IV. Indefinite integral

IV.1. Antiderivative, indefinite integral

IV.1.1. Antiderivative. Let I be an interval with end points a, b (where $a < b$; numbers a, b can also be infinite). A function F is called an <u>*antiderivative*</u> to function f on interval I if

a) $F'(x) = f(x)$ for all $x \in (a, b)$,

b) $F'_+(a) = f(a)$ (if point a belongs to interval I),

c) $F'_-(b) = f(b)$ (if point b belongs to interval I).

IV.1.2. Remark. Instead of "F is an antiderivative to function f on interval I", we shall sometimes write only "$F' = f$ in I".

Almost the whole part IV of this textbook deals with the question how to find an antiderivative to a given function f. The antiderivative is important mainly because it enables to evaluate a so called definite integral of function f. (See part V.) Moreover, the solution of many differential equations also leads to the problem of finding the antiderivative. (See section IV.7.)

To find an antiderivative to a given function f is an "inverse" problem to the differentiation of function f. We already know that not every function has a derivative. Similarly, it is also not true that an antiderivative automatically exists to every given function. The following theorem gives a sufficient condition for the existence of an antiderivative to f on interval I.

IV.1.3. Theorem (on the existence of an antiderivative). *If function f is continuous on interval I then there exists an antiderivative to f on I.*

IV.1.4. Remark. The other question is whether an antiderivative to function f on interval I (if it exists) is unique. The following theorem says that it is not unique. There exist infinitely many antiderivatives, but any two of them differ at most in an additive constant.

IV.1.5. Theorem. a) *If F is an antiderivative to function f on interval I and C is an arbitrary real constant then function G defined by the formula $G(x) = F(x) + C$ (for $x \in I$) is also an antiderivative to function f on interval I.*

b) If F and G are two antiderivatives to function f on interval I then there exists $C \in \mathbb{R}$ such that $G(x) = F(x) + C$ for all $x \in I$.

P r o o f : a) If $F' = f$ in I, then also $G' = (F + C)' = F' + C' = F' = f$ in I.

b) If $F' = f$ in I and $G' = f$ in I, then $(G - F)' = G' - F' = f - f = 0$ in I. Hence the function $H = G - F$ has a zero derivative in interval I and so it is a constant function in interval I. (See theorem III.5.4, item e).) Thus, there exists $C \in \mathbb{R}$ such that $H(x) = G(x) - F(x) = C$ for all $x \in I$. This means that $G(x) = F(x) + C$ for all $x \in I$.

IV.1.6. Theorem. *If F and G are antiderivatives to functions f and g on interval I, then $F + G$ is an antiderivative to $f + g$ on interval I.*

If F is an antiderivative to f on interval I and α is a real number, then $\alpha \cdot F$ is an antiderivative to $\alpha \cdot f$ on interval I.

P r o o f : This follows immediately from theorem III.4.9, parts a) and b).

IV.1.7. Indefinite integral. The <u>*indefinite integral*</u> of function f on interval I is the set of all antiderivatives to function f on interval I. This set is denoted by

$$\int f(x)\,\mathrm{d}x, \quad x \in I \qquad \text{or briefly only} \qquad \int f(x)\,\mathrm{d}x, \quad \int f\,\mathrm{d}x.$$

If F is an antiderivative to function f on interval I, then the set $\int f(x)\,\mathrm{d}x$ contains only function F and all other functions that differ from F at most in an additive constant on I. This can be expressed by the equation

$$\int f(x)\,\mathrm{d}x = F(x) + C, \qquad x \in I.$$

The function f is called the <u>*integrand*</u> and the constant C is called the <u>*constant of integration*</u>.

It follows from the definition of the indefinite integral that the integral $\int f(x)\,\mathrm{d}x$ in interval I exists if and only if function f has an antiderivative in interval I. Due to theorem IV.1.3, the continuity of function f in I is a sufficient condition for the existence.

IV.1.8. Theorem. *If the indefinite integrals of functions f and g exist on interval I and $\alpha \in \mathbb{R}$ then*

$$\int [\,f(x) + g(x)\,]\,\mathrm{d}x = \int f(x)\,\mathrm{d}x + \int g(x)\,\mathrm{d}x, \qquad x \in I \tag{IV.1.1}$$

$$\int \alpha f(x)\,\mathrm{d}x = \alpha \int f(x)\,\mathrm{d}x, \qquad x \in I. \tag{IV.1.2}$$

This theorem directly follows from theorem IV.1.6. Equality (IV.1.1) is in fact the equality of sets of functions. It says:

a) An arbitrary function from the set on the left hand side, i.e. from $\int [f(x) + g(x)]\,\mathrm{d}x$, can be expressed as a sum of two functions from the sets on the right hand side, i.e. a function from the set $\int f(x)\,\mathrm{d}x$ and a function from the set $\int g(x)\,\mathrm{d}x$.

b) On the other hand, the sum of two arbitrary functions from the sets on the right hand side, i.e. a function from $\int f(x)\,\mathrm{d}x$ and a function from $\int g(x)\,\mathrm{d}x$, belongs to the set on the left hand side, i.e. to $\int [f(x) + g(x)]\,\mathrm{d}x$.

The precise sense of equality (IV.1.2) can be explained similarly.

IV.1.9. Corollary. Theorem IV.1.8 can be generalized: *If functions $f_1, \ldots, f_n$ have indefinite integrals in interval I and $\alpha_1, \ldots, \alpha_n$ are real numbers then*

$$\int [\,\alpha_1 f_1(x) + \ldots + \alpha_n f_n(x)\,]\,\mathrm{d}x = \alpha_1 \int f_1(x)\,\mathrm{d}x + \ldots + \alpha_n \int f_n(x)\,\mathrm{d}x, \quad x \in I.$$

IV.1.10. Table of basic indefinite integrals. Basic formulas for some indefinite integrals can simply be obtained by "inverting" the formulas for differentiation of elementary functions (see paragraphs III.4.10 and III.4.15). Each of these formulas holds on an arbitrary interval I which is contained in the domain of the integrand.

a) $\int 0 \, dx = C$, b) $\int x^\alpha \, dx = \frac{x^{\alpha+1}}{\alpha+1} + C \quad (\alpha \neq -1)$,

c) $\int \cos x \, dx = \sin x + C$, d) $\int \sin x \, dx = -\cos x + C$,

e) $\int \frac{1}{(\cos x)^2} \, dx = \tan x + C$, f) $\int \frac{1}{(\sin x)^2} \, dx = -\cot x + C$,

g) $\int \frac{1}{\sqrt{1-x^2}} \, dx = \arcsin x + C$, h) $\int \frac{1}{1+x^2} \, dx = \arctan x + C$,

i) $\int a^x \, dx = \frac{a^x}{\ln a} + C \quad (a > 0,\ a \neq 1)$, j) $\int e^x \, dx = e^x + C$,

k) $\int \frac{1}{x} \, dx = \ln |x| + C$.

IV.1.11. Remark. It is not quite clear at first sight how formula k) follows from the formula for differentiation of the function $\ln x$ (i.e. from formula f) in paragraph III.4.15). $|x| = x$ in the interval $(0, +\infty)$, so the validity of k) is obvious in this interval. However, since $|x| = -x$ in the interval $(-\infty, 0)$, we also have

$$[\ln |x|]' = \frac{1}{|x|} \cdot |x|' = \frac{1}{(-x)} \cdot (-x)' = \frac{1}{(-x)} \cdot (-1) = \frac{1}{x} \qquad \text{for } x \in (-\infty, 0).$$

This confirms the validity of formula k) in the interval $(-\infty, 0)$, too.

IV.1.12. Remark. Using the formulas for derivatives of functions arc sine and arc cosine (from paragraph III.4.15), we can also express the integrals from items g) and h) in paragraph IV.1.10 as follows:

g)* $\int \frac{1}{\sqrt{1-x^2}} \, dx = -\arccos x + C$, h)* $\int \frac{1}{1+x^2} \, dx = -\text{arccot}\, x + C$.

This means that the difference $[\arcsin x - (-\arccos x)]$ is a constant function in the interval $(-1, +1)$ and similarly, the difference $[\arctan x - (-\text{arccot}\, x)]$ is a constant function in the interval $(-\infty, +\infty)$.

IV.1.13. Remark. *Integration* is the name given to the procedure leading to an antiderivative or to an indefinite integral of a given function f (which is practically the same thing, since in order to write down an indefinite integral it is sufficient to know any of the antiderivatives to f).

While the differentiation of a given function can be more or less well algorithmized (it usually suffices to choose and apply appropriate differentiation formulas), this cannot be said of the integration. There exists no general algorithm which will always lead to a desired result and which will provide a rule for the integration of any

function. Conversely, there exist a lot of relatively simple functions whose antiderivatives exist, but they cannot be expressed in a "closed form", i.e. by a formula that prescribes the performance of a finite number of operations and allows us to use only elementary functions in them. (An example of such a function is $f(x) = e^{-x^2}$ in the interval $(-\infty, +\infty)$.) A large range of various methods have been developed which enable us to find indefinite integrals for various classes of "similar" functions. This was an object of interest of many mathematicians especially in the 18th and 19th Centuries. Many of these methods are simple, while others are quite complicated, involving many relatively artificial substitutions, etc. Nowadays, these complicated procedures have rather lost their importance. This has been caused mainly by the development of computers, which make it possible to compute so called definite integrals and to solve differential equations (which are the fields where indefinite integrals are mostly used) by various approximate methods. Nevertheless, this does not mean that we can entirely rely on computers and need not learn to integrate at all! (Children still learn – and will continue to learn – their multiplication tables, even though computers can multiply numbers faster and more reliably.) This is why we show several simple methods for integration in the rest of this chapter.

IV.1.14. Example. $\int (x^3 - 3x^2 + 4x - 1) \, dx = \int x^3 \, dx - 3\int x^2 \, dx + 4\int x \, dx - \int 1 \, dx = \frac{1}{4}x^4 - x^3 + 2x^2 - x + C.$

The method used in this example is called *term by term integration* or *integration by decomposition*. The integrand was decomposed into the sum (or the difference) of more simpler functions, each of which could be integrated individually. The correctness of this approach follows from theorem IV.1.8 and from corollary IV.1.9.

Constants of integration need not be added separately to each indefinite integral in the decomposition shown above. All such constants can be associated to one which can be denoted e.g. by C and added at the end.

IV.2. Integration by parts

IV.2.1. Motivation. We have derived the rule for differentiation of a product of two functions in paragraph III.4.6. This rule can be symbolically written as follows:

$$(u \cdot v)' = u' \cdot v + u \cdot v'.$$

This implies:
$$u' \cdot v = (u \cdot v)' - u \cdot v'.$$

IV.2.2. Theorem (on integration by parts). *If functions u and v have continuous derivatives on interval I then the following formula holds on I:*

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

IV.2.3. Example. $\int x \cdot \sin x \, dx =^{*)} -x \cdot \cos x - \int (-\cos x) \, dx = -x \cdot \cos x + \sin x + C$ (for $x \in \mathbb{R}$).

*) We have used: $u'(x) = \sin x,\ v(x) = x,\ u(x) = -\cos x,\ v'(x) = 1.$

IV.2.4. Example. $\int (x^2-3x+1)\cdot e^x\, dx = {}^{*)}\ (x^2-3x+1)\cdot e^x - \int (2x-3)\cdot e^x\, dx =$
$^{**)} = (x^2-3x+1-\cdot e^x - (2x-3)\cdot e^x + \int 2e^x\, dx = (x^2-5x+6)\cdot e^x + C$ (for $x \in \mathbb{R}$).

*) We have used: $u'(x) = e^x,\ v(x) = x^2-3x+1,\ u(x) = e^x,\ v'(x) = 2x-3.$

**) We have used: $u'(x) = e^x,\ v(x) = 2x-3,\ u(x) = e^x,\ v'(x) = 2.$

IV.2.5. Example. $\int \ln x\, dx = \int 1\cdot \ln x\, dx = {}^{*)}\ x\cdot \ln x - \int x\cdot \frac{1}{x}\, dx =$
$x\cdot\ln x - \int 1\, dx = x\cdot \ln x - x + C$ (for $x > 0$).

*) We have used: $u'(x) = 1,\ v(x) = \ln x,\ u(x) = x,\ v'(x) = 1/x.$

IV.2.6. Example. $\int e^x\cdot \sin x\, dx = {}^{*)}\ e^x\cdot\sin x - \int e^x\cdot \cos x\, dx = {}^{**)}\ e^x\cdot \sin x -$
$- e^x\cdot\cos x + \int e^x\cdot(-\sin x)\, dx = e^x\cdot\sin x - e^x\cdot\cos x - \int e^x\cdot \sin x\, dx.$

*) We have used: $u'(x) = e^x,\ v(x) = \sin x,\ u(x) = e^x,\ v'(x) = \cos x.$

**) We have used: $u'(x) = e^x,\ v(x) = \cos x,\ u(x) = e^x,\ v'(x) = -\sin x.$

If we denote the integral we wish to express by $\mathcal{I}$, then the above computation gives:

$$\mathcal{I} = e^x\cdot\sin x - e^x\cdot\cos x - \mathcal{I}.$$

Thus, we obtain: $\mathcal{I} = \int e^x\cdot\sin x\, dx = \frac{1}{2}[e^x\cdot\sin x - e^x\cdot\cos x] + C$ (for $x\in\mathbb{R}$).

IV.2.7.* Example. $\int \frac{1}{(1+x^2)^n}\, dx = \int 1\cdot\frac{1}{(1+x^2)^n}\, dx = {}^{*)}\ x\cdot\frac{1}{(1+x^2)^n} -$

$$-\int x\cdot\Big[\frac{1}{(1+x^2)^n}\Big]'\, dx = \frac{x}{(1+x^2)^n} - \int x\cdot\Big[\frac{-n}{(1+x^2)^{n+1}}\cdot 2x\Big]\, dx =$$

$$= \frac{x}{(1+x^2)^n} + 2n\int\frac{1+x^2}{(1+x^2)^{n+1}}\, dx - 2n\int\frac{1}{(1+x^2)^{n+1}}\, dx.$$

*) We have used: $u'(x) = 1,\ v(x) = \frac{1}{(1+x^2)^n},\ u(x) = x,\ v'(x) = \frac{(-n)\cdot 2x}{(1+x^2)^{n+1}}.$

Denote $\mathcal{I}_n = \int\frac{1}{(1+x^2)^n}\, dx.$ Applying integration by parts, we obtain the equation

$$\mathcal{I}_n = \frac{x}{(1+x^2)^n} + 2n\cdot\mathcal{I}_n - 2n\cdot\mathcal{I}_{n+1}.$$

This leads to the formula $\mathcal{I}_{n+1} = \frac{1}{2n}\cdot\frac{x}{(1+x^2)^n} + \frac{2n-1}{2n}\mathcal{I}_n.$

Writing this for example with $n = 1$ and substituting for $\mathcal{I}_1$ from IV.1.9 h), we get:

$$\mathcal{I}_2 = \frac{1}{2}\frac{x}{1+x^2} + \frac{1}{2}\int\frac{1}{1+x^2}\, dx = \frac{1}{2}\frac{x}{1+x^2} + \frac{1}{2}\arctan x + C; \qquad x\in\mathbb{R}.$$

IV.2.8. Problems. Find the integrals:

a) $\int x\cdot e^x\, dx,$ b) $\int x\cdot\cos x\, dx,$ c) $\int x^2\cdot\cos x\, dx,$

d) $\int x^2\cdot\ln x\, dx,$ e) $\int x\cdot \operatorname{arccot} x\, dx,$ f) $\int e^x\cdot\cos x\, dx.$

Results: a) $e^x\cdot(x-1) + C$ (for $x\in\mathbb{R}$),

b) $x\cdot\sin x + \cos x + C$ (for $x\in\mathbb{R}$),

c) $x^2\cdot\sin x + 2x\cos x - 2\sin x + C$ (for $x\in\mathbb{R}$),

d) $\frac{1}{3}x^3[\ln x - \frac{1}{3}] + C$ (for $x > 0$),

e) $\frac{1}{2}(x^2+1)\cdot\operatorname{arccot} x + \frac{1}{2}x + C$ (for $x\in\mathbb{R}$),

f) $\frac{1}{2}e^x(\sin x + \cos x) + C$ (for $x\in\mathbb{R}$).

