Mathematics II

Prague, 1998

Contents

IV. Line Integrals *(Jiří Neustupa)*

V. Surface Integrals *(Jiří Neustupa)*

VI. Potential and Solenoidal Vector Field *(Jiří Neustupa)*

Introduction

The presented text approximately coincides with the contents of the Mathematics II course, taught at the Faculty of Mechanical Engineering in the second term. It deals with multi–variable calculus (i.e. with differential calculus of functions of more variables and with multiple, line and surface integrals). Functions of more variables appear more often in science than functions of one variable and their analysis leads to a large variety of applications. Since the problems we deal with have a more–dimensional character, their understanding requires not only a computational skill, but also a good space imagination. We therefore consider the text not as an independent textbook, but as a complementary material to the lectures and exercises where all the topics will be explained and commented in detail.

The text contains many well known theorems of applied mathematics, like the Green theorem, the Gauss–Ostrogradsky theorem, the Stokes theorem, etc. The conclusions of these theorems are certain integral formulas. The students often identify the theorems with these formulas. However, you should be aware that the important parts of all theorems are also their assumptions. It would be naive to think that the conclusive formulas hold in all cases. The opposite is true – they hold only in certain specific situations. The assumptions of the theorems represent the brief and simplest description of these situations and they are as important as the conclusions of the theorems.

Each chapter contains the section "Exercises" at the end. Many further exercises and solved examples can be found in the textbooks [1] and [3].

The authors wish to express their thanks to Mr. Robin Healey for carefully reading the text and correcting the language. If you still find some misprints or incorrect formulations in the text then they are only the authors who are responsible.

We believe that this text will be a useful study aid not only for students who attend the lectures and the exercises in English, but also for all other students who study in Czech.

I. Functions of Several Real Variables

Functions of several real variables very often appear in mathematics and in science. You already know many formulas which can be understood as definitions of functions of several variables. For example, the formula $V = \pi r^2 h$, which determines the volume V of a circular cylinder from its radius r and height h can be taken as the definition of function V, which depends on two real variables $r > 0, h > 0$.

I.1. Euclidean space $\mathbf{E}_n$.

In this section, we will recall some notions that you know from the Mathematics I course. We will deal here with Euclidean space $\mathbf{E}_n$, subsets of $\mathbf{E}_n$ and its properties.

I.1.1. n-dimensional arithmetic space. If n is a natural number (we use the notation: $n \in \mathbf{N}$) then the set of all ordered n–tuples of real numbers is denoted by $\mathbf{R}^n$. Let us define the sum of any two elements $[x_1, x_2, ..., x_n]$, $[y_1, y_2, ..., y_n]$ from $\mathbf{R}^n$ by the formula:

$$[x_1, x_2, ..., x_n] + [y_1, y_2, ..., y_n] = [x_1 + y_1, x_2 + y_2, ..., x_n + y_n]$$

and the product of any element $[x_1, x_2, ..., x_n]$ from $\mathbf{R}^n$ and any real number λ by the formula

$$\lambda \cdot [x_1, x_2, ..., x_n] = [\lambda x_1, \lambda x_2, ..., \lambda x_n].$$

The set $\mathbf{R}^n$ with these two operations is called the *n-dimensional arithmetic space*. Its elements are called *arithmetic vectors*.

I.1.2. Euclidean space $\mathbf{E}_n$ – definition. Let us define the distance ρ of any two elements $X = [x_1, x_2, ..., x_n]$, $Y = [y_1, y_2, ..., y_n]$ from $\mathbf{R}^n$ by the formula

$$\rho(X, Y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + ... + (x_n - y_n)^2}.$$

The set $\mathbf{R}^n$ with this distance ρ defined for all pairs of elements of $\mathbf{R}^n$ is called *n-dimensional Euclidean space*. This is denoted by $\mathbf{E}_n$.

I.1.3. Zero element of $\mathbf{E}_n$. The point $[0, 0, ..., 0]$ is called the *zero element of* $\mathbf{E}_n$ or the *origin* of $\mathbf{E}_n$. The zero element is denoted $\mathcal{O}$.

I.1.4. Remark. Elements from $\mathbf{E}_n$ are often called *points*, because $\mathbf{E}_1$ can be imagined as a straight line, $\mathbf{E}_2$ as a plane, etc.

The distance between the zero element $\mathcal{O}$ and an arbitrary point X of $\mathbf{E}_n$ is denoted $|X|$, i.e.

$$|X| = \rho(\mathcal{O}, X).$$

From this it follows that the distance between $X, Y \in \mathbf{E}_n$ can be expressed in the following way

$$\rho(X, Y) = |X - Y|,$$

where the difference $X - Y$ is understood in the sense of n-dimensional arithmetic space as $X + (-1) \cdot Y$.

In the following paragraphs, we will define some properties of subsets of $\mathbf{E}_n$, which play an important role in particular in definitions of continuity and limits of functions.

I.1.5. Neighbourhoods in $\mathbf{E}_n$. If $X \in \mathbf{E}_n$, then a *neighbourhood* of the point X is any subset $\{Y \in \mathbf{E}_n : \rho(X,Y) < \varepsilon\}$ where $\varepsilon > 0$. The neighbourhood is denoted by $U_\varepsilon(X)$ or simply $U(X)$.

A *reduced neighbourhood* of the point $X \in \mathbf{E}_n$ is every set of the type $\{Y \in \mathbf{E}_n : 0 < \rho(X,Y) < \varepsilon\}$ where $\varepsilon > 0$, i.e. $U_\varepsilon(X) - X$. This neighbourhood will be denoted by $R_\varepsilon(X)$ or only $R(X)$.

I.1.6. Interior point. Let $M \subset \mathbf{E}_n$. A point $X \in M$ is called an *interior point of M* if there exists a neighbourhood $U(X)$ such that $U(X) \subset M$.

I.1.7. Accumulation point. Let $M \subset \mathbf{E}_n$. A point $X \in \mathbf{E}_n$ is called an *accumulation point of M* or a *point of accumulation of M* if in every reduced neighbourhood $R(X)$ there exists at least one point Y which belongs to M.

I.1.8. Remark. If you read this definition carefully, you can see that if X is an accumulation point of M then in every neighbourhood $U(X)$ there exist an infinite number of points which belong to M.

From the definition it also follows that even if X is an accumulation point of M then it is possible: $X \notin M$.

I.1.9. Isolated point. Let $M \subset \mathbf{E}_n$. A point $X \in M$ is called an *isolated point of M* if in some reduced neighbourhood $R(X)$ there is no point which belongs to M.

I.1.10. Remark. The following assertion holds:

If M is a subset in $\mathbf{E}_n$, X is any point of M and X is not an isolated point of M, then X is an accumulation point of M.

If M is a subset in $\mathbf{E}_n$, X is any point of M and X is not an accumulation point of M, then X is an isolated point of M.

I.1.11. Boundary point. A point $X \in \mathbf{E}_n$ is called a *boundary point of a subset M* if in every $U(X)$ there exists at least one point which belongs to M and at least one point which does not belong to M.

I.1.12. Open set. A subset M of $\mathbf{E}_n$ is called an *open set in* $\mathbf{E}_n$ (or shortly: an *open set*), if every point $X \in M$ is an interior point of M.

I.1.13. Examples. See Fig. 1. The following sets are open in $\mathbf{E}_2$. Sketch the fourth one:

$$\emptyset, \qquad \mathbf{E}_2, \qquad \{X \in \mathbf{E}_2 \,:\, \rho(\mathcal{O}, X) < 5\},$$

$$\{[x,y] \in \mathbf{E}_2 \,:\, \rho([1,0],[x,y]) < 1\} \cup \{[x,y] \in \mathbf{E}_2 \,:\, x \in (-2;-1)\}$$

I.1.14. Closure of a set. Let $M \subset \mathbf{E}_n$. A *closure of subset M in $\mathbf{E}_n$* (or shortly: a *closure of M*) is the name given to the union of M with the set of all accumulation points of M. The closure of subset M in $\mathbf{E}_n$ is denoted by $\overline{M}$.

I.1.15. Closed set. A subset M of $\mathbf{E}_n$ is called a *closed set in $\mathbf{E}_n$* (or shortly: a *closed set*), if $\overline{M} = M$.

I.1.16. Examples. The following sets are closed in $\mathbf{E}_2$:

$$\emptyset, \qquad \mathbf{E}_2, \qquad \{X \in \mathbf{E}_2 \,:\, \rho(\mathcal{O},X) \leq 5\},$$

$$\{[x,y] \in \mathbf{E}_2 \,:\, \rho([1,2],[x,y]) \leq 1\} \cup \{[x,y] \in \mathbf{E}_2 \,:\, x \in [-2;-1], y \in [0;1]\}$$

I.1.17. Remark. We can prove that the complement of an open set in $\mathbf{E}_n$ is a closed set, and vice versa.

I.1.18. Boundary of a set. Let $M \subset \mathbf{E}_n$. A *boundary of M* is the name given to a set of all boundary points of M. The boundary of M is denoted by ∂M.

I.1.19. Examples.

The boundary of $\{X \in \mathbf{E}_2 \,:\, \rho(\mathcal{O},X) \leq 5\}$ *is* $\{X \in \mathbf{E}_2 \,:\, \rho(\mathcal{O},X) = 5\}$

The boundary of $\{X \in \mathbf{E}_2 \,:\, \rho(\mathcal{O},X) < 5\}$ *is* $\{X \in \mathbf{E}_2 \,:\, \rho(\mathcal{O},X) = 5\}$

The boundary of $\{[x,y] \in \mathbf{E}_2 \,:\, x \in (-2;-1)\}$ is $\{[x,y] \in \mathbf{E}_2 \,:\, x = -2 \vee x = -1\}$.

