{P_n'''(x_0)}{3!}.$$

Continuing the process of differentiation, we finally obtain a derivative of order n, which contains a single term:

$$P_n^{(n)}(x) = 2 \cdot 3 \cdots (n-1)nc_n.$$

Thus, for the coefficient c_n, we obtain the formula

$$c_n = \frac{P_n^{(n)}(x_0)}{n!}.$$

Taking into account that $0! = 1$, all the formulas obtained for the coefficients c_k can be written in the following general form:

$$c_k = \frac{P_n^{(k)}(x_0)}{k!}, \quad k = 0, 1, \ldots, n.$$

Substituting the representations for the coefficients c_k into formula (1), we obtain the *Taylor's formula for the polynomial* $P_n(x)$:

$$P_n(x) = \sum_{k=0}^{n} \frac{P_n^{(k)}(x_0)}{k!}(x - x_0)^k. \tag{2}$$

This formula allows us to obtain the expansion of the polynomial $P_n(x)$ in powers of $(x - x_0)$ using the values of the derivatives of the polynomial of order 0 up to n at the point x_0. Note that all derivatives of the polynomial $P_n(x)$ of higher orders ($n+1$, $n+2$, $\ldots$) vanish.

The version of the Taylor's formula (2) when $x_0 = 0$ is also called the *Maclaurin formula*.

Taylor's formula for arbitrary differentiable functions

19B/00:00 (07:55)

The Taylor's formula found earlier for polynomials (2) is an exact equality: both left-hand and right-hand sides in this equality contain polynomials of degree n. This equality holds for all $x \in \mathbb{R}$.

Suppose now that, instead of the polynomial $P_n(x)$, we consider an arbitrary function $f(x)$ defined on some interval containing the point x_0. Also suppose that the function f is differentiable at the point x_0 up to the order n.

Then we can no longer write relation (2) in the form of equality, but we can introduce into consideration the quantity $r_n(x_0, x)$, by which the function $f(x)$ differs from the sum given on the right-hand side of (2):

$$r_n(x_0, x) \overset{\text{def}}{=} f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k.$$

The value $r_n(x_0, x)$ is called the *remainder term of the Taylor's formula* for the function f.

Using the remainder term, we can write the *Taylor's formula for an arbitrary differentiable function* f as follows:

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + r_n(x_0, x). \tag{3}$$

If for some $n \in \mathbb{N}$, $x_0 \in \mathbb{R}$, $x \in \mathbb{R}$, the value $r_n(x_0, x)$ is small, then this means that the function f can be approximated in the point x by a polynomial of degree n according to the Taylor's formula, that is, we can obtain an approximation for the function f in the form of a simpler function (a polynomial).

We have reason to expect that the remainder term $r_n(x_0, x)$ will be small, at least in a situation where the point x is close to the point x_0. Indeed, taking $n = 1$, we obtain

$$r_n(x_0, x) = f(x) - \sum_{k=0}^{1} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k =$$
$$= f(x) - \big(f(x_0) + f'(x_0)(x - x_0)\big).$$

Since $n = 1$ and, therefore, the function f is differentiable at the point x_0, we obtain, by the definition of differentiability, that the right-hand side of the last equality is $o(x-x_0)$ as $x \to x_0$. This means that the remainder term in this case approaches 0, as $x \to x_0$, faster than the linear function $(x - x_0)$, that is, it is small enough for x close to x_0.

We can also expect that while increasing n, that is, in a situation where the function is differentiable more times, the rate of approach to 0 of the remainder term, as $x \to x_0$, will be even higher. However, in order to prove this, we need to study the properties of the remainder term in more detail.

Various representations of the remainder term in the Taylor's formula

Representation of the remainder term in the form of Lagrange

19B/25:40 (12:51)

We give without proof one of the representations the remainder term in the Taylor's formula:

$$r_n(x_0, x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}.$$

This representation of the remainder term is called the *Lagrange form of the remainder term*. It is interesting in that it is similar to the term in the Taylor's formula corresponding to $k = n + 1$, except that the derivative is found not at the point x_0, but at some point ξ from the interval (x_0, x).

Let's write the Taylor's formula with the remainder term in the Lagrange form:

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}. \tag{4}$$

Representation of the remainder term in the Peano form

20A/00:00 (20:31)

We noted earlier that for $n = 1$ the remainder term of the Taylor's formula has the form $o(x - x_0)$, $x \to x_0$. It turns out that similar representations for the remainder term in the form of little-o can also be obtained for other values of n.

Theorem (on the Taylor's formula with the remainder term in the Peano form).

Let the function f be n times continuously differentiable on the segment $[x_0, x]$. Then the following expansion of the function f by the Taylor's formula takes place:

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + o((x - x_0)^n), \quad x \to x_0. \tag{5}$$

The representation of the remainder term $r_n(x_0, x) = o((x - x_0)^n)$, $x \to x_0$, used in this formula is called the *remainder term in the Peano form*. Thus, when expanding the function f by the Taylor's formula up to the derivative of order n, the remainder term decreases, as $x \to x_0$, faster than the function $(x - x_0)^n$.

Expansions of elementary functions by the Taylor's formula in a neighborhood of zero

Expansions of functions e^x, $\sin x$, $\cos x$

20A/20:31 (19:50)

FUNCTION e^x.

Since $(e^x)^{(n)} = e^x$, $n = 0, 1, 2, \ldots$, we obtain that the derivatives of this function of any order are equal to 1 at the point 0. Therefore, the expansion of the function e^x by the Taylor's formula at the point $x_0 = 0$ with the remainder term in the Peano form will be as follows:

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + o(x^n) =$$

$$= \sum_{k=0}^{n} \frac{x^k}{k!} + o(x^n), \quad x \to 0.$$

From the obtained formula, the previously proved equivalence $e^x \sim 1 + x$, $x \to 0$, follows, since for $n = 1$ the expansion takes the form $e^x = 1 + x + o(x)$, $x \to 0$.

REMARK.

This and all subsequent expansions of elementary functions are valid not only for positive, but also for negative values of x belonging to the domain of definition of the function.

FUNCTION $\sin x$.

Let's sequentially find the derivatives of the function $\sin x$ at the point 0. The function $\sin x$ itself takes the value 0 at the point 0. Its first derivative is $\cos x$, so it equals 1 at the point 0. The second derivative is $-\sin x$, it equals 0 at the point 0. Finally, the third derivative is $-\cos x$, it equals -1 at the point 0. The fourth derivative coincides with the original function $\sin x$, therefore, starting from it, the set of values at the point 0 will be repeated: 0, 1, 0, -1, ...

So, we obtain that even-order derivatives are 0 at the point 0, and odd-order derivatives take alternating values of 1 and -1, starting from 1. Therefore, the expansion of the function $\sin x$ by the Taylor's formula at the point $x_0 = 0$ with the remainder term in the Peano form will be as follows:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + \frac{(-1)^n x^{2n+1}}{(2n+1)!} + o(x^{2n+2}) =$$

$$= \sum_{k=0}^{n} \frac{(-1)^k x^{2k+1}}{(2k+1)!} + o(x^{2n+2}), \quad x \to 0.$$

The remainder term has the form $o(x^{2n+2})$, since we can take into account one more (zero-valued) term corresponding to $k = 2n+2$ in the sum.

It should be noted that in the obtained expansion of the function $\sin x$ there are only odd powers of x, starting at x in the first power, and their signs alternate.

From the obtained formula, the previously proved equivalence $\sin x \sim x$, $x \to 0$, follows, since for $n = 0$ the expansion takes the form $\sin x = x + o(x^2)$, $x \to 0$.

FUNCTION $\cos x$.

With successive differentiation of the function $\cos x$, we will obtain the following functions (starting with the zero derivative): $\cos x$, $-\sin x$, $-\cos x$, $\sin x$, $\cos x$, $-\sin x, \dots$ At the point 0, these functions take the following values: 1, 0, -1, 0, 1, 0, -1, 0, ... In this case, the derivatives of odd order are 0 at 0, and the derivatives of even order take alternating values of 1 and -1, starting from 1. Therefore, the expansion of the function $\cos x$ by the Taylor's formula at the point $x_0 = 0$ with the remainder term in the Peano form will be as follows:

$$\cos x = 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \dots + \frac{(-1)^n x^{2n}}{(2n)!} + o(x^{2n+1}) =$$
$$= \sum_{k=0}^{n} \frac{(-1)^k x^{2k}}{(2k)!} + o(x^{2n+1}), \quad x \to 0.$$

The remainder term has the form $o(x^{2n+1})$, since we can take into account one more (zero-valued) term corresponding to $k = 2n+1$ in the sum.

It should be noted that the expansion of the function $\cos x$ contains only even powers of x, starting with $x^0 = 1$, and their signs alternate.

From this formula, the previously proved equivalence $\cos x \sim 1 - x^2/2$, $x \to 0$, follows, since for $n = 1$ the expansion takes the form

$$\cos x = 1 - x^2/2 + o(x^3), \quad x \to 0.$$

Expansions of the functions $\ln(1+x)$ and $(1+x)^\alpha$

20B/00:00 (15:52)

FUNCTION $\ln(1+x)$.

In this case, we find the expansion of the logarithm function in a neighborhood of point 1; moreover, the estimate $x > -1$ must be satisfied for x.

Let's calculate several initial derivatives of the function $\ln(1+x)$ at the point 0 and substitute them in the corresponding terms of the Taylor's formula:

$$\big(\ln(1+x)\big)^{(0)}\Big|_{x=0} = \ln 0 = 0, \quad f(0) = 0;$$

$$\big(\ln(1+x)\big)'\Big|_{x=0} = \frac{1}{1+x}\Big|_{x=0} = 1, \quad f'(0)x = x;$$

$$\big(\ln(1+x)\big)''\Big|_{x=0} = -\frac{1}{(1+x)^2}\Big|_{x=0} = -1, \quad \frac{f''(0)x^2}{2} = -\frac{x^2}{2};$$

$$\big(\ln(1+x)\big)'''\Big|_{x=0} = \frac{2}{(1+x)^3}\Big|_{x=0} = 2, \quad \frac{f'''(0)x^3}{3!} = \frac{x^3}{3};$$

$$\big(\ln(1+x)\big)^{(4)}\Big|_{x=0} = -\frac{2\cdot 3}{(1+x)^4}\Big|_{x=0} = -3!, \qquad \frac{f^{(4)}(0)x^4}{4!} = -\frac{x^4}{4}.$$

Thus, the terms have alternating signs in this expansion. In addition, in the denominator, instead of the factorial, only one factor remains, since all other factors are cancelled out with the coefficients of the corresponding derivatives:

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots + \frac{(-1)^{n-1}x^n}{n} + o(x^n) =$$
$$= \sum_{k=1}^{n} \frac{(-1)^{k-1}x^k}{k} + o(x^n), \quad x \to 0.$$

The resulting expansion of the function $\ln(1+x)$ contains all powers of x, starting from the first power, and their signs alternate. Moreover, the denominator does not have factorials.

From this formula, the previously proved equivalence $\ln(1+x) \sim x$, $x \to 0$, follows, since for $n = 1$ the expansion takes the form $\ln(1+x) = x + o(x)$, $x \to 0$.

Function $(1+x)^\alpha$, $\alpha \neq 0$.

In this case, we also find the expansion of the function in a neighborhood of point 1; moreover, any real number, except 0, can be taken as α.

Let's use the formula for the derivative of a power function of order n:

$$\big((1+x)^\alpha\big)^{(n)} = \alpha(\alpha-1)\cdots(\alpha-n+1)(1+x)^{\alpha-n}, \quad n = 0, 1, 2, \dots$$

For derivatives at the point 0, we will sequentially obtain the values 1, α, $\alpha(\alpha-1)$, $\alpha(\alpha-1)(\alpha-2)$, … Therefore, the expansion of the function $(1+x)^\alpha$ by the Taylor's formula at the point $x_0 = 0$ with the remainder term in the Peano form will be as follows:

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \ldots +$$
$$+ \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n + o(x^n) =$$
$$= \sum_{k=0}^{n} \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}x^k + o(x^n), \quad x \to 0.$$

Note that if $\alpha \in \mathbb{N}$, then, starting from some order, all derivatives vanish, and we obtain the version of the Taylor's formula for polynomials in which the remainder term equals 0.

From this formula, the previously proved equivalence $(1+x)^\alpha \sim 1+\alpha x$, $x \to 0$, follows, since for $n = 1$ the expansion takes the form

$$(1+x)^\alpha = 1 + \alpha x + o(x), \quad x \to 0.$$

3. Application of differential calculus to the study of functions

Local extrema of functions

A necessary condition for the existence of a local extremum

`20B/38:53 (06:18)`

The necessary condition for the existence of an interior local extremum at the point x_0 for the function f can be obtained using previously proved Fermat's theorem if we change its formulation as follows.

Theorem (a necessary condition for the existence of a local extremum).

Let the function f be defined and continuous in some neighborhood U_{x_0} of the point x_0 (this, in particular, means that the point x_0 is the interior point of the domain of the function f). Let the point x_0 be the point of the interior local extremum of the function f. Then either the function f is not differentiable at the point x_0, or the function f is differentiable at the given point and $f'(x_0) = 0$.

Proof.

The theorem follows immediately from Fermat's theorem. □

It follows from this theorem that if the function f is differentiable at the point x_0, but its derivative at this point does not vanish, then this point cannot be a point of the interior local extremum of the function f.

Interior points at which the function is non-differentiable or the derivative vanishes are called *points suspected for a local extremum.* However, a point suspected for an extremum may not be an extremum point.

For example, for the function $f(x) = x^3$, the point 0 is a point suspected for an extremum, since $f'(0) = 2x|_{x=0} = 0$, but it is not a local extremum point, since the inequality $f(x) < f(0)$ holds for all $x < 0$, and the inequality $f(x) > f(0)$ holds for all $x > 0$.

Thus, the formulated necessary condition for the existence of a local extremum is not a sufficient condition.

The first sufficient condition for the existence of a local extremum

`21A/17:34 (07:51)`

In a sufficient condition considered in this section, the differential properties of the function are analyzed not at the point of a local extremum, but in its neighborhood.

Theorem (first sufficient condition for the existence of a local extremum).

Let the function f be differentiable in some punctured neighborhood $\mathring{U}_{x_0}$ of the point x_0. Then the presence or absence of a local extremum at the point x_0 is determined by the signs of the derivative $f'(x)$ in the left-hand and the right-hand neighborhoods ($U^-_{x_0}$ and $U^+_{x_0}$) of the point x_0 as shown in table 3.

Table 3. Signs of a derivative and local extrema

$U_{x_0}^-$	$U_{x_0}^+$	Local extremum
$f'(x) > 0$	$f'(x) > 0$	no extremum
$f'(x) > 0$	$f'(x) < 0$	strict local maximum
$f'(x) < 0$	$f'(x) > 0$	strict local minimum
$f'(x) < 0$	$f'(x) < 0$	no extremum

Proof.