IV.3. Integration by substitution

IV.3.1. Motivation. Analogously, as integration by parts represents the reverse procedure to the differentiation of a product of two functions, integration by substitution is a method which arose by inverting the Chain Rule (see theorem III.4.14).

IV.3.2. Theorem. *Suppose that function $f(x)$ is continuous on interval I. Suppose further that function $x = g(t)$ is differentiable on interval J and it maps J onto I. Then*

(IV.3.1) $$\int f(x)\, dx = \int f(g(t))\cdot g'(t)\, dt \qquad \text{for } x\in I,\ t\in J,\ x = g(t).$$

Remark IV.3.3. The indefinite integral on the left hand side of (IV.3.1) is the set of all primitive functions to function $f(x)$ on interval I. Similarly, the indefinite integral on the right hand side is the set of all primitive functions to function $f(g(t))\cdot g'(t)$ on interval J. Thus, equality (IV.3.1) is the equality between two sets of functions. It looks strange at the first sight, because the set on the left hand side contains functions defined on interval I, while the set on the right hand side contains functions defined on interval J. However, it is necessary to understand the equality in such a way that if one replaces x by $g(t)$ in every function in the set on the left hand side, then both sets coincide.

Formula (IV.3.1) can be used in two situations:

1) We wish to evaluate the integral on the right hand side and we transform it to the integral on the left hand side. (This has a sense if the integral on the right

hand side cannot be computed directly and the integral on the left hand side is simpler.)

2) We wish to find the integral on the left hand side and we transform it to the integral on the right hand side. (This has a sense if the integral on the right hand side is simpler.)

We describe both possibilities in so called first and second substitution methods.

IV.3.4. The 1st method of substitution. Assume that all assumptions of theorem IV.3.2 are fulfilled. Suppose that we wish to evaluate the integral $\int f(g(t)) \cdot g'(t)\, dt$ on the right hand side of (IV.3.1). We put $g(t) = x$ and $g'(t)\, dt = dx$. So we obtain the integral appearing on the left hand side of (IV.3.1). We evaluate this integral (on interval I). Finally, replacing x by $g(t)$, we obtain the desired integral $\int f(g(t)) \cdot g'(t)\, dt$.

IV.3.5. Example. Evaluate $\int (\sin t)^3 \cdot \cos t\, dt$. Using the substitution $\sin t = x$ and the equality $(\sin t)'\, dt = \cos t\, dt = dx$, we obtain:

$$\int (\sin t)^3 \cdot \cos t\, dt = \int x^3\, dx = \tfrac{1}{4} x^4 + C = \tfrac{1}{4} (\sin t)^4 + C.$$

The assumptions of theorem IV.3.2 are fulfilled, because the function $g(t) = \sin t$ is differentiable in $J = (-\infty, +\infty)$, maps this interval onto $I = [-1, +1]$ and function $f(x) = x^3$ is continuous on interval I.

IV.3.6. Remark. When we use the 1st method of substitution then the exact specification of interval I which is the range of $g|_J$ is not necessary. In many cases, it is easier to verify that g maps J into (and not necessarily onto) some interval I' where f is continuous. Then f is also continuous in the interval $I = R(g|_J)$, because $I \subset I'$.

Thus, as a part of the verification of the assumptions of theorem IV.3.2 in example IV.3.5, it is sufficient to find out that function g maps interval J into the interval $\mathcal{I} = (-\infty, +\infty)$ where the function $f(x) = x^3$ is continuous.

IV.3.7. Remark. We can use various denotations for the variable in integration. It is usually denoted by x, but this is not a rule. Thus, if we start with an integral of the type $\int f(g(x)) \cdot g'(x)\, dx$ then the new variable can be denoted for example by s. Using the substitution $s = g(x)$, $ds = g'(x)\, dx$, we transform this integral to $\int f(s)\, ds$. Get used to various denotations of variables!

IV.3.8. The 2nd method of substitution. Suppose that the assumptions of theorem IV.3.2 are fulfilled and moreover, function g is strictly monotone in interval J. Assume that we wish to compute the integral $\int f(x)\, dx$ in interval I. We put $x = g(t)$ and $dx = g'(t) \cdot dt$. Thus, we transform the problem to the computation of the integral $\int f(g(t)) \cdot g'(t)\, dt$ in interval J. We evaluate this integral, replace t by $g_{-1}(x)$ and we obtain the expression of the original integral $\int f(x)\, dx$.

IV.3.9. Remark. The application of the 2nd method requires the inverse substitution $t = g_{-1}(x)$ at the very end of the computation. The assumption about the strict monotonicity of function g on interval J guarantees the existence of the inverse function g_{-1} to g.

IV.3.10. Example. Evaluate $\int \sqrt{9 - x^2}\, dx$. The integrand is defined and continuous in the interval $[-3, 3]$. Hence we will evaluate the integral on this interval. We use the substitution $x = g(t) = 3 \sin t$ (for $t \in [-\pi/2, \pi/2]$). Function g is increasing (and therefore strictly monotone) on the interval $[-\pi/2, \pi/2]$, it is differentiable here and it maps this interval onto the interval $[-3, 3]$. Thus, function g satisfies the assumptions of the second substitution method. We also replace dx by $g'(t)\, dt$, i.e. by $\cos t\, dt$ in the integral. We obtain:

$$\int \sqrt{9 - x^2}\, dx = \int \sqrt{9 - 9 \sin^2 t} \cdot 3 \cos t\, dt = 9 \int \sqrt{\cos^2 t} \cdot \cos t\, dt =$$

$$= 9 \int |\cos t| \cdot \cos t\, dt \overset{*)}{=} 9 \int \cos^2 t\, dt = 9 \int \frac{1 + \cos(2t)}{2}\, dt =$$

$$= \frac{9}{2} \left[t + \frac{\sin(2t)}{2} \right] + C = \frac{9}{2} [t + \sin t \cdot \cos t] + C =$$

$$\overset{**)}{=} \tfrac{9}{2} \left[\arcsin\left(\tfrac{1}{3} x\right) + \tfrac{1}{3} x \cdot \sqrt{1 - \tfrac{1}{9} x^2} \right] + C.$$

*) Function cosine is nonnegative in the interval $[-\pi/2, \pi/2]$, hence $|\cos t| = \cos t$ for all t from this interval.

**) For $t \in [-\pi/2, \pi/2]$, we have: $x = 3 \sin t \iff t = \arcsin\left(\tfrac{1}{3} x\right)$.

IV.3.11. Example. Evaluate the integral $\int \frac{1}{1 - x}\, dx$. The domain of the function $1/(1 - x)$ is the union $(-\infty, 1) \cup (1, +\infty)$. The function $1/(1 - x)$ is continuous on the two intervals $(-\infty, 1)$ and $(1, +\infty)$, hence an antiderivative and the indefinite integral exist on each of them.

We use the substitution $x = 1 - u$. Further, we replace dx by $(1 - u)'\, du = -du$. We obtain

$$\int \frac{1}{1 - x}\, dx = -\int \frac{1}{u}\, du = -\ln|u| + C = -\ln|1 - u| + C.$$

Verify for yourself that all assumptions of the 2nd method of substitution are fulfilled here.

IV.3.12. Problems.

a) $\int \frac{x}{\sqrt[3]{1 - x^2}}\, dx$ b) $\int \frac{\cos x}{(\sin x)^3}\, dx$ c) $\int \frac{\exp \sqrt{x}}{\sqrt{x}}\, dx$ d) $\int \frac{\sqrt{\arctan x}}{1 + x^2}\, dx$

e) $\int \frac{\ln x - 3}{x \sqrt{\ln x}}\, dx$ f) $\int e^{-x}\, dx$ g) $\int \frac{e^x}{1 + e^{2x}}\, dx$ h) $\int (1 - 2x)^{1000}\, dx$

i) $\int \frac{1}{\sqrt{1 - 2x^2}}\, dx$ j) $\int \frac{1}{\sqrt{2^x + 1}}\, dx$ k) $\int \frac{\sqrt{1 - x}}{x}\, dx$ l) $\int \frac{dx}{(1 + x^2) \arctan x}$

Results: a) $-\frac{3}{4}(1-x^2)^{2/3}+C$ (on intervals $(-\infty,-1)$, $(-1,1)$, $(1,+\infty)$),
b) $-1/(2\sin^2 x)+C$ (on each interval $(k\pi,/k+1)\pi)$, k being an integer),
c) $2\exp\sqrt{x}+C$ (on $(-\infty,+\infty)$), d) $\frac{2}{3}(\arctan x)^{3/2}+C$ (on $(-\infty,+\infty)$),
e) $\frac{2}{3}\sqrt{\ln x}(\ln x-9)+C$ (on $(1,+\infty)$), f) $-e^{-x}+C$ (on $(-\infty,+\infty)$),
g) $\arctan(e^x)+C$ (on $(-\infty,+\infty)$), h) $-\frac{1}{2002}(1-2x)^{1001}+C$ (on $(-\infty,+\infty)$),
i) $\arcsin(\sqrt{2}x)/\sqrt{2}+C$ (on $(-1/\sqrt{2},1/\sqrt{2})$), j) $\frac{1}{\ln 2}\ln\left[\frac{\sqrt{2^x+1}-1}{\sqrt{2^x+1}+1}\right]+C$
(on $(-\infty,+\infty)$), k) $2\sqrt{1-x}+\ln\left|\frac{\sqrt{1-x}-1}{\sqrt{1-x}+1}\right|+C$ (on $(-\infty,1)$),
l) $\ln|\arctan x|+C$ (on intervals $(-\infty,0)$ and $(0,+\infty)$).

IV.4. Integration of simple rational functions

IV.4.1. Rational function. A *rational function* is a function of the type P/Q, where P and Q are polynomials and the degree of polynomial Q is greater than or equal to one.

IV.4.2. Remark. We are going to explain how it is possible to integrate rational functions which have a polynomial of the degree at most three in the denominator. The procedure of integration of other rational functions is similar, though it can be technically more complicated. Example IV.6.6 involves the integration of a rational function with a polynomial of the fourth degree in the denominator.

The basic idea is to write the rational function as a sum of terms, so called partial fractions, each of which can be relatively simply integrated. The procedure starts with the decomposition of the polynomial in the denominator (i.e. polynomial Q) to the product of factors, which are linear, in some cases also quadratic, polynomials.

Linear factors appearing in the decomposition have a close relation to the roots of polynomial Q: Q contains the linear factor $(x-\alpha)$ if and only if α is the root of Q.

A quadratic factor appears in the decomposition of Q only if it has no real factors, i.e. if it has no real roots and its discriminant is negative – see cases Ia) and IIa) in the following paragraph.

IV.4.3. The decomposition of a polynomial of at most the third degree.
I) Suppose first that Q is a quadratic polynomial, i.e.

$$Q(x) = q_0x^2+q_1x+q_2 \qquad (\text{where } q_0\neq 0).$$

There are three possibilities:

- Ia) Q has no real root and so it has no real linear factor.
- Ib) Q has a unique (repeated) real root α and it can be decomposed as follows: $Q(x) = q_0(x-\alpha)^2$.
- Ic) Q has two distinct real roots α, β and it can be decomposed as follows: $Q(x) = q_0(x-\alpha)(x-\beta)$.

II) Suppose now that Q is a cubic polynomial, i.e.

$$Q(x) = q_0x^3+q_1x^2+q_2x+q_3 \qquad (\text{where } q_0\neq 0).$$

There are four possibilities:

- IIa) Q has a unique single real root α. In this case Q can be written in the form $Q(x) = q_0(x-\alpha)(x^2+rx+s)$, where (x^2+rx+s) is a quadratic polynomial which is has no real factors (its discriminant is negative).
- IIb) Q has a unique triple real root α. Then $Q(x) = q_0(x-\alpha)^3$.
- IIc) Q has a single real root α and a double real root β. Then Q can be written as follows: $Q(x) = q_0(x-\alpha)(x-\beta)^2$.
- IId) Q has three distinct real roots α, β, γ. Then $Q(x) = q_0(x-\alpha)(x-\beta)(x-\gamma)$.

Sketch for yourself the graph of polynomial Q in all the above cases!

IV.4.4. Example. Decompose the polynomial

$$Q(x) = 2x^3-4x^2-20x-50.$$

to the product of linear or quadratic factors (which are further irreducible in the real domain).

First we need to find the real roots of this polynomial. There exist general formulas (so called Cardan's formulas) which enable us to express all the roots of a cubic polynomial (see e.g. [Re]). However, these formulas are not simple and so we try to find at least one root simply by guessing it. (School examples are usually chosen so that cubic polynomials have at least one "nice" root – like 0, ±1, ±2, etc., and so you have a good chance to guess it.) Substituting successively the numbers $0, \pm1\ \pm2, \dots$ to Q, we find out that $Q(5)=0$. Thus, $\alpha=5$ is the real root of Q. Hence Q is divisible by the factor $(x-5)$. The division gives:

$$Q(x):(x-5) = (2x^3-4x^2-20x-50):(x-5) = 2(x^2+3x+5).$$

The discriminant of the quadratic polynomial x^2+3x+5 is negative (it equals -11), therefore the polynomial has only complex roots and it has no real linear factors. Thus, the desired decomposition of polynomial Q is:

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

This corresponds to case IIa) in the previous paragraph IV.4.3.

IV.4.5. The partial fraction decomposition of a rational function with a quadratic polynomial in the denominator. Suppose that P is a polynomial of the degree less than 2 (i.e. it is either a constant function or it is a linear polynomial). Suppose further that polynomials P and Q are not divisible by the same linear factor. (Otherwise the rational function P/Q could be cancelled by this linear factor and the resulting function would be of another, simpler, type.) In this paragraph, we show how it is possible to write P/Q in the form of a sum of simpler, so called *partial fractions*. The form of the sum depends on the factors of Q. That is why we successively discuss the cases Ia), Ib) and Ic) from paragraph IV.4.3 and we always show the corresponding decomposition of the rational function P/Q.

Ia) $\dfrac{P(x)}{Q(x)} = \dfrac{P(x)}{q_0x^2 + q_1x + q_2}$ .... In this case the rational function $P(x)/Q(x)$ is already a partial fraction.

Ib) $$\frac{P(x)}{Q(x)} = \frac{P(x)}{q_0\,(x-\alpha)^2} = \frac{A_1}{(x-\alpha)} + \frac{A_2}{(x-\alpha)^2}$$

Ic) $$\frac{P(x)}{Q(x)} = \frac{P(x)}{q_0\,(x-\alpha)\,(x-\beta)} = \frac{A}{(x-\alpha)} + \frac{B}{(x-\beta)}$$

There are several methods for calculating coefficients A_1 and A_2 (respectively A and B). One of them is shown in the following example.

IV.4.6. Example. Find the partial fraction decomposition of the rational function

$$\frac{P(x)}{Q(x)} = \frac{2x-7}{x^2-5x+6}.$$

Polynomial Q can be written as a product of two linear factors as follows: $Q(x) = x^2 - 5x + 6 = (x-2)\,(x-3)$. Using formula Ic) from the preceding paragraph, we can write:

$$\frac{P(x)}{Q(x)} = \frac{2x-7}{x^2-5x+6} = \frac{A}{x-2} + \frac{B}{x-3}.$$

The sum on the right hand side can be transformed to the common denominator:

$$\frac{A}{x-2} + \frac{B}{x-3} = \frac{A\,(x-3) + B\,(x-2)}{(x-2)\,(x-3)} = \frac{(A+B)\,x - (3A+2B)}{(x-2)\,(x-3)}.$$

The last fraction is identical with $P(x)/Q(x)$. Comparing the numerators, we obtain: $(A+B)\,x - (3A+2B) = 2x - 7$. Two polynomials are equal if and only if all their coefficients at the corresponding powers of the independent variable (i.e. x) are equal. In our case, the comparison of the coefficients yields:

$$A + B = 2, \qquad 3A + 2B = -7.$$

This is a system of two linear algebraic equations for two unknowns A and B. Solving this system, we obtain: $A = 3$, $B = -1$. Thus, the decomposition we wished to find is:

$$\frac{2x-7}{x^2-5x+6} = \frac{3}{x-2} - \frac{1}{x-3}.$$

The method, explained in this example, is called "the method of undetermined coefficients". (The undetermined coefficients were A and B.)