I.1.20. Remark. We can prove that the boundary of an arbitrary set M in $\mathbf{E}_n$ is a closed set.

I.1.21. Line segment in $\mathbf{E}_n$. Let $A, B \in \mathbf{E}_n$ and $A \neq B$. The set of points X such that $X = A + t(B - A)$, $t \in [0; 1]$ is called the *line segment in* $\mathbf{E}_n$ and it is denoted it by $\overline{AB}$.

I.1.22. Remark. The formula $X = A + t(B - A)$ should be understood in the sense of n-dimensional arithmetic space as $X = A + t \cdot (B + (-1) \cdot A)$.

I.1.23. Polygonal line in $\mathbf{E}_n$. Let $A_1,\ A_2, \ldots\ A_r \in \mathbf{E}_n$, r be a natural number $r \geq 2$ and $A_i \neq A_{i+1}, i = 1, 2, \ldots, r - 1$. The union of the line segments

$$\overline{A_1 A_2} \cup \overline{A_2 A_3} \cup \ldots \cup \overline{A_{r-1} A_r}$$

is called a *polygonal line connecting points* A_1, A_r.

I.1.24. Domain in $\mathbf{E}_n$. Let D be an arbitrary open set in $\mathbf{E}_n$. If for an arbitrary pair of points of D there exists a polygonal line connecting these two points and entirely belonging to D, then D is called a *domain in* $\mathbf{E}_n$.

I.1.25. Examples.
The set $\{X \in \mathbf{E}_2 \,:\, \rho(\mathcal{O}, X) < 5\}$ is a domain in $\mathbf{E}_2$.
The set $\{X \in \mathbf{E}_2 \,:\, \rho(\mathcal{O}, X) \leq 5\}$ is not a domain in $\mathbf{E}_2$.
The set $\{X \in \mathbf{E}_2 \,:\, \rho(\mathcal{O}, X) < 5\} \cup \{X \in \mathbf{E}_2 \,:\, \rho([-10, 0], X) < 5\}$ is not a domain in $\mathbf{E}_2$.
The set $\{X \in \mathbf{E}_2 \,:\, \rho(\mathcal{O}, X) < 5\} \cup \{[5, 0]\}$ is not a domain in $\mathbf{E}_2$.

I.1.26. Bounded set in $\mathbf{E}_n$. A subset M of E_n is called *bounded* if there exists $r > 0$ such that $\forall X \in M : \rho(\mathcal{O}, X) \leq r$.

I.1.27. Examples.
The set $\{[x, y] \in \mathbf{E}_2 \,:\, x \in (-2; -1)\}$ is not bounded.
The set $M = \{[x, y] \in \mathbf{E}_2 \,:\, x \in (-2; -1),\ y \in [0; 1]\}$ is bounded, because it holds for instance $\forall X \in M : \rho(\mathcal{O}, X) \leq 3$.

I.1.28. Limit of a sequence in $\mathbf{E}_n$. The element $A \in \mathbf{E}_n$ is called the *limit of a sequence* $\{A^{(i)}\}$, $A^{(i)} \in \mathbf{E}_n$, for $i = 1, 2, \ldots$ if

$$\forall\, U_\varepsilon(A) \ \ \exists\, n_0 \in \mathbf{N} \ \ \forall\, i \in \mathbf{N} : \quad (i \geq n_0) \Rightarrow A^{(i)} \in U_\varepsilon(A).$$

(We read it: For every neighbourhood $U_\varepsilon(A)$ of the point A there exists $n_0 \in \mathbf{N}$ so that for all $i \in \mathbf{N}$ it holds: If $i \geq n_0$, then $A^{(i)} \in U_\varepsilon(A)$. The fact that A is the limit of the sequence $\left\{A^{(i)}\right\}$ is written down in this way: $\lim A^{(i)} = A$ or $A^{(i)} \to A$. We also say that the sequence $\left\{A^{(i)}\right\}$ is convergent or the sequence $\left\{A^{(i)}\right\}$ converges to point A.

I.1.29. Remark. The definition of the limit of a sequence in $\mathbf{E}_n$ uses the notion of a neighbourhood of a point in $\mathbf{E}_n$ but formally it is very similar to the definition of a limit of a sequence in $\mathbf{R}^*$, see [4], III.1.4.

On the other hand, the definition of the limit can be overwritten in the following way:

$$\forall \varepsilon > 0 \;\; \exists n_0 \in \mathbf{N} \;\; \forall i \in \mathbf{N} : \quad (i \geq n_0) \Rightarrow \rho(A^{(i)}, A) < \varepsilon$$

This means that the limit of a sequence in $\mathbf{E}_n$ can also be defined in another way: The element $A \in \mathbf{E}_n$ is called the limit of a sequence $\left\{A^{(i)}\right\}$, $A^{(i)} \in \mathbf{E}_n$, for $i = 1, 2, ...$ if the sequence of real numbers $\left\{\rho(A^{(i)}, A)\right\}$ converges to $0 \in \mathbf{R}$.

I.1.30. Theorem. *Every sequence in $\mathbf{E}_n$ has at most one limit.*

The proof is analogous to the proof of Theorem III.1.8, see [4]. The neighbourhoods in $\mathbf{R}^*$ must be replaced by the neighbourhoods in $\mathbf{E}_n$, but the scheme of the proof is the same.

I.1.31. Remark. Note that $A^{(i)} = [a_1^{(i)}, a_2^{(i)}, ..., a_n^{(i)}]$. The question is what is the relation between the convergency of the sequence $\left\{A^{(i)}\right\}$ and the convergency of sequences $\left\{a_1^{(i)}\right\}$, $\left\{a_2^{(i)}\right\}$, ... $\left\{a_n^{(i)}\right\}$. The next theorem shows that this relation is very natural.

I.1.32. Theorem. *The sequence $\left\{A^{(i)}\right\}$, $A^{(i)} = [a_1^{(i)}, a_2^{(i)}, ..., a_n^{(i)}] \in \mathbf{E}_n$, for $i = 1, 2, ...$ converges to the point $A = [a_1, a_2, ... a_n] \in \mathbf{E}_n$ if and only if every sequence $\left\{a_r^{(i)}\right\}$, converges to the number a_r, for $r = 1, 2, ..., n$.*

I.1.33. Remark. The proof of this theorem is based on inequality:

$$|a_r^{(i)} - a_r| \leq \rho(A^{(i)}, A) \leq \sqrt{n} \cdot \max_{s \in \{1,2,\ldots,n\}} |a_s^{(i)} - a_s| \qquad \text{for } i = 1, 2, ...$$

I.1.34. Example. Find the limit of the sequence $\{X^{(k)}\}$ in $\mathbf{E}_3$ where

$$X^{(k)} = \left[\frac{\sin(k)}{k}, \frac{k^2 - 7k}{6 - 5k - 2k^2}, \frac{4}{k}\right].$$

Solution: First, we find the limits of each coordinate:

$$\lim \frac{\sin(k)}{k} = 0, \quad \lim \frac{k^2 - 7k}{6 - 5k - 2k^2} = -\frac{1}{2}, \quad \lim \frac{4}{k} = 0$$

Further, using Theorem $I.1.32$ we get $\lim X^{(k)} = [0, -\frac{1}{2}, 0]$.

I.2. Real functions of several real variables.

I.2.1. Real functions of n real variables – the definition. If $M \subset \mathbf{E}_n$, $n \in \mathbf{N}$, then each mapping of M to $\mathbf{E}_1$ is called a <u>*real function of n real variables*</u> (shortly: a <u>*function*</u>).

I.2.2. Domain of definition, range, graph. A function is a special case of a mapping and the notions "domain of definition of a mapping" and a "range of mapping" are known from secondary school. Hence, the notions "domain of definition of a function" (shortly: domain of a function) and "range of a function" (shortly: range of a function) can be regarded as known. In accordance with the notation which is used in connection with general mappings, $D(f)$ will be the domain of definition and $R(f)$ will be range of function f.

A *graph* of function f of variables $x_1, x_2, ..., x_n$ is the set

$$G(f) = \{[X, f(X)] \in \mathbf{R}^{n+1} \,:\, X = [x_1, x_2, ..., x_n] \in D(f)\}$$

I.2.3. Remark. Let f be a function of n variables and let $X = [x_1, x_2, ..., x_n]$ belong to its domain. The value of function f is denoted by $f(X)$ or by $f(x_1, x_2, ..., x_n)$.

I.2.4. Example.

Let f be a function of two variables x, y, which is defined for all $[x, y] \in D(f) = [3; +\infty) \times \mathbf{R}$ by its function value: $f(x, y) = \sqrt{x-3}$.

Let g be a function of a single variable x, which is defined for all $x \in D(g) = [3; +\infty)$ by its function value: $g(x) = \sqrt{x-3}$.

Although the values of functions are defined in both cases by the same formula, f and g are different functions with different domains of definition $D(f) \subset \mathbf{E}_2$, $D(g) \subset \mathbf{E}_1$.

I.2.5. Operations with functions. Let f, g be functions of variables $x_1, x_2, ..., x_n$, $D(f), D(g) \subset \mathbf{E}_n$. A *sum of functions* f and g is a function h such that $h(X) = f(X) + g(X)$ for $X = [x_1, x_2, ..., x_n] \in D(f) \cap D(g)$. Thus $D(h) = D(f) \cap D(g)$. We use the notation $h = f + g$.

We define a *difference of functions* and a *product of functions* f and g by analogy. A *quotient of functions* f and g can also be defined similarly – however, its domain is the set $[D(f) \cap D(g)] - \{X \in D(g) \,:\, g(X) = 0\}$.