This statement follows from the second corollary of the Lagrange theorem: a function increases on an interval on which its derivative is positive, and decreases on an interval on which its derivative is negative. Thus, if the derivative changes sign when passing through the point x_0, then this point is a point of strict local extremum, and if the sign does not change, then there is no extremum. $\square$

Convex functions

Definitions of convex functions

21B/00:00 (16:10)

Definition 1 of a convex function.

Let the function f be defined on the interval (a, b). A function f is called *convex upwards* (or *concave*) on the interval (a, b) if the graph of the secant drawn through the points $\bigl(x_1, f(x_1)\bigr)$ and $\bigl(x_2, f(x_2)\bigr)$ lies below the function graph on the interval (x_1, x_2) for any points $x_1, x_2 \in (a, b)$.

Similarly, a function f is called *convex downwards* (or just *convex*) on the interval (a, b) if the graph of the secant drawn through the points $\bigl(x_1, f(x_1)\bigr)$ and $\bigl(x_2, f(x_2)\bigr)$ lies above the function graph in the interval (x_1, x_2) for any points $x_1, x_2 \in (a, b)$.

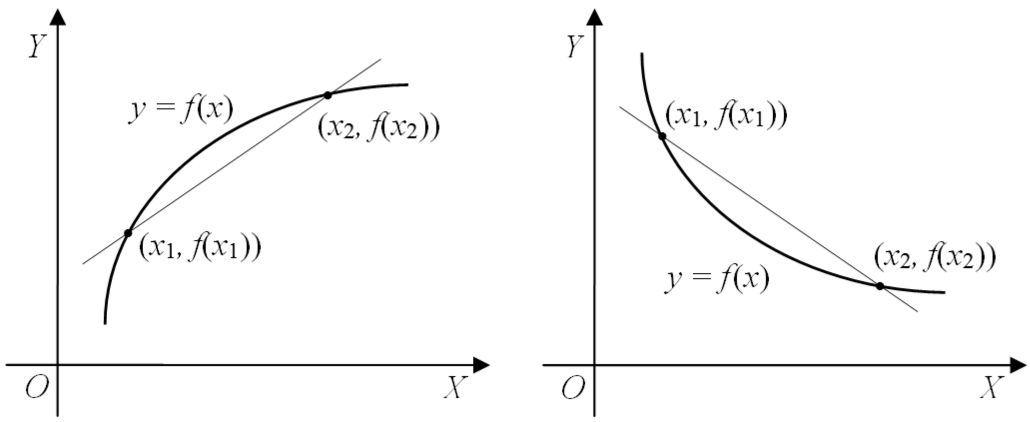

Fig. 1. Convex upwards and convex downwards functions

The left-hand part of Fig. 1 shows an example of a function that is convex upwards, and the right-hand part of Fig. 1 shows an example of a function that is convex downwards.

Sufficient condition for convexity

21B/16:10 (15:42)

Theorem (a sufficient condition for convexity).

Let a function f be differentiable up to the second order on the interval (a, b). Then if $f''(x) > 0$ for any $x \in (a, b)$, then f is convex downwards on the interval (a, b), and if $f''(x) < 0$ for any $x \in (a, b)$, then f is convex upwards on the interval (a, b).

Remark.

We have previously established that the positive or negative first derivative means an increase or, accordingly, a decrease of the function. Thus, the increase and decrease of the function are associated with the properties of the first derivative, and its convexity is associated with the properties of the second derivative.

Inflection points of a function

Definition of an inflection point

22A/00:00 (07:37)

In studying the properties of functions associated with increasing and decreasing, we introduced the notion of a local extremum point, that is, a point located between the intervals of increasing and decreasing of a function.

Similarly, we can introduce a special notion for a point located between the intervals at which the function is convex downwards and convex upwards.

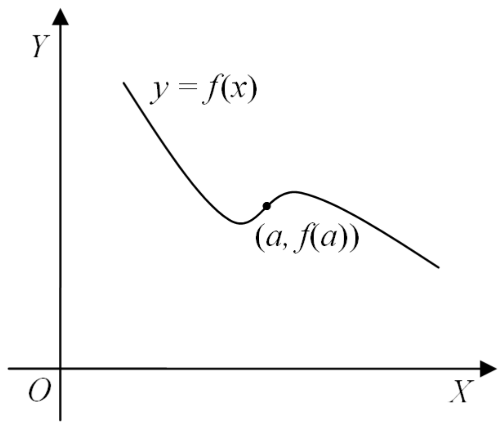

Fig. 2. Inflection point of a function

Definition.

Let the function f be defined in some neighborhood of the point a and be continuous at this point. The point a is called the *inflection point* of the function f if there exist intervals (b, a) and (a, c) such that on one of them the function f is convex downwards and on the other the function f is convex upwards (see Fig. 2).

A necessary condition for the existence of an inflection point

22A/07:37 (10:20)

Theorem (a necessary condition for the existence of an inflection point).

Let a be the inflection point of the function f and let the function f be twice differentiable in some neighborhood of the point a and its second derivative be continuous at the point a. Then $f''(a) = 0$.

Points at which the second derivative is continuous and vanishes are called *points suspected for inflection*. However, a point suspected for inflection is not necessarily an inflection point.

For example, for the functions $f_1(x) = x^3$ and $f_2(x) = x^4$, the point 0 is a point suspected for inflection, since $f_1''(0) = 6x|_{x=0} = 0$, $f_2''(0) = 12x^2|_{x=0} = 0$. However, the point 0 is an inflection point for the function x^3 and it is not an inflection point for the function x^4 (these facts will be strictly proved later, by means of sufficient conditions for the existence of an inflection point).

Thus, the formulated necessary condition for the existence of an inflection point is not a sufficient condition.

The first sufficient condition for the existence of an inflection point $\boxed{\texttt{22A/17:57 (07:04)}}$

In the sufficient condition considered in this section (as in the first sufficient condition for the existence of a local extremum), the differential properties of the function are analyzed not at the inflection point itself, but in its neighborhood.

Theorem (first sufficient condition for the existence of an inflection point).

Let the function f be continuous at the point a and twice differentiable in some punctured neighborhood $\overset{\circ}{U}_a$ of the point a. If the second derivative of the function f has different signs in the left-hand and the right-hand neighborhoods (U_a^- and U_a^+) of the point a, then the point a is an inflection point, and if the second derivative has the same signs, then the point a is not an inflection point.

Proof.

This statement immediately follows from the theorem on a sufficient condition for convexity. If, for example, the function f'' is positive in the left-hand neighborhood of the point a and negative in the right-hand neighborhood, then this means that the function f is convex downwards in the left-hand neighborhood and it is convex upwards in the right-hand neighborhood, therefore, the point a is an inflection point. A similar statement is also true if the function f'' is negative in the left-hand neighborhood and positive in the right-hand neighborhood. If the second derivative has the same signs in the left-hand and right-hand neighborhood of the point a, then the function f has the same convexity type to the left and right of the point a, therefore, the point a is not an inflection point. $\square$

Using this sufficient condition, one can easily prove that the point 0 is an inflection point for the function x^3, but it is not an inflection point for the function x^4. Indeed, $(x^3)'' = 6x$ and, therefore, the second derivative takes values of different signs to the left and right of the point 0, while $(x^4)'' = 12x^2$ takes positive values both to the left and to the right of the point 0.

Asymptotes

Definition.

Let the function f be defined in a punctured neighborhood of the point x_0 (the neighborhood can be one-sided). The line $x = x_0$ is called a *vertical asymptote* of the graph of the function $y = f(x)$ if at least one of the following conditions is true:

$$\lim_{x\to x_0-0} f(x) = \infty, \qquad \lim_{x\to x_0+0} f(x) = \infty.$$

Let the function f be defined in a neighborhood of $+\infty$. The line $y = kx + b$ is called a *non-vertical asymptote* of the graph of the function $y = f(x)$, as $x \to +\infty$, if

$$\lim_{x\to+\infty}\bigl(f(x) - (kx + b)\bigr) = 0. \tag{2}$$

If $k = 0$, then the asymptote is called a *horizontal asymptote*, and if $k \neq 0$, then it is called an *oblique asymptote* or *slant asymptote*.

EXAMPLES OF ASYMPTOTES.

The graph of the function $y = 1/x$ has a vertical asymptote $x = 0$ and a horizontal asymptote $y = 0$ (see the left-hand part of Fig. 3). The graph of the function $y = \sqrt{x^2 + 1}$ has two oblique asymptotes: $y = x$ and $y = -x$ (see the right-hand part of Fig. 3). Note that both graphs are hyperbole branches.

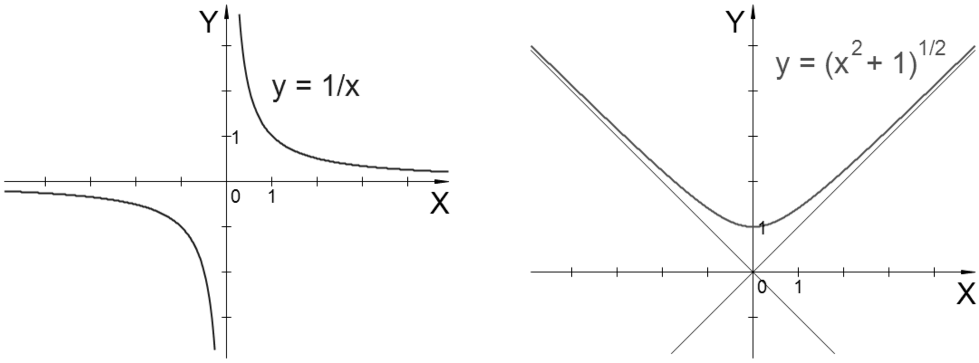

Fig. 3. Examples of asymptotes

The graph of the function $y = \operatorname{tg} x$ has an infinite number of vertical asymptotes of the form $y = \pi/2 + \pi k$, where $k \in \mathbb{Z}$. The graph of the function $y = \operatorname{arctg} x$ has two horizontal asymptotes $y = -\pi/2$ and $y = \pi/2$, and the graph of the function $y = \operatorname{artanh} x$ has two horizontal asymptotes $y = -1$ and $y = 1$.

4. Antiderivative and indefinite integral

Definition of an antiderivative and indefinite integral

Definition. Let the function f be defined on the interval (a, b). Let the function F be a differentiable function on this interval, with $F'(x) = f(x)$ for $x \in (a, b)$. Then the function F is called the *antiderivative* (or *primitive function*) of the function f on a given interval.

The process of finding an antiderivative is called *indefinite integration* (or *antidifferentiation*). If a function has an antiderivative on (a, b), then it is called *integrable* on (a, b).

Hereinafter we, as a rule, will not specify interval on which the function is integrable.

The question arises: how many different antiderivatives exist? Let F_1 be the antiderivative of the function f, that is, $F_1'(x) = f(x)$. Let $F_2(x) = F_1(x) + C$, where C is a constant. Then the function F_2 is also the antiderivative of the function f, since

$$F_2'(x) = \big(F_1(x) + C\big)' = F_1'(x) = f(x).$$

Therefore, if we add a constant to some antiderivative, then we will also get a primitive. So, there exists an infinite number of antiderivatives, differing from each other by a constant term.

There are no other antiderivatives: *all possible antiderivatives can be obtained by adding a constant to some selected antiderivative.* Let us formalize this fact as a theorem.

Theorem (on antiderivatives of a given function). Let F_1 and F_2 be antiderivatives of f. Then there exists a constant $C \in \mathbb{R}$ such that $F_2(x) = F_1(x) + C$.

So, knowing one antiderivative, we can obtain all the other antiderivatives, since they all differ from the chosen antiderivative by a constant term.

Definition. The *indefinite integral* $\int f(x)\,dx$ of the function f is the set of all its antiderivatives: if F_1 is some antiderivative of the function f (that is, $F_1'(x) = f(x)$), then

$$\int f(x)\,dx \stackrel{\text{def}}{=} \{F_1(x) + C, C \in \mathbb{R}\}.$$

The symbol $\int$ is called the *integral sign*, the function $f(x)$ is called the *integrand*, and the expression $f(x)\,dx$ under the integral sign is called the *element of integfation.*

Usually curly braces are not used and, moreover, it is not indicated that C is an arbitrary real constant:

$$\int f(x)\,dx = F_1(x) + C.$$

Table of indefinite integrals

$$\int 0\,dx = C.$$

$$\int A\,dx = Ax + C.$$

$$\int x^\alpha\,dx = \frac{x^{\alpha+1}}{\alpha+1} + C, \quad x > 0, \quad \alpha \in \mathbb{R} \setminus \{-1\}.$$

$$\int \frac{1}{x}\,dx = \ln|x| + C, \quad x \neq 0.$$

To prove the last formula, it suffices to differentiate the superposition $\ln|x| = \ln y \circ |x|$ for $x \neq 0$:

$$(\ln|x|)' = (\ln y)'\big|_{y=|x|} \cdot (|x|)' = \frac{1}{y}\bigg|_{y=|x|} \cdot \operatorname{sign} x = \frac{\operatorname{sign} x}{|x|} = \frac{1}{x}.$$

$$\int e^x\,dx = e^x + C.$$

$$\int a^x\,dx = \frac{a^x}{\ln a} + C, \quad a > 0, \quad a \neq 1.$$

$$\int \sin x\,dx = -\cos x + C.$$

$$\int \cos x\,dx = \sin x + C.$$

$$\int \frac{1}{\cos^2 x}\,dx = \operatorname{tg} x + C.$$

$$\int \frac{1}{\sqrt{1-x^2}}\,dx = \arcsin x + C.$$

$$\int \frac{1}{1+x^2}\,dx = \operatorname{arctg} x + C.$$

$$\int \sinh x\,dx = \cosh x + C.$$

$$\int \cosh x\,dx = \sinh x + C.$$

The simplest properties of an indefinite integral

1. If the function f is integrable, then

$$\Big(\int f(x)\,dx\Big)' = f(x).$$

Proof. Let $F(x)$ be the antiderivative of the function $f(x)$, then

$$\Big(\int f(x)\,dx\Big)' = (F(x) + C)' = F'(x) = f(x). \ \square$$

2. If the function f is differentiable, then

$$\int f'(x)\,dx = f(x) + C.$$

Proof. In this case, $f(x)$ is one of the antiderivatives of the function $f'(x)$, whence the formula to be proved follows. $\square$

3. Additivity of the indefinite integral. Let f and g be integrable, then the function $f+g$ is also integrable, and the formula holds:

$$\int (f+g)\,dx = \int f\,dx + \int g\,dx. \tag{1}$$

4. Homogeneity of the indefinite integral. Let f be integrable, $\alpha \in \mathbb{R}$, $\alpha \neq 0$. Then the function αf is integrable, and the formula holds:

$$\int \alpha f\,dx = \alpha \int f\,dx. \tag{2}$$

Formula (2) means that the constant factor can be taken out of the integral sign.