IV.4.7. The partial fraction decomposition of a rational function with a cubic polynomial in the denominator. Suppose that P is a polynomial of the degree less than 3 (i.e. it is either a constant function or a linear polynomial or a quadratic polynomial) and Q is a polynomial of the degree 3. Assume further that polynomials P and Q have no common factor which could be used to cancel the rational function P/Q. We show how it is possible to decompose P/Q to the sum of partial fractions. The form of the partial fractions is closely related to the decomposition of polynomial Q, explained in paragraph IV.4.3. This is why we

discuss all four cases IIa), IIb), IIc) and IId) from paragraph IV.4.3 and we always show the corresponding partial fraction decomposition of the rational function P/Q.

IIa) $$\frac{P(x)}{Q(x)} = \frac{P(x)}{q_0\,(x-\alpha)\,(x^2+rx+s)} = \frac{A}{x-\alpha} + \frac{Bx+C}{x^2+rx+s}$$

IIb) $$\frac{P(x)}{Q(x)} = \frac{P(x)}{q_0\,(x-\alpha)^3} = \frac{A_1}{x-\alpha} + \frac{A_2}{(x-\alpha)^2} + \frac{A_3}{(x-\alpha)^3}$$

IIc) $$\frac{P(x)}{Q(x)} = \frac{P(x)}{q_0\,(x-\alpha)\,(x-\beta)^2} = \frac{A}{x-\alpha} + \frac{B_1}{x-\beta} + \frac{B_2}{(x-\beta)^2}$$

IId) $$\frac{P(x)}{Q(x)} = \frac{P(x)}{q_0\,(x-\alpha)\,(x-\beta)\,(x-\gamma)} = \frac{A}{x-\alpha} + \frac{B}{x-\beta} + \frac{C}{x-\gamma}$$

The coefficients A, B, C (respectively A_1, A_2, A_3, respectively A, B_1, B_2) can be evaluated similarly as in example IV.4.6 in particular cases. You can also see this in the following example.

IV.4.8. Example. Decompose to the sum of partial fractions the rational function

$$\frac{P(x)}{Q(x)} = \frac{10x^2 + 4x}{2x^3 - 4x^2 - 20x - 50}.$$

The decomposition of polynomial Q is: $Q(x) = 2\,(x-5)\,(x^2+3x+5)$ (see example IV.4.4). Using formula IIa) from the previous paragraph, we get:

$$\frac{P(x)}{Q(x)} = \frac{10x^2 + 4x}{2x^3 - 4x^2 - 20x - 50} = \frac{A}{x-5} + \frac{Bx+C}{x^2+3x+5}.$$

Transforming the sum on the right hand side to the same denominator and making simple rearrangements, we obtain:

$$\frac{A}{x-5} + \frac{Bx+C}{x^2+3x+5} = \frac{(A+B)\,x^2 + (3A-5B+C)\,x + (5A-5C)}{(x-5)\,(x^2+3x+5)}.$$

The last fraction is identical with $P(x)/Q(x)$. The comparison of the numerators gives: $(A+B)\,x^2 + (3A-5B+C)\,x + (5A-5C) = 10x^2 + 4x$. If we further compare the coefficients at the same powers of x, we obtain a system of three linear algebraic equations for three unknowns A, B, C:

$$A + B = 5, \qquad 3A - 5B + C = 2, \qquad 5A - 5C = 0.$$

The system has a unique solution: $A = 3$, $B = 2$, $C = 3$. Thus, the desired decomposition is:

$$\frac{10x^2 + 4x}{2x^3 - 4x^2 - 20x - 50} = \frac{3}{x-5} + \frac{2x+3}{x^2+3x+5}.$$

IV.4.9. The integration of a rational function. If we have to integrate a rational function P/Q (where Q is at most a cubic polynomial), we follow these steps:

a) If the degree of P is greater than or equal to the degree of Q then we divide P by Q. Suppose that the resulting polynomial is P_1 and the remainder is P_2. Hence P/Q can be written in the form $P_1 + P_2/Q$, where the degree of P_2 is already less than the degree of Q. The integration of polynomial P_1 is already simple and the quotient P_2/Q can be integrated in the way described in the next item.

b) If the degree of P is less than the degree of Q then we decompose the rational function P/Q to the sum of partial fractions and we integrate this sum "term by term". The integration of the two types of partial fractions is explained in paragraphs IV.4.11 and IV.4.12.

IV.4.10. Remark. Let us discuss in this paragraph a simple case when polynomial Q in the denominator of the rational function P/Q is linear, i.e. its degree equals 1. If the degree of polynomial P equals zero (i.e. P is a constant) then the rational function P/Q is already a partial fraction. If the degree of P is greater than zero, then dividing P by Q we transform P/Q to the form $P_1 + P_2/Q$, where P_1 is a polynomial and P_2 is a constant. P_2/Q is a partial fraction. The integration of P_1 is obvious and the integration of P_2/Q is described in paragraph IV.4.11.

Thus, the case when the degree polynomial Q equals one is simple and therefore we do not further pay more attention to this case.

IV.4.11. The integration of partial fractions of the type $A/(x-\alpha)^n$. To compute the integrals of the type $\int A/(x-\alpha)^n \, dx$, we can use the substitution $x - \alpha = t$, i.e. $x = t + \alpha$. The above integral is thus transformed to the integral $\int A/t^n \, dt$, which can be evaluated by means of formula b) from paragraph IV.1.9 (for $n \neq 1$) or by means of formula k) from the same paragraph (for $n = 1$). Replacing t by $x - \alpha$ in the result, we obtain the expression of the integral $\int A/(x-\alpha)^n \, dx$ on two intervals: $(-\infty, \alpha)$ and $(\alpha, +\infty)$.

IV.4.12. The integration of partial fractions of the type $(Bx+C)/(x^2+rx+s)$. Suppose that the quadratic polynomial x^2+rx+s cannot be decomposed in the real domain (otherwise $(Bx+C)/(x^2+rx+s)$ would not be a partial fraction). Bear in mind that we can recognize this for example in such a way that the discriminant of the polynomial $x^2 + rx + s$ is negative. Before we explain the integration of a general function $(Bx+C)/(x^2+rx+s)$, we will illustrate the procedure on a concrete example. Since C already denotes one of the coefficients in the partial fraction, the constant of integration will be denoted by c.

$$\int \frac{2x+3}{x^2+2x+5}\,dx \stackrel{*)}{=} \int \frac{2x+3}{(x^2+2x+1)+4}\,dx = \int \frac{2x+3}{(x+1)^2+4}\,dx =$$

$$\stackrel{**)}{=} \int \frac{2u+1}{u^2+4}\,du \stackrel{***)}{=} \frac{2}{4}\int \frac{4v+1}{v^2+1}\,dv = \int \frac{2v}{v^2+1}\,dv + \frac{1}{2}\int \frac{1}{v^2+1}\,dv =$$

$$\stackrel{+)}{=} \ln(v^2+1) + \tfrac{1}{2}\arctan v + c = \ln\left(\tfrac{1}{4}u^2+1\right) + \tfrac{1}{2}\arctan\left(\tfrac{1}{2}u\right) + c =$$

$$= \ln\left[\tfrac{1}{4}(x+1)^2+1\right] + \tfrac{1}{2}\arctan\left[\tfrac{1}{2}(x-1)\right] + c.$$

$^{*)}$ We have completed x^2+2x to x^2+2x+1, which is a second power of $(x+1)$.

$^{**)}$ We have used the substitution $x+1=u$, i.e. $x=u-1$, and we have also used the equality $dx = du$ which follows from the substitution.

$^{***)}$ We have used the substitution $u = 2v$ and the following equality $du = 2\,dv$.

$^{+)}$ We have used the fact that $2v = (v^2+1)'$, which means that $2v/(v^2+1)$ is the logarithmic derivative of the function $\ln(v^2+1)$.

The partial fraction $(2x+3)/(x^2+2x+5)$ is a continuous function on the interval $(-\infty,+\infty)$, hence its antiderivative and the indefinite integral exist on this interval.

Let us now return to the general partial fraction $(Bx+C)/(x^2+rx+s)$. The fraction can be written in the form

$$\frac{Bx+C}{x^2+rx+s} = R\,\frac{2x+r}{x^2+rx+s} + S\,\frac{1}{x^2+rx+s}.$$

The values of R and S can be specified by comparing the coefficients at the same powers of x, similarly as the values of A, B and C in paragraph IV.4.8. The first fraction on the right hand side is of the type f'/f and so its integral is:

$$\int \frac{2x+r}{x^2+rx+s}\,dx = \ln|x^2+rx+s| + c \qquad (\text{for } x \in (-\infty,+\infty)).$$

(Hint: You can use the substitution $u = x^2+rx+s$.) The second fraction can be integrated in the way which was already explained on a concrete example in this paragraph. We obtain:

$$\int \frac{dx}{x^2+rx+s}\,dx = \frac{2}{\sqrt{-D}}\arctan\frac{2x+r}{\sqrt{-D}} + c \qquad (\text{for } x \in (-\infty,+\infty)).$$

(We remind that the discriminant $D = r^2 - 4s$ is supposed to be negative in this paragraph.) You can either remember the procedure of integration (illustrated on the concrete example) or the above final formula.

IV.4.13. Problems. Find the integrals a) $\displaystyle\int \frac{x^3-3x^2+2}{2x+1}\,dx$,

b) $\displaystyle\int \frac{x}{x^3-1}\,dx$, c) $\displaystyle\int \frac{x^3+1}{x^3-x^2}\,dx$, d) $\displaystyle\int \frac{5x-4}{x^3-x^2-2x}\,dx$.

Results: a) $\frac{1}{6}x - \frac{7}{8}x^2 + \frac{7}{8}x + \frac{9}{16}\ln|2x+1| + c$ (for $x \in (-\infty,-\frac{1}{2})$ and $x \in (-\frac{1}{2},+\infty)$),

b) $\dfrac{1}{3}\ln\dfrac{|x-1|}{\sqrt{x^2+x+1}} + \dfrac{1}{\sqrt{3}}\arctan\dfrac{2x+1}{\sqrt{3}} + c$ (for $x \in (-\infty,1)$ and $x \in (1,+\infty)$),

c) $x + \dfrac{1}{x} + \ln\dfrac{(x-1)^2}{|x|} + c$ (for $x \in (-\infty,0)$, $x \in (0,1)$ and $x \in (1,+\infty)$),

d) $\ln\dfrac{x^2\,|x-2|}{|x+1|^3} + c$ (for $x \in (-\infty,-1)$, $x \in (-1,0)$, $x \in (0,2)$ and $x \in (2,+\infty)$).

IV.5. Integration of functions of the type $\sin^n x \cdot \cos^m x$

IV.5.1. The integration of functions $\sin^n x \cdot \cos^m x$, where at least one of the numbers m, n is odd. Let m, n be such integers that at least one of them is odd. Suppose for example that m is odd. (If m is even and n is odd then the procedure is analogous.) m can be expressed in the form $m = 2k+1$ where k is an appropriate integer. Then the function $\sin^n x \cdot \cos^m x$ can be written in the form:

$$\sin^n x \cdot \cos^m x \;=\; \sin^n x \cdot \cos^{2k+1} x \;=\; \sin^n x \cdot (\cos^2 x)^k \cdot \cos x \;=$$
$$= \sin^n x \cdot (1-\sin^2 x)^k \cdot \cos x.$$

The integral $\int \sin^n x \cdot \cos^m x \; dx$ can be evaluated by means of the substitution $\sin x = t$:

$$\int \sin^n x \cdot \cos^m x \; dx \;=\; \int \sin^n x \cdot (1-\sin^2 x)^k \cdot \cos x \; dx \;=\; \int t^n \cdot (1-t^2)^k \; dt.$$

IV.5.2. Example. $\int \sin^2 x \cdot \cos^5 x \; dx \;=\; \int \sin^2 x \cdot \cos^4 x \cdot \cos x \; dx \;=$

$$= \int \sin^2 x \cdot (1-\sin^2 x)^2 \cdot \cos x \; dx \;=\; \int \sin^2 x \cdot (1 - 2\sin^2 x + \sin^4 x) \cdot \cos x \; dx \;=$$
$$= \int \sin^2 x \cdot \cos x \; dx \;-\; 2\int \sin^4 x \cdot \cos x \; dx \;+\; \int \sin^6 x \cdot \cos x \; dx \;=$$
$$=^{*)} \int t^2 \; dt \;-\; 2\int t^4 \; dt \;+\; \int t^6 \; dt \;=\; \tfrac{1}{3}t^3 \;-\; \tfrac{2}{5}t^5 \;+\; \tfrac{1}{7}t^7 \;+\; C \;=$$
$$= \tfrac{1}{3}\sin^3 x \;-\; \tfrac{2}{5}\sin^5 x \;+\; \tfrac{1}{7}\sin^7 x \;+\; C \quad (\text{for } x \in (-\infty, +\infty))$$

*) We have used the substitution $\sin x = t$.

IV.5.3. Example. $\int \frac{1}{\sin x} \, dx \;=\; \int \frac{\sin x}{\sin^2 x} \, dx \;=\; \int \frac{\sin x}{1-\cos^2 x} \, dx \;=$

$$=^{*)} -\int \frac{1}{1-t^2} \, dt \;=\; -\int \frac{1}{(1-t)(1+t)} \, dt \;=\; -\tfrac{1}{2}\int \frac{1}{1-t} \, dt \;-\; \tfrac{1}{2}\int \frac{1}{1+t} \, dt \;=$$
$$= \tfrac{1}{2}\ln|1+t| \;-\; \tfrac{1}{2}\ln|1+t| \;+\; C \;=\; \tfrac{1}{2}\ln|1-\cos x| \;-\; \tfrac{1}{2}\ln|1+\cos x| \;+\; C \;=$$
$$= \frac{1}{2}\ln\left|\frac{1-\cos x}{1+\cos x}\right| \;+\; C$$

(for $x \in (k\pi, (k+1)\pi)$; k is an arbitrary integer)
*) We have used the substitution $\cos x = t$.

IV.5.4. The integration of functions $\sin^n x \cdot \cos^m x$, where both the numbers m, n are even. The integrand can be rewritten by means of the formulas

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

We also use the substitution $2x = t$. Thus, we decompose the integrand to the sum of functions "of type IV.5.1" (i.e. functions satisfying the assumptions from paragraph IV.5.1) and functions "of type IV.5.4" (i.e. functions satisfying the assumptions from this paragraph). The integration of the functions "of type IV.5.1" is already known. The functions "of type IV.5.4" can be rewritten again by means of the above formulas. We can continue doing this until all remaining functions are "of type IV.5.1".

IV.5.5. Example.

$$\int \sin^4 x \; dx \;=\; \int \left(\frac{1-\cos(2x)}{2}\right)^2 dx \;=\; \int \frac{1-2\cos(2x)-\cos^2(2x)}{4} \, dx \;=$$
$$=^{*)} \int \frac{1-2\cos t - \cos^2 t}{8} \, dt \;=\; \int \tfrac{1}{8} \, dt \;-\; \int \tfrac{1}{4}\cos t \; dt \;+\; \int \tfrac{1}{8}\cos^2 t \; dt \;=$$
$$= \tfrac{1}{8}t \;-\; \tfrac{1}{4}\sin t + \frac{1}{8}\int \frac{1+\cos(2t)}{2} \, dt \;=^{**)}\; \tfrac{1}{8}t \;-\; \tfrac{1}{4}\sin t + \frac{1}{8}\int \frac{1+\cos u}{4} \, du \;=$$
$$= \tfrac{1}{8}t - \tfrac{1}{4}\sin t + \tfrac{1}{32}u + \tfrac{1}{32}\sin u + C = \tfrac{1}{8}t - \tfrac{1}{4}\sin t + \tfrac{1}{16}t + \tfrac{1}{32}\sin(2t) + C =$$
$$= \tfrac{3}{8}x \;-\; \tfrac{1}{4}\sin(2x) \;+\; \tfrac{1}{32}\sin(4x) \;+\; C \qquad (\text{for } x \in \mathbb{R}).$$

*) We have used the substitution $t = 2x$.
**) We have used the substitution $u = 2t$.