I.2.6. Restriction of a function. Suppose that f is a function and $M \subset D(f)$. A function which is defined only on M and which assigns to each $X \in M$ the same value as function f is called the *restriction* of function f to set M, and it is denoted by $f|_M$. The set of all function values of function f on set M can be denoted by $R(f|_M)$ or by $f(M)$.

I.2.7. Composite function. We assume that function f of n variables $y_1, y_2, ..., y_n$ is defined for each $Y = [y_1, y_2, ...y_n] \in D \subset \mathbf{E}_n$ and functions $\phi_1, \phi_2, ..., \phi_n$ of m variables $x_1, x_2, ..., x_m$ are defined for each $X = [x_1, x_2, ...x_m] \in \Omega \subset \mathbf{E}_m$. Let

$$[\phi_1(X), \phi_2(X), ..., \phi_n(X)] \in D \quad \text{for } X \in \Omega.$$

Then the function

$$F(X) = f(\phi_1(X), \phi_2(X), ..., \phi_n(X))$$

defined for each $X \in \Omega$ is called a *composite* function.

I.2.8. Remark. We denote as ϕ the mapping defined by

$$\phi(X) = [\phi_1(X), \phi_2(X), ..., \phi_n(X)] \text{ for } X \in \Omega.$$

The mapping ϕ is called a *vector valued function of m variables*, with $D(\phi) \subset \mathbf{E}_m$ and $R(\phi) \subset \mathbf{E}_n$.

The fact that F is defined as a composite function by the relation

$$F(X) = f(\phi_1(X), \phi_2(X), ..., \phi_n(X)) \text{ for } X \in \Omega$$

is denoted

$$F = f \circ \phi.$$

I.2.9. A bounded function. Function f is called *bounded above* (or *upper bounded*) if there exists a number $K \in \mathbf{R}$ such that $\forall X \in D(f) : f(X) \leq K$. We can by analogy define the function *bounded below* (or *lower bounded*). Function f is called *bounded* if it is bounded above and bounded below.

Assume further that $M \subset D(f)$. Function f is called *bounded above on set M* if there exists a number $K \in \mathbf{R}$ such that $\forall X \in M : f(X) \leq K$. We can similarly define the function *bounded below on set M* and the notion of a function *bounded on set M*.

I.2.10. Extremes of a function. We say that function f has its *maximum* at the point $A \in D(f)$ if $\forall X \in D(f) \ : \ f(A) \geq f(X)$. We write:

$$\max f = f(A).$$

Analogously, we can define the *minimum* of a function f. We denote it $\min f$.

Suppose that $M \subset D(f)$. We say, that function f has its *maximum on set M at point $A \in M$* if $\forall X \in M \ : \ f(A) \geq f(X)$. We write: $\max_M f = f(A)$. Other often used notations of the maximum of function f on a set M are

$$\max_{M} f, \ \max_{X \in M} f(X).$$

Analogously, we can also define the *minimum of function f on set M*. We denote it:

$$\min\nolimits_M f, \ \min_{M} f, \ \min_{X \in M} f(X).$$

The maxima and minima of function f are called the *extremes* (or *extrema*) of function f.

The maxima and minima of function f on a set are called the *extremes on a set* of function f.

I.2.11. Remark. Obviously, the extremes of function f are special cases of the extremes of f on a set.

I.3. Limits and continuity.

I.3.1. Limit of a function. Assume that $C \in \mathbf{E}_n$, $y \in \mathbf{R}^*$ and the domain of definition of a function f contains some reduced neighbourhood $R(C)$ of point C. If for each sequence $\left\{X^{(i)}\right\}$ in $R(C)$ the implication

$$\left\{X^{(i)}\right\} \to C \Rightarrow f(X^{(i)}) \to y$$

is true, then we say that function f has the *limit at point C* equal to y. We write

$$\lim_{X \to C} = y.$$

Assume that $y \in \mathbf{R}^*$ and the domain of definition of a function f contains the following set:

$$R_r(\infty) = \{X \in \mathbf{E}_n,\ |X| > r\} \quad \text{for some } r > 0$$

If for each sequence $\left\{X^{(i)}\right\}$ in $R_r(\infty)$ the implication

$$\left\{|X^{(i)}|\right\} \to +\infty \Rightarrow f(X^{(i)}) \to y$$

is true, then we say that function f has the *limit at infinity* equal to y. We write

$$\lim_{|X| \to \infty} = y.$$

Further, we generalize the definition of the limit of a function at a point in order to be able to define a limit not only at point X for which there exists a reduced neighbourhood $R(X)$ such that $R(X) \subset D(f)$.

I.3.2. Limit of a function with respect to a set. Assume that the domain of definition of function f contains some $M \subset \mathbf{E}_n$, $C \in \mathbf{E}_n$ is an accumulation point of M, and $y \in \mathbf{R}^*$. If for each sequence $\left\{X^{(i)}\right\}$ in M the implication

$$\left\{X^{(i)}\right\} \to C \Rightarrow f(X^{(i)}) \to y$$

is true, then we say that function f has the *limit at point C with respect to set M* equal to y. We write

$$\lim_{\substack{X \in M \\ X \to C}} f(X) = y.$$

Assume that the domain of definition of function f contains some $M \subset \mathbf{E}_n$, such that each set

$$R_r(\infty) = \{X \in \mathbf{E}_n,\ |X| > r\} \quad r > 0$$

contains at least one point of M, and assume that $y \in \mathbf{R}^*$. If for each sequence $\left\{X^{(i)}\right\}$ in M the implication

$$\left\{|X^{(i)}|\right\} \to +\infty \Rightarrow f(X^{(i)}) \to y$$

is true, then we say that function f has the *limit at infinity with respect to set M* equal to y. We write

$$\lim_{\substack{X \in M \\ |X| \to +\infty}} = y.$$

I.3.3. Remark.

From the definition it follows that if there exists $\lim\limits_{X\to C} f(X)$ then there exists $\lim\limits_{\substack{X\in M\\ X\to C}} f(X)$ for each $M \subset D(f)$ with accumulation point C and

$$\lim_{\substack{X\in M\\ X\to C}} f(X) = \lim_{X\to C} f(X).$$

The next theorem is an easy consequence of Theorem III.1.8, see [4].

I.3.4. Theorem.

Function f can have at any point $C \in \mathbf{E}_n$ at most one limit.

Function f can have at infinity at most one limit.

Function f can have at any point $C \in \mathbf{E}_n$ at most one limit with respect to a set M.

Function f can have at infinity at most one limit with respect to a set M.

I.3.5. Example. Let f be a function of two variables defined by

$$f(x,y) = \begin{cases} 1 & \text{if } x\cdot y = 0 \\ 0 & \text{if } x\cdot y \neq 0 \end{cases} \quad \text{for } [x,y] \in \mathbf{E}_2$$

Prove that the function has no limit at point $[0,0]$.

Solution: Assume the sequence $\left\{\left[\frac{1}{n},\frac{1}{n}\right]\right\}$. This sequence converges to point $[0,0]$, see I.1.32. The function value $f\left(\frac{1}{n},\frac{1}{n}\right) = 0$ for all $n = 1,2,...,$ hence $f\left(\frac{1}{n},\frac{1}{n}\right) \to 0$.

Assume the sequence $\left\{\left[\frac{1}{n},0\right]\right\}$. This sequence also converges to point $[0,0]$. The function value $f\left(\frac{1}{n},0\right) = 1$ for all $n = 1,2,...,$ hence $f\left(\frac{1}{n},0\right) \to 1$.
Thus, the function has no limit at point $[0,0]$.

I.3.6. Example. Prove that function f from the previous example has a limit at point $[0,0]$ with respect to the subset $M = \{[x,y] \in \mathbf{E}_2 : x > 0 \wedge y > 0\}$.

Solution: We can see that $[0,0]$ is an accumulation point of M. It is clear from the definition of f that for all points $X \in M$ we get $f(X) = 0$. Hence, for *every* sequence $\{X^{(i)}\}$, $X^{(i)} \in M$ it holds $f(X^{(i)}) \to 0$. Thus, the function f has limit 0 at point $[0,0]$ with respect to set M.

The following theorem is fully analogous with the theorem about the limit of the functions of one real variable. It concerns the limit of the sum, difference, product and quotient of two functions of several real variables, and it can easily be proved by means of Theorem III.2.13, see [4].

We use the symbol "#," which has the meaning of any of the symbols "$+$", "$-$", "$\cdot$", "$/$" here.

I.3.7. Theorem. *Let $C \in \mathbf{E}_n$, $a, b \in \mathbf{R}^*$. Let* $\lim\limits_{X \to C} f(X) = a$, $\lim\limits_{X \to C} g(X) = b$.

Then $\lim\limits_{X \to C} [f(X) \# g(X)] = a \# b$ *(if the expression has a sense).*

Let $a, b \in \mathbf{R}^*$. Let $\lim\limits_{|X| \to +\infty} f(X) = a$, $\lim\limits_{|X| \to +\infty} g(X) = b$. *Then*

$$\lim_{|X| \to +\infty} [f(X) \# g(X)] = a \# b \quad \textit{(if the expression has a sense).}$$

I.3.8. Continuity of a function at a point. We say that function f is *continuous at the point* $C \in D(f) \subset \mathbf{E}_n$ if

$$\lim_{X \to C} f(X) = f(C).$$

I.3.9. Continuity of a function. We say that function f is continuous if f is continuous at each point $C \in D(f)$.