Remark. In the case of $\alpha = 0$, formula (2) turns out to be incorrect, as we noted earlier that $\int 0\,dx = C$.

If we combine the properties of additivity and homogeneity, then we get the property of linearity.

5. Linearity of the indefinite integral. Let f and g be integrable, $\alpha, \beta \in \mathbb{R}$, with α and β not turning into 0 at the same time: $|a|+|b| \neq 0$. Then the function $\alpha f+\alpha g$ is also integrable, and the formula holds:

$$\int (\alpha f+\beta g)\,dx=\alpha\int f\,dx+\beta\int g\,dx.$$

Example. Using the simplest properties of the indefinite integral and the table of indefinite integrals, one can find the integrals of linear combinations of functions, for example:

$$\int (5e^x+6\cos x)\,dx=5\int e^x\,dx+6\int \cos x\,dx=5e^x+6\sin x+C.$$

To verify the resulting relation, it suffices to differentiate the expression on the right-hand side.

Change of variables in an undefined integral

Theorem (on the change of variables). Let $f(x)$ be an integrable function on (a,b), and one of its antiderivatives is the function $F(x)$. Let $\varphi(t)$ be a differentiable function on the interval (α,β), and $\varphi(t)\in(a,b)$ as $t\in(\alpha,\beta)$. Then

$$\int f\big(\varphi(t)\big)\varphi'(t)\,dt=F\big(\varphi(t)\big)+C. \tag{3}$$

Remark. Considering that the expression $\varphi'(t)\,dt$ is the differential of the function φ, the left-hand side of equality (3) can be written as $\int f(\varphi)d\varphi$.

If we assume that φ is an independent variable, then equality (3) turns into the definition of an indefinite integral:

$$\int f(\varphi)d\varphi=F(\varphi)+C. \tag{4}$$

However, the proved theorem means that equality (4) also holds for the case when φ is a *dependent variable*, that is, it represents a differentiable function of some independent variable (for example, t). In this case, the expression $d\varphi$ must be understood as the *differential* of the function.

The noted circumstance is an additional justification for the inclusion of the expression dx in the notation of an indefinite integral. It should be noted that this notation is also convenient for calculating the integrals by changing variables.

An example of applying the variable changing theorem.

Find the integral $\int \operatorname{tg} x\,dx$:

$$\int \operatorname{tg} x\,dx=\int\frac{\sin x\,dx}{\cos x}.$$

We introduce a new variable: $y=\cos x$. The variable y is the function φ from the variable changing theorem, that is, we can assume that y depends on x. Then dy is the differential of the function $\cos x$, therefore $dy=-\sin x\,dx$. Thus, by virtue of the remark on the variable changing theorem, the expression in the numerator of the original integral can be replaced with $-dy$, and the expression in the denominator can be replaced with y. As a result of changing the variable $y=\cos x$, the original integral is significantly simplified and can now be found using the table of indefinite integrals:

$$\int\frac{\sin x\,dx}{\cos x}=\int\frac{-dy}{y}=-\int\frac{dy}{y}=-\ln|y|+C.$$

It remains for us to return to the original variable x. Finally we obtain:

$$\int \operatorname{tg} x\,dx = -\ln|\cos x| + C.$$

Remarks. 1. When finding the last integral, we actually applied the formula (3), representing the original integral in the form:

$$\int \operatorname{tg} x\,dx = \int f(\cos x)\cdot(\cos x)'\,dx, \quad f(y) = -\frac{1}{y}.$$

However, when performing a variable change in an indefinite integral, formula (3) usually is not used. Instead, in the integral, both the original variable x and its differential dx are replaced, as was done in the above example. The result of successfully changing the variable is to simplify the integrand.

2. Using a similar change of variable, one can also find the integral $\int \frac{dx}{\sin x}$. To do this, take into account that $\sin x = 2\sin\frac{x}{2}\cos\frac{x}{2}$, and additionally transform the resulting expression to obtain the result of differentiation of the function $\operatorname{tg}\frac{x}{2}$ in it.

3. The resulting formula for the integral $\int \operatorname{tg} x\,dx$ makes sense on any interval that does not contain points $\frac{\pi}{2} + \pi k$, $k \in \mathbb{Z}$, that is, points at which the tangent function is not defined.

Formula of integration by parts

Unfortunately, the integral of the product of functions is not equal to the product of the integrals. This is due to the more compicated form of the formula for differentiating the product, compared with the formula for differentiating the sum:

$$(uv)' = u'v + uv'. \tag{5}$$

Nevertheless, using formula (5) for differentiating the product, we can obtain the formula of integration by parts, which in some cases allows us to simplify the calculation of the integral of the product.

Let us express the product uv' from equality (5):

$$uv' = (uv)' - u'v.$$

Integrating the last equality and using the linearity of the indefinite integral (the simplest property 5), we obtain:

$$\int uv'\,dx = \int \big((uv)' - u'v\big)\,dx = \int (uv)'\,dx - \int u'v\,dx. \tag{6}$$

Given the simplest property 2 of the indefinite integral, we have:

$$\int (uv)'\,dx = uv + C.$$

Since the remaining term $\int u'v\,dx$ on the right-hand side of equality (6) also contains an arbitrary constant, we can add the constant C to this arbitrary constant and not specify it explicitly. Finally we obtain the following formula:

$$\int uv'\,dx = uv - \int u'v\,dx.$$

This formula is called the *formula of integration by parts*. It holds if the functions u and v are differentiable and there exists at least one of the integrals included in it (in this case, there necessarily exists another integral).

So, the formula of integration by parts allows us to express the integral of the product of the functions u and v' in terms of the integral of the product of u' and v. It is used in situations where the integral on its right-hand side is easier to find than the integral on the left-hand side.

The formula of integration by parts can also be written in the following form:

$$\int udv = uv - \int vdu.$$

EXAMPLE OF APPLYING THE FORMULA OF INTEGRATION BY PARTS.

Let us find the integral $\int \ln x\, dx$. We put $u(x) = \ln x$, $dv = dx$, whence $v(x) = x$. Then

$$\int \ln x\, dx = x \ln x - \int x(\ln x)'\, dx =$$
$$= x \ln x - \int x \cdot \frac{1}{x}\, dx = x \ln x - \int dx = x \ln x - x + C.$$

5. Integration of rational functions

Auxiliary information from algebra

The *rational function* $R(x)$ is the ratio of two polynomials:

$$R(x) = \frac{P_m(x)}{Q_n(x)}.$$

In studying the question of integrating rational functions, the following facts from the course of algebra are used.

Theorem 1 (on the factorization of a real polynomial).

A polynomial $Q_n(x)$ of degree n with real coefficients can be decomposed into the following irreducible factors:

$$Q_n(x) = a_0(x-c_1)^{\alpha_1}\dots(x-c_k)^{\alpha_k}(x^2+p_1x+q_1)^{\beta_1}\dots(x^2+p_lx+q_l)^{\beta_l}. \tag{1}$$

Here a_0 is the coefficient of the highest degree of the polynomial $Q_n(x)$, $c_1,\dots,c_k$ are the real roots of the polynomial $Q_n(x)$ of multiplicity $\alpha_1,\dots,\alpha_k$, quadratic factors of the form $x^2+p_ix+q_i$ correspond to the complex conjugate roots of the polynomial $Q_n(x)$ of multiplicity β_i, $i=1,\dots,l$. This, in particular, means that all quadratic factors $x^2+p_ix+q_i$ have a negative discriminant: $p_i^2-4q_i<0$. In addition, the following relation holds:

$$\alpha_1+\dots+\alpha_k+2(\beta_1+\dots+\beta_l)=n.$$

Theorem 2 (on the partial fraction decomposition of a real rational function).

Let $R(x)$ be a rational function of the form $\frac{P_m(x)}{Q_n(x)}$, and decomposition (1) takes place for the polynomial $Q_n(x)$. Then $R(x)$ can be represented as follows:

$$R(x) = \tilde{P}(x) + \sum_{i=1}^{k}\sum_{j=1}^{\alpha_i}\frac{A_{ij}}{(x-c_i)^j} + \sum_{i=1}^{l}\sum_{j=1}^{\beta_i}\frac{B_{ij}x+D_{ij}}{(x^2+p_ix+q_i)^j}. \tag{2}$$

The term $\tilde{P}(x)$ appears if the degree m of the polynomial $P_m(x)$ is greater than or equal to the degree n of the polynomial $Q_n(x)$. This term $\tilde{P}(x)$ is a polynomial of degree $m-n$ obtained by dividing the polynomial $P_m(x)$ by the polynomial $Q_n(x)$.

For all remaining terms in formula (2) (called *partial fractions*), the degree of the numerator is less than the degree of the denominator.

Integration of terms in the partial fraction decomposition of a rational function

After finding the partial fraction decomposition of the rational function, it remains to integrate separately all the obtained terms.

1. The integral of the polynomial $\tilde{P}(x)$. This integral will also be a polynomial This integral will also be a polynomial, its degree will be 1 more than the degree of $\tilde{P}(x)$.

2. The integral of a partial fraction of the form $\frac{A}{(x-c)^k}$ corresponding to the real root c of multiplicity k.

For $k \neq 1$, we have:

$$\int \frac{A}{(x-c)^k}\,dx = \frac{A}{(1-k)(x-c)^{k-1}} + C.$$

For $k = 1$, we have:

$$\int \frac{A}{x-c}\,dx = A\ln|x-c| + C.$$

3. The integral of a partial fraction of the form $\frac{Bx+D}{(x^2+px+q)^k}$, provided that the discriminant of the quadratic polynomial is less than zero: $p^2 - 4q < 0$ (this fraction corresponds to complex conjugate roots of multiplicity k).

For this case we give a formula only for the simplest situation as follows:

$$\int \frac{dt}{t^2+\Delta^2} = \frac{1}{\Delta}\operatorname{arctg}\frac{t}{\Delta} + C.$$

Thus, we have shown that all the integrals arising during the integration of a rational function are expressed in terms of elementary functions, and the following theorem holds.

Theorem (on the integration of a rational function).

Any rational function can be integrated in elementary functions.

This is an important fact, since there are elementary functions whose integrals are not expressed in terms of elementary functions. Examples of such functions are $\frac{e^x}{x}$, $\frac{\sin x}{x}$, $\frac{\cos x}{x}$. In particular, the integral $\int \frac{\sin x}{x}\,dx$ is called the *sine integral*, and this function cannot be expressed in terms of elementary functions.

Having proved the theorem on the integration of a rational function, we can use it to study the integrability of other types of functions. If we can reduce (for example, by changing a variable) a certain integrand to a rational function, then we can state that the original function is also integrated in elementary functions.

6. Definite integral, Darboux sums, and integrability criterion

The problem of finding the area of a curvilinear trapezoid

The basic concepts related to a definite integral can be considered by the example of the geometric problem of finding the area of a curvilinear trapezoid.

Let a function $f(x)$ be defined on the segment (closed interval) $[a, b]$ and taking positive values on this segment: $f(x) > 0$, $x \in [a, b]$. It is required to find the area of the figure G bounded by the OX axis, the vertical lines $x = a$ and $x = b$, and the graph of the function $y = f(x)$. Such a figure is called a curvilinear trapezoid with the base $[a, b]$ (see Fig. 4).

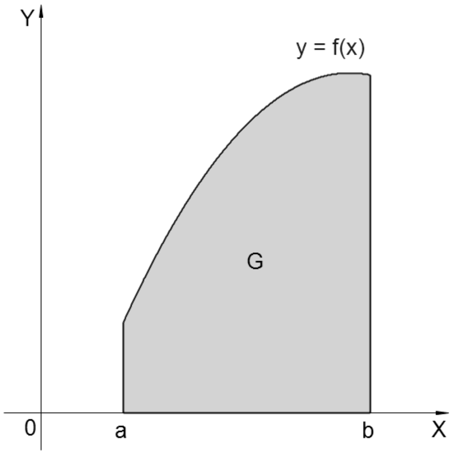

Fig. 4. Curvilinear trapezoid

How to find the approximate area of a curvilinear trapezoid?

Let us divide the segment $[a, b]$ into smaller segments (not necessarily of equal length) with endpoints $a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$. For brevity, we denote the obtained segments as follows: $\Delta_i = [x_{i-1}, x_i]$, $i = 1, \ldots, n$. We also introduce the notation for the length of the segment Δ_i: $\Delta x_i = x_i - x_{i-1}$, $i = 1, \ldots, n$.

Choose a point ξ_i on each of the segments Δ_i: $\xi_i \in \Delta_i$, $i = 1, \ldots, n$.

Provided that the function f has sufficiently "good" properties, we can assume that the area of the curvilinear trapezoid with the base Δ_i will be close to the area of the rectangle with the same base Δ_i and a height equal to the value of the function f at the point ξ_i. The area of this rectangle is $f(\xi_i)\Delta x_i$.

Summing up the areas of all such rectangles, we get the approximate value of the area of the original curvilinear trapezoid: $\sum_{i=1}^{n} f(\xi_i)\Delta x_i$ (see the left-hand part of Fig. 5).

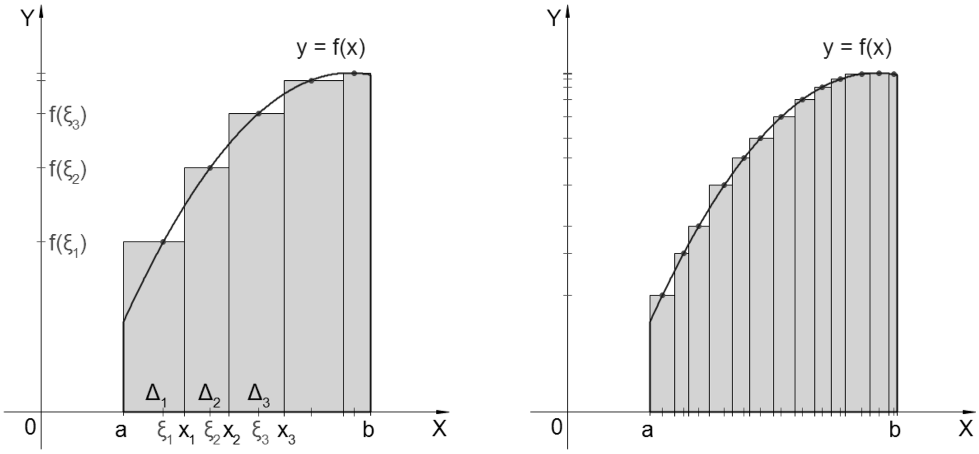

Fig. 5. Curvilinear trapezoid approximation by a set of rectangles

As the number of points x_i increases, we can expect that we will get a more accurate approximation to the area of a curvilinear trapezoid (see the right-hand part of Fig. 5).