IV.5.6. Problems. Evaluate the integrals

$$\text{a)} \int \sin^7 x \; dx, \qquad \text{b)} \int \cos^6 x \; dx, \qquad \text{c)} \int \frac{\cos^3 x}{\sin^2 x} \, dx.$$

Results: a) $-\cos x + \cos^3 x - \frac{3}{5}\cos^5 x + \frac{1}{7}\cos^7 x + C$) (for $x \in (-\infty, +\infty)$),
b) $\frac{5}{16}x + \frac{1}{4}\sin(2x) + \frac{3}{64}\sin(4x) - \frac{1}{48}\sin^3(6x) + C$ (for $x \in (-\infty, +\infty)$),
c) $-\dfrac{1}{\sin x} - \sin x + C$ (for $x \in (k\pi, (k+1)\pi)$; k is an arbitrary integer).

IV.6. Integration of some other types of functions

IV.6.1. Rational function of two variables. We shall several times use the symbol $R(u,v)$ in this section. It will denote a so called <u>rational function of two variables</u> u and v, i.e. a ratio of two polynomials $P(u,v)$ and $Q(u,v)$.

You will learn more about functions of several variables in the second term in Mathematics. For our purposes, it will be sufficient to know that a polynomial of two variables u and v is a sum of a finite number of terms, each of which has the form $k\,u^m v^n$ where k is a real constant and m, n are nonnegative integers.

Examples of rational functions of two variables are:

$$R(u,v) \;=\; \frac{2u + v^2 + 2}{u-1}, \qquad R(u,v) \;=\; \frac{1}{2+u}, \qquad R(u,v) \;=\; \frac{u-v}{u^2+v^2}.$$

IV.6.2.* Integrals of the type $\int R(\sin x, \cos x)\,dx$. Integrals of this type can be transformed by means of the substitution

$$t = \tan\left(\frac{x}{2}\right) \tag{IV.6.1}$$

to integrals of rational functions. Let us explain this in greater detail. Using the known formulas for trigonometric functions, we obtain:

$$\cos^2(x/2) = \frac{1}{1+\tan^2(x/2)}, \qquad \sin^2(x/2) = 1-\cos^2(x/2) = \frac{\tan^2(x/2)}{1+\tan^2(x/2)},$$

$$\cos x = \cos^2(x/2) - \sin^2(x/2) = \frac{1-\tan^2(x/2)}{1+\tan^2(x/2)} = \frac{1-t^2}{1+t^2},$$

$$\sin x = 2\sin(x/2)\cdot\cos(x/2) = \frac{2\tan(x/2)}{1+\tan^2(x/2)} = \frac{2t}{1+t^2}.$$

Differentiating the equation $\tan(x/2) = t$, we further get:

$$\frac{1}{2\cos^2(x/2)}\,dx = dt, \qquad \text{i.e. } \tfrac{1}{2}\left[1+\tan^2(x/2)\right]dx = dt.$$

This implies: $dx = \dfrac{2}{1+\tan^2(x/2)}\,dt = \dfrac{2}{1+t^2}\,dt.$

If we now replace $\sin x$, $\cos x$ and dx in the integral $\int R(\sin x, \cos x)\,dx$ by the above expressions, we obtain an integral of a rational function (of variable t). Integrals of rational functions are dealt with in Section IV.4 in this text.

The substitution (IV.6.1) can be used only on an interval where the function $\tan(x/2)$ is defined, i.e. on each interval $(-\pi+2k\pi, \pi+2k\pi)$, where k is an integer. (Sometimes we evaluate the integral on an even smaller interval. This depends on a concrete form of an integrated rational function.)

IV.6.3.* Example. $\displaystyle\int \frac{1-\sin x}{1+\cos x}\,dx = \int \frac{1-2t/(1+t^2)}{1+(1-t^2)/(1+t^2)}\,\frac{2}{1+t^2}\,dt =$

$$= \int \frac{1-2t+t^2}{1+t^2}\,dt = \int\left[1-\frac{2t}{1+t^2}\right]dt = t - \ln(1+t^2) + C =$$

$$= \tan(x/2) - \ln\left[1+\tan^2(x/2)\right] + C.$$

The integrand is a continuous function on each of the intervals $(-\pi+2k\pi, \pi+2k\pi)$, where k is an integer. Hence a primitive function and the indefinite integral exist on each of these intervals. The intervals are identical with those where we can use substitution (IV.6.1). Thus, the obtained expression of the integral is valid on each of the intervals $(-\pi+2k\pi, \pi+2k\pi)$.

IV.6.4. Remark. Before starting to compute any integral which is not quite simple, we recommend that you think at least a short while over the possibilities you have and then to try to choose the simplest one. For instance, the integral

$$\int \frac{\cos x}{1-\sin x}\,dx$$

is an integral "of type IV.6.2", so it can be evaluated by means of substitution (IV.6.1). However, there also exists another substitution which is much simpler. Can you discover it?

IV.6.5. Integrals of the type $\displaystyle\int R\left(x, \sqrt{\frac{ax+b}{cx+d}}\right)dx$. Suppose that $ad-bc \neq 0$. (This inequality guarantees that the linear polynomials $ax+b$ and $cx+d$ have no common factor. Otherwise the fraction $(ax+b)/(cx+d)$ could be cancelled by this factor and the result would be a constant function.) We can use the substitution

$$t = \sqrt{\frac{ax+b}{cx+d}}. \tag{IV.6.2}$$

This can be used to express x and dx. Substituting this to the integral we want to compute, we obtain an integral of a rational function (of variable t). The procedure is shown in example IV.6.6. Substitution (IV.6.2) can be used on each interval which is a subset of the domain of the integrand. In addition to others, the following inequality must hold on such an interval: $(ax+b)/(cx+d) \geq 0$.

The integral $\int R(x, \sqrt{ax+b})\,dx$ is a special case of the above integral. (It corresponds to the values $c = 0$, $d = 1$.) To evaluate it, we can thus use the substitution $t = \sqrt{ax+b}$.

IV.6.6. Example. Evaluate $\displaystyle\int \frac{1}{x}\sqrt{\frac{1-x}{1+x}}\,dx.$

The fraction under the square root must be nonnegative. Solving the inequality $(1-x)/(1+x) \geq 0$, we get the condition: $x \in (-1, 1]$. Moreover, there is also a function $1/x$, involved in the integral. This implies another condition: $x \neq 0$. Hence the integrand is defined on intervals $(-1, 0)$ and $(0, 1]$. Since it is continuous on each of these intervals, an antiderivative and the indefinite integral exist on each of them. Using the substitution

$$t = \sqrt{\frac{1-x}{1+x}}, \qquad \text{i.e. } x = -\frac{t^2-1}{t^2+1}, \qquad dx = -\left(\frac{t^2-1}{t^2+1}\right)' dt = -\frac{4t}{(t^2+1)^2}\,dt,$$

we get: $\displaystyle\int \frac{1}{x}\sqrt{\frac{1-x}{1+x}}\,dx = \int \frac{t^2+1}{t^2-1}\,t\,\frac{4t}{(t^2+1)^2}\,dt =$

$$= \int \frac{4t^2\,dt}{(t-1)(t+1)(t^2+1)} = \int\frac{dt}{t-1} - \int\frac{dt}{t+1} + 2\int\frac{dt}{t^2+1} =$$

$$= \ln\left|\frac{t-1}{t+1}\right| + 2\arctan t + C = \ln\left|\frac{\sqrt{1+x}-\sqrt{1-x}}{\sqrt{1+x}+\sqrt{1-x}}\right| + 2\arctan\sqrt{\frac{1-x}{1+x}} + C$$

(for $x \in (-1, 0)$ or for $x \in (0, 1]$). The rational function of the variable t, which we obtained after the application of the substitution, contains the polynomial of the fourth degree in the denominator. This goes beyond the scope of Section IV.4. Nevertheless, its decomposition is quite analogous:

$$\frac{4t^2}{(t-1)(t+1)(t^2+1)} = \frac{A}{t-1} + \frac{B}{t+1} + \frac{C_1 t + C_2}{t^2+1}.$$

The coefficients A, B, C_1 and C_2 can be evaluated by the same method as the numbers A, B and C in paragraph IV.4.8. We obtain: $A = 1$, $B = -1$, $C_1 = 0$, $C_2 = 2$.

IV.6.7.* Integrals of the type $\int R(x, \sqrt{ax^2+bx+c})\,dx$. We suppose that $a \neq 0$, otherwise the integral would be a special case of the integral solved in paragraph IV.6.4. We shall distinguish between two cases:

a) $\underline{a < 0}$ Since the term under the square root must be nonnegative, the quadratic polynomial ax^2+bx+c must have two different real roots α, β. Let us choose their denotation in such a way that $\alpha < \beta$. Then $ax^2+bx+c \geq 0$ for $x \in [\alpha, \beta]$. For those $x \in (\alpha, \beta]$, we can write:

$$\sqrt{ax^2+bx+c} = \sqrt{a\,(x-\alpha)(x-\beta)} = \sqrt{-a\,(x-\alpha)(\beta-x)} = \\ = \sqrt{-a}\,(x-\alpha)\sqrt{\frac{\beta-x}{x-\alpha}}.$$

The integral of this function is the integral treated in paragraph IV.6.4. It can be transformed to the integral of a rational function by means of the substitution (IV.6.2), i.e.

$$t = \sqrt{\frac{\beta-x}{x-\alpha}}.$$

b) $\underline{a > 0}$ In this case, we can use a so called Euler substitution

(IV.6.3) $$\sqrt{ax^2+bx+c} = t + \sqrt{a}\,x.$$

Raising the two sides of the equality (IV.6.3) to the second power, we can express x and dx. The Euler substitution transforms the integral we wish to evaluate to an integral of a rational function of variable t. The procedure is shown in example IV.6.8. The substitution (IV.6.3) can be used on each interval which is a subset of the domain of the integrand.

IV.6.8.* Example. Evaluate $\displaystyle\int \frac{dx}{x\sqrt{x^2+2x-1}}\,dx$.

Since the quadratic polynomial x^2+2x-1 is under the square root in the denominator, the inequality $x^2+2x-1 > 0$ must be satisfied in the domain of the integrand. Moreover, the integrand also contains x in the denominator. This gives the condition $x \neq 0$. It can easily be verified that maximum intervals where the two conditions are fulfilled are $(-\infty, -1-\sqrt{2})$ and $(-1+\sqrt{2}, +\infty)$. The integrand is continuous on each of them, hence a primitive function and the indefinite integral also exist on each of them.

We use the substitution $\sqrt{x^2+2x-1} = x+t$. Raising it to the second power, we obtain:

$$x = \frac{t^2+1}{2\,(1-t)}, \qquad dx = \frac{-t^2+2t+1}{2\,(1-t)^2}\,dt.$$

Substituting this to the integral we wish to evaluate, we obtain:

$$\int \frac{dx}{x\sqrt{x^2+2x-1}}\,dx = \int \frac{2\,(1-t)}{t^2+1}\;\frac{1}{\frac{t^2+1}{2\,(1-t)}+t}\;\frac{-t^2+2t+1}{2\,(1-t)^2}\,dt =$$

$$= \int \frac{2}{t^2+1}\,dt = 2\arctan t + C = 2\arctan\left(\sqrt{x^2+2x-1}-x\right) + C$$

(for $x \in (-\infty, -1-\sqrt{2})$ or $x \in (-1-\sqrt{2}, +\infty)$).

IV.7. Separable differential equations

Differential equations form an extensive part of mathematics, with a lot of applications in engineering, physics, chemistry and other scientific disciplines. We introduce some basic notions and show the so called "method of separation of variables" for the solution. Since the method is based mainly on integration, we present it as a part of Chapter IV dealing with indefinite integrals, and we regard it as an illustrative example of applications of theory of indefinite integrals.

You will learn more about differential equations in Mathematics III in your third term of study.

IV.7.1. The differential equation of the 1st order. The equation

(IV.7.1) $$F(x, y, y') = 0$$

is called the *differential equation of the 1st order*. The function $y(x)$ is the unknown in equation (IV.7.1).

A function of three variables x, y and y' appears on the right hand side of equation (IV.7.1). As was mentioned in paragraph IV.6.1, you will learn more about such functions in Mathematics II in the second term. In this section, it will be sufficient for our purposes to regard $F(x,y,y')$ as an algebraic expression which can contain both x, y and y'. The example is $F(x,y,y') = y' - 2y\sqrt{x^2-1}$.

Equation (IV.7.1) is called "differential" because it contains not only the unknown function y, but also its derivative. The denotation "of the 1st order" means that the highest derivative of the unknown function appearing in the equation is the derivative of the 1st order.

IV.7.2. A solution, a maximum solution. The *solution* of differential equation (IV.7.1) in interval I is a function y which satisfies equation (IV.7.1) at each point of interval I. This means that if I is an interval with end points a, b ($a < b$,) then

a) $y'(x) = f(x, y(x))$ for all $x \in (a, b)$,

b) $y'_+(a) = f(a, y(a))$, if a belongs to interval I,

c) $y'_-(b) = f(b, y(b))$, if b belongs to interval I.

If y is a solution of differential equation (IV.7.1) in interval I, z is a solution of the same equation in interval J, $I \subset J$, $I \neq J$ and $y(x) = z(x)$ for all $x \in I$, then the solution z is called the *extension* of the solution y. A solution which cannot be extended any more (i.e. no extension of it exists) is called the *maximum solution*.

IV.7.3. Example. The function

$$y(x) = 2\sqrt{x-1} \qquad (\text{for } x \in (2, +\infty))$$

is a solution of the differential equation

$$y' = \frac{2}{y} \tag{IV.7.2}$$

in interval $(2, +\infty)$. It is not a maximum solution, because the function

$$z(x) = 2\sqrt{x-1} \qquad (\text{for } x \in (1, +\infty))$$

is the extension of the solution y. Function z is already the maximum solution, because as a solution of equation (IV.7.2) it cannot be extended any more.

IV.7.5. Remark. Function z from the preceding example is not the only maximum solution of equation (IV.7.2). It can easily be verified that if A is an arbitrary real number then the functions

$$u(x) = 2\sqrt{x-A} \qquad (\text{for } x \in (A, +\infty)),$$
$$v(x) = -2\sqrt{x-A} \qquad (\text{for } x \in (A, +\infty))$$

are also the maximum solutions of (IV.7.2). Thus, equation (IV.7.2) has infinitely many maximum solutions. (Sketch their graphs for yourself.)

Choose for instance $x_0 = 5$ and look at the values of maximum solutions of equation (IV.7.2) at this point. If $A \geq 5$ then point x_0 does not belong to the domain of u and v. If $A = 4$, $A = 3$, etc., we obtain:

$$A = 4 \;\ldots\; u(5) = 2,\ v(5) = -2, \qquad A = 3 \;\ldots\; u(5) = 2\sqrt{2},\ v(5) = -2\sqrt{2}, \qquad \text{etc.}$$

We can see that we can distinguish between different maximum solutions by means of their values at a chosen point x_0. (This is not a general rule, but it is true in the case of differential equation (IV.7.2).) When solving a differential equation, it is therefore natural to seek for a solution which takes on a chosen (or a given) value y_0 at a chosen (or a given) point x_0.

IV.7.6. The Cauchy problem. The problem of finding a solution y of differential equation (IV.7.1) which satisfies the condition $y(x_0) = y_0$ is called the *Cauchy problem* for differential equation (IV.7.1).

IV.7.7. Remark. The condition $y(x_0) = y_0$ is called the *initial condition* for the following reason: Differential equations can be used for modelling processes running in time and then the independent variable x plays the role of time. Thus, we often solve problems of the following type: *What is the behavior of the process described by equation (IV.7.1) at times $x > x_0$, if we know its state at an initial time instant x_0?* Certainly, we can also ask what preceded the state at time x_0, i.e. we can study the behavior of the process at times $x < x_0$. However, problems of this type are not very frequent (or were not so frequent in the past when the terminology was created).