I.3.10. Remark. If you read the definition $I.3.8$ and definition $I.3.1$ carefully, you will see that function f can be continuous at point C only if it is defined in some neighbourhood of C (i.e. if $D(f)$ contains some neighbourhood $U(C)$. But this condition is satisfied for each point of $D(f)$ only if $D(f)$ is an open set. We will now study a more general situation.

I.3.11. Continuity of a function at a point with respect to a set. Let $M \subset D(f) \subset \mathbf{E}_n$ and C be an accumulation point of M. We say that function f is *continuous at point C with respect to set M* if

$$\lim_{\substack{X \in M \\ X \to C}} f(X) = f(C).$$

Let C be an isolated point of M. Then we also say that function f is continuous at point C with respect to set M.

I.3.12. Continuity of a function on a set. Let $M \subset D(f) \subset \mathbf{E}_n$. We say that function f is *continuous on set M* if f is continuous at each point $C \in M$ with respect to set M.

I.3.13. Remark. You can see from the definition that if a function f is continuous on a set M and $M_1 \subset M$ then function f is continuous on set M_1.

I.3.14. Theorem (on continuity of the sum, difference, product, quotient, and absolute value). *If functions f and g are continuous at point C, then also the functions $f + g$, $f - g$, $f \cdot g$, and $|f|$ are continuous at C. If, in addition $g(C) \neq 0$ then the function f/g is also continuous at point C.*

(This part of the theorem is also valid when we replace "continuity at point C" by "continuity at point C with respect to set M".)

If functions f and g are continuous on set M, then also the functions $f+g$, $f-g$, $f\cdot g$, and $|f|$ are continuous on set M. If, in addition $g(X)\neq 0$ for all $X\in M$ then the function f/g is also continuous on set M.

I.3.15. Example. Let f be a function of two variables defined by

$$g(x,y)=\begin{cases}\dfrac{2xy}{x^2+y^2} & \text{if } [x,y]\neq[0,0]\\ 0 & \text{if } [x,y]=[0,0]\end{cases} \quad \text{for } [x,y]\in\mathbf{E}_2$$

a) Prove that function g is continuous on $\mathbf{E}_2-[0,0]$.

b) Prove also that function g is continuous on axis x, i.e. on the set $\{[x,y]\in\mathbf{E}_2 : x\in\mathbf{R},\ y=0\}$.

Solution:

a) From the previous theorem it follows that function g is continuous at any point $[x,y]\neq[0,0]$. At point $[0,0]$ the value of function g is defined, but g, we claim, has no limit at point $[0,0]$. We show the proof by contradiction. We suppose that there exists $\lim\limits_{X\to[0,0]} g(X)$. We denote $M=\{[x,y]\in\mathbf{E}_2 : y=mx\}$. Then for every M we have

$$\lim_{\substack{X\in M\\ X\to[0,0]}} g(X)=\lim_{X\to[0,0]} g(X).$$

We calculate the limit at point $[0,0]$ with respect to each defined set M:

$$\lim_{\substack{X\in M\\ X\to[0,0]}} g(X)=\lim_{\substack{y=mx\\ [x,y]\to[0,0]}}\frac{2xy}{x^2+y^2}=\lim_{[x,mx]\to[0,0]}\frac{2xy}{x^2+y^2}=$$

$$=\lim_{x\to0}\frac{2xmx}{x^2+m^2x^2}=\frac{2m}{1+m^2}.$$

Because the value depends on m, we have a contradiction, see Remark $I.3.3$. Hence, the limit of g at $[0,0]$ does not exist, and function g is not continuous at $[0,0]$.

b) The set $\{[x,y]\in\mathbf{E}_2 : x\in\mathbf{R}\wedge y=0\}$ is a set M from a) with $m=0$. From the same calculations (where now $m=0$) we get

$$\lim_{\substack{X\in M\\ X\to[0,0]}} g(X)=0.$$

We have g(0,0)=0, so g is continuous at $[0,0]$ with respect to $\{[x,y] : x\in\mathbf{R}\wedge y=0\}$. (In other points is g continuous.) Hence, g is continuous on axis x.

I.3.16. Theorem (on continuity of a composite function). *We assume that function f is continuous at point $B=[b_1,b_2,...b_n]$, functions $\phi_1,\phi_2,...,\ \phi_n$ are continuous at point $A=[a_1,a_2,...a_m]$ and $B=[\phi_1(A),\phi_2(A),...,\phi_n(A)]$. Let us denote*

$\phi = [\phi_1, \phi_2, ..., \phi_n]$, *i.e.* ϕ *is a vector valued function defined by coordinate functions* $\phi_1, \phi_2, ..., \phi_n$. *Then composite function* $F = f \circ \phi$ *is continuous at point* A.

We assume now that function f *is continuous on set* D, *functions* $\phi_1, \phi_2, ..., \phi_n$ *are continuous on set* Ω *and* $\phi(X) = [\phi_1(X), \phi_2(X), ..., \phi_n(X)] \in D$ *for* $X \in \Omega$. *Then the composite function* $F = f \circ \phi$ *is continuous on set* Ω.

I.3.17. Theorem (Darboux' property). *If function* f *is continuous on a domain* M *and* A, B *are any two points from* M, *then to any given number* Y *between* $f(A)$ *and* $f(B)$ *and to any polygonal line* $L \subset M$ *connecting* A, B *there exists a point* $X \in L$ *such that* $f(X) = Y$

I.3.18. Theorem. *Function* f, *which is continuous on a bounded closed set* M, *has its maximum and minimum on this set M. (Thus,* $\max_{X \in M} f(X)$ *and* $\min_{X \in M} f(X)$ *exist.)*

I.4. Partial derivatives, differentials.

When we hold all but one of the independent variables constant and derive with respect to that one variable, we get a partial derivative. For example, the partial derivative of a function $f(x, y)$ with respect to x at point $[x_0, y_0]$ is the value of the derivative of the function of one real variable $f(x, y_0)$ at point x_0.

I.4.1. Partial derivative of a function. Let f be a function of n real variables $x_1, x_2, ..., x_n$ and $A = [a_1, a_2, ..., a_n] \in \mathbf{E}_n$. If there exists a finite limit

$$\lim_{h \to 0} \frac{f(a_1, a_2, ..., a_{k-1}, a_k + h, a_{k+1}, ..., a_n) - f(a_1, a_2, ..., a_{k-1}, a_k, a_{k+1}, ..., a_n)}{h}$$

then its value is called the *partial derivative of* f *with respect to* x_k *at point* A. It is denoted by

$$\frac{\partial f}{\partial x_k}(A) \quad or \quad \left.\frac{\partial f}{\partial x_k}\right|_A .$$

Let us assume the set of all points A for which $\frac{\partial f}{\partial x_k}(A)$ exists. The function defined by its function value $\frac{\partial f}{\partial x_k}(A)$ in this set is called the *partial derivative of* f *with respect to* x_k. This function is denoted by

$$\frac{\partial f}{\partial x_k}.$$

From the definition it follows that

$$D\left(\frac{\partial f}{\partial x_k}\right) \subset D(f).$$

I.4.2. Remark. Let g be a function of one real variable x_k defined in the following way:

$$g(x_k) = f(a_1, a_2, ..., a_{k-1}, x_k, a_{k+1}, ..., a_n)$$

From the definition it follows that the partial derivative of f with respect to x_k at the point $A = [a_1, a_2, ..., a_n]$ is defined as the derivative of function g at point a_k. Hence, to calculate the partial derivative with respect to x_k we assume other variables to be constant and calculate the derivative of a function of one variable x_k. Thus, all theorems about calculation of derivatives also hold for partial derivatives.

Let now $f : f(x, y)$ then $\frac{\partial f}{\partial x}(a, b) = \frac{\mathrm{d}(f(x,b))}{\mathrm{d}x}\Big|_a = \tan\alpha$, see Fig. 3:

I.4.3. Example. We calculate the partial derivatives of function f of three variables x, y, z given by the formula $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$, $[x, y, z] \in \mathbf{E}_3$. Deriving the expression with respect to x we regard y and z as constants and we get (using the formula about the derivatives of composite functions of one variable):

$$\frac{\partial f}{\partial x}(x, y, z) = \frac{1}{2}\frac{1}{\sqrt{x^2 + y^2 + z^2}}2x = \frac{x}{\sqrt{x^2 + y^2 + z^2}}.$$

Analogously, we get

$$\frac{\partial f}{\partial y}(x, y, z) = \frac{y}{\sqrt{x^2 + y^2 + z^2}}, \quad \frac{\partial f}{\partial z}(x, y, z) = \frac{z}{\sqrt{x^2 + y^2 + z^2}}.$$

$D(f) = \mathbf{E}_3, \quad D(\frac{\partial f}{\partial x}) = D(\frac{\partial f}{\partial y}) = D(\frac{\partial f}{\partial z}) = \mathbf{E}_3 - \{[0, 0, 0]\}.$

I.4.4. Remark. Let us suppose that the function $f(x, y)$ has partial derivatives $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ at point $[x_0, y_0]$. What can we say about the behaviour of the function in the neighbourhood of point $[x_0, y_0]$? For example, is this function continuous at this point?

I.4.5. Example.

$$f(x,y) = \begin{cases} 1 & \text{if } x \cdot y = 0 \\ 0 & \text{if } x \cdot y \neq 0 \end{cases}$$

The partial derivatives at point $[0,0]$ exist, but the function is not continuous at [0,0], see $I.3.5$. (Because $f(x,0) = f(0,y) = 1$ for $x, y \in \mathbf{R}$, we get $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0$.)

In the next paragraph we will solve the question of the functions that can be well approximated by a linear function in the neighbourhood of some point.