If the expression $\sum_{i=1}^{n} f(\xi_i)\Delta x_i$ has a limit as the number of points x_i unlimitedly increases (and, accordingly, as the length of all segments Δ_i unlimitedly decreases), then it is natural to consider this limit as the area of the initial curvilinear trapezoid.

It is this limit that is called the definite integral of the function f over the segment $[a, b]$.

Definition of a definite integral

Definition.

Let the function f be defined on the segment $[a, b]$. The *partition* T of the segment $[a, b]$ is the set of points x_i, $i = 0, \ldots, n$, which has the following property:

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

For the segments $[x_{i-1}, x_i]$ with endpoints at adjacent points of the partition T, as well as for their lengths $x_i - x_{i-1}$, we will use the notation introduced above:

$$\Delta_i \stackrel{\text{def}}{=} [x_{i-1}, x_i], \quad \Delta x_i \stackrel{\text{def}}{=} x_i - x_{i-1}, \quad i = 1, \ldots, n.$$

Obviously, $\Delta x_i > 0$.

The *mesh* of the partition T (notation $l(T)$) is the maximum of the lengths of the segments Δ_i:

$$l(T) \stackrel{\text{def}}{=} \max_{i=1,\ldots,n} \Delta x_i.$$

A *sample* ξ constructed on the basis of a partition T is a set of points $\xi_i \in \Delta_i$, $i = 1, \ldots, n$.

The *integral sum* $\sigma_T(f, \xi)$ of the function f by the partition T and the sample ξ is the following expression:

$$\sigma_T(f, \xi) \stackrel{\text{def}}{=} \sum_{i=1}^{n} f(\xi_i)\Delta x_i.$$

A function f is called *Riemann integrable* on the segment $[a, b]$ if there exists a number I such that

$$\forall\,\varepsilon > 0 \quad \exists\,\delta > 0 \quad \forall\,T, l(T) < \delta, \quad \forall\,\xi \quad |\sigma_T(f,\xi) - I| < \varepsilon. \tag{1}$$

Briefly, condition (1) can be written using the limit notation:

$$I = \lim_{l(T)\to 0, \forall\,\xi} \sigma_T(f,\xi).$$

The number I is called the *Riemann integral* of the function f over the segment $[a,b]$, and it is denoted as follows: $\int_a^b f(x)\,dx$.

So, the Riemann integral of the function f is the limit of the integral sums $\sigma_T(f,\xi)$ as $l(T) \to 0$, $\forall\,\xi$, if this limit exists:

$$\int_a^b f(x)dx = \lim_{l(T)\to 0, \forall\,\xi} \sigma_T(f,\xi).$$

In what follows, Riemann integrability and the Riemann integral will be called simply integrability and integral, respectively.

Remark.

A necessary condition for integrability

Theorem (a necessary condition for integrability).

If the function is integrable on a segment, then it is bounded on this segment.

Remarks.

1. Taking into account this theorem, we will consider only bounded functions hereinafter, not always noting this condition.

2. The converse of the proved theorem is false: if the function is bounded, then this does not follow that it is integrable. We give a corresponding example at the end of this chapter.

Darboux sums and Darboux integrals

Definition of Darboux sums

Definition.

Let the function f be defined and bounded on the segment $[a,b]$.

We choose some partition T of this segment and introduce the following notation:

$$M_i = \sup_{x\in\Delta_i} f(x), \quad m_i = \inf_{x\in\Delta_i} f(x), \quad i = 1,\ldots,n.$$

Since the function f is bounded on $[a,b]$, the values M_i and m_i exist for all $i = 1,\ldots,n$.

The *upper Darboux sum* $S_T^+(f)$ and the *lower Darboux sum* $S_T^-(f)$ are defined as follows:

$$S_T^+(f) \stackrel{\text{def}}{=} \sum_{i=1}^n M_i\Delta x_i, \quad S_T^-(f) \stackrel{\text{def}}{=} \sum_{i=1}^n m_i\Delta x_i.$$

If it is clear which function f is associated with Darboux sums, then the short notation S_T^+ and S_T^- can be used for them.

Remark.

The main difference between Darboux sums and integral sums is that the notion of sample ξ is not used in the definition of Darboux sums: Darboux sums depend only on the function f itself and the partition T of the original segment.

Properties of Darboux sums

In the formulations of all properties, it is assumed that the function f is defined and bounded on the segment $[a, b]$.

1. Let T be some partition of the segment $[a, b]$, ξ be an arbitrary sample associated with this partition. Then the following double inequality holds:

$$S_T^-(f) \le \sigma_T(f, \xi) \le S_T^+(f). \tag{2}$$

2. For a fixed partition T of the segment $[a, b]$, the following relations hold:

$$S_T^+(f) = \sup_\xi \sigma_T(f, \xi), \quad S_T^-(f) = \inf_\xi \sigma_T(f, \xi).$$

Before stating the next property, we introduce the concept of refinement of a partition.

Definition.

The partition T_2 is called the *refinement* of the partition T_1 if any element of the partition T_1 belongs to the partition T_2, i. e., $T_1 \subset T_2$. In other words, the refinement T_2 of the partition T_1 contains all points of the partition T_1 and, possibly, some other points of the original segment.

3. If the partition T_2 is a refinement of the partition T_1, then the following chain of inequalities holds:

$$S_{T_1}^- \le S_{T_2}^- \le S_{T_2}^+ \le S_{T_1}^+. \tag{5}$$

Without proof.

4. If T' and T'' are some partitions of the interval $[a, b]$, then the estimate holds:

$$S_{T'}^- \le S_{T''}^+. \tag{6}$$

Thus, any lower Darboux sum of the function f is less than or equal to any of its upper Darboux sums.

5. There exist values $I^-(f) = \sup_T S_T^-(f)$, $I^+(f) = \inf_T S_T^+(f)$, and the following estimate holds for them:

$$I^-(f) \le I^+(f). \tag{8}$$

Definition.

The values $I^-(f)$ and $I^+(f)$ are called the *lower and upper Darboux integrals* for the function f on the segment $[a, b]$, respectively. Thus, by virtue of property 5, any bounded function has the lower and upper Darboux integrals, and inequality (8) holds for them.

Integrability criterion in terms of Darboux sums

Theorem (integrability criterion in terms of Darboux sums).

The function f is integrable on the segment $[a, b]$ if and only if two conditions are satisfied:

1) f is bounded on $[a, b]$,

2) $\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall T, l(T) < \delta, \quad S_T^+(f) - S_T^-(f) < \varepsilon$.

Remark.

Condition 2 of the theorem can be written as follows:

$$\lim_{l(T)\to 0} \big(S_T^+(f) - S_T^-(f)\big) = 0.$$

Corollary.

If the function f is integrable on the segment $[a,b]$, then its upper and lower Darboux integrals coincide and, moreover, they are equal to the integral of the function f over the segment $[a,b]$.

Remark.

It follows from the corollary that if the upper and lower Darboux integrals are different, then the function is not integrable.

An example of a bounded function that is not integrable.

We define the following function, called the *Dirichlet function*:

$$D(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

Thus, the Dirichlet function is equal to 1 at rational points and is equal to 0 at irrational points.

Obviously, this function is bounded. However, it is not integrable on any of the segments $[a,b]$ of nonzero length. We show this for the segment $[0,1]$.

Let us accept without proof the following fact: in any arbitrarily small neighborhood of any rational number there is some irrational number, and vice versa, in any arbitrarily small neighborhood of any irrational number there is some rational number. Therefore, any segment of nonzero length necessarily contains both irrational and rational numbers. This means that for any partition T of the segment $[0,1]$ the following relations hold:

$$m_i = \inf_{x \in \Delta_i} D(x) = 0, \quad M_i = \sup_{x \in \Delta_i} D(x) = 1.$$

Then for Darboux sums of the Dirichlet function over any partition T of the segment $[0,1]$ we have:

$$S_T^+(D) = \sum_{i=1}^n M_i \Delta x_i = \sum_{i=1}^n 1 \cdot \Delta x_i = \sum_{i=1}^n \Delta x_i = 1,$$

$$S_T^-(f) = \sum_{i=1}^n m_i \Delta x_i = \sum_{i=1}^n 0 \cdot \Delta x_i = 0.$$

Similar relations hold for Darboux integrals:

$$I^+(D) = \inf_T S_T^+(D) = \inf_T 1 = 1,$$

$$I^-(D) = \sup_T S_T^-(D) = \sup_T 0 = 0.$$

We proved that $I^-(D) \neq I^+(D)$, therefore, the Dirichlet function is not integrable on the interval $[0,1]$.

Classes of integrable functions

The simplest example of an integrable function

Consider the constant function $f(x) = c$ and show that it is integrable on any segment $[a,b]$.

To do this, we calculate the integral sum for this function on this segment:

$$\sigma_T(\xi) = \sum_{i=1}^n f(x) \Delta x_i = c \sum_{i=1}^n \Delta x_i = c(b-a).$$

Thus, for any partition T and any sample ξ, the integral sum takes the same value, therefore, when passing to the limit as $l(T) \to 0$, $\forall\, \xi$, this value will not change.

We have proved that

$$\int_a^b c\, dx = c(b-a).$$

Integrability of continuous functions

Theorem (integrability theorem for continuous functions).

If the function is continuous on a segment, then it is integrable on this segment.

Remark.

Continuity is not a necessary condition for integrability. An integrable function may have points of discontinuity.

Integrability of monotone functions

Theorem (integrability theorem for monotone functions).

If the function is monotone on the segment, then it is integrable on this segment.

Remark.

This fact does not follow from the previous theorem, since a monotone function can have a finite or even infinite number of discontinuity points (of the first kind).

7. Properties of a definite integral

Properties associated with integrands

Theorem 1 (on linearity of a definite integral with respect to the integrand).

Let the functions f and g be integrable on the segment $[a, b]$, $\alpha, \beta \in \mathbb{R}$. Then the function $\alpha f + \beta g$ is also integrable on $[a, b]$, and the following equality holds:

$$\int_a^b \bigl(\alpha f(x) + \beta g(x)\bigr)\,dx = \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx. \tag{1}$$

Theorem 2 (on integrability of the product of integrable functions).

Let the functions f and g be integrable on the segment $[a, b]$. Then the function fg is also integrable on $[a, b]$.

Remark.

In this case, we can only establish the fact of integrability, since there is no formula expressing the integral of the product of functions in terms of the integrals of the factors.

Properties associated with integration segments

Theorem 3 (on integrability on an embedded segment).

If the function f is integrable on the segment $[a, b]$, then it is integrable on any segment $[c, d] \subset [a, b]$.

Theorem 4 (the first theorem on the additivity of a definite integral with respect to the integration segment).

Let the function f be integrable on $[a, b]$, $c \in (a, b)$ (note that, by virtue of theorem 3, this function is integrable on the segments $[a, c]$ and $[c, b]$). Then the following equality holds:

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx. \tag{3}$$

Remark.

The converse statement is also true: if the function is integrable on the segments $[a, c]$ and $[c, b]$, then it is integrable on the segment $[a, b]$, and equality (3) holds.

Definition.

We assume that the integral of any function defined at *a over a segment of zero length* $[a, a]$ is 0:

$$\int_a^a f(x)\,dx \stackrel{\text{def}}{=} 0.$$

In addition, we define the integral *from b to a for $a < b$* as follows:

$$\int_b^a f(x)\,dx \stackrel{\text{def}}{=} -\int_a^b f(x)\,dx.$$

This is a quite natural definition, which follows from the initial definition of a definite integral, if we allow the situation $x_{i-1} > x_i$ (for which $\Delta x_i < 0$).

So, we can say that if we swap the limits of integration, then the sign of the integral changes to the opposite.

Theorem 5 (the second theorem on the additivity of a definite integral with respect to the integration segment).

Let the function f be integrable on $[a, b]$, $c_1, c_2, c_3 \in [a, b]$. Then the equality holds

$$\int_{c_1}^{c_3} f(x)\,dx = \int_{c_1}^{c_2} f(x)\,dx + \int_{c_2}^{c_3} f(x)\,dx. \tag{4}$$

Estimates for integrals

Theorem 6 (on the non-negativity of the integral of a non-negative function).

If the function f is integrable on $[a, b]$ and $\forall\, x \in [a, b]$ $f(x) \geq 0$, then

$$\int_a^b f(x)\,dx \geq 0. \tag{5}$$

Theorem 7 (on the comparison of integrals).

If the functions f and g are integrable on $[a, b]$ and $\forall\, x \in [a, b]$ $f(x) \leq g(x)$, then

$$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx. \tag{6}$$

Proof.

We use the previously proved theorems 1 and 6. Let us introduce the auxiliary function $h(x) = g(x) - f(x)$. Obviously, this function is non-negative. In addition, by virtue of theorem 1, this function is integrable; moreover,

$$\int_a^b h(x)\,dx = \int_a^b g(x)\,dx - \int_a^b f(x)\,dx.$$

According to theorem 6, the left-hand side of the resulting equality is non-negative:

$$\int_a^b h(x)\,dx \geq 0.$$

Therefore, the right-hand side is also non-negative, therefore, estimate (6) holds. $\square$

Corollary.

If the function f is integrable on $[a, b]$ and $\forall\, x \in [a, b]$ $m \leq f(x) \leq M$ for some $m, M \in \mathbb{R}$, then

$$m(b-a) \leq \int_a^b f(x)\,dx \leq M(b-a). \tag{7}$$

Proof.

Earlier, we established that the constant function $f(x) = c$ is integrable on any interval, and

$$\int_a^b c\,dx = c(b-a).$$

We apply theorem 7 to the double inequality $m \leq f(x) \leq M$:

$$\int_a^b m\,dx \leq \int_a^b f(x)\,dx \leq \int_a^b M\,dx.$$

Given the formula for the integral of the constant, we obtain relation (7). □

Theorem 8 (on the integral of a non-negative function continuous at some point).

Let the function f be integrable and non-negative on $[a, b]$. Also suppose that the function f is continuous at the point $c \in [a, b]$, and moreover, $f(c) > 0$. Then

$$\int_a^b f(x)\,dx > 0.$$

Theorem 9 (on the integral of the absolute value of a function).

If the function f is integrable on $[a, b]$, then its absolute value $|f|$ is also integrable on $[a, b]$, and the estimate holds:

$$\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\,dx.$$

Remark.

The integrability of the absolute value of a function does not imply the integrability of the function itself. To prove this statement, it suffices to give an example. Consider the following function (which can be obtained from the Dirichlet function by stretching and shifting):

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ -1, & x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

This function, like the Dirichlet function, is not integrable on any segment of positive length, because for any segment $[a, b]$ its upper Darboux integral is $(b - a)$, and it differs from the lower Darboux integral equal to $-(b - a)$. At the same time, the absolute value of this function is a constant: $|f(x)| = 1$, and the constant is integrable on any interval.

Mean value theorems for definite integrals

Theorem 10 (the first mean value theorem).