IV.7.8. Separable differential equations, solution by separation of variables. In the rest of this chapter, we will deal with a special case of differential equation (IV.7.1), i.e. the case when the right hand side is equal to the product of a function depending only on x and a function depending only on y:

$$y' = g(x) \cdot h(y). \tag{IV.7.3}$$

Differential equation (IV.7.3) is called a *differential equation with separable variables* or shortly a *separable differential equation*.

Assume that function $g(x)$ is continuous on interval I and function $h(y)$ is continuous on interval J. Differential equation (IV.7.3) will be solved for $x \in I$.

Denote by M a set of those points y^* from interval J, where $h(y^*) = 0$. We can easily check (by the substitution to equation (IV.7.3)) that the constant function defined by the equation $y(x) = y^*$ (for all $x \in I$) is a solution of equation (IV.7.3) on interval I. Thus, the computation of zero points of function $h(y)$ in interval J leads to constant solutions of equation (IV.7.3).

Suppose further that interval J was chosen so that it contains no zero points of function $h(y)$, i.e. $h(y) \neq 0$ for all $y \in J$. Hence we can divide equation (IV.7.3) by $h(y)$. Moreover, we write the derivative y' in the form $\mathrm{d}y/\mathrm{d}x$ and we multiply equation (IV.7.3) formally by $\mathrm{d}x$. We obtain:

$$\frac{\mathrm{d}y}{h(y)} = g(x)\,\mathrm{d}x.$$

The left hand side of the above equation depends only on y and the right hand side depends only on x. Thus, we have separated the variables. This fact provided the name for the whole method that we are just explaining: the *method of separation of variables*. Further, we add signs of integrals in front of the left hand side as well as in front of the right hand side and we compute the obtained integrals. If $H(y)$ is a primitive function to $1/h(y)$ on interval J and $G(x)$ is a primitive function to $g(x)$ on interval I, then the evaluation of the indefinite integrals leads to the equation:

$$H(y) = G(x) + C.$$

(Constants which arise on both the sides can be coupled to one which can be written for instance on the right hand side only.) The function $H(y)$ is strictly monotone on interval J because it has a continuous and non–zero derivative $1/h(y)$. Therefore an inverse function to $H(y)$ in interval J exists. Using the inverse function, we get:

$$y = H_{-1}(G(x) + C). \tag{IV.7.4}$$

Function y is a solution of differential equation (IV.7.3) in each interval $I_1 \subset I$ such that $G(x) + C$ belongs to the domain of function H_{-1} for all $x \in I_1$.

The explained procedure is not quite correct from the formal point of view because we separate the "non–separable" symbol $\mathrm{d}y/\mathrm{d}x$ denoting the derivative of function y, we treat it as a fraction and we multiply equation (IV.7.3) by the "auxiliary" term $\mathrm{d}x$. Nevertheless, this procedure leads to good results and you can easily remember it.

Different constants C in formula (IV.7.4) lead to different solutions of differential equation (IV.7.3). Although it does not hold generally, in many cases formula (IV.7.4) involves all solutions of equation (IV.7.3) on interval I. (This means that every solution of equation (IV.7.3) on interval I can be represented by formula (IV.7.4) for some appropriate $C \in \mathbb{R}$.) For this reason it is often said that formula (IV.7.4) represents a so called *general solution* of differential equation (IV.7.3) on interval I. To distinguish between the general solution and a concrete solution (corresponding to a concrete value of the constant C in (IV.7.4)), the latter is called a *particular solution*.

IV.7.9. Remark. Note that the problem of evaluating the indefinite integral $\int g(x)\,dx$ in interval I can also be formulated as follows: *Find all solutions of the differential equation* $y' = g(x)$ *on interval* I.

IV.7.10. Example. Find a maximum solution of the Cauchy problem given by the differential equation $y' = 5xy$ and by the initial condition $y(0) = 2$.

The right hand side is the product of the function $g(x) = 5x$ and the function $h(y) = y$. Both the functions are continuous on the interval $(-\infty, +\infty)$. Solving the equation $h(y) = 0$, we obtain the unique zero point of the function $h(y)$: $y^* = 0$. Thus the differential equation $y' = 5xy$ has only one constant solution: $y(x) = 0$ which is defined in the whole interval $I = (-\infty, +\infty)$.

The function $h(y)$ is different from zero in the intervals $J_1 = (-\infty, 0)$ and $J_2 = (0, +\infty)$. We can therefore take y from J_1 or from J_2 and we can solve the differential equation by separation of variables. We write the derivative y' in the form dy/dx and we multiply the equation formally by dx. Further, we divide the equation by y. This leads to an equation with separated variables x and y. Then we add the signs of integrals. Successively, we obtain:

$$\frac{dy}{dx} = 5xy, \qquad \frac{dy}{y} = 5x\,dx, \qquad \int \frac{dy}{y} = \int 5x\,dx.$$

Evaluating the integrals on the left hand side and the right hand side of the equation, we get:

$$\ln|y| = \tfrac{5}{2}x^2 + C.$$

C can be written in the form $\ln|K|$, where $K \in \mathbb{R}$, $K \neq 0$. The last equation gives:

$$|y| = \exp\left(\tfrac{5}{2}x^2 + \ln|K|\right) = |K \cdot \exp(\tfrac{5}{2}x^2)|.$$

This yields: $y = \pm K \cdot \exp(\frac{5}{2}x^2)$. Since K can take on positive as well as negative values, the symbol $\pm$ can be omitted.

We have obtained the solution in the form $y = K \cdot \exp(\frac{5}{2}x^2)$ under the assumption that $K \neq 0$. However, for $K = 0$ the formula gives the already discovered zero solution. Thus, we can see that the condition $K \neq 0$ was connected only with the method we used for solving the equation, but it need not be taken into account in the final formula expressing the solution. Thus, the general solution of the differential equation is:

$$y = K \cdot \exp(\tfrac{5}{2}x^2); \qquad K \in \mathbb{R}.$$

Now we use the general solution and we write it with the values $x = 0$ and $y = 2$, following from the initial condition. We obtain: $2 = K$. Substituting this value of K back to the general solution, we obtain the desired solution of the Cauchy problem:

$$y = 2\exp(\tfrac{5}{2}x^2); \qquad x \in (-\infty, +\infty).$$

IV.7.11. Remark. Differential equations we have to solve are rarely written in the form (IV.7.3). For instance, equation $xy' - y = 0$ does not have this form although the equation can easily be transformed to the form (IV.7.3): $y' = y/x$.

If we solve the equation $y' = y/x$ by the separation of variables, we obtain the general form of a maximum solution: $y = K \cdot x$ for $x \in (-\infty, 0)$ or for $x \in (0, +\infty)$ $(K \in \mathbb{R})$. The point $x = 0$ must be omitted because zero cannot appear in the denominator of the fraction on the right hand side.

Let us now return to the original equation $xy' - y = 0$. Here we have no reason to exclude the point $x = 0$ from the domain of the solution and so the general form of a maximum solution is $y = K \cdot x$ for $x \in (-\infty, +\infty)$ $(K \in \mathbb{R})$.

You can derive the conclusion from the above example: If the differential equation does not have the form (IV.7.3) and you transform it to this form by means of arithmetic operations, you must pay attention to the domain of the solution.

IV.7.12. Example. A body whose mass is m moves on a straight line. The resistance of the environment is proportional to the speed of motion, the coefficient of proportionality is k. Moreover, a constant force f, parallel with the motion, acts onto the body. The known velocity at the time $t = 0$ is v_0. What is the velocity $v(t)$ of the motion at a general time t?

Solution: Let $a(t)$ be the acceleration of the body. Second Newton's law says that

(IV.7.5) $$m\,a(t) = F(t)$$

where $F(t)$ is a total force acting on the body, parallel with the direction of the motion. This force can be expressed: $F(t) = f - k\,v(t)$. (The term $k\,v(t)$ must be written with the "$-$" sign because the resistance, e.g. the friction, acts opposite the direction of the motion.)

You already know, e.g. from paragraphs III.4.2 and III.4.6, that the velocity $v(t)$ is the derivative of the position $s(t)$. Analogously, the acceleration $a(t)$ can be defined to be the derivative of the velocity $v(t)$. If we denote the derivative with respect to time, as it is usual in physics, by the dot, we can write: $a(t) = \dot{v}(t)$. Substituting this and the expression of the force $F(t)$ to equation (IV.7.5), we obtain the differential equation

(IV.7.6) $$m\,\dot{v}(t) = f - k\,v(t).$$

This is the differential equation with separable variables. You can easily verify that the general solution is

(IV.7.7) $$v(t) = \frac{f}{k} - \frac{C}{k}\,e^{-kt/m}.$$

Using the formula (IV.7.7) with the concrete values $t = 0$ and $v(0) = v_0$, we can obtain the equation which enables to evaluate C: $C = f - kv_0$. Substituting this C back to the general solution (IV.7.7), we get the desired particular solution:

$$v(t) = \frac{f}{k} + \left(v_0 - \frac{f}{k}\right)e^{-kt/m}.$$

IV.7.13. Problems. Find maximum solutions of the Cauchy problems:

a) $x^2y' + y^2 = 0$, $y(-1) = 1$; b) $2y'\sqrt{x} = y$, $y(4) = 1$;

c) $y' = (2y+1)\cot x$, $y(\pi/4) = \frac{1}{2}$; d) $y' = x^2 + 3\sin^5 x \cos x$, $y(0) = 2$.

Results: a) $y = -x$ for $x \in (-\infty, +\infty)$

b) $y = \exp(\sqrt{x} - 2)$ for $x \in (0, +\infty)$

c) $y = 2\sin^2 x - \frac{1}{2}$ for $x \in (0, \pi)$

d) $y = \frac{1}{3}x^3 + \frac{1}{2}\sin^6 x + 2$ for $x \in (-\infty, +\infty)$. (The variables are already separated in the given equation. Thus, in order to solve it, you only need to integrate it.)

V. Definite (Riemann's) integral

V.1. Historical approach

V.1.1. Motivation, introduction. Imagine that f is a continuous non–negative function in a closed interval $[a, b]$. We want to evaluate the area of region R which is bounded by the graph of function f, by the x–axis and by the straight lines $x = a$ and $x = b$. (See Fig. 24.)

If function f is non–negative and constant in the whole interval $[a, b]$, taking the value h, then the problem is simple. Region R is a rectangle whose sides have the lengths $b - a$ and h. The area of such a rectangle is $s(R) = (b - a) \cdot h$.

Similarly, if function f is non–negative and linear (i.e. $f: \; y = kx + q$) in the interval $[a, b]$ then the problem is also simple. Region R is a trapezoid whose base has the length $b - a$ and the two vertical sides have the lengths $ka + q$ and $kb + q$. Sketch a picture and evaluate the area of R for yourself.

V.1.2. Historical introduction of the definite integral. The general problem is more complicated and it was solved in the 17th century by Isaac Newton (1642–1727) and Gottfried Wilhelm Leibniz (1646–1716). Their fundamental idea was to divide the interval $[a, b]$ into infinitely many "infinitely short" sub–intervals. A typical sub–interval has the end points x and $x + \mathrm{d}x$ where $x \in [\, a, b)$ and $\mathrm{d}x$ is an "infinitely small" positive number. The region, bounded above by the graph of function f on the interval $[\, x, x + \mathrm{d}x \,]$ and bounded below by the segment which coincides with this interval on the x–axis, is "infinitely narrow" and function f can therefore be considered to be constant, taking the value $f(x)$, on the whole interval $[\, x, x + \mathrm{d}x \,]$. Thus, the region is an "infinitely narrow" rectangle whose sides have the lengths $\mathrm{d}x$ and $f(x)$. Its area is $\mathrm{d}s = f(x)\,\mathrm{d}x$.

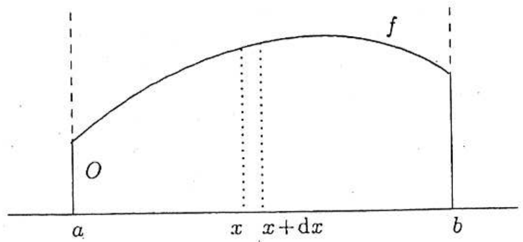

Fig. 24

The area of the whole region R equals the sum of all "infinitely small" numbers $\mathrm{d}s$. The sum was originally denoted by S, however this letter was later stretched and replaced by the sign of integral $\int$. The area of region R can be written as follows:

$$s(R) = \int_a^b \mathrm{d}s = \int_a^b f(x)\,\mathrm{d}x. \qquad \text{(V.1.1)}$$

The expression on the right hand side is called the *definite integral* of function f on the interval $[a, b]$ and the numbers a and b are called the *limits* of the integral: a is the *lower limit* and b is the *upper limit*.

The formulas for the area of a square, a rectangle, a trapezoid, a parallelogram, a disc, etc., are well known. However, an analogous simple formula, expressing the area of a general region of type R (see Fig. 24) does not exist. We must therefore understand formula (V.1.1) so that it defines the area of region R and not only tells us how to evaluate something that we have already defined in another way.

The definite integral also has a sense for functions whose values at some (or at all) points are negative. However, the integral expresses the area between the graph of function f and the x–axis with the minus sign on those intervals where f is negative.

V.1.3. A remark on the physical sense of the definite integral. Many physical problems naturally also lead to the definite integral. For instance, imagine that a bar covers the interval $[a, b]$ on the x–axis. The longitudinal density of the bar is ρ. (The longitudinal density expresses the mass per unit of length. It is given in $\mathrm{kg\,m^{-1}}$.) ρ can vary along the bar, e.g. due to the variable intersection of the bar or because the material the bar is made of varies. Thus, ρ is generally a function of x. We wish to express the total mass of the bar.

As in paragraph V.1.2, we divide the interval $[a, b]$ into infinitely many "infinitely small" parts. A typical part is the interval $[\, x, x + \mathrm{d}x \,]$. Since this interval is "infinitely short", the density ρ can be considered to be constant and equal to $\rho(x)$ on the whole interval $[\, x, x + \mathrm{d}x \,]$. Then the mass of the part of the bar which covers this interval is $\mathrm{d}m = \rho(x)\,\mathrm{d}x$. The total mass of the bar is the sum of all masses $\mathrm{d}m$ of all "infinitely short" parts of the bar whose length is $\mathrm{d}x$:

$$m = \int_a^b \mathrm{d}m = \int_a^b \rho(x)\,\mathrm{d}x. \qquad \text{(V.1.2)}$$

V.1.4. A remark on the introduction of the definite integral. The conception of the definite integral described in the preceding paragraphs is not sufficiently rigorous and clear from the today's point of view. The main reason is that it uses the notion of an "infinitely small", but at the same time positive, number $\mathrm{d}x$. In fact, such a number does not exist. (Assume for a while that $\mathrm{d}x$ with the required properties exists. What value it can have? For example, $\mathrm{d}x = 10^{-2}$? Definitely not, because 10^{-3} is positive and less. Or $\mathrm{d}x = 10^{-6}$? But 10^{-7} is also positive and less than 10^{-6}. You can see that the idea of an infinitely small positive number always leads to a contradiction.) The next reason why the introduction of the definite integral from paragraph V.1.2 is not quite correct is that it does not explain what one can understand under the "sum of infinitely many infinitely small numbers".

Thus, in order to give a correct sense to the definite integral, it is necessary

a) to write a rigorous definition,

b) to derive main properties of the integral (to specify for which functions its construction works; in other words: for which functions the integral exists, etc.)

c) and the last but not least question is how the value of the integral can be evaluated.