I.4.6. Differentials. Let function f be defined in a neighbourhood $U(A)$ of point $A = [a_1, a_2, ..., a_n] \in \mathbf{E}_n$ and for every $X \in U(A)$ let the following relation be satisfied:

$$f(X) - f(A) = [\alpha_1(x_1 - a_1) + \alpha_2(x_2 - a_2) + ... + \alpha_n(x_n - a_n)] + \varepsilon(X) \qquad (I.4.1.)$$

where $\alpha_1, \alpha_2, ..., \alpha_n$ are some real numbers, $\varepsilon(X)$ is a function continuous at point A, $\varepsilon(A) = 0$, and

$$\lim_{X \to A} \frac{\varepsilon(X)}{\rho(X,A)} = 0. \qquad (I.4.2.)$$

Then the function is called *differentiable* at point A and the linear expression

$$[\alpha_1(x_1 - a_1) + \alpha_2(x_2 - a_2) + ... + \alpha_n(x_n - a_n)]$$

is called the *total differential of function f at point A*, and is denoted by $df(A)$.

I.4.7. Remark. If the function is differentiable at point A, it follows from the definition that it must be defined in some neighbourhood of this point.

Relation $(I.4.1)$ means that the function value $f(X)$ can be approximated by the linear function

$$f(A) + df(A),$$

$$\text{i.e.} \quad f(A) + [\alpha_1(x_1 - a_1) + \alpha_2(x_2 - a_2) + ... + \alpha_n(x_n - a_n)]\,. \qquad (I.4.3.)$$

The "error" function of this approximation equals $\varepsilon(X)$. From the relation

$$\lim_{X \to A} \frac{\varepsilon(X)}{\rho(X,A)} = 0$$

it follows that this "error" is essentially less than the distance between X and A.

I.4.8. Geometrical meaning. If $n = 2$ then the graph of $f(X)$ in the neighbourhood $U(A)$ is a surface in $\mathbf{E}_3$ which contains the point $[A, f(A)]$. The graph of function $(I.4.3)$ is a plane which contains the point $[A, f(A)]$. Relations $(I.4.1)$, $(I.4.2)$ mean that the plane is the tangent plane to the graph of function $f(X)$ at the point $[A, f(A)]$.

We will formulate two theorems which state the relation between partial derivatives and the differentiability of a function. Differentiability is an important property of a function, being the condition in a number of theorems.

I.4.9. Theorem. Necessary condition of differentiability at a point. *If function $f(X)$ is differentiable at point A then f is continuous at A, and there exist partial derivatives at this point*

$$\frac{\partial f}{\partial x_1}(A), \frac{\partial f}{\partial x_2}(A), ..., \frac{\partial f}{\partial x_n}(A)$$

and the constants from the definition of a differential $\alpha_1, \alpha_2, ..., \alpha_n$ are equal to these derivatives, i.e.

$$df(A) = \frac{\partial f}{\partial x_1}(A)(x_1 - a_1) + \frac{\partial f}{\partial x_2}(A)(x_2 - a_2) + ... + \frac{\partial f}{\partial x_n}(A)(x_n - a_n).$$

I.4.10. Theorem. Sufficient condition of differentiability at a point. *If the function f of n variables $x_1, x_2, ..., x_n$ has its partial derivatives*

$$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_n}$$

in a neighbourhood $U(A)$ and the derivatives all are continuous at point A, then function f is differentiable at point A.

The next assertion is an easy consequence of the previous theorem.

I.4.11. Theorem. Sufficient condition of differentiability on an open set. *If function f has partial derivatives*

$$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_n}$$

in an open set M which all are continuous on this set M, then function f is differentiable at every point of M.

I.4.12. Example.

We find the equation of the tangent plane to the graph of function f given by the formula $f(x, y) = x^2 + y^2 + 2x - y - 7$ at the point $T = [A, f(A)],\ A = [3, 4]$.
Solution: We easily get $f(A) = 3^2 + 4^2 + 6 - 4 - 7 = 20$. There exist partial derivatives $\frac{\partial f}{\partial x}(x, y) = 2x + 2$, $\frac{\partial f}{\partial y}(x, y) = 2y - 1$. We can see that $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ are defined and continuous on $\mathbf{E}_2$, and therefore on some neighbourhood of $A = [3, 4]$. Hence, f is differentiable at A and

$$df(A) \ = \ \frac{\partial f}{\partial x}\bigg|_A (x - a_1) + \frac{\partial f}{\partial y}\bigg|_A (y - a_2) \ = \ 8(x - 3) + 7(y - 4).$$

The tangent plane is the graph of the function $f(A) + df(A)$. The equation of the graph of this function is:

$$z = f(A) + df(A) \quad \text{i.e.} \quad z = f(A) + \left.\frac{\partial f}{\partial x}\right|_A (x - a_1) + \left.\frac{\partial f}{\partial y}\right|_A (y - a_2).$$

Hence we get:

$$z = 20 + 8(x-3) + 7(y-4)$$

We can calculate the partial derivatives of composite functions of several variables by using the so called Chain rule. You already know the analogy of this rule for functions of one variable. Compare the two formulations.

I.4.13. Derivatives of composite functions - Chain rule. If functions $\phi_1(X)$, $\phi_2(X), ..., \phi_n(X)$ are differentiable at the point $A = [a_1, a_2, ..., a_m]$ and function f is differentiable at the point $B = [\phi_1(A), \phi_2(A), ..., \phi_n(A)]$ then the composite function F defined by the formula

$$F(X) = f(\phi_1(X), \phi_2(X), ..., \phi_n(X))$$

in some neighbourhood of point A is differentiable at this point and for $k = 1, 2, ..., m$

$$\frac{\partial F}{\partial x_k}(A) = \frac{\partial f}{\partial y_1}(B)\frac{\partial \phi_1}{\partial x_k}(A) + \frac{\partial f}{\partial y_2}(B)\frac{\partial \phi_2}{\partial x_k}(A) + ... + \frac{\partial f}{\partial y_n}(B)\frac{\partial \phi_n}{\partial x_k}(A). \qquad (I.4.4.)$$

I.4.14. Examples.

Let f be a function of two variables u, v and let ϕ_1, ϕ_2 be functions of two variables x, y. A composite function F is defined by its function value:

$$F(x,y) = f(\phi_1(x,y), \phi_2(x,y))$$

Find the expressions of $\frac{\partial F}{\partial x}$ and $\frac{\partial F}{\partial y}$. (Functions f, ϕ_1, ϕ_2 are assumed differentiable and $[\phi_1(x,y), \phi_2(x,y)] \in D(f)$ if $[x,y] \in D(\phi_1) \cap D(\phi_2)$.)
Solution: To simplify the expressions, let us denote as ϕ the vector valued function defined by its function value:

$$\phi(x,y) \equiv [\phi_1(x,y), \phi_2(x,y)] \quad \text{for} \quad [x,y] \in D(\phi) \equiv D(\phi_1 \cap \phi_2)$$

Using the Chain rule we get:

$$\frac{\partial F}{\partial x}(x,y) = \frac{\partial f}{\partial u}(\phi(x,y))\frac{\partial \phi_1}{\partial x}(x,y) + \frac{\partial f}{\partial v}(\phi(x,y))\frac{\partial \phi_2}{\partial x}(x,y)$$

$$\frac{\partial F}{\partial y}(x,y) = \frac{\partial f}{\partial u}(\phi(x,y))\frac{\partial \phi_1}{\partial y}(x,y) + \frac{\partial f}{\partial v}(\phi(x,y))\frac{\partial \phi_2}{\partial y}(x,y)$$

Shortly:

$$\frac{\partial F}{\partial x} = \frac{\partial f}{\partial u}\frac{\partial \phi_1}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial \phi_2}{\partial x}$$

$$\frac{\partial F}{\partial y} = \frac{\partial f}{\partial u}\frac{\partial \phi_1}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial \phi_2}{\partial y}$$

Let g be a function of two variables u, v and let ϕ, ψ be functions of one variable x, and $\phi(x) \equiv x$. A composite function G is defined by its function value:

$$G(x) = f(x, \psi(x))$$

Find the expression of $\frac{dG}{dx}$. (Functions g, ψ are assumed differentiable and $[x, \psi(x)] \in D(g)$ if $x \in D(\psi)$.)
Solution: Using the Chain rule we get:

$$\frac{dG}{dx}(x) = \frac{\partial f}{\partial u}(x, \psi(x)) \cdot 1 + \frac{\partial f}{\partial v}(x, \psi(x))\frac{d\psi}{dx}(x)$$

i.e.

$$G'(x) = \frac{\partial f}{\partial u}(x, \psi(x)) + \frac{\partial f}{\partial v}(x, \psi(x))\psi'(x)$$

or shortly:

$$G' = \frac{\partial f}{\partial u} + \frac{\partial f}{\partial v}\psi'$$

I.4.15. Higher order derivatives. Let the function f of n variables $x_1, x_2, ..., x_n$ have a partial derivative $\frac{\partial f}{\partial x_k}$ in subset M_1, $k \in \{1, 2, ..., n\}$. This partial derivative is also a function of n variables. If there exists a partial derivative of this function, i.e.

$$\frac{\partial \left(\frac{\partial f}{\partial x_k}\right)}{\partial x_l}, \quad l \in \{1, 2, ..., n\}$$

in some set $M_2 \subset M_1$ then it is called a <u>*partial derivative of the second order*</u>, and it is denoted by

$$\frac{\partial^2 f}{\partial x_l \partial x_k}, \qquad \frac{\partial^2 f}{\partial x_k^2} \quad (\text{if } k = l).$$

I.4.16. Example. Find the all partial derivatives of the first and second order of the function $f : f(x, y) = e^{xy^2}$.
Solution:

$$\frac{\partial f}{\partial x}(x, y) = e^{xy^2} y^2, \qquad \frac{\partial f}{\partial y}(x, y) = e^{xy^2} 2xy,$$

$$\frac{\partial^2 f}{\partial x^2}(x,y) = e^{xy^2} y^4, \qquad \frac{\partial^2 f}{\partial y^2}(x,y) = e^{xy^2} 2xy\, 2xy + e^{xy^2} 2x = 2xe^{xy^2}(2xy^2+1),$$

$$\frac{\partial^2 f}{\partial y \partial x}(x,y) = e^{xy^2} 2xy\, y^2 + e^{xy^2} 2y = 2ye^{xy^2}(xy^2+1),$$

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = e^{xy^2} y^2\, 2xy + e^{xy^2} 2y = 2ye^{xy^2}(xy^2+1)$$

Function f and all partial derivatives of the first and second order are defined and continuous in $\mathbf{E}_2$.