Suppose that the functions f and g are integrable on $[a, b]$ and the following conditions are satisfied for them:

1) for the function f, a double estimate holds: $m \leq f(x) \leq M$ for $x \in [a, b]$;

2) the function g preserves the sign on $[a, b]$, i. e., either $g(x) \geq 0$ for $x \in [a, b]$, or $g(x) \leq 0$ for $x \in [a, b]$.

Then there exists a value $\mu \in [m, M]$ such that the following equality holds:

$$\int_a^b f(x)g(x)\,dx = \mu \int_a^b g(x)\,dx. \tag{8}$$

Theorem 11 (the second mean value theorem).

Suppose that the functions f and g are defined on $[a, b]$ and the following conditions are satisfied for them:

1) the function f is continuous on $[a, b]$ (this condition immediately implies the integrability of the function f on $[a, b]$);

2) the function g is integrable on $[a, b]$ and preserves the sign on this segment, i. e., either $g(x) \geq 0$ for $x \in [a, b]$, or $g(x) \leq 0$ for $x \in [a, b]$.

Then there exists a point $c \in [a, b]$ such that the following equality holds:

$$\int_a^b f(x)g(x)\,dx = f(c) \int_a^b g(x)\,dx. \tag{10}$$

THEOREM 12 (THE THIRD MEAN VALUE THEOREM).
Let the function f be continuous on $[a, b]$. Then there exists a point $c \in [a, b]$ such that the following equality holds:

$$\int_a^b f(x)\,dx = f(c)(b-a). \tag{11}$$

REMARK (GEOMETRIC INTERPRETATION OF THE THIRD MEAN VALUE THEOREM).
Assume that $f(x) > 0$ for $x \in [a, b]$. We noted earlier that the value of a definite integral $\int_a^b f(x)\,dx$ can be interpreted as the area of a curvilinear trapezoid bounded by the graph $y = f(x)$ and the segments of the axis OX and the lines $x = a$ and $x = b$ (this fact will be proved later when we give a rigorous definition of area). Formula (11) means that there exists a point $c \in [a, b]$ for which a rectangle with the base $[a, b]$ and the height $f(c)$ has the same area as the curvilinear trapezoid.

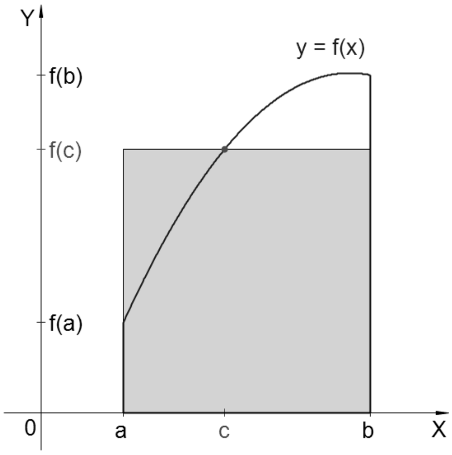

Fig. 6. Geometric sense of the third mean value theorem

PROOF.
It is enough to use the second mean value theorem (theorem 11) by putting $g(x) \equiv 1$ in it. Obviously, in this case the function $g(x)$ preserves the sign. Then

$$\int_a^b g(x)\,dx = \int_a^b dx = b-a.$$

Substituting the function $g(x) \equiv 1$ and the found value of the integral of this function into formula (9), we obtain (11). $\square$

8. Integral with variable upper limit. Newton-Leibniz formula

Integral with variable upper limit

Definition.

Let the function f be integrable on the segment $[a, b]$. Then, by the integrability theorem on the embedded segment, it is integrable on the segment $[a, x]$ for any $x \in [a, b]$. Therefore, for any $x \in [a, b]$ there exists an integral $\int_a^x f(t)\,dt$. Denote this integral by $F(x)$:

$$F(x) \stackrel{\text{def}}{=} \int_a^x f(t)\,dt.$$

The function $F(x)$ is called an *integral with a variable upper limit.* Obviously, $F(a) = 0$ as an integral with the same integration limits.

Theorem 1 (on the continuity of an integral with a variable upper limit).

For any function f integrable on the interval $[a, b]$, its integral with a variable upper limit F is a continuous function on this interval.

Theorem 2 (on the differentiability of an integral with a variable upper limit and a continuous integrand).

If the function f is integrable on the segment $[a, b]$ and continuous at the point $x_0 \in (a, b)$, then its integral with a variable upper limit F is a differentiable function at the point x_0, and the formula holds

$$F'(x_0) = f(x_0).$$

Remark.

Theorems 1 and 2 indicate that the integration operation "improves" the properties of functions: if the original function is integrable, then its integral with a variable upper limit is a continuous function, and if the original function is continuous, then its integral with a variable upper limit is a differentiable function.

Newton-Leibniz formula

Theorem 3 (on the existence of an antiderivative for a continuous function).

Any function f continuous on $[a, b]$ has an antiderivative on (a, b), which is an integral with a variable upper limit: $F(x) = \int_a^x f(t)\,dt$.

Proof.

Since f is continuous on $[a, b]$, it follows from theorem 2 that its integral with a variable upper limit F is a differentiable function on (a, b), and for any point $x \in (a, b)$ the equality $F'(x) = f(x)$ is true. We have obtained that $F(x)$ satisfies the definition of the antiderivative of the function $f(x)$ for $x \in (a, b)$. $\square$

Corollary.

If f is a continuous function on $[a, b]$, and $\Phi(x)$ is its antiderivative on (a, b), then this antiderivative can be represented in the following form, where C is some constant:

$$\Phi(x) = \int_a^x f(t)\,dt + C. \tag{1}$$

Proof.

By theorem 3, we obtain that the integral with a variable upper limit $\int_a^x f(t)\,dt$ is an antiderivative of the function f. The theorem on antiderivatives of a given function states that any two antiderivatives of the function f are different by some constant term C. □

Theorem 4 (the fundamental theorem of calculus).

If the function f is continuous on $[a, b]$, and $\Phi(x)$ is a continuous function on $[a, b]$, which is the antiderivative of the function f on (a, b) (a function $\Phi(x)$ with the indicated properties exists by virtue of theorem 3), then

$$\int_a^b f(x)\,dx = \Phi(b) - \Phi(a). \tag{2}$$

Formula (2) is called the *Newton-Leibniz formula.*

Remarks.

1. The antiderivative (and the indefinite integral) is defined by means of the differentiation operation, and the definite integral is defined by means of the limit of integral sums and, therefore, its definition is not related with the differentiation operation. Nevertheless, there is a deep relation between the operations of differentiation (that is, finding the derivative) and integration (that is, finding the definite integral), which is established by the Newton-Leynnitz formula. That is why theorem 4 is called the *fundamental theorem of calculus.*

2. The Newton-Leibniz formula (2) allows us to reduce the problem of finding a definite integral to the problem of finding the antiderivative of an integrand over a given interval.

3. Formula (2) remains valid for the case $a \geq b$.

4. Formula (2) is often written in the following form:

$$\int_a^b f(x)\,dx = \Phi(x)\Big|_a^b.$$

Proof.

By the corollary of theorem 3, there exists a constant $C \in \mathbb{R}$ such that the antiderivative $\Phi(x)$ of the function $f(x)$ is representable in the form (1). Given this form, we find the values of the antiderivative $\Phi(x)$ at the endpoints of the interval $[a, b]$:

$$\Phi(a) = \int_a^a f(t)\,dt + C = C, \quad \Phi(b) = \int_a^b f(t)\,dt + C.$$

The difference $\Phi(b) - \Phi(a)$ is $\int_a^b f(t)\,dt + C - C = \int_a^b f(t)\,dt$. Thus, the Newton-Leibniz formula is proved, since the value of the integral does not depend on the choice of a letter for the integration parameter (in this case x or t). □

Some methods for calculating definite integrals based on the Newton-Leibniz formula

Theorem 5 (on the change of variables in a definite integral).

Let the function $f(x)$ be continuous on $[a,b]$, and the function $\varphi(t)$ act from $[\alpha,\beta]$ to $[a,b]$ and be continuously differentiable on $[\alpha,\beta]$ (this means that the derivative $\varphi'(t)$ is defined and continuous on $[\alpha,\beta]$). Let, in addition, $\varphi(\alpha)=a$, $\varphi(\beta)=b$. Then

$$\int_a^b f(x)\,dx = \int_\alpha^\beta f\bigl(\varphi(t)\bigr)\varphi'(t)\,dt. \tag{3}$$

Corollaries.

1. Let the function f be an *odd* function defined and continuous on the segment $[-a,a]$. Then $\int_{-a}^a f(t)\,dt = 0$.

2. Let the function f be an *even* function defined and continuous on the interval $[-a,a]$. Then $\int_{-a}^a f(t)\,dt = 2\int_0^a f(t)\,dt$.

3. Let the function f be a continuous *periodic* function with period T. Then

$$\forall\, a\in\mathbb{R}\quad \int_a^{a+T} f(t)\,dt = \int_0^T f(t)\,dt. \tag{5}$$

Thus, the integral of a periodic function over any segment whose length is equal to its period T is equal to the integral over the segment $[0,T]$.

Theorem (on integration by parts of a definite integral).

Let the functions u, v be continuously differentiable on the segment $[a,b]$. Then the following formula holds:

$$\int_a^b uv'\,dx = uv\big|_a^b - \int_a^b u'v\,dx. \tag{7}$$

Formula (7) is called the *integration formula by parts for a definite integral*. Recall that the expression $uv\big|_a^b$ means the difference $u(b)v(b)-u(a)v(a)$.

9. Improper integrals

Tasks leading to the notion of an improper integral

Starting to study a definite integral, we considered the problem of finding the area of a curvilinear trapezoid defined using some continuous function f on the segment $[a, b]$. We further proved that to solve this problem, it is necessary to calculate the integral $\int_a^b f(x)\,dx$.

Now suppose that the function f is defined and continuous on the entire positive semiaxis OX, and on this semiaxis it is positive and decreasing. Is it possible to determine the area of the infinite region D bounded by the positive semiaxis OX, the line $x = 0$, and the graph $y = f(x)$?

Let us choose some point $c > 0$ and consider the part of the region D located to the left of the line $x = c$. This part is a curvilinear trapezoid defined on the segment $[0, c]$, and its area is $\Phi(c) = \int_0^c f(x)\,dx$.

As the value of c increases, the area of $\Phi(c)$ will increase too. If there exists a limit $\Phi(c)$ as $c \to +\infty$, then it is natural to consider this limit as the area of an infinite region D.

Consider another example. Suppose now that the function f is defined and continuous on the half-interval $(0, b]$, takes positive values on it and increases unlimitedly as $x \to +0$.

In this case, we get an infinite region D bounded by the segment $[0, b]$ of the axis OX, the lines $x = 0$ and $x = b$, and the graph $y = f(x)$. To determine the area of the region D, we can choose the point $c \in (0, b)$ and consider the part of the region D bounded by the vertical lines $x = c$ and $x = b$. This part is a curvilinear trapezoid, and its area is $\Phi(c) = \int_c^b f(x)\,dx$.

If there exists a limit $\Phi(c)$ as $c \to +0$, then this limit can be considered as the area of the infinite region D.

These examples show that improper integrals can be of two types: integrals over an infinite integration interval of a bounded function and integrals over a finite interval, but of a function that is unbounded on a given interval. In any of these cases, the passing to limit is used to determine the improper integral.

Definitions of an improper integral

Definition 1 (definition of an improper integral over a semi-infinite interval).

Let a function f be defined on the set $[a, +\infty)$ and integrable on any segment $[a, c]$, $c > a$. If there exists a finite limit of the integral $\int_a^c f(x)\,dx$ as $c \to +\infty$, then they say that there exists an *improper integral* $\int_a^{+\infty} f(x)\,dx$, and its value is assumed to be equal to this limit:

$$\int_a^{+\infty} f(x)\,dx \stackrel{\text{def}}{=} \lim_{c\to+\infty} \int_a^c f(x)\,dx.$$

In this case, they also say that the improper integral $\int_a^{+\infty} f(x)\,dx$ *converges*.

If the limit $\lim_{c\to+\infty}\int_a^c f(x)\,dx$ does not exist or is equal to infinity, then they say that the improper integral $\int_a^{+\infty} f(x)\,dx$ *diverges.*

An improper integral over a semi-infinite interval of the form $(-\infty, b]$ is defined in a similar way.

Examples.

1. Consider the integral $\int_1^{+\infty}\frac{dx}{x^\alpha}$ when $\alpha\in\mathbb{R}$.

We choose the value $c>0$ and find the integral over a finite segment:

$$\int_1^{+\infty}\frac{dx}{x^\alpha}=\begin{cases}\left.\dfrac{x^{-\alpha+1}}{-\alpha+1}\right|_1^c=\dfrac{c^{-\alpha+1}}{-\alpha+1}-\dfrac{1}{-\alpha+1}, & \alpha\neq 1,\\[2ex] \ln x|_1^c=\ln c, & \alpha=1.\end{cases}$$

The function $\ln c$ approaches infinity as $c\to+\infty$. The function $c^{-\alpha+1}$ approaches infinity as $c\to+\infty$ if $\alpha<1$, and aproaches 0 if $\alpha>1$. Consequently, the initial improper integral diverges for $\alpha\le 1$ and converges for $\alpha>1$, and for the converging integral the formula holds:

$$\int_1^{+\infty}\frac{dx}{x^\alpha}=\lim_{c\to+\infty}\left(\frac{c^{-\alpha+1}}{-\alpha+1}-\frac{1}{-\alpha+1}\right)=\frac{1}{\alpha-1},\quad \alpha>1.$$

2. Consider the integral $\int_0^{+\infty}e^{-x}\,dx$. In this case, we have for the segment $[0,c]$:

$$\int_0^c e^{-x}\,dx=-e^{-x}\big|_0^c=-e^{-c}+1.$$

Hence,

$$\int_1^{+\infty}e^{-x}\,dx=\lim_{c\to+\infty}(-e^{-c}+1)=1.$$

Definition 2 (definition of an improper integral for an unbounded function).

Let the function f be defined on the half-interval $[a,b)$ and integrable on any segment $[a,c]$, $a<c<b$. If there exists a finite limit of the integral $\int_a^c f(x)\,dx$ as $c\to b-0$, then they say that there exists an *improper integral* $\int_a^b f(x)\,dx$, and its value is assumed to be equal to this limit:

$$\int_a^b f(x)\,dx\stackrel{\text{def}}{=}\lim_{c\to b-0}\int_a^c f(x)\,dx.$$

In this case, they also say that the improper integral $\int_a^b f(x)\,dx$ *converges.*

If the limit $\lim_{c\to b-0}\int_a^c f(x)\,dx$ does not exist or is equal to infinity, then they say that the improper integral $\int_a^b f(x)\,dx$ *diverges.*

The improper integral for the function defined on the half-interval $(a,b]$ is defined in a similar way.

Example.