V.1.5. The definite integral as a function of the upper limit, the Newton–Leibniz formula. Answers to the above questions were paradoxically found in the reverse order. Newton and Leibniz were again the first to recognize that the evaluation of the area of the region between the graph of function f and the x–axis is an inverse problem to the differentiation:
Let f be a continuous function in $[a, b]$ and $x \in [a, b]$. Denote by $A(x)$ the area of the region between the graph of function f and the x–axis on the interval $[a, x]$, i.e.

$$A(x) = \int_a^x f(t)\,\mathrm{d}t. \tag{V.1.3}$$

Then, and this is very important, the derivative of A equals f in $[a, b]$. Thus, A is an antiderivative to f in $[a, b]$! Moreover, it follows directly from (V.1.3) that $A(a) = 0$ and $A(b) = \int_a^b f(x)\,\mathrm{d}x$. Hence we obtain the formula

$$\int_a^b f(x)\,\mathrm{d}x = A(b) - A(a). \tag{V.1.4}$$

Formula (V.1.4) is known as the Newton–Leibniz formula. If we take into account that any other antiderivative F to f on $[a, b]$ has a form $F = A + C$ where C is a constant then we can easily verify that $F(b) - F(a) = A(b) - A(a)$. This says that formula (V.1.4) remains valid even if the right hand side equals $F(b) - F(a)$ where F is an arbitrary antiderivative to f on $[a, b]$ and not only that particular one which was defined by (V.1.3). We shall again deal with the Newton–Leibniz formula in greater detail in Section V.4.

V.1.6. Linearity of the definite integral. It is natural to expect from the way in which the definite integral was introduced (we avoid the word "defined", because no precise definition has been shown yet) that apart from other things, it will have the following properties:

1. $\displaystyle\int_a^b [\,f(x) + g(x)\,]\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x,$
2. $\displaystyle\int_a^b c \cdot f(x)\,\mathrm{d}x = c\int_a^b f(x)\,\mathrm{d}x$ (if c is a constant) and
3. $\displaystyle\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x$ (if $c \in [a, b]$).

We shall discuss these and other properties in greater detail in Section V.3.

V.1.7. Who invented calculus? Though calculus is the result of a long development that began already in ancient Greece, and many scientists had worked on finding tangents or calculating areas (let us mention e.g. P. Fermat 1601–1665, R. Descartes 1596–1650 and B. Cavalieri 1598–1647), I. Newton (1642–1727) and G. W. Leibniz (1646–1716) were probably the first who recognized that the processes of differentiation (i.e. calculation of the slope of the tangent) and integration (i.e. evaluation of the area) are mutually inverse. However, since they lived at the same time, the question of priority led to bitter controversies. Newton probably arrived at his results earlier, but Leibniz published his conception first. Leibniz was accused of copying Newton's work, and this led to a split in the mathematical world for more than 150 years.

While Newton's disciples, mostly British, pursued his approach, Leibniz's followers, mostly Germans and French, were developing his methods. Due to Leibniz's precise notation and rigorous formalism, their contribution to the further development of calculus and its applications was perhaps greater than that of their British counterparts. Today's historians and mathematicians agree with the opinion that Newton's and Leibniz's inventions were simultaneous but independent.

V.1.8. A remark on definitions of the definite integral. The introduction of the definite integral on the intuitive level, explained in paragraph V.1.2, was considered to be relatively satisfactory for a certain time. However, the increasing importance of calculus and the range of its applications in physics and other disciplines required a correct and rigorous definition of the definite integral. Let us mention three such definitions that have appeared during the last several centuries.

The first definition is due to Newton. Newton took formula (V.1.4) as a base and suggested that the definite integral of function f on the interval $[a, b]$ would be understood to be the difference $F(b) - F(a)$, where F is an antiderivative to f on $[a, b]$. However, this definition rather hazes the original geometric and physical problems which had initially motivated the introduction of the definite integral.

A further definition which was very successful was given by George Riemann (1822–1866). We shall explain this definition in section V.2. The definite integral defined in the way proposed by Riemann is usually called the *Riemann integral*. We shall deal only with this type of definite integral in this textbook.

Today's applied mathematics mostly uses the so called Lebesgue definite integral. This integral represents a generalization of the Riemann integral, however its theory requires a longer explication and goes beyond the scope of this text.

Although we have explained that the notion of an infinitely small positive number $\mathrm{d}x$ does not coincide with today's conception of real numbers, we are not going to release this notion completely. By analogy with our derivation of the formula for the mass of an inhomogeneous bar in paragraph V.1.3, we can further follow the same approach and derive, e.g., the formulas for the length of the graph of a function $y = f(x)$, for the position of the center of mass or for the static moment of a mass curve, etc. These applications of the definite integral will be shown in Section V.7.

V.2. Definition of the Riemann integral

Suppose that

a) $[a, b]$ is a closed, bounded and non–empty interval,

b) f is a bounded function in the interval $[a, b]$.

You will see in this section that the definition of the Riemann integral of function f on the interval $[a, b]$ is based on the partition of $[a, b]$ to many smaller sub–intervals, on the construction of a Riemann sum which is an approximation of the area of the region R between the graph of function f, the x–axis and the straight lines $x = a$ and $x = b$ and finally on the consideration of the limit of the sum as the partition of $[a, b]$ becomes finer and finer.

V.2.1. The partition of an interval. Let $a < b$. A system of points $x_0, x_1, \ldots,$ x_n in the interval $[a, b]$ such that $a = x_0 < x_1 < \ldots < x_n = b$ is called the *partition* of $[a, b]$. If the partition is named P then we write:

$$P: \quad a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b. \tag{V.2.1}$$

The *norm* of partition P is the number $\|P\| = \max_{i=1,\ldots,n}(x_i - x_{i-1})$. (In other words, $\|P\|$ is the length of the largest of the sub-intervals $[x_0, x_1]$, $[x_1, x_2]$, $\ldots,$ $[x_{n-1}, x_n]$. $\|P\|$ provides information on how fine the partition P is.)

V.2.2. Riemann's sum and its limit. Suppose that f is a bounded function in the interval $[a, b]$ and P is a partition of $[a, b]$ given by (V.2.1). Denote by Δx_i the length of the i-th sub-interval $[x_{i-1}, x_i]$. (I.e. $\Delta x_i = x_i - x_{i-1}$.)

Let us choose a point ζ_i in each of the intervals $[x_{i-1}, x_i]$. Denote by V the system of all chosen points $\zeta_1 \in [x_0, x_1]$, $\zeta_2 \in [x_1, x_2]$, $\ldots,$ $\zeta_n \in [x_{n-1}, x_n]$.

Then the sum

$$s(f, P, V) = \sum_{i=1}^{n} f(\zeta_i) \cdot \Delta x_i$$

is called *Riemann's sum* of function f on the interval $[a, b]$ associated with partition P and system V.

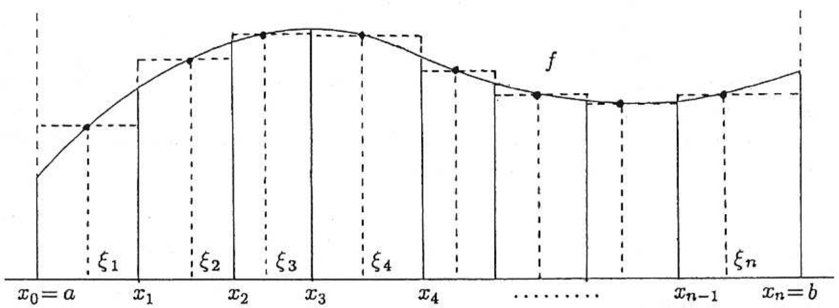

Fig. 25

We say that number S is a *limit of Riemann's sums* $s(f, P, V)$ *as* $\|P\| \to 0+$ if to each $\epsilon > 0$ there exists $\delta > 0$ such that for all partitions P of $[a, b]$ and for all systems V of the points $\zeta_i \in [x_{i-1}, x_i]$ the implication

$$\|P\| < \delta \implies |s(f, P, V) - S| < \epsilon$$

holds. We write:

$$\lim_{\|P\| \to 0+} s(f, P, V) = S. \tag{V.2.2}$$

V.2.3. The Riemann integral. If the limit (V.2.2) exists then function f is said to be *integrable* in the interval $[a, b]$ and the number S is called the *Riemann integral* of function f in $[a, b]$. The integral is denoted

$$\int_a^b f(x)\,\mathrm{d}x \quad \text{or} \quad \int_a^b f\,\mathrm{d}x.$$

Numbers a and b are called the *limits* of the integral. a is the *lower limit* and b is the *upper limit*. The integrated function is called the *integrand*.

Instead of saying that function f is integrable in $[a, b]$, we can also say that the *Riemann integral* $\int_a^b f(x)\,\mathrm{d}x$ *exists*.

The Riemann integral is often called the *definite integral*. (We know from paragraph V.1.8 that there exist several types of definite integrals. However, this cannot lead to a misunderstanding because the only type of definite integral we deal with in this text is the Riemann integral.)

The variable in the Riemann integral can also be denoted by other letters than x and so the integral can be written down as $\int_a^b f(t)\,\mathrm{d}t$, $\int_a^b f(s)\,\mathrm{d}s$ etc.

V.2.4. The area of the region between the graph of function f and the x–axis. If f is a non-negative integrable function in the interval $[a, b]$ then the *area* of region R that is bounded by the graph of f, the x–axis and the straight lines $x = a$ and $x = b$ is defined to be the value of the integral $\int_a^b f(x)\,\mathrm{d}x$.

By analogy, if function f is non-positive and integrable in $[a, b]$ then the area of region R between the graph of f, the x–axis and the straight lines $x = a$ and $x = b$ is defined to be equal to $-\int_a^b f(x)\,\mathrm{d}x$.

Think for yourself over this fact: In a general case when f is an integrable function in $[a, b]$ which can be positive as well as negative, the integral $\int_a^b f(x)\,\mathrm{d}x$ is equal to the sum of the areas of all parts of the plane between the graph of f and the x–axis where the contributions of the parts below the x–axis are taken with the minus sign. (Sketch a picture.)

V.2.5. The extension of the definition of the Riemann integral. If f is an integrable function in the interval $[a, b]$ then we put

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

Specially, we also put $\displaystyle\int_a^a f(x)\,\mathrm{d}x = 0.$

V.2.6. The mean value of function f in the interval $[a, b]$. Suppose that f is an integrable function in the interval $[a, b]$. Then the number

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

is said to be the *mean value* of function f in the interval $[a, b]$.

The geometric interpretation of the mean value is easy: Suppose for simplicity that function f is non-negative in $[a, b]$. Then the mean value of f in $[a, b]$ is such a number μ that region R between the graph of f, the x–axis and the straight lines $x = a$ and $x = b$ has the same area as the rectangle whose sides are $b - a$ and μ. (Sketch a picture.)

V.3. Important properties of the Riemann integral

Let us have in mind that the Riemann integral has been constructed from the beginning only for bounded functions and on a bounded interval $[a, b]$.

Most of the bounded functions you will meet in applications will be integrable. Nevertheless, there also exist "bad" functions for which the limit of the Riemann sums (V.2.2) in the interval $[a, b]$ does not exist. Such functions are not integrable in $[a, b]$, which means that their Riemann integral in $[a, b]$ does not exist. The border between integrable and non–integrable functions is, in general, not easy to describe. In order to have a tool which can enable us to recognize some integrable functions, we show in the next theorem and in remark V.3.2 sufficient conditions which guarantee the integrability of a function f in $[a, b]$.

Bear in mind that that the statements "function f is integrable in the interval $[a, b]$" and "the Riemann integral $\int_a^b f(x)\,dx$ exists" say exactly the same.

V.3.1. Theorem (on the existence of the Riemann integral). *Let function f be continuous in the interval $[a, b]$. Then f is integrable in $[a, b]$.*

V.3.2. Remark. Theorem V.3.1 is satisfactory in most practical cases. However, its generalization is also sometimes useful:

Let function f be bounded and piecewise continuous in the interval $[a, b]$. Then f is integrable in $[a, b]$.

(Function f is said to be <u>*piecewise continuous*</u> in the interval $[a, b]$ if $[a, b]$ can be divided into a finite number of sub–intervals such that f is continuous in the interior of each of them.)

V.3.3. Theorem. a) *If function f is integrable in the interval $[a, b]$ and $[c, d] \subset [a, b]$ then f is integrable in the interval $[c, d]$, too.*

b) *If functions f and g are both integrable in the interval $[a, b]$ then their product $f \cdot g$ is also an integrable function in $[a, b]$.*

c) *If f is an integrable function in the interval $[a, b]$ and a function g differs from f in $[a, b]$ only at a finite number of points then g is also an integrable function in $[a, b]$ and*

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

Assertion a) easily follows from the definition of the Riemann integral.

Assertion b) represents a statement about the integrability of the product of the two functions f and g. However, bear in mind that this <u>does not mean</u> that $\int_a^b f \cdot g\,dx = \left(\int_a^b f\,dx\right) \cdot \left(\int_a^b g\,dx\right)$!

Assertion c) states that the change of function values of function f at a finite number of points has no influence on the existence or on the value of the integral $\int_a^b f\,dx$. In other words: <u>Neither the existence nor the value of the integral $\int_a^b f\,dx$ depends on the values of f at a finite number of points.</u> Thus, function f need not even be defined at a finite number of points of the interval $[a, b]$ and this has no

influence on the existence or on the value of the integral $\int_a^b f\,dx$. In particular, it plays no role whether the integral is considered on an open or on a closed interval.

V.3.4. Upper and lower bounds of the Riemann integral. It can easily be deduced from the definition of the Riemann integral and its geometric interpretation (see paragraph V.2.4) that if function f, integrable in $[a, b]$, satisfies

$$m \le f(x) \le M \quad \text{for all } x \in [a, b],$$

then

$$m \cdot (b - a) \le \int_a^b f(x)\,dx \le M \cdot (b - a).$$

Since this holds with an arbitrary lower bound m and an arbitrary upper bound M of function f on $[a, b]$, we can replace m by the greatest lower bound (i.e. by $\inf_{x\in[a,b]} f(x)$) and M by the least upper bound (i.e. by $\sup_{x\in[a,b]} f(x)$) and the inequality still holds.

In particular, if f has its minimum and maximum in the interval $[a, b]$ then

$$\min_{x\in[a,b]} f(x) \cdot (b - a) \le \int_a^b f(x)\,dx \le \max_{x\in[a,b]} f(x) \cdot (b - a). \tag{V.3.1}$$

The next theorem also easily follows from the definition of the Riemann integral and its geometric sense.

V.3.5. Theorem. *If functions f and g are both integrable in $[a, b]$ and $g(x) \le f(x)$ for all $x \in [a, b]$ then*

$$\int_a^b g(x)\,dx \le \int_a^b f(x)\,dx.$$

Particularly, if $f(x) \ge 0$ for all $x \in [a, b]$ then $\int_a^b f(x)\,dx \ge 0.$

The next theorem confirms the validity of the formulas from paragraph V.1.6.

V.3.6. Theorem. (Linearity of the Riemann integral.) *If f and g are integrable functions in $[a, b]$ and $\alpha \in \mathbf{R}$ then*

$$\int_a^b [f(x) + g(x)]\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx \quad \text{and} \quad \int_a^b \alpha \cdot f(x)\,dx = \alpha \int_a^b f(x)\,dx.$$

(The same property is already known from the theory of the indefinite integral – see theorem IV.1.8.)

V.3.7. Theorem. (Additivity of the Riemann integral with respect to the interval.) *If the integrals $\int_a^c f\,dx$ and $\int_c^b f\,dx$ exist then*

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

V.3.8. Theorem. (The Riemann integral as a function of the upper limit.) *Suppose that function f is integrable in the interval $[a, b]$. Then*

a) *the function $A(x) = \int_a^x f(t)\,dt$ is continuous in $[a, b]$,* *b) the identity*

$$A'(x) = \frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(t)\,\mathrm{d}t = f(x) \qquad \text{(V.3.2)}$$

holds at all points $x \in (a,b)$ where function f is continuous.

The correctness of assertion a) follows (at least intuitively) from the geometric sense of the Riemann integral (see paragraph V.2.4). Try to sketch a picture and think it over.