I.4.17. Remark. By analogy, we can define partial derivatives of the third order etc.

In the previous example we derived $\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$, but in general $\frac{\partial^2 f}{\partial x_l \partial x_k} \neq \frac{\partial^2 f}{\partial x_k \partial x_l}$. The next theorem states the sufficient conditions which ensure that partial derivatives differing by the order of differentiation define the same functions.

I4.18. Theorem. *Let function f have partial derivatives $\frac{\partial f}{\partial x_k}$, $\frac{\partial f}{\partial x_l}$, $k, l = 1, 2, ..., n$, $k \neq l$ in a neighbourhood $U(A)$. Let $\frac{\partial^2 f}{\partial x_l \partial x_k}$ be continuous at point A. Then there exists $\frac{\partial^2 f}{\partial x_k \partial x_l}(A)$ and*

$$\frac{\partial^2 f}{\partial x_k \partial x_l}(A) = \frac{\partial^2 f}{\partial x_l \partial x_k}(A).$$

I.5. Gradients, directional derivatives.

I.5.1. Gradients. If function f of n variables denoted by $x_1, x_2, ... x_n$ has all its partial derivatives at point A, then the vector

$$\left[\frac{\partial f}{\partial x_1}(A), \frac{\partial f}{\partial x_2}(A), ..., \frac{\partial f}{\partial x_n}(A)\right] \in \mathbf{E}_n$$

is called the underbar*gradient of function f at point A*, and it is denoted by

$$(\operatorname{grad} f)(A), \qquad (\nabla f)(A), \qquad (\operatorname{grad} f)|_A, \qquad \nabla f|_A.$$

If the gradient of a function f exists at points of some set M, the vector function given by the relation $\Phi(X) = (\operatorname{grad} f)(X)$, $X \in M$ is called the *gradient of a function f* and it is denoted

$$\operatorname{grad} f \quad \text{or} \quad \nabla f.$$

I.5.2. Directional derivatives. In this paragraph we will generalize the notion of a partial derivative.

Let f be a function of n real variables $x_1, x_2, ..., x_n$, $A = [a_1, a_2, ..., a_n]$ be some given point $A \in \mathbf{E}_n$ and $\vec{s} = (s_1, s_2, ..., s_n) \neq \mathcal{O}$ a given vector. We denote by $\vec{S}$

$$\vec{S} = (S_1, S_2, ...S_n) = \frac{\vec{s}}{|\vec{s}|} = \left(\frac{s_1}{|\vec{s}|}, \frac{s_2}{|\vec{s}|}, ..., \frac{s_n}{|\vec{s}|}\right)$$

If there exists the limit

$$\lim_{t\to 0} \frac{f(A + \vec{S}\,t) - f(A)}{t} = \lim_{t\to 0} \frac{f(a_1 + S_1 t, a_2 + S_2 t, ..., a_n + S_n t) - f(a_1, a_2, .., a_n)}{t}$$

it is called the *directional derivative of f in direction $\vec{s}$ at point A*, and it is denoted by

$$\frac{\partial f}{\partial \vec{s}}(A).$$

I.5.3. Remark. It is clear that this definition is a generalization of the notion of a partial derivative. Indeed, if we choose in this definition for instance $\vec{s} = (1, 0, 0, ..., 0)$ we will get the definition of $\frac{\partial f}{\partial x_1}(A)$.

I.5.4. Remark. It is also clear that this definition is identical with the definition of the derivative of the following function of one real variable t at point 0 :

$$f(a_1 + S_1 t, a_2 + S_2 t, ..., a_n + S_n t)$$

This is a composite function of function f and a vector function Φ, the value of which is defined by the formula

$$\Phi(t) = [a_1 + S_1 t, a_2 + S_2 t, ..., a_n + S_n t].$$

Assuming differentiability of f and using formula $(I.4.4)$ we get

$$\frac{\partial f}{\partial \vec{s}}(A) = \frac{d\Phi}{dt}(0) =$$

$$= \frac{\partial f}{\partial x_1}(A)\frac{d(a_1 + S_1\, t)}{dt}(0) + \frac{\partial f}{\partial x_2}(A)\frac{d(a_2 + S_2\, t)}{dt}(0) + ... + \frac{\partial f}{\partial x_n}(A)\frac{d(a_n + S_n\, t)}{dt}(0) =$$

$$= \frac{\partial f}{\partial x_1}(A)\, S_1 + \frac{\partial f}{\partial x_2}(A)\, S_2 + ... + \frac{\partial f}{\partial x_n}(A)\, S_n \;\; = \;\; \vec{S} \cdot (\mathrm{grad}\, f)(A).$$

Hence, assuming the differentiability of a function f at point A we derive a formula for the directional derivative at point A :

$$\frac{\partial f}{\partial \vec{s}}(A) = \frac{\vec{s} \cdot (\mathrm{grad}\, f)(A)}{|\vec{s}|}$$

I.5.5. Remark. Because $\vec{a} \cdot \vec{b} = |\vec{a}||\vec{b}| \cos\alpha$, where α is the angle between $\vec{a}$, $\vec{b}$ we get:

$$\frac{\partial f}{\partial \vec{s}}(A) = \frac{\vec{s} \cdot (\mathrm{grad}\, f)(A)}{|\vec{s}|} = |(\mathrm{grad}\, f)(A)| \; \cos\alpha$$

Thus, if the angle α between $(\mathrm{grad}\, f)(A)$ and $\vec{s}$ equals zero, the directional derivative is maximal, i.e. the gradient at a point is the direction in which the increment of the function (in a sufficiently small neighbourhood) is maximal.

I.5.6. Example. Find the directional derivative $\frac{\partial f}{\partial \vec{s}}(A)$ if $A \equiv [1, 2]$, $\vec{s} = (1, 1)$ and $f \,:\, f(x, y) = x^2 + xy$.

Solution: We define the unit vector $\vec{S}$:

$$|\vec{s}| = \sqrt{1^2 + 1^2} = \sqrt{2}, \qquad \vec{S} = \frac{\vec{s}}{|\vec{s}|} = \frac{(1, 1)}{\sqrt{2}} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

Further, we get expressions for the partial derivatives of f

$$\frac{\partial f}{\partial x}(x, y) = 2x + y, \qquad \frac{\partial f}{\partial y}(x, y) = x.$$

These partial derivatives are defined and continuous in $\mathbf{E}_2$. Thus, function f is differentiable at each point of $\mathbf{E}_2$, in particular at the point $A \equiv [1, 2]$. Hence,

$$\frac{\partial f}{\partial \vec{s}}(A) \; = \; (\mathrm{grad}\, f)|_A \cdot \vec{S} \; = \; \left.\frac{\partial f}{\partial x}\right|_A S_1 + \left.\frac{\partial f}{\partial y}\right|_A S_2 \; = \; 4\,\frac{1}{\sqrt{2}} + 1\,\frac{1}{\sqrt{2}} \; = \; \frac{5}{\sqrt{2}}.$$

I.6. Implicit functions.

I.6.1. Examples. Let us assume the equation $x^2 + y^2 = 1$ and the point $X_0 \equiv [x_0, y_0] \equiv [\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}]$. (Draw the sketch.) It is clear that the equation defines some function $f : y = f(x)$ in the neighborhood of the point $x_0 = \frac{\sqrt{2}}{2}$, such that $f(\frac{\sqrt{2}}{2}) = \frac{\sqrt{2}}{2}$. Indeed, we get from the equation $y = +\sqrt{1-x^2}$ or $y = -\sqrt{1-x^2}$. Taking into account the condition $f(\frac{\sqrt{2}}{2}) = \frac{\sqrt{2}}{2}$ we get $y = f(x) = +\sqrt{1-x^2}$. This function has the following properties:
$f(\frac{\sqrt{2}}{2}) = \frac{\sqrt{2}}{2}$
The function is defined in some neighbourhood of x_0, i.e. in the interval $(\frac{\sqrt{2}}{2} - \delta; \frac{\sqrt{2}}{2} + \delta)$, (here $\delta = 1 - \frac{\sqrt{2}}{2}$).
If we substitute $f(x)$ into the relation $x^2 + y^2 = 1$ we get an identity: $x^2 + (f(x))^2 = 1 \Rightarrow 1 = 1$.
There is at most one such function. (There is no such function for instance if we choose the point $A \equiv [1, 0]$ or $A \equiv [-1, 0]$).
The graph of the function f locally coincides with the "graph" of the equation, i.e. there exists $\delta > 0$, such that

$$\left\{[x,y] \in (\frac{\sqrt{2}}{2} - \delta; \frac{\sqrt{2}}{2} + \delta) \times (\frac{\sqrt{2}}{2} - \delta; \frac{\sqrt{2}}{2} + \delta) \, : \, y = \sqrt{1-x^2}\right\} =$$

$$= \left\{[x,y] \in (\frac{\sqrt{2}}{2} - \delta; \frac{\sqrt{2}}{2} + \delta) \times (\frac{\sqrt{2}}{2} - \delta; \frac{\sqrt{2}}{2} + \delta) \, : \, x^2 + y^2 = 1\right\}.$$

This section deals with the conditions that ensure the existence of such a function, even if we are not able to express it explicitly from the originally given equation.