Consider the integral $\int_0^1\frac{dx}{x^\alpha}$ when $\alpha\in\mathbb{R}$. Obviously, for $\alpha\le 0$ this integral is an usual (proper) integral, since the function $\frac{1}{x^\alpha}$ in this case is defined and continuous on the entire segment $[0,1]$. The value of the integral for $\alpha<0$ is equal to

$$\int_0^1\frac{dx}{x^\alpha}=\left.\frac{x^{-\alpha+1}}{-\alpha+1}\right|_0^1=\frac{1}{1-\alpha}.$$

The formula $\int_0^1 \frac{dx}{x^\alpha} = \frac{1}{1-\alpha}$ is also valid for the case $\alpha = 0$.

For $\alpha > 0$ we get an improper integral, since the function $\frac{1}{x^\alpha}$ is unbounded in a neighborhood of the point 0. Therefore, we choose the value $c \in (0,1)$ and find the integral over the finite segment:

$$\int_c^1 \frac{dx}{x^\alpha} = \begin{cases} \left.\dfrac{x^{-\alpha+1}}{-\alpha+1}\right|_c^1 = \dfrac{1}{-\alpha+1} - \dfrac{c^{-\alpha+1}}{-\alpha+1}, & \alpha \neq 1, \\ \ln x|_c^1 = -\ln c, & \alpha = 1. \end{cases}$$

The function $-\ln$ c approaches infinity as $c \to +0$. The function $c^{-\alpha+1}$ approaches infinity as $c \to +0$ if $\alpha > 1$, and aporoaches 0 if $\alpha < 1$. Consequently, the initial improper integral diverges for $\alpha \geq 1$ and converges for $0 < \alpha < 1$, and for the converging integral the formula holds:

$$\int_0^1 \frac{dx}{x^\alpha} = \lim_{c \to +0} \left(\frac{1}{-\alpha+1} - \frac{c^{-\alpha+1}}{-\alpha+1} \right) = \frac{1}{1-\alpha}, \quad 0 < \alpha < 1.$$

So, the integral $\int_0^1 \frac{dx}{x^\alpha}$ exists for $\alpha < 1$, and it must be understood in an improper sense for $0 < \alpha < 1$.

It will be convenient for us to simultaneously consider improper integrals over semi-infinite intervals and improper integrals of unbounded functions. So, let us give a general definition of an improper integral.

Definition 3 (the definition of an improper integral in the general case).

Let the function f be defined on the half-interval $[a,b)$ and integrable on any segment $[a,c]$, $a < c < b$. The point b is either finite or equal to $+\infty$. If there exists a finite limit of the integral $\int_a^c f(x)\,dx$ as $c \to b-0$, then they say that there exists an *improper integral* $\int_a^b f(x)\,dx$, and its value is assumed to be equal to this limit:

$$\int_a^b f(x)\,dx \stackrel{\text{def}}{=} \lim_{c \to b-0} \int_a^c f(x)\,dx.$$

In this case, they also say that the improper integral $\int_a^b f(x)\,dx$ *converges.*

If the limit $\lim_{c \to b-0} \int_a^c f(x)\,dx$ does not exist or is equal to infinity, then they say that the improper integral $\int_a^b f(x)\,dx$ *diverges.*

An improper integral with a singularity at the left endpoint a of the integration interval is defined in a similar way; the left endpoint may be equal to $-\infty$.

If an improper integral has a singularity at both endpoints of the integration interval (a,b), then it is considered as the sum of the integrals over the intervals $(a,d]$ and $[d,b)$ for some point $d \in (a,b)$ and is convergent if and only if improper integrals converge over each of the intervals $(a,d]$ and $[d,b)$. We return to the discussion of integrals with several singularities at the end of the next chapter.

Properties of improper integrals

Theorem 1 (on the linearity of an improper integral with respect to integrands).

Let the functions f and g be defined on $[a,b)$, $\alpha, \beta \in \mathbb{R}$. Let there exist improper integrals $\int_a^b f(x)\,dx$ and $\int_a^b g(x)\,dx$. Then there exists an improper integral $\int_a^b \big(\alpha f(x) + \beta g(x)\big)\,dx$, and the following formula holds:

$$\int_a^b \big(\alpha f(x) + \beta g(x)\big)\,dx = \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx. \tag{1}$$

Theorem 2 (on the additivity of an improper integral with respect to the integration interval).

Let the function f be defined on $[a, b)$, and there exists an improper integral $\int_a^b f(x)\,dx$. Then for any point $d \in (a, b)$ the improper integral $\int_d^b f(x)\,dx$ converges, and the equality holds:

$$\int_a^b f(x)\,dx = \int_a^d f(x)\,dx + \int_d^b f(x)\,dx.$$

Theorem 3 (on the change of variables in an improper integral).

Let the function f be defined on $[a, b)$, and there exists an improper integral $\int_a^b f(x)\,dx$. Let the function φ act from $[\alpha, \beta)$ on $[a, b)$, be continuously differentiable on $[\alpha, \beta)$, $\varphi'(t) > 0$ for $t \in [\alpha, \beta)$, $\varphi(\alpha) = a$ and $\lim_{t\to\beta-0} \varphi(t) = b$. Then there exists an improper integral $\int_\alpha^\beta f\big(\varphi(t)\big)\varphi'(t)\,dt$, and the equality holds:

$$\int_a^b f(x)\,dx = \int_\alpha^\beta f\big(\varphi(t)\big)\varphi'(t)\,dt.$$

Theorem 4 (on integration by parts of an improper integral).

Let the functions u and v be defined and continuously differentiable on $[a, b)$, and there exists a limit $\lim_{x\to b} u(x)v(x)$. Then the improper integrals $\int_a^b uv'\,dx$ and $\int_a^b u'v\,dx$ either both converge or both diverge, and if they converge, then the following relation holds, which is called the *integration formula by parts for improper integrals*:

$$\int_a^b u'v\,dx = (uv)\big|_a^b - \int_a^b uv'\,dx.$$

In this formula, the notation $(uv)\big|_a^b$ means the following difference: $\lim_{x\to b} u(x)v(x) - u(a)v(a)$.

Absolute convergence of improper integrals

Definition.

Let the function f be defined on the interval $[a, b)$. The improper integral $\int_a^b f(x)\,dx$ is called *absolutely convergent* if the integral $\int_a^b |f(x)|\,dx$ converges.

Theorem (on the convergence of an absolutely convergent integral).

If the improper integral $\int_a^b f(x)\,dx$ absolutely converges, then it converges.

Remark.

The converse is not true: we will show later that a convergent improper integral is not necessarily absolutely convergent. Thus, the property of absolute convergence is stronger than the property of usual convergence.

Properties of improper integrals of non-negative functions

In this section, we consider improper integrals of non-negative functions. Since the absolute value of the function is non-negative, all the results obtained in this section can also be used to study the absolute convergence of improper integrals of functions taking both negative and positive values.

Theorem (the comparison test for improper integrals of non-negative functions).

Let the functions f and g be defined on the interval $[a,b)$, and for any $x \in [a,b)$ the double inequality holds: $0 \le f(x) \le g(x)$. Suppose that for any $c \in [a,b)$ there exist integrals $\int_a^c f(x)\,dx$ and $\int_a^c g(x)\,dx$. Then the following two statements are true.

1. If the improper integral $\int_a^b g(x)\,dx$ converges, then the integral $\int_a^b f(x)\,dx$ also converges.

2. If the improper integral $\int_a^b f(x)\,dx$ diverges, then the integral $\int_a^b g(x)\,dx$ also diverges.

Corollary.

Let the functions f and g be defined on the interval $[a,b)$ and be non-negative on this interval. Let $f(x) \sim g(x)$ as $x \to b-0$. Then the integrals $\int_a^b f(x)\,dx$ and $\int_a^b g(x)\,dx$ either both converge or both diverge.

Examples.

1. Consider the integral $\int_1^{+\infty} \frac{\sin x}{x^2}\,dx$.

For the absolute value of the integrand, the following estimate holds:

$$\left|\frac{\sin x}{x^2}\right| \le \frac{1}{x^2}.$$

Earlier, we proved that the integral $\int_1^{+\infty} \frac{1}{x^2}\,dx$ converges. Therefore, the integral $\int_1^{+\infty} \left|\frac{\sin x}{x^2}\right| dx$ also converges by the comparison test. And this, in turn, means that the initial integral converges absolutely.

In a similar way, one can prove that absolute convergence holds for the integral $\int_1^{+\infty} \frac{\sin x}{x^\alpha}\,dx$ for any $\alpha > 1$.

2. Consider the integral $\int_2^{+\infty} \frac{1}{\ln x}\,dx$.

For any $x > 0$, the estimate $\ln x < x$ is valid. It follows that $\frac{1}{x} < \frac{1}{\ln x}$. Earlier, we proved that the integral $\int_1^{+\infty} \frac{1}{x}\,dx$ diverges. Obviously, the integral $\int_2^{+\infty} \frac{1}{x}\,dx$ also diverges. Then the initial integral $\int_2^{+\infty} \frac{1}{\ln x}\,dx$ also diverges by the comparison test.

3. Consider the integral $\int_1^{+\infty} \frac{1}{x^\alpha + \sin x}\,dx$, $\alpha > 1$.

Let us show that the integrand is equivalent to the function $\frac{1}{x^\alpha}$ as $x \to +\infty$:

$$\lim_{x\to+\infty} \frac{\frac{1}{x^\alpha+\sin x}}{\frac{1}{x^\alpha}} = \lim_{x\to+\infty} \frac{x^\alpha}{x^\alpha + \sin x} = \lim_{x\to+\infty} \frac{1}{1 + \frac{\sin x}{x^\alpha}} = 1.$$

So, we have proved that $\frac{1}{x^\alpha+\sin x} \sim \frac{1}{x^\alpha}$, $x \to +\infty$.

Since the integral $\int_1^{+\infty} \frac{1}{x^\alpha}\,dx$ converges for $\alpha > 1$, we obtain from the corollary of the comparison test that the initial integral also converges.

Conditional convergence of improper integrals

Definition.

The improper integral $\int_a^b f(x)\,dx$ is called *conditionally convergent* if this integral converges, and the integral $\int_a^b |f(x)|\,dx$ diverges. In other words, the integral converges conditionally if it converges, but it is not absolutely convergent.

It is clear that conditional convergence may hold only for integrals whose integrands change sign.

Example.

It can be proved that the integral $\int_1^{+\infty} \frac{\sin x}{x}\,dx$ converges conditionally.

Dirichlet test for conditional convergence of an improper integral

Theorem (Dirichlet test for conditional convergence of an improper integral).

Let the functions f and g be defined on the interval $[a, b)$ and satisfy the following conditions:

1) the function f is continuous on $[a, b)$, and the integral $\int_a^c f(x)\,dx$ is uniformly bounded for all $c \in (a, b)$, i.e.,

$$\exists\, M > 0 \quad \forall\, c \in (a, b) \quad \left|\int_a^c f(x)\,dx\right| \le M;$$

2) the function g is continuously differentiable on $[a, b)$, and $g(c)$ monotonically approaches 0 as $c \to b - 0$ (the monotonicity condition means that $g'(c)$ preserves the sign for all $c \in (a, b)$).

Then the improper integral $\int_a^b f(x)g(x)\,dx$ converges (generally speaking, conditionally).

10. Numerical series

Definition and examples

Recall how the finite sum of terms is written using the summation symbol $\sum$:

$$\sum_{k=1}^{n} a_k = a_1 + a_2 + \dots + a_n.$$

If the symbol ∞ is indicated in the notation of the sum instead of the finite number n, then this notation can be considered as the formal sum of an infinite number of terms:

$$\sum_{k=1}^{\infty} a_k = a_1 + a_2 + \dots + a_k + \dots$$

The expression $\sum_{k=1}^{\infty} a_k$ is called a *numerical series*, and the value a_k is called a *common term* of the series. We can assume that the numerical series $\sum_{k=1}^{\infty} a_k$ is the formal sum of all elements of the sequence $\{a_k\}$. In this case, however, it is necessary to determine the sense of the sum of an infinite number of terms.

Consider the finite sum

$$S_n = \sum_{k=1}^{n} a_k.$$

This sum is called the *partial sum of the series*; it exists for any number $n \in \mathbb{R}$. Thus, we get a sequence of partial sums $\{S_n\}$.

If there exists a finite limit S of the sequence $\{S_n\}$ as $n \to \infty$, then the numerical series $\sum_{k=1}^{\infty} a_k$ is called *convergent*, and the limit S is called the *sum* of this numerical series. In this case, the notation $\sum_{k=1}^{\infty} a_k$ also means the value of the sum, i. e., the limit S (just as the notion of an improper integral means the limit value of usual proper integrals):

$$\sum_{k=1}^{\infty} a_k \overset{\text{def}}{=} \lim_{n\to\infty} \sum_{k=1}^{n} a_k.$$

If the sequence of partial sums $\{S_n\}$ has no limit or has an infinite limit, then the numerical series $\sum_{k=1}^{\infty} a_k$ is called *divergent*, in which case its sum is not defined (as well as the value of the divergent improper integral).

We emphasize that, in any case, the notation $\sum_{k=1}^{\infty} a_k$ can be considered as a formal sum of an infinite number of terms, regardless of whether this sum corresponds to some finite value or not.

As a summation parameter, the symbols i and j are often used along with the symbol k.

The initial value of the summation parameter does not have to be 1. Series with a summation parameter starting with 0 are often considered. Obviously, if the series $\sum_{k=1}^{\infty} a_k$ converges, then the series $\sum_{k=n_0}^{\infty} a_k$ also converges for any $n_0 \in \mathbb{N}$.

Example.

Let $q \neq 0$ be an arbitrary real number. Consider a series with the common term q^k:

$$\sum_{k=0}^{\infty} q^k = 1 + q + q^2 + \dots + q^k + \dots$$

This series is the formal sum of all terms of the geometric progression with the first term 1 and the ratio q.

Recall the formula for the sum of the initial terms of such a geometric progression (provided that $q \neq 1$):

$$S_n = \sum_{k=0}^{n} q^k = \frac{1 - q^{n+1}}{1 - q}.$$

In this case, S_n denotes the sum of $(n+1)$ initial terms of the geometric progression. It is clear that if $q = 1$, then $S_n = n + 1$.

If $|q| < 1$, then $\lim_{n\to\infty} S_n = \frac{1}{1-q}$.

If $|q| \geq 1$, then the limit of the sequence $\{S_n\}$ as $n \to \infty$ is either infinite or (for $q = -1$) does not exist (since for $q = -1$ the sequence $\{q^n\}$ has the form $\{1, -1, 1, -1, \dots\}$ and, therefore, the sequence $\{S_n\}$ is equal to $\{1, 0, 1, 0, \dots\}$).