Formula (V.3.2) can be proved in this way:

$$A'(x) = \lim_{h\to 0}\frac{A(x+h)-A(x)}{h} = \lim_{h\to 0}\frac{1}{h}\left[\int_a^{x+h} f(t)\,\mathrm{d}t - \int_a^x f(t)\,\mathrm{d}t\right] =$$

$$= \lim_{h\to 0}\frac{1}{h}\int_x^{x+h} f(t)\,\mathrm{d}t = \lim_{h\to 0}\mu(h)$$

where $\mu(h)$ is the mean value of function f in the interval with the end points x and $x+h$. The continuity of function f at point x and (V.3.1) imply that $\mu(h)\to f(x)$ for $h\to 0$. So we obtain (V.3.2).

V.3.9. Remark. We can deduce from theorem V.3.8 that if function f is continuous in the interval I and $a\in I$ then $A(x) = \int_a^x f(t)\,dt$ is an antiderivative to f in I.

V.3.10. Remark. The following generalization of part b) of theorem V.3.8 is often useful in applications of the definite integral:

If function f is continuous in interval I and $a(x)$, $b(x)$ are differentiable functions in interval J such that their values belong to the interval I then

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_{a(x)}^{b(x)} f(t)\,\mathrm{d}t = f(b(x))\cdot b'(x) - f(a(x))\cdot a'(x) \quad \text{for } x\in J.$$

V.3.11. Problems. Do the following Riemann integrals exist?

a) $\displaystyle\int_{-2}^{1}\frac{x+1}{x^2-x-6}\,\mathrm{d}x$ b) $\displaystyle\int_1^2\frac{\ln x}{x}\,\mathrm{d}x$ c) $\displaystyle\int_0^1\frac{\sin x}{x}\,\mathrm{d}x$

d) $\displaystyle\int_{-1}^{5}\mathrm{e}^{-x}\,\mathrm{d}x$ e) $\displaystyle\int_{-2.5}^{3}\frac{x}{\ln(x+3)}\,\mathrm{d}x$ f) $\displaystyle\int_{-2}^{-1}\frac{x^2+1}{x^3-2x^2+x}\,\mathrm{d}x$

R e s u l t s : a) no, b) yes, c) yes, d) yes, e) no, f) yes.

V.3.12. Problems. Find $G'(x)$ if function G is defined by the following integrals.

a) $\displaystyle G(x) = \int_{1/x}^{\sqrt{x}}\cos(t^2)\,\mathrm{d}t;\ x>0$ b) $\displaystyle G(x) = \int_0^{2x}\frac{\sin t}{t}\,\mathrm{d}t$

c) $\displaystyle G(x) = \int_x^0\sqrt{1+t^4}\,\mathrm{d}t$ d) $\displaystyle G(x) = \int_{x^2}^{x^3}\ln t\,\mathrm{d}t;\ x>0$

R e s u l t s : a) $\displaystyle G'(x) = \frac{\cos x}{2\sqrt{x}} + \frac{\cos 1/x^2}{x^2}$ (for $x>0$), b) $\displaystyle 2\,\frac{\sin 2x}{2x}$,
c) $-\sqrt{1+x^4}$, d) $(9x^2-4x)\ln x$ (for $x>0$).

V.4. Evaluation of the Riemann integral

We arrive at an important point in the theory of the Riemann integral which is the question of its evaluation. The next theorem is often called, due to its importance, the Fundamental Theorem of Integral Calculus.

V.4.1. Theorem. *If f is a continuous function in interval $[a,b]$ and F is an antiderivative to f in $[a,b]$ then*

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a). \qquad \text{(V.4.1)}$$

Formula (V.4.1) is called the *Newton–Leibniz formula*. The difference $F(b) - F(a)$ is usually written down in a shorter way: $F(b) - F(a) = [F]_a^b$.

The proof of theorem V.4.1 is easy: Function $A(x) = \int_a^x f(t)\,\mathrm{d}t$ is an antiderivative to f in $[a,b]$, too. Two antiderivatives can differ at most by an additive constant (theorem IV.1.5). Hence there exists a constant C such that $F = G + C$ in $[a,b]$. This means that $F(a) = A(a) + C = C$ (because $A(a) = 0$) and $F(b) = A(b) + C = A(b) + F(a)$. This implies: $\int_a^b f(t)\,\mathrm{d}t = A(b) = F(b) - F(a)$.

The Newton-Leibniz formula connects the definite integral with the indefinite integral: If you know the indefinite integral of function f in the interval $[a,b]$ then you know all antiderivatives to f in $[a,b]$. You can therefore choose any of them, use it in the Newton–Leibniz formula and you obtain the value of the definite integral of function f in $[a,b]$. The fact that the indefinite integral and the antiderivative play such an important role in the evaluation of the definite integral is one of the main reasons why you learned to calculate indefinite integrals in Chapter IV.

We have already several times repeated that all antiderivatives to f in $[a,b]$ differ at most by an additive constant. Thus, if you choose for instance the antiderivative $F+c$ (where c is a constant) instead of F and use it in the Newton–Leibniz formula, you obtain

$$\int_a^b f(x)\,\mathrm{d}x = [F+c]_a^b = (F(b)+c) - (F(a)+c) = F(b) - F(a).$$

The result is the same as in (V.4.1). This confirms that you can indeed use an arbitrary antiderivative to f in order to evaluate the definite integral $\int_a^b f\,dx$.

V.4.2. Example. $\displaystyle\int_0^\pi \sin x\,\mathrm{d}x = [-\cos x]_0^\pi = (-\cos\pi) - (-\cos 0) = 2.$

The next two theorems show that the two methods of integration, by parts and by substitution, known from the theory of the indefinite integral, can also be directly applied to the definite integral.

V.4.3. Theorem. (Integration by parts.) *Suppose that functions u and v have continuous derivatives in the interval $[a,b]$. Then*

$$\int_a^b u'\cdot v\,\mathrm{d}x = [u\cdot v]_a^b - \int_a^b u\cdot v'\,\mathrm{d}x. \qquad \text{(V.4.2)}$$

V.4.4. Example. $\int_0^2 e^{2x} \cdot x \, dx = {}^{*)} \left[\frac{1}{2} e^{2x} \cdot x \right]_0^2 - \int_0^2 \frac{1}{2} e^{2x} \, dx =$

$$= \tfrac{1}{2} e^4 \cdot 2 - \tfrac{1}{2} e^0 \cdot 0 - \left[\tfrac{1}{4} e^{2x} \right]_0^2 = e^4 - \tfrac{1}{4} e^4 + \tfrac{1}{4} e^0 = \tfrac{3}{4} e^4 + \tfrac{1}{4}.$$

*) We have put $u'(x) = e^{2x}$, $u(x) = \frac{1}{2} e^{2x}$, $v(x) = x$ and $v'(x) = 1$.

V.4.5. Theorem. (Integration by substitution.) *Let function g have a continuous derivative in the interval $[a, b]$ and map $[a, b]$ into an interval J. Let function f be continuous in J. Then*

$$\int_a^b f(g(x))\, g'(x) \, dx = \int_{g(a)}^{g(b)} f(s) \, ds. \qquad \text{(V.4.3)}$$

Formula (V.4.3) can be used in two ways:

a) We want to evaluate the integral on the left hand side and we transform it to the integral on the right hand side (if the integral on the right hand side is simpler) or

b) we want to evaluate the integral on the right hand side and we transform it to the integral on the left hand side (if the integral on the left hand side is simpler).

V.4.6. Example. Evaluate the integral $\int_0^{\pi/2} \sin^2 x \cdot \cos x \, dx$.

If we put $[a, b] = [0, \pi/2]$, $s = g(x) = \sin x$, $f(s) = s^2$, $J = (-\infty, +\infty)$ then all the assumptions of theorem V.4.5 are fulfilled. Moreover, $g(0) = \sin 0 = 0$ and $g(\pi/2) = \sin(\pi/2) = 1$. Applying formula (V.4.3), we obtain:

$$\int_0^{\pi/2} \sin^2 x \cdot \cos x \, dx = \int_0^1 s^2 \, ds = \left[\tfrac{1}{3} s^3 \right]_0^1 = \tfrac{1}{3}.$$

V.4.7. Example. Evaluate the integral $\int_0^2 \sqrt{4 - x^2} \, dx$.

This integral can be considered to be the integral on the right hand side of (V.4.3) (with the variable denoted by x instead of s). The function $f(x) = \sqrt{4 - x^2}$ is continuous in $[0, 2]$, hence the integral exists. Put $x = g(t) = 2 \sin t$, $dx = g'(t)\, dt = 2 \cos t \, dt$. Then $g(a) = 2 \sin a = 0$ and $g(b) = 2 \sin b = 2$. We can therefore choose $a = 0$ and $b = \pi/2$. All the assumptions of theorem V.4.5 are now fulfilled and so, applying formula (V.4.3), we get:

$$\int_0^2 \sqrt{4 - x^2} \, dx = \int_0^{\pi/2} \sqrt{4 - 4 \sin^2 t} \cdot 2 \cos t \, dt = \int_0^{\pi/2} 4 \cos^2 t \, dt =$$

$$= \int_0^{\pi/2} 2(1 + \cos 2t) \, dt = \left[2t + \sin 2t \right]_0^{\pi/2} = \pi.$$

V.4.8. Remark. If you need to evaluate the Riemann integral of function f in the interval $[a, b]$ and you consider applying integration by parts or by a substitution then you have two possibilities:

1) You can use theorem V.4.3 or theorem V.4.5. Thus, you transform the integral to another (simpler) integral and you also deal with the limits of the integral.

This approach was explained in examples V.4.5, V.4.6 a V.4.7.

2) You can first calculate the integral as an indefinite integral and then you can apply the Newton–Leibniz formula (V.4.1) on the interval $[a, b]$.

In order to show this approach in greater detail, we evaluate the integral from example V.4.6 once again, this time by the method we are just explaining. Thus, let us begin with the indefinite integral $\int \sin^2 x \cos x \, dx$. We can use the substitution $s = \sin x$. Then $ds = \cos x \, dx$ and

$$\int \sin^2 x \cdot \cos x \, dx = \int s^2 \, ds = \tfrac{1}{3} s^3 + C = \tfrac{1}{3} \sin^3 x + C.$$

Using formula (V.4.1), we obtain: $\int_0^{\pi/2} \sin^2 x \cos x \, dx = \left[\frac{1}{3} \sin^3 x \right]_0^{\pi/2} = \frac{1}{3}$.

You will see after having solved a large number of problems individually that the first method, based on the direct application of the integration by parts or the integration by substitution is usually technically easier and less laborious.

V.4.9. Problems. Evaluate the following integrals.

a) $\int_{-1}^1 (3x^2 - 4x + 7) \, dx$ b) $\int_0^1 (8t^3 - 12t^2 + 5) \, dt$ c) $\int_1^2 \frac{4}{s^2} \, ds$

d) $\int_1^{27} x^{-4/3} \, dx$ e) $\int_0^2 \sqrt{4 - s^2} \, ds$ f) $\int_0^1 u \arctan u \, du$

g) $\int_0^{\pi} \cos^2 x \, dx$ h) $\int_0^1 9w\, e^{3w} \, dw$ i) $\int_{1/e}^{e} |\ln r| \, dr$

j) $\int_0^1 \sqrt{\frac{x}{4 - x}} \, dx$ k) $\int_0^{\pi/2} \sin^3 \varphi \cos^2 \varphi \, d\varphi$ l) $\int_0^{\pi/2} \frac{d\theta}{3 + 2 \sin \theta}$

Evaluate the area of the domain between the graph of function f and the x–axis.

m) $f(x) = x^2 - 4x + 3$, $0 \le x \le 3$ n) $f(x) = 1 - (x^2/4)$, $-2 \le x \le 3$

o) $f(x) = 5 - 5x^{2/3}$, $-1 \le x \le 8$ p) $f(x) = 1 - \sqrt{x}$, $0 \le x \le 4$

Find the mean value of function f on a given interval.

q) $f(x) = \sqrt{3x}$ on $[0, 3]$ r) $f(x) = \sqrt{ax}$ on $[0, a]$

s) $f(x) = mx + b$ on $[-1, 1]$ t) $f(x) = mx + b$ on $[-k, k]$

Results: a) 18, b) 4, c) 2, d) 2, e) π, f) $\frac{1}{4}\pi - \frac{1}{2}$, g) $\frac{1}{2}\pi$, h) $2e^3 + 1$, i) $2(1 - e^{-1})$, j) $2\pi/3 - \sqrt{3}$, k) $\frac{2}{15}$, l) $2/\sqrt{5} \cdot \arctan(1/\sqrt{5})$, m) 0, n) $\frac{41}{12}$, o) -58, p) $-\frac{4}{3}$, q) 2, r) $\frac{2}{3} a$, s) b, t) b.

V.5. Numerical integration

It is well known from the theory of the indefinite integral that an antiderivative to a given function f often exists, but it cannot be obtained by standard integration procedures and it cannot be expressed in a "closed form", i.e. by a formula which prescribes a finite number of operations. By analogy, it often happens that the Riemann integral $\int_a^b f\,dx$ exists, but it cannot be evaluated by a standard integration based on the application of the antiderivative and the Newton–Leibniz formula. However, there exist <u>approximate methods</u> (often also called <u>numerical methods</u>), which enable us to evaluate the integral only approximately, but with an error as small as we wish. We shall explain two such methods in this section. All these methods require the performance of a relatively high number of arithmetic operations in order to achieve higher accuracy (i.e. a smaller error). They can therefore be practically realized only on computers.

Both the methods are based on the partition

$$D: \quad a = x_0 < x_1 < x_2 \ldots < x_{n-1} < x_n = b$$

of the interval $[a,b]$ onto n sub–intervals $[x_{k-1}, x_k]$ $(k = 1, 2, \ldots, n)$ of the same length h. Hence

$$h = \frac{b-a}{n} \quad \text{a} \quad x_k = a + k \cdot h \quad (k = 1, 2, \ldots, n).$$

We shall denote $y_k = f(x_k)$.

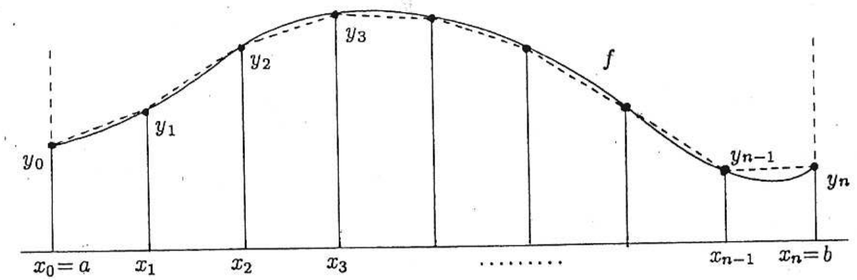

Fig. 26

V.5.1. Trapezoidal Rule. We approximate function f by a linear function on each sub–interval $[x_{k-1}, x_k]$. The linear function is uniquely determined by the requirement that its graph (a straight line) passes through two chosen points. Let us choose the points $[x_{k-1}, y_{k-1}]$ and $[x_k, y_k]$. Then the considered linear function is: $y = y_{k-1} + (y_k - y_{k-1})/h \cdot (x - x_{k-1})$. Its integral in the interval $[x_{k-1}, x_k]$ (let us denote it by I_k) represents the area of the trapezoid (see Fig. 26) and so $I_k = h \cdot (y_{k-1} + y_k)/2$. If we add all the numbers $I_1, I_2, \ldots, I_n$, we obtain

(V.5.1) $$L_n = \frac{h}{2} \cdot \left[y_0 + 2y_1 + 2y_2 + \ldots + 2y_{n-1} + y_n \right].$$

L_n is the approximate value of the Riemann integral $\int_a^b f\,dx$. The geometric sense of L_n is shown in Fig. 26 – it is the sum of the areas of n trapezoids constructed on the intervals $[x_0, x_1], [x_1, x_2] \ldots, [x_{n-1}, x_n]$.

It can naturally be expected that the finer the partition of the interval $[a,b]$, the better is the approximation of the Riemann integral $\int_a^b f\,dx$. In other words, the accuracy of the approximation should increase with increasing n (i.e. with decreasing h). This expectation is correct. It can be proved that if function f has a continuous second derivative f'' in $[a,b]$ and M_2 is the maximum of $|f''|$ in $[a,b]$ then the following <u>error estimate</u> holds:

(V.5.2) $$\left| L_n - \int_a^b f(x)\,dx \right| \le \frac{b-a}{12}\, h^2 M_2.$$

(Estimate (V.5.2) is called an error estimate because it provides an upper estimate of the error we make if we replace the exact value of the integral $\int_a^b f\,dx$ by its approximate value L_n.)