Let us assume an another example. There is given the equation

$$e^x + x - 10 = y + \tan(y) \qquad (I.6.1.)$$

It is easy to see that $f(x) = e^x + x - 10$ is a continuous increasing function for $x \in (-\infty; +\infty)$, $R(f) = (-\infty; +\infty)$ and function $g(y) = y + \tan y$ is also a continuous increasing function on each interval $y \in (-\frac{\pi}{2} + k\pi; \frac{\pi}{2} + k\pi), k \in \mathbf{N}$, $R(g) = (-\infty; +\infty)$. Due to these properties of the functions f, g it is clear that for every $x \in (-\infty; +\infty)$ there exists the unique $y \in (-\frac{\pi}{2}; \frac{\pi}{2})$ such that equation $(I.6.1)$ is satisfied. Hence, by means of $(I.6.1)$ a function ϕ_0 is defined with the domain of definition $D(\phi_0) = (-\infty; +\infty)$ and the range $R(\phi_0) = (-\frac{\pi}{2}; \frac{\pi}{2})$. Since the function value of ϕ_0 is defined as a solution of some equation and the analytic expression of the function value is not known, the function is called an *implicit* function.

If we substitute function ϕ_0 into relation $(I.6.1)$ we get an identity:

$$\forall x \in (-\infty; +\infty) : e^x + x - 10 = \phi_0(x) + \tan(\phi_0(x)).$$

We can keep repeating $k \in \mathbf{N}$: For every $x \in (-\infty; +\infty)$ there exists the unique $y \in (-\frac{\pi}{2} + k\pi; \frac{\pi}{2} + k\pi)$ such that equation $(I.6.1)$ is satisfied. Hence, by means of this relation we can define a function ϕ_k with the domain of definition $D(\phi_k) = (-\infty; +\infty)$ and the range $R(\phi_k) = (-\frac{\pi}{2} + k\pi; \frac{\pi}{2} + k\pi)$.

If some point $[x_0, y_0]$ satisfying relation $(I.6.1)$ is given, then this relation defines a unique function ϕ_k, such that $y_0 = \phi_k(x_0)$.

Relation $(I.6.1)$) can be written in the form $F(x, y) = 0$. The following theorem states sufficient conditions which ensure that the relation $F(x, y) = 0$ defines an implicit function.

I.6.2. Theorem. *Let F be a function of two variables which are denoted x, y. We suppose that F and partial derivatives $\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}$ are continuous in some neighbourhood $U(A)$ of the point $A = [a, b]$. We assume that $F(A) = 0$ and $\frac{\partial F}{\partial y}(A) \neq 0$. Then there are $\delta > 0, \varepsilon > 0$ such that the unique function f is defined in a way that satisfies the following properties:*

a) $b = f(a)$

b) $\forall x \in (a - \delta; a + \delta) : \ f(x) \in (b - \varepsilon; b + \varepsilon)$ and $F(x, f(x)) = 0$.

c) f, f' are continuous in $(a - \delta; a + \delta)$

d) $\forall x \in (a - \delta; a + \delta)$

$$f'(x) = -\frac{\frac{\partial F}{\partial x}(x, f(x))}{\frac{\partial F}{\partial y}(x, f(x))} \qquad (I.6.2.)$$

Moreover, if all partial derivatives of F are continuous in a neighbourhood $U(A)$ up to the k-th order, then $f, f', ...f^{(k)}$ are continuous in $(a - \delta; a + \delta)$.

I.6.3. Remark. It is very simple to derive formula $(I.6.2.)$ from a) - c). Indeed, deriving the two sides of identity (see $(I.4.4)$

$$F(x, f(x)) = 0$$

we get

$$\frac{\partial F}{\partial x}(x, f(x)) + \frac{\partial F}{\partial y}(x, f(x)) \cdot f'(x) = 0. \qquad (I.6.3.)$$

(The left hand side is derived by means of the Chain rule for composite functions of several variables.) If we calculate $f'(x)$ from this relation we get $(I.6.2)$. Taking into account a) in Theorem $I.6.2$ we get

$$f'(a) = -\frac{\frac{\partial F}{\partial x}(A)}{\frac{\partial F}{\partial y}(A)}.$$

Deriving $(I.6.3)$ we get

$$\frac{\partial^2 F}{\partial x^2} + 2\frac{\partial^2 F}{\partial x \partial y} \cdot f'(x) + \frac{\partial^2 F}{\partial y^2} \cdot (f'(x))^2 + \frac{\partial F}{\partial y} \cdot f''(x) = 0. \qquad (I.6.4.)$$

From this relation we can express $f''(x)$.

We can calculate higher order derivatives of an implicit function in a similar way.

The next theorem states sufficient conditions which ensure that the relation $F(x, y, z) = 0$ defines an implicit function.

I.6.4. Theorem. *Let F be a function of three variables which are denoted x, y, z. We suppose that F and partial derivatives $\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z}$ are continuous in some neighbourhood $U(A)$ of point $A = [a, b, c]$. We assume that $F(A) = 0$ and $\frac{\partial F}{\partial z}(A) \neq 0$. Then there are $\delta > 0, \varepsilon > 0$ such that the unique function f is defined in a way that satisfies the following properties:*

a) $c = f(a, b)$

b) $\forall [x, y] \in (a-\delta; a+\delta) \times (b-\delta; b+\delta) \ : \ f(x, y) \in (c-\varepsilon; c+\varepsilon)$ *and* $F(x, y, f(x, y)) = 0$.

c) $f, \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ *are continuous in* $(a-\delta; a+\delta) \times (b-\delta; b+\delta)$

d) $\forall [x, y] \in (a-\delta; a+\delta) \times (b-\delta; b+\delta)$

$$\frac{\partial f}{\partial x}(x, y) = -\frac{\frac{\partial F}{\partial x}(x, y, f(x, y))}{\frac{\partial F}{\partial z}(x, y, f(x, y))}, \qquad \frac{\partial f}{\partial y}(x, y) = -\frac{\frac{\partial F}{\partial y}(x, y, f(x, y))}{\frac{\partial F}{\partial z}(x, y, f(x, y))} \qquad (I.6.5.)$$

Moreover, if all partial derivatives of F are continuous in a neighbourhood $U(A)$ up to the k-th order, then all partial derivatives of f are continuous in $(a-\delta; a+\delta) \times (b-\delta; b+\delta)$.

I.6.5. Example. Prove that the equation $F(x, y, z) \equiv z^3 - xy + yz + y^3 - 2 = 0$ in some neighbourhood of the point $A \equiv [1, 1, 1]$ defines function f such that $F(x, y, f(x, y)) = 0$ in some neighbourhood of point $[1, 1]$, and calculate the partial derivatives at this point.

Solution: We use the previous theorem. Function $F(x, y, z) \equiv z^3 - xy + yz + y^3 - 2$ is polynomial, so it is defined and continuous in $\mathbf{E}_3$ and all (first order) partial derivatives are also defined and continuous in $\mathbf{E}_3$. Substituting A into the equation we get $F(1, 1, 1) = 0$. For the partial derivatives we get the following expressions:

$$\frac{\partial F}{\partial x}(x, y, z) = -y, \qquad \frac{\partial F}{\partial y}(x, y, z) = -x + z + 3y^2, \qquad \frac{\partial F}{\partial z}(x, y, z) = 3z^2 + y$$

Substituting the point $A \equiv [1, 1, 1]$ into these expressions we get:

$$\left.\frac{\partial F}{\partial x}(x, y, z)\right|_A = -1, \qquad \left.\frac{\partial F}{\partial y}(x, y, z)\right|_A = 3, \qquad \left.\frac{\partial F}{\partial z}(x, y, z)\right|_A = 4 \neq 0$$

Thus, all conditions of the theorem are satisfied. The unique function $f(x, y)$ defined and continuous in some neighbourhood of $[1, 1]$ exists, such that $f(1, 1) = 1$, $F(x, y, f(x, y)) = 0$ in some neighbourhood of $[1, 1]$. Function f has continuous partial derivatives in some neighbourhood of $[1, 1]$. Using the formulas from *d)* of the theorem for partial derivatives, substituting $[x, y] = [1, 1]$, taking into account $f(1, 1) = 1$, we get:

$$\frac{\partial f}{\partial x}(1,1) = -\frac{\frac{\partial F}{\partial x}(1,1,f(1,1))}{\frac{\partial F}{\partial z}(1,1,f(1,1))} = -\frac{\frac{\partial F}{\partial x}(1,1,1)}{\frac{\partial F}{\partial z}(1,1,1)} = \left.\frac{-(-y)}{3z^2+y}\right|_{[1,1,1]} = \frac{1}{4},$$

$$\frac{\partial f}{\partial y}(1,1) = -\frac{\frac{\partial F}{\partial y}(1,1,f(1,1))}{\frac{\partial F}{\partial z}(1,1,f(1,1))} = -\frac{\frac{\partial F}{\partial y}(1,1,1)}{\frac{\partial F}{\partial z}(1,1,1)} = \left.\frac{-(-x+z+3y^2)}{3z^2+y}\right|_{[1,1,1]} = -\frac{3}{4}$$

I.7. Local extremes.