So, if $|q| \geq 1$, then the series $\sum_{k=0}^{\infty} q^k$ diverges, and if $|q| < 1$, then the series $\sum_{k=0}^{\infty} q^k$ converges, and its sum is $\frac{1}{1-q}$:

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}, \quad |q| < 1, q \neq 0.$$

This formula is called the *formula of the sum of an infinitely decreasing geometric progression.*

THEOREM (A NECESSARY CONDITION FOR THE CONVERGENCE OF A NUMERICAL SERIES).

If the series $\sum_{k=1}^{\infty} a_k$ converges, then its common term a_k approaches 0:

$$\lim_{k\to\infty} a_k = 0.$$

REMARK.

This condition means that if the common term of a series does not approach 0, then the series is not convergent. Thus, it makes it easy to prove the divergence of many series. However, it should be emphasized that this condition is not a sufficient condition for convergence: from the fact that the common term of a series approaches 0, it does not follow that the series converges (we will give the corresponding examples later).

Absolutely convergent numerical series and their arithmetic properties

DEFINITION.

It is said that the series $\sum_{k=1}^{\infty} a_k$ *absolutely converges* if the series $\sum_{k=1}^{\infty} |a_k|$ converges.

THEOREM (ON THE CONVERGENCE OF AN ABSOLUTELY CONVERGENT NUMERICAL SERIES).

If the series absolutely converges, then it is convergent.

THEOREM (ON ARITHMETIC PROPERTIES OF CONVERGENT NUMERICAL SERIES).

Let $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} b_k$ be convergent series whose sums are S_a and S_b respectively. Let $\alpha, \beta \in \mathbb{R}$, $|\alpha| + |\beta| \neq 0$.

Then the series $\sum_{k=1}^{\infty} (\alpha a_k + \beta b_k)$ also converges, and its sum is $\alpha S_a + \beta S_b$.

Thus, for convergent series, the same arithmetic transformations can be used as for finite sums:

$$\sum_{k=1}^{\infty}(\alpha a_k + \beta b_k) = \alpha \sum_{k=1}^{\infty} a_k + \beta \sum_{k=1}^{\infty} b_k. \tag{4}$$

In addition, if the initial series converge absolutely, then the series $\sum_{k=1}^{\infty}(\alpha a_k + \beta b_k)$ also converges absolutely.

Comparison test and integral test of convergence

Theorem (comparison test for numerical series).
Let $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} b_k$ be series for which the following condition holds:

$$\exists\, m \in \mathbb{N} \quad \forall\, k \geq m \quad 0 \leq a_k \leq b_k.$$

Then two statements are valid.
1. If the series $\sum_{k=1}^{\infty} b_k$ converges, then the series $\sum_{k=1}^{\infty} a_k$ also converges.
2. If the series $\sum_{k=1}^{\infty} a_k$ diverges, then the series $\sum_{k=1}^{\infty} b_k$ also diverges.

Theorem (integral test of convergence).
Let the function f be defined on the set $[1, +\infty)$, be non-negative and non-increasing, and $\lim_{x\to+\infty} f(x) = 0$.

Then the improper integral $\int_1^{+\infty} f(x)\,dx$ and the series $\sum_{k=1}^{\infty} f(k)$ either both converge or both diverge.

Earlier, we found that the improper integral $\int_1^{+\infty} \frac{1}{x^\alpha}\,dx$ converges for $\alpha > 1$ and diverges for $\alpha \leq 1$. Now we can extend this result to the corresponding series. For $\alpha > 0$, the function $f(x) = \frac{1}{x^\alpha}$ satisfies all the conditions of the previous theorem (it is non-negative and monotonically approaches 0 as $x \to +\infty$), therefore, by virtue of of the previous theorem, the series $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$ converges for $\alpha > 1$ and diverges for $\alpha \in (0, 1]$. For $\alpha \leq 0$, the series $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$ also diverges, since, in this case, its common term $\frac{1}{k^\alpha}$ does not approach 0 as $k \to \infty$, and therefore the necessary convergence condition is not satisfied for the series. Thus, we have proved the following result.

Corollary.
The numerical series $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$ converges for $\alpha > 1$ and diverges for $\alpha \leq 1$. In particular, the series $\sum_{k=1}^{\infty} \frac{1}{k}$, called the harmonic series, diverges.

D'Alembert's test and Cauchy's test of convergence of numerical series

The tests considered in this section have no analogues for improper integrals.

Theorem (D'Alembert's test for convergence of a numerical series).
Let $\sum_{k=1}^{\infty} a_k$ be a series with positive terms: $\forall\, k \in \mathbb{N}$ $a_k > 0$.
1. Let the following condition be satisfied:

$$\exists\, q \in (0, 1) \quad \exists\, m \in \mathbb{N} \quad \forall\, k \geq m \quad \frac{a_{k+1}}{a_k} \leq q.$$

Then the series $\sum_{k=1}^{\infty} a_k$ converges.
2. Let the following condition be satisfied:

$$\exists\, m \in \mathbb{N} \quad \forall\, k \geq m \quad \frac{a_{k+1}}{a_k} \geq 1.$$

Then the series $\sum_{k=1}^{\infty} a_k$ diverges.

Corollary (the limit D'Alembert test).

Let $\sum_{k=1}^{\infty} a_k$ be a series with positive terms: $\forall\, k \in \mathbb{N}\; a_k > 0$. Suppose that there exists a limit $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = q$. If $q < 1$, then the series $\sum_{k=1}^{\infty} a_k$ converges, and if $q > 1$, then the series diverges.

Remarks.

1. If the limit $\lim_{k\to\infty} \frac{a_{k+1}}{a_k}$ is 1, then nothing can be said about the convergence or divergence of the series, and further investigation is required.

2. If the limit $\lim_{k\to\infty} \frac{a_{k+1}}{a_k}$ is equal to $+\infty$, then, by similar reasoning, we can prove that the series diverges.

Example.

Consider the series $\sum_{k=0}^{\infty} \frac{x^k}{k!}$. Recall that by definition it is supposed that $0! = 1$. Here x is an arbitrary real number. Denote $a_k = \frac{x^k}{k!}$ and consider the following limit:

$$\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = \lim_{k\to\infty} \frac{\frac{x^{k+1}}{(k+1)!}}{\frac{x^k}{k!}} = \lim_{k\to\infty} \frac{x^{k+1}k!}{x^k(k+1)!} = \lim_{k\to\infty} \frac{x}{k+1} = 0.$$

The limit exists, and its value is less than 1, therefore, due to the limit D'Alembert test, this series converges for any value of the parameter $x \in \mathbb{R}$.

Theorem (Cauchy's test for convergence of a numerical series).

Let $\sum_{k=1}^{\infty} a_k$ be a series with non-negative terms: $\forall\, k \in \mathbb{N}\; a_k \geq 0$.

1. Let the following condition be satisfied:

$$\exists\, q \in (0,1) \quad \exists\, m \in \mathbb{N} \quad \forall\, k \geq m \quad \sqrt[k]{a_k} \leq q.$$

Then the series $\sum_{k=1}^{\infty} a_k$ converges.

2. Let the following condition be satisfied:

$$\exists\, m \in \mathbb{N} \quad \forall\, k \geq m \quad \sqrt[k]{a_k} \geq 1.$$

Then the series $\sum_{k=1}^{\infty} a_k$ diverges.

Corollary (the limit Cauchy test).

Let $\sum_{k=1}^{\infty} a_k$ be a series with non-negative terms: $\forall\, k \in \mathbb{N}\; a_k \geq 0$. Suppose that there exists a limit $\lim_{k\to\infty} \sqrt[k]{a_k} = q$. If $q < 1$, then the series $\sum_{k=1}^{\infty} a_k$ converges, and if $q > 1$, then the series diverges.

Remarks.

1. If the limit $\lim_{k\to\infty} \sqrt[k]{a_k}$ is 1, then nothing can be said about the convergence or divergence of the series, and further investigation is required.

2. If the limit $\lim_{k\to\infty} \sqrt[k]{a_k}$ is equal to $+\infty$, then, by similar reasoning, we can prove that the series diverges.

11. Alternating series and conditional convergence

Alternating series

Definition.

The series $\sum_{k=1}^{\infty} a_k$ is called *conditionally convergent* if it converges, and the series $\sum_{k=1}^{\infty} |a_k|$ diverges. Thus, a convergent series is called conditionally convergent if it does not converge absolutely.

Such a situation is possible only when the terms of a series have different signs.

Definition.

A series of the form $\sum_{k=1}^{\infty}(-1)^{k+1}a_k$ is called an *alternating series*, or *Leibniz series*, if $\forall\, k \in \mathbb{N}$ $a_k \geq 0$, the sequence $\{a_k\}$ is non-increasing, i. e., $a_1 \geq a_2 \geq a_3 \geq a_4 \geq \ldots$, and $a_k \to 0$ as $k \to \infty$.

Remark.

The series obtained from the Leibniz series by multiplying by (-1) is also called alternating. This series can be represented as $\sum_{k=1}^{\infty}(-1)^k a_k$ under the same conditions on a_k as for the Leibniz series.

Theorem (on the convergence of the Leibniz series).

The Leibniz series $\sum_{k=1}^{\infty}(-1)^{k+1}a_k$ converges.

Remark.

The theorem on the convergence of the Leibniz series guarantees only its conditional convergence. For example, the series $\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}$ is a Leibniz series, however, we previously established that the series $\sum_{k=1}^{\infty} \frac{1}{k}$, consisting of absolute values of terms of the initial series, is divergent. In what follows, we prove that the sum of the series $\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}$ is equal to $\ln 2$.

Theorem (on the estimation of the Leibniz series in terms of its partial sums).

Let $\sum_{k=1}^{\infty}(-1)^{k+1}a_k = S$ be the Leibniz series, and $S_n = \sum_{k=1}^{n}(-1)^{k+1}a_k$ be its partial sums. Then for any $k \in \mathbb{N}$ the following estimate holds:

$$|S - S_k| \leq a_{k+1}.$$

Dirichlet's test and Abel's test of conditional convergence of a numerical series

Theorem (Dirichlet's test for conditional convergence of a numerical series).

Let the following conditions be satisfied for the series $\sum_{k=1}^{\infty} a_k b_k$:

1) $\exists\, M \quad \forall\, n \in \mathbb{N} \quad \left|\sum_{k=1}^{n} a_k\right| \leq M$;

2) $b_k \to 0$ as $k \to \infty$, and approaching 0 is monotone.

Then the series $\sum_{k=1}^{\infty} a_k b_k$ converges (generally speaking, conditionally).

Without proof.

Examples.

1. Once again, let us turn to the Leibniz series and write it in the following form: $\sum_{k=1}^{\infty}(-1)^{k+1}b_k$. By the definition of the Leibniz series, for the sequence $\{b_k\}$ two conditions are satisfied: $\{b_k\}$ is non-increasing and $b_k \to 0$ as $k \to \infty$. Thus, for $\{b_k\}$ the condition 2 of Dirichlet's test is satisfied. Also we can take the sequence $\left\{(-1)^{k+1}\right\}$ as the sequence $\{a_k\}$. Obviously, this sequence satisfies condition 1 of Dirichlet's test:

$$\forall\, n \in \mathbb{N} \quad \left|\sum_{k=1}^{n} a_k\right| = |1 - 1 + 1 - 1 + ...| \le 1.$$

Thus, the fact that the Leibniz series converges follows directly from Dirichlet's test.

2. Consider the following series: $\sum_{k=1}^{\infty} \frac{\sin kx}{k^\alpha}$, $x \in \mathbb{R}$, $\alpha > 0$. If $\alpha > 1$, then this series converges absolutely for any $x \in \mathbb{R}$, since in this case the absolute value of its common term can be estimated as follows:

$$\left|\frac{\sin kx}{k^\alpha}\right| \le \frac{1}{k^\alpha}.$$

Earlier, when discussing the integral convergence test, we established that the series $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$ converges for $\alpha > 1$. Therefore, using the comparison test, we obtain that the series $\sum_{k=1}^{\infty}\left|\frac{\sin kx}{k^\alpha}\right|$ also converges, which means that the series $\sum_{k=1}^{\infty} \frac{\sin kx}{k^\alpha}$ converges absolutely.

Consider the case $\alpha \in (0, 1]$ and show that in this case all conditions of Dirichlet's test are satisfied for the series $\sum_{k=1}^{\infty} \frac{\sin kx}{k^\alpha}$.

First, we discard the case of $x = 2\pi m$, $m \in \mathbb{Z}$, since in this case all terms of the series turn to 0, and therefore the sum of the series is also 0.

We take $\frac{1}{k^\alpha}$ as b_k, since it is obvious that the sequence $\left\{\frac{1}{k^\alpha}\right\}$ is monotonic (decreasing) and approaches zero as $k \to \infty$. We take $\sin kx$ as a_k and show that for partial sums $\sum_{k=1}^{n} \sin kx$, condition 1 of Dirichlet's test is satisfied.

The following estimate for partial sums $\sum_{k=1}^{n} \sin kx$ can be proved for all $n \in \mathbb{N}$:

$$\left|\sum_{k=1}^{n} \sin kx\right| \le \frac{1}{\left|\sin \frac{x}{2}\right|}.$$

Thus, condition 1 of Dirichlet's test is also satisfied, and the series $\sum_{k=1}^{\infty} \frac{\sin kx}{k^\alpha}$ is convergent for $\alpha \in (0, 1]$. However, for these values of α convergence is conditional.

Theorem (Abel's test for conditional convergence of a numerical series).

Let the following conditions be satisfied for a series $\sum_{k=1}^{\infty} a_k b_k$:

1) the series $\sum_{k=1}^{\infty} a_k$ converges;

2) the sequence $\{b_k\}$ is monotone and bounded.

Then the series $\sum_{k=1}^{\infty} a_k b_k$ converges (generally speaking, conditionally).

Remark.

If we compare Dirichlet's test and Abel' test, then it can be noted that in Abel's test, condition 1 is stronger (since the convergence of the corresponding series is required instead of uniformly boundedness of its partial sums), and condition 2 is weaker (since it is not necessary that the sequence $\{b_k\}$ had a zero limit).

Chapter 7

FIRST-ORDER ORDINARY DIFFERENTIAL EQUATIONS

7.1. Introduction

Let x be the independent variable, and let y be the dependent variable.

A **differential equation** is an equation, which involves the derivative of a function $y(x)$. The equation may also contain the function itself as well as the independent variable.

The general form of a differential equation of the first order is

$$F(x, y, y') = 0. \tag{1}$$

The solution procedure consists in finding the unknown function $y(x)$, which obeys equation (1) on a given interval.

The general solution of equation (1) is a function $y = \varphi(x, C)$, which is the solution of (1) for any values of a parameter C. By setting $C = const$ we obtain a particular solution of equation (1).

Sometimes the solution can be found in the implicit form only. If the equation

$$\Phi(x, y, C) = 0, \tag{2}$$

determines the general solution of (1), then it is called the general integral of the differential equation.