V.5.2. Simpson's Rule. Let us now choose an integer n so that it is <u>even</u> and let us approximate function f by a quadratic polynomial on each of the sub–intervals $[x_0, x_2], [x_2, x_4], \ldots, [x_{n-2}, x_n]$. The quadratic polynomial on the sub–interval $[x_{k-2}, x_k]$ $(k = 2, 4, \ldots, n)$ is uniquely determined if we require that its graph (a parabola) passes through three chosen points. Let the three chosen points be $[x_{k-2}, y_{k-2}], [x_{k-1}, y_{k-1}], [x_k, y_k]$. The coefficients and the integral of such a quadratic polynomial in the interval $[x_{k-2}, x_k]$ can be relatively easily evaluated – you can verify for yourself that the integral is $I_k = h \cdot (y_{k-2} + 4y_{k-1} + y_k)/3$. Summing all the numbers $I_2, I_4, \ldots, I_n$, we obtain

(V.5.3) $$S_n = \frac{h}{3} \cdot \left[y_0 + 4y_1 + 2y_2 + 4y_3 + \ldots + 2y_{n-2} + 4y_{n-1} + y_n \right].$$

S_n is an approximate value of the Riemann integral $\int_a^b f\,dx$. It can be proved that if the fourth derivative $f^{(4)}$ of function f is continuous in $[a,b]$ and M_4 is the maximum of $|f^{(4)}|$ in $[a,b]$ then the following <u>error estimate</u> holds:

(V.5.4) $$\left| S_n - \int_a^b f\,dx \right| \le \frac{b-a}{180}\, h^4 M_4.$$

V.6. The improper Riemann integral

We have assumed that

(V.6.1) $[a,b]$ is a bounded interval and

(V.6.2) f is a bounded function in this interval

in the definition of the Riemann integral $\int_a^b f\,dx$ (see Section V.2). However, there often arises the necessity to evaluate definite integrals where either the domain of integration (i.e. an interval) or the integrand (i.e. a function), or both, are unbounded. Such integrals are called <u>improper integrals</u>. We will explain the definition of the improper Riemann integral in this section. Let us begin with a simple example.

V.6.1. Example. The Riemann integral $\displaystyle\int_1^{+\infty} \frac{1}{x}\,dx$ does not exist because the domain of integration is not a bounded interval. However, the question "what is the area of the region $R = \{[x,y];\ x \in [1,+\infty),\ 0 \le y \le 1/x\}$" is quite reasonable. ($R$ is a subset of plane $\mathbb{E}_2$, corresponding to $x \in [1,+\infty)$, bounded above by the graph of the function $1/x$ and bounded below by the x–axis. Sketch a picture.) The answer can be found in this way: Let us choose $t \in [1,+\infty)$. The area of the region $R_t = \{[x,y];\ x \in [1,t],\ 0 \le y \le 1/x\}$, i.e. of the part of the plane which corresponds to $x \in [1,t]$ and which is bounded above by the graph of the function $1/x$ and bounded below by the x–axis, is

$$s(R_t) = \int_1^t \frac{1}{x}\,\mathrm{d}x = [\ln x]_1^t = \ln t - \ln 1 = \ln t.$$

The area of the whole region R can now be naturally defined by the equation

$$s(R) = \lim_{t\to+\infty} s(R_t) = \lim_{t\to+\infty} \ln t = +\infty.$$

This approach is the motivation for the following definition.

V.6.2. The improper Riemann integral with the singular upper limit. Assume that function f is defined in the interval $[a,b)$ and it is integrable in each interval $[a,t]$ (for $a \le t < b$). If

a) at least one of the conditions (V.6.1) and (V.6.2) is not fulfilled and

b) the limit
$$\lim_{t\to b-} \int_a^t f(x)\,\mathrm{d}x,$$
exists

then the value of the limit is called the *improper Riemann integral* with the *singular upper limit*.

The improper Riemann integral is denoted in the same way as the "usual" Riemann integral, i.e. $\int_a^b f(x)\,\mathrm{d}x$. Thus, we can write:

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{t\to b-} \int_a^t f(x)\,\mathrm{d}x.$$

The improper Riemann integral with the *singular lower limit* can be defined by analogy. (Try to write the definition for yourself.)

While the value of the "proper" Riemann integral can only be a finite number, the value of an improper Riemann integral can be finite or infinite. If the improper integral $\int_a^b f\,\mathrm{d}x$ is finite then we say that the integral *converges*. If $\int_a^b f\,\mathrm{d}x = \pm\infty$, we say that the integral *diverges*.

V.6.3. Example. Let us evaluate the improper Riemann integral $\displaystyle\int_{-5}^{2} \frac{1}{\sqrt{2-x}}\,\mathrm{d}x.$

The function $f(x) = 1/\sqrt{2-x}$ is unbounded in the interval $[-5,2]$. (This is due to its behavior in the neighborhood of the upper limit 2, namely $\lim_{x\to 2-} 1/\sqrt{2-x} = +\infty$.) Hence the Riemann integral of function f in $[-5,2]$ does not exist. However, f is a continuous function in the interval $[-5,2)$ and so it is integrable in each interval $[-5,t]$ (where $-5 \le t < 2$). The antiderivative to f in $[-5,2)$ is $F(x) = -2\sqrt{2-x}$. Thus,

$$\int_{-5}^t \frac{1}{x}\,\mathrm{d}x = F(t) - F(1) = -2\sqrt{2-t} + 2\sqrt{2-(-5)} = -2\sqrt{2-t} + 2\sqrt{7}.$$

Since $\lim_{t\to 2-} -2\sqrt{2-t} + 2\sqrt{7} = 2\sqrt{7}$, we have: $\displaystyle\int_{-5}^2 \frac{1}{\sqrt{2-x}}\,\mathrm{d}x = 2\sqrt{7}.$

V.6.4. The improper Riemann integral with both singular limits. The definition of the improper Riemann integral can be extended to the cases when both the limits are singular: Assume that $c \in (a,b)$ and the two integrals $\int_a^c f\,\mathrm{d}x$ and $\int_c^b f\,\mathrm{d}x$ exist (the first as the improper integral with the singular lower limit a and the second as the improper integral with the upper limit b). If the sum of the integrals has a sense (i.e. if it is not for instance $-\infty + \infty$) then we put

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x.$$

The integral $\int_a^b f\,\mathrm{d}x$ is an improper integral with both singular limits.

The definition can be further extended to situations when some points inside the interval (a,b) are singular. However, we do not deal with such cases in this textbook.

V.6.5. Remark. If the improper integral $\int_a^b f\,\mathrm{d}x$ exists and the interval (a,b) is unbounded then we sometimes speak about the *improper integral due to the limit*. If function f is unbounded then we speak about the *improper integral due to the function*. These two effects can also be combined (the function is unbounded in the neighborhood of one limit and the second limit is infinite) or can be cumulated (the function is unbounded in the neighborhood of an infinite limit).

V.6.6. The evaluation of the improper Riemann integral. We shall assume for simplicity that f is a continuous function in the interval (open or closed) with the end points a, b (where $a < b$) in this paragraph.

According to theorem IV.1.3, function f has an antiderivative in the interval between a and b. Let us denote it by F. (There exist in fact infinitely many antiderivatives and they all differ by additive constants – see theorem IV.1.5. However, we shall need only one of them.) It can be shown that the integral $\int_a^b f\,\mathrm{d}x$ exists as an improper integral and

(V.6.1) $$\int_a^b f(x)\,\mathrm{d}x = \lim_{t\to b-} F(t) - \lim_{t\to a+} F(t),$$

if the right hand side has a sense (i.e. if both the limits exist and their difference has a sense).

V.6.7. Example. The integral $\displaystyle\int_0^{+\infty} \frac{1}{\sqrt{x}}\,\mathrm{d}x$ is an improper integral with both singular limits: 0 is a singular limit because the function $1/\sqrt{x}$ is not bounded in its right neighborhood and $+\infty$ is a singular limit because it is not a finite number. The antiderivative to $1/\sqrt{x}$ in $(0,+\infty)$ is for instance the function $2\sqrt{x}$. Then

$$\int_0^{+\infty} \frac{1}{\sqrt{x}}\,\mathrm{d}x = \lim_{t\to+\infty} 2\sqrt{t} - \lim_{t\to 0+} 2\sqrt{t} = +\infty - 0 = +\infty.$$

V.6.8. Problems. Evaluate the improper integrals.

a) $\int_0^{+\infty} \frac{dx}{1+x^2}$ b) $\int_{-\infty}^{+\infty} \frac{dx}{4+x^2}$ c) $\int_{-\infty}^{-2} \frac{dx}{x^2}$

d) $\int_2^{+\infty} \frac{dx}{x^2-1}$ e) $\int_0^{+\infty} x^2 e^{-x}\, dx$ f) $\int_0^5 \frac{1}{\sqrt{x}}\, dx$

g) $\int_{-\infty}^{+\infty} \frac{dx}{x^2+4x+9}$ h) $\int_0^{\infty} x\, e^{-x^2}\, dx$ i) $\int_1^5 (x-1)^a\, dx;\ a > -1$

Results: a) $\pi/2$, b) $\pi/2$, c) $\frac{1}{2}$, d) $-\frac{1}{2}\ln\frac{1}{3}$, e) 2, f) $2\sqrt{5}$, g) $\pi/\sqrt{20}$, h) $\frac{1}{2}$, i) $4^{a+1}/(a+1)$.

V.7. Some geometric and physical applications of the definite integral

The simple geometric and physical sense of the definite integral has already been discussed in paragraphs V.1.1, V.1.3 and V.2.4. We will show more possible applications of the definite integral in geometry and physics in this section.

We assume that f is a non-negative and continuous function in the interval $[a, b]$ everywhere in this section.

V.7.1. The volume of a circular body. Let us denote, as in paragraph V.1.1, by R the region which is bounded by the graph of function f, by the x-axis and by the straight lines $x = a$, $x = b$. (See Fig. 24.) The revolution of region R about the x-axis leads to a circular body. Its volume is:

(V.7.1) $$V = \pi \int_a^b f^2(x)\, dx.$$

We can arrive at formula (V.7.1) by a similar consideration as in paragraph V.1.2, where we have derived the expression of the area of region R by the definite integral of function f. The interval $[a, b]$ can be divided into infinitely many "infinitely short" parts. A typical part is the interval with the end points x and $x+dx$ where $x \in [a, b)$. The region between the graph of function f on the interval $[x, x+dx]$ and the line segment with the end points x and $x+dx$ on the x-axis is an "infinitely narrow" rectangle whose sides have the lengths dx and $f(x)$. Its revolution about the x-axis leads to an "infinitely thin" cylinder whose volume is $dV = \pi f^2(x)\, dx$. The volume of the whole circular body is the sum of infinitely many "infinitely small" numbers dV. Thus, we obtain formula (V.7.1).

Other formulas, shown in this section, can be obtained by analogy. This is why we do not show the ideas which stand in the background.

V.7.2. The area of a circular surface. The revolution of the graph of function f about the x-axis creates a circular surface. Its area is

$$s = 2\pi \int_a^b f(x)\sqrt{[f'(x)]^2+1}\, dx.$$

V.7.3. The length of a curve. The graph of function f in the interval $[a, b]$ is a curve. Its length is

$$l = \int_a^b \sqrt{[f'(x)]^2+1}\, dx.$$

V.7.4. The static moments, the center of mass and the moments of inertia of a curve. Suppose that the curve from the preceding paragraph (i.e. the graph of function f) is covered by a mass with a constant longitudinal density ρ. The total mass of the curve is in this case equal to the product of ρ with the length of the curve:

$$m = \rho l = \rho \int_a^b \sqrt{[f'(x)]^2+1}\, dx.$$

The so called static moments of the curve with respect to the x– and y–axes are the integrals

$$S_x = \rho \int_a^b f(x)\sqrt{[f'(x)]^2+1}\, dx, \qquad S_y = \rho \int_a^b x\sqrt{[f'(x)]^2+1}\, dx.$$

The coordinates x_C and y_C of the center of mass of the considered curve satisfy:

$$x_C = \frac{S_x}{m}, \qquad y_C = \frac{S_y}{m}.$$

The so called moments of inertia of the curve according to the x– and y–axes are

$$J_x = \rho \int_a^b f^2(x)\sqrt{[f'(x)]^2+1}\, dx, \qquad J_y = \rho \int_a^b x^2\sqrt{[f'(x)]^2+1}\, dx.$$

V.7.5. The static moments, the center of mass and the moments of inertia of a planar region. Let R be a region from paragraph V.7.1. Suppose that R is covered by a mass with a constant planar density ρ. The total mass of region R is

$$m = \rho \int_a^b f(x)\, dx.$$

The static moments of region R with respect to the x– and y–axes are

$$S_x = \frac{1}{2}\rho \int_a^b f^2(x)\, dx, \qquad S_y = \rho \int_a^b x\, f(x)\, dx.$$

The coordinates x_C and y_C of the center of mass of region R satisfy:

$$x_C = \frac{S_x}{m}, \qquad y_C = \frac{S_y}{m}.$$

The moments of inertia of the region with respect to the x– and y–axes are

$$J_x = \frac{1}{3}\rho \int_a^b f^3(x)\, dx, \qquad J_y = \frac{1}{2}\rho \int_a^b x^2 f(x)\, dx.$$

V.7.6. The static moments and the center of mass of a circular body. Suppose that the circular body created by region R revolving about the x-axis contains a mass which is distributed with a constant volume density ρ. The total mass of the body is equal to the product of ρ and the volume of the body:

$$m = \rho V = \rho \int_a^b f^2(x)\,dx.$$

The static moments of the body according to the xy–, xz– and yz–planes are

$$S_{xy} = 0, \qquad S_{xz} = 0, \qquad S_{yz} = \pi\rho \int_a^b x\,f^2(x)\,dx.$$

The coordinates x_C, y_C and z_C of the center of mass of the body satisfy:

$$x_C = \frac{S_{yz}}{m}, \qquad y_C = \frac{S_{xz}}{m} = 0, \qquad z_C = \frac{S_{xy}}{m} = 0.$$

The moment of inertia of the body with respect to the axis of revolution x is

$$J_x = \frac{1}{2}\rho\pi \int_a^b f^4(x)\,dx.$$

V.7.7. Remark. You will see in the Mathematics II course how we can evaluate the lengths of more general curves, the areas of more general surfaces and the volumes of more general bodies. You will also learn how we can evaluate the mass, the static moments, the moments of inertia and the coordinates of the center of mass of curves, surfaces or bodies in the case of a generally non–constant density.

Bibliography

[BC] B.Budínský, J.Charvát: *Matematika I.* SNTL/Alfa, Praha 1987. (A detailed and understandable textbook, originally written for faculties of civil engineering. In Czech.)

[FT] R.L.Finney, G.B.Thomas, Jr.: *Calculus.* Addison–Wesley Publishing Company, New York, 1990, 1994, etc. (A basic textbook used especially at American universities.)

[J1] V.Jarník: *Diferenciální počet I.* Nakladatelství ČSAV, Praha 1946, 1951, 1953, 1955, 1963, etc. (A fundamental and detailed textbook on differential calculus. In Czech.)

[J2] V.Jarník: *Integrální počet I.* Nakladatelství ČSAV, Praha 1948, 1954, 1956, 1963, etc. (A fundamental and detailed textbook on integral calculus. In Czech.)

[NK] J.Neustupa, S.Kračmar: *Problems in Mathematics I.* Editorial Center of the CTU, Praha 1998. (Book of problems, useful for the exercises and recommended seminars in Mathematics I and for individual study.)

[Ne] J.Neustupa: *Matematika I.* Editorial Center of the CTU, Praha 2002. (A Czech version of this textbook.)

[Re] K.Rektorys: *Survey of Applied Mathematics.* SNTL, Praha 1998 (An extensive survey of classical parts of applied mathematics.)

Index + Vocabulary