If there given an initial condition $y(x_0) = y_0$ in addition to equation (1), then it is necessary to find the particular solution, which obeys the initial condition.

Here we consider only such classes of first-order differential equations, which can be solved analytically.

7.2. Directly Integrable Equations

A directly integrable differential equation has the following form:

$$y' = f(x), \tag{3}$$

where $f(x)$ is a given function.

From this equation follows that the function $y(x)$ is a primitive of $f(x)$ and hence

$$y(x) = \int f(x)dx + C. \tag{4}$$

A constant C can be determined from the initial condition, if the one is given.

Example: Find the solution of the equation

$$y'(x) = x + \cos x$$

with the initial condition $y(0) = 1$.

Solution: In view of (4) the general solution is

$$y(x) = \int (x + \cos x)dx + C = \frac{x^2}{2} + \sin x + C.$$

Taking into account the initial condition, we find: $1 = 0 + C$, that is, $C = 1$. Therefore, the function $y(x) = x^2/2 + \sin x + 1$ being the solution of the given equation, satisfies the initial condition.

7.3. Separable Equations

A separable differential equation is an equation of the form

$$y' = f(x)g(y), \tag{5}$$

that is, $y'(x)$ equals the product of given functions, $f(x)$ and $g(y)$, each of which is a function of one variable only.

We can not integrate equation (5) directly because the right-hand side contains an unknown function $y(x)$ together with the variable x.

To separate the variables we rewrite the equation in the form:

$$\frac{dy}{g(y)} = f(x)dx \tag{5a}$$

and then integrate both sides:

$$\int \frac{dy}{g(y)} = \int f(x)dx + C. \tag{6}$$

Thus, the general integral of equation (5) is found.

A differential equations of the form

$$y' = f(ax + by + c) \tag{7}$$

can be reduced to a separable equation by introducing of a new dependent variable $u(x)$ instead of y:

$$u = ax + by + c. \tag{8}$$

Next we have to derive the equation for the variable $u(x)$. By differentiating (8), we obtain $u' = a + by'$, which implies the equation

$$u' = a + b\,f(u)$$

being the separable equation.

Then we obtain $\quad \dfrac{du}{b\,f(u) + a} = dx \quad \Rightarrow \quad \displaystyle\int \frac{du}{b\,f(u) + a} = x + C.$

Example 1: Solve the equation

$$y' = e^{2x-3y}.$$

Solution: The variables can be easily separated:

$$e^{3y}dy = e^{2x}dx.$$

By integrating, we obtain a general integral of the given equation:

$$\frac{1}{3}e^{3y} = \frac{1}{2}e^{2x} + C.$$

By means of simple formula manipulations we can also write the general solution in the explicit form:

$$y = \frac{1}{3}\ln(\frac{3}{2}e^{2x} + C),$$

where the constant $3C_1$ is denoted by C.

Example 2: Find the solution of the equation

$$y' = \cos(x+y), \qquad (9)$$

which obeys the initial condition $y(0) = \pi/2$.

Solution: Let us introduce a new variable:

$$u = x + y.$$

Then from (9) we obtain the separable equation for $u(x)$

$$u' = 1 + \cos u.$$

By separating the variables and integrating, we have:

$$\int \frac{du}{1+\cos u} = x + C.$$

Using the formula $1 + \cos u = 2\cos^2 u/2$ we obtain the algebraic equation

$$\tan(u/2) = x + C,$$

which implies

$$u = 2\arctan(x + C).$$

Since $y = u - x$, the general solution of the given equation is the following one: $y = 2\arctan(x+C) - x$.

The initial condition yields: $\pi/2 = 2\arctan C$, so that $C = 1$.

Finally we obtain:

$$y = 2\arctan(x+1) - x.$$

7.4. Homogeneous Equations

If some differential equation can be represented in the following form:

$$y' = f(\frac{y}{x}), \tag{10}$$

then it is called a homogeneous equation.
One of the main methods of solving differential equations is based on introducing a new dependent variable $u(x)$ instead of y. There is no general rule to make the right choice of u because it depends on the form of the equation. That is why it is necessary to consider different classes of equations separately. One of typical techniques of such a kind is illustrated below by solving an homogeneous equation.
The right-hand side of equation (10) suggests the substitution $u = y/x$. Then we have to derive the equation for the new dependent variable u.
To find the derivative of $y = ux$, we use the rule of differentiation of the product:

$$y' = u'x + u.$$

From (10) we obtain the equation

$$u'x + u = f(u),$$

which being rewritten in the form

$$u' = \frac{1}{x}(f(u) - u) \tag{11}$$

is a separable equation. Then the problem of integration is solved just in the same way as above. (See equation (5).)

Example: Solve the equation

$$y' = \frac{y}{x - \sqrt{xy}}. \tag{12}$$

Solution: Since

$$\frac{y}{x - \sqrt{xy}} = \frac{y/x}{1 - \sqrt{y/x}} = f(\frac{y}{x}),$$

the given equation is the homogeneous equation.
To solve this problem, we introduce the variable $u = y/x$ instead of y and derive a differential equation for $u(x)$.
First, $y = ux$, so $y' = u'x + u$. Therefore, by (12),

$$u'x + u = \frac{u}{1 - \sqrt{u}} \qquad \Rightarrow \qquad u'x = \frac{\sqrt{u^3}}{1 - \sqrt{u}} \qquad \Rightarrow$$

$$\frac{1-\sqrt{u}}{\sqrt{u^3}}du = \frac{dx}{x} \quad \Rightarrow \quad \int (u^{-3/2} - \frac{1}{u})du + C = \int \frac{dx}{x} \quad \Rightarrow$$

$$-2/\sqrt{u} - \ln|u| = \ln|x| - C.$$

Replacing u by y/x we obtain the general integral of equation (12):

$$\ln|y| + 2\sqrt{x/y} = C. \tag{13}$$

7.5. Linear Equations

A linear differential equation is an equation, which can be represented as

$$y' + P(x)y = Q(x), \tag{14}$$

where $P(x)$ and $Q(x)$ are given functions.

To solve the equation, we introduce a new dependent variable $u(x)$ instead of y by the equality

$$y = u(x)\,v(x), \tag{15}$$

keeping in mind to determine a function $v(x)$ later.

To derive the differential equation for $u(x)$ we find the derivative $y' = u'v + uv'$ and substitute it into original equation (14):

$$u'v + v'u + P(x)\;u\,v = Q(x).$$

Next we group the terms and take out the common factor:

$$u'v + u(v' + P(x)\,v) = Q(x). \tag{16}$$

Now we are ready to determine the function $v(x)$. Let $v(x)$ be a function such that

$$v' + P(x)\,v = 0. \tag{17}$$

By separating the variables, we obtain the solution of equation (17):

$$\int \frac{dv}{v} = -\int P(x)dx \quad \Rightarrow \quad \ln|v| = -\int P(x)dx \quad \Rightarrow$$

$$v = e^{-\int P(x)dx}. \tag{18}$$

A constant of integration is chosen to be equal to zero because it is enough to have one function only, which obeys condition (17).

In view of (18), equation (16) is reduced to the directly integrable equation of the form

$$u' = Q(x)e^{f(x)}, \tag{19}$$

where $f(x) = \int P(x)dx$ is one of primitives of P.

Therefore,

$$u(x) = \int Q(x)e^{f(x)}dx + C. \qquad (20)$$

Thus, equation (14) has the following general solution:

$$y(x) = e^{-f(x)}(\int Q(x)e^{f(x)}dx + C). \qquad (21)$$

Example: Find the general solution of the equation

$$y' = 3y/x + x. \qquad (22)$$

Solution: Let $y = uv$. Then $y' = u'v + uv'$.
Substituting these expressions into the original equation, we obtain

$$u'v + v'u = 3uv/x + x \quad \Rightarrow$$

$$u'v + u(v' - 3v/x) = x. \qquad (23)$$

Then we find the function $v(x)$ by solving of the equation

$$v' - 3v/x = 0.$$

The variables are easily separated and we have

$$\int \frac{dv}{v} = 3\int \frac{dx}{x} \quad \Rightarrow \quad \ln|v| = 3\ln|x| \quad \Rightarrow \quad v = x^3.$$

Now we come back to (23), which is reduced to the separable equation

$$u'x^3 = x.$$

Therefore, $$u = \int \frac{dx}{x^2} + C = -\frac{1}{x} + C.$$

Finally, we obtain

$$y = u\,v = (-\frac{1}{x} + C)x^3 = -x^2 + Cx^3.$$

7.6. The Bernoulli Equations

The Bernoulli Equation is an equation of the form

$$y'(x) + P(x)y = Q(x)y^n, \qquad (24)$$

where n is any rational number except 0 and 1.

The technique of solving the Bernoulli equations is just the same as for linear equations: A new dependent variable $u(x)$ is introduced by means of the equality

$$y = u(x)\,v(x). \qquad (25)$$

This variable satisfies the equation

$$u'v + u(v' + P(x)\,v) = Q(x)\,u^n v^n, \qquad (26)$$

where the function $v(x)$ is a partial solution of the equation

$$v' + P(x)v = 0 \tag{27}$$

and hence,

$$v = e^{-\int P(x)dx}. \tag{28}$$

Therefore, equation (26) is transformed to the form

$$u'v = Q(x)u^n v^n$$

and can be rewritten as a separable equation:

$$u^{-n}du = Q(x)v^{n-1}dx.$$

By integrating, we obtain

$$\frac{1}{-n+1}u^{-n+1} = \int Q(x)v^{n-1}dx + C. \tag{29}$$

Thus,

$$u(x) = \left((1-n)\int Q(x)v^{n-1}dx + C\right)^{\frac{1}{1-n}}. \tag{30}$$

The general solution of (24) is $y(x) = u(x)v(x)$.

Example: Find the general solution of the equation

$$y' + 4xy = 2xe^{-x^2}\sqrt{y}. \tag{31}$$

Solution: Let $y = uv$. Since the derivative of y is $y' = u'v + uv'$, then (31) can be transformed to the equation with respect to the variable $u(x)$:

$$u'v + v'u + 4xuv = 2xe^{-x^2}\sqrt{uv} \qquad \Rightarrow$$

$$u'v + u(v' + 4xv) = 2xe^{-x^2}\sqrt{uv}. \tag{32}$$

To find the function $v(x)$, we solve the equation

$$v' + 4vx = 0.$$

This is the separable equation, and its partial solution is

$$v = e^{-2x^2}. \tag{33}$$

From (32) we have

$$u'e^{-2x^2} = 2xe^{-x^2}\sqrt{ue^{-2x^2}} \quad \Rightarrow \quad u' = 2x\sqrt{u} \quad \Rightarrow$$

$$\int \frac{du}{\sqrt{u}} = \int 2xdx + C \quad \Rightarrow \quad 2\sqrt{u} = x^2 + C \quad \Rightarrow$$

$$u = (x^2 + C)^2/4. \tag{34}$$

Therefore, the general solution of the given equation is

$$y(x) = \frac{1}{4}(x^2 + C)^2 e^{-2x^2}. \tag{35}$$

7.1. Exact Differential Equations

An exact differential equation has the following form

$$P(x,y)dx + Q(x,y)dy = 0, \qquad (36)$$

where the partial derivatives of $P(x,y)$ and $Q(x,y)$ obey the condition

$$P'_y = Q'_x. \qquad (37)$$

Due to condition (37), the expression on the left-hand side of (36) is the total differential of some function $u(x,y)$ by the theorem of a total differential (See Chapter 2, page 35.):

$$du(x,y) = P(x,y)dx + Q(x,y)dy = 0.$$

Therefore,

$$\frac{\partial u(x,y)}{\partial x} = P(x,y), \qquad (38)$$

$$\frac{\partial u(x,y)}{\partial y} = Q(x,y). \qquad (39)$$

If we hold fixed y, then by integrating of (38) with respect to x we obtain

$$u(x,y) = \int P(x,y)dx + \varphi(y). \qquad (40)$$

Note that a constant of integration may be a function of y because y is fixed during integration.

To find the function $\varphi(y)$, we substitute this expression for $u(x,y)$ into (39):

$$\frac{\partial}{\partial y}\int P(x,y)dx + \varphi'(y) = Q(x,y) \quad \Rightarrow$$

$$\varphi'(y) = Q(x,y) - \frac{\partial}{\partial y}\int P(x,y)dx. \qquad (41)$$

This is an ordinary differential equation for the function $\varphi(y)$. Note also that the expression on the right-hand side is a function of y only. Otherwise, the equation (36) is not an exact differential equation.

By solving equation (41), we find a partial solution $\varphi(y)$ and hence, the general solution: $u(x,y) = C$.

Example: Find the general solution of the equation

$$(y + 2x^{-2})dx + (x - 3y^{-2})dy = 0. \qquad (42)$$

Solution: Here $P(x,y) = y + 2x^{-2}$ and $Q(x,y) = x - 3y^{-2}$.

Let us check whether $P'_y(x,y) = Q'_x(x,y)$.

$$P'_y(x,y) = \frac{\partial}{\partial y}\left(y + \frac{2}{x^2}\right) = 1 \quad \text{and} \quad Q'_x(x,y) = \frac{\partial}{\partial x}\left(x - \frac{3}{y^2}\right) = 1.$$

Therefore, equation (42) is the exact differential equation of the form $du(x, y) = 0$.
The general solution of this equation is $u(x, y) = C$.
All we need to write the answer is the function $u(x, y)$.
By formula (40),

$$u(x, y) = \int (y + \frac{2}{x^2})dx + \varphi(y) = yx - \frac{2}{x} + \varphi(y). \qquad (43)$$

In view of (41), we have

$$\varphi'(y) = x - \frac{3}{y^2} - \frac{\partial}{\partial y}(yx - \frac{2}{x}) \Rightarrow$$

$$\varphi'(y) = x - 3y^{-2} - x = -3y^{-2}.$$

This is the directly integrable equation with a partial solution $\varphi(y) = 3/y$.
Thus, by formula (44), we find the required function $u(x, y)$:

$$u(x, y) = yx - \frac{2}{x} + \frac{3}{y}.$$

Therefore, equation (42) has the following general integral:

$$yx - \frac{2}{x} + \frac{3}{y} = C.$$

In conclusion, let us note that any equation (36) can be transformed to the exact differential equation by multiplying both sides by some integrating factor $\mu(x, y)$. It is known that such a factor exists, however there is no general rule to find this factor but the following two cases:

1) If the expression $(Q'_x - P'_y)/P$ depends on the variable y only, then the integrating factor is also a function of y only, which obeys the equation
$$\frac{d \ln \mu(y)}{dy} = \frac{1}{P(x, y)}(Q'_x - P'_y).$$
2) If the expression $(Q'_x - P'_y)/Q$ depends on the variable x only, then the integrating factor is also a function of x only, which obeys the equation
$$\frac{d \ln \mu(x)}{dx} = -\frac{1}{Q(x, y)}(Q'_x - P'_y).$$