I.7.1. Remark. In order to distinguish between extremes of function f on a set and local extremes, an extreme of f on a set is often called a *global extreme of f on a set* or an *absolute extreme of f on a set*. A maximum on a set is therefore called a *global maximum of f on a set* or an *absolute maximum of f on a set*. Analogously, we can define a *global minimum of f on a set* or an *absolute minimum of f on a set*.

I.7.2. Local maxima and local minima. We suppose that f is a function of n variables $x_1, x_2, ..., x_n$ defined in some subset D of $\mathbf{E}_n$ and A is an interior point of D.

If there exists a reduced neighbourhood $R(A) \subset D$ such that $\forall X : X \in R(A) \Rightarrow f(A) \geq f(X)$, then we say that function f has a *local maximum at point A*. Moreover, if there exists a reduced neighbourhood $R(A) \subset D$ such that $\forall X : X \in R(A) \Rightarrow f(A) > f(X)$, then we say that function f has a *strict local maximum at point A*.

A local minimum at a point and a strict local minimum is defined by analogy. If there exists a reduced neighbourhood $R(A) \subset D$ such that $\forall X : X \in R(A) \Rightarrow f(A) \leq f(X)$, then we say that function f has a *local minimum at point A*. Moreover, if there exists a reduced neighbourhood $R(A) \subset D$ such that $\forall X : X \in R(A) \Rightarrow f(A) < f(X)$, then we say that function f has a *strict local minimum at point A*.

Local maxima and local minima are called *local extremes*. It is assumed in these definitions that point A is an interior point of function f. These definitions can be extended in some sense to other cases.

I.7.3. Local maxima and local minima with respect to a set. We suppose that f is a function of n variables $x_1, x_2, ..., x_n$ defined in some subset D of $\mathbf{E}_n$ and point $A \in D$. If there exists a reduced neighbourhood $R(A)$ such that $\forall X : X \in R(A) \cap D \Rightarrow f(A) \geq f(X)$, then we say that function f has a *local maximum with respect to set D at point A*.

Moreover, if there exists a reduced neighbourhood $R(A)$ such that $\forall X : X \in R(A) \cap D \Rightarrow f(A) > f(X)$, then we say that function f has a *strict local maximum with respect to set D at point A*.

By analogy, we can define a local minimum with respect to a set at a point, and a strict local minimum with respect to a set at a point.

I.7.4. Theorem. Necessary condition of local extremes of differentiable functions. *If function f is differentiable at point $A \in \mathbf{E}_n$ and f has a local extreme at point A then*

$$(\operatorname{grad} f)(A) = \mathcal{O}.$$

I.7.5. Critical points. An interior point A of the domain of a function f where

$$(\operatorname{grad} f)(A) = \mathcal{O}$$

or where at least one partial derivative at point A does not exist is a so called *critical point of f.*

An interior point A of a set G which is contained in the domain of a function f where

$$(\operatorname{grad} f)(A) = \mathcal{O}$$

or where at least one partial derivative at point A does not exist is called a *critical point of f on set G.*

I.7.6. Remark. Theorem $I.7.4$ implies that the only points where a function f can ever have a global extreme on a set G are critical the points of function f on set G or the boundary points of set G.

I.7.7. Theorem. Sufficient condition of local extremes of differentiable functions of two variables. *Let f be a function of two variables, and let f be differentiable at point A and* $(\operatorname{grad} f)(A) = \mathcal{O}$. *We assume that there exist all partial derivatives of the second order in a neighbourhood $U(A)$ which are continuous at point A. Denoting*

$$\Delta_2(A) = \begin{vmatrix} \dfrac{\partial^2 f}{\partial x^2}(A), & \dfrac{\partial^2 f}{\partial x \partial y}(A) \\ \dfrac{\partial^2 f}{\partial x \partial y}(A), & \dfrac{\partial^2 f}{\partial y^2}(A) \end{vmatrix}, \quad \Delta_1(A) = \frac{\partial^2 f}{\partial x^2}(A),$$

we have:

a) If $\Delta_2(A) > 0$ and $\Delta_1(A) > 0$ then function f has a strict local minimum at point A.

b) If $\Delta_2(A) > 0$ and $\Delta_1(A) < 0$ then function f has a strict local maximum at point A.

c) If $\Delta_2(A) < 0$ then function f has no local extreme at point A.

I.7.8. Theorem. Sufficient condition of local extremes of differentiable functions of n variables. *Let f be a function of n variables, and let f be differentiable at point A and* $(\operatorname{grad} f)(A) = \mathcal{O}$. *We assume that there exist all partial* *derivatives of the second order in a neighbourhood $U(A)$ which are continuous at point A. We use the following notation:*

$$\Delta_k(A) = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2}(A), & \frac{\partial^2 f}{\partial x_1 \partial x_2}(A), & ..., & \frac{\partial^2 f}{\partial x_1 \partial x_k}(A) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(A), & \frac{\partial^2 f}{\partial x_2^2}(A), & ..., & \frac{\partial^2 f}{\partial x_2 \partial x_k}(A) \\ & & \cdots & \\ \frac{\partial^2 f}{\partial x_k \partial x_1}(A), & \frac{\partial^2 f}{\partial x_k \partial x_2}(A), & ..., & \frac{\partial^2 f}{\partial x_k^2}(A) \end{vmatrix}, \quad \textit{for } k = 1, 2, ..., n$$

Assuming $\Delta_k(A) \neq 0$ for $k = 1, 2, ..., n$ we have:

a) If $\Delta_k(A) > 0$ for $k = 1, 2, ..., n$ then function f has a strict local minimum at point A.

b) If $(-1)^k \Delta_k(A) > 0$ for $k = 1, 2, ..., n$ then function f has a strict local maximum at point A.

c) In other cases function f has no local extreme at point A.

I.7.9. Example. Find all local extremes of function $f\,:\, f(x, y, z) = x^2 + 3z^2 + 3y - xz - xy$.

Solution: Function f is defined in $\mathbf{E}_3$. We find all critical points of f. We calculate the partial derivatives:

$$\frac{\partial f}{\partial x}(x, y, z) = 2x - y - z, \qquad \frac{\partial f}{\partial y}(x, y, z) = 3 - x, \qquad \frac{\partial f}{\partial z}(x, y, z) = 6z - x$$

The partial derivatives are defined and continuous in $\mathbf{E}_3$. Using the necessary condition of a local extreme, we solve the system $(\operatorname{grad} f)(X) = \mathcal{O}$, i.e.

$$\begin{array}{ccccccc} 2x & - & y & - & z & = & 0 \\ 3 & - & x & & & = & 0 \\ 6z & - & x & & & = & 0 \end{array}$$

From the second equation we get $x = 3$, substituting this value into the third equation we get $z = \frac{1}{2}$ and, finally, the first equation implies $y = \frac{11}{2}$. Thus, the unique critical point of f is the point $A = [3, \frac{11}{2}, \frac{1}{2}]$.

Now we will use the sufficient condition of the existence of an extreme. We calculate all partial derivatives of the second order:

$$\frac{\partial^2 f}{\partial x^2}(x, y, z) = 2, \qquad \frac{\partial^2 f}{\partial y \partial x}(x, y, z) = \frac{\partial^2 f}{\partial z \partial x}(x, y, z) = -1,$$

$$\frac{\partial^2 f}{\partial y^2}(x, y, z) = \frac{\partial^2 f}{\partial z \partial y}(x, y, z) = 0, \qquad \frac{\partial^2 f}{\partial z^2}(x, y, z) = 6$$

Hence,

$$\Delta_3 = \begin{vmatrix} 2, & -1, & -1 \\ -1, & 0, & 0 \\ -1, & 0, & 6 \end{vmatrix} = -6, \qquad \Delta_2 = \begin{vmatrix} 2, & -1 \\ -1, & 0 \end{vmatrix} = -1, \qquad \Delta_1 = |\,2\,| = 2.$$

The condition in *a*) of the previous theorem is not satisfied, and the condition in *b*) is also not satisfied. Using *c*) we can conclude that the function f has no local extreme in $\mathbf{E}_2$. From this it also follows that the function f has no (global) extreme on $\mathbf{E}_2$.

In the next example we pay attention to a procedure for finding global extremes of a function on a set.

I.7.10. Example. Find the global extremes of function $f : f(x,y) = \dfrac{x^3}{3} + xy^2 - 4xy$ on the set $G = \left\{[x,y] \in \mathbf{E}_2 \,:\, y \geq \dfrac{x^2}{3} \wedge y \leq 3x\right\}$. (Draw the sketch of G.)

Solution: function f is a function defined and continuous in $\mathbf{E}_2$, and set G is a bounded closed subset of $\mathbf{E}_2$, so the (global) extremes of f on G exist, see Theorem *I*.3.18.

A function can have global extremes at critical points or at boundary points only, see Remark *I*.7.6.

A) Firstly, we find all critical points – interior points of G where $(\operatorname{grad} f)(X) = \mathcal{O}$ or where the function is not differentiable. We calculate the partial derivatives:

$$\frac{\partial f}{\partial x}(x,y) \;=\; x^2 + y^2 - 4y = x^2 + y(y-4), \qquad \frac{\partial f}{\partial y}(x,y) \;=\; 2xy - 4x = 2x(y-2)$$

Partial derivatives are defined and continuous in $\mathbf{E}_2$, so f is differentiable. Using the necessary condition of a local extreme we solve the system $(\operatorname{grad} f)(X) = \mathcal{O}$, i.e.

$$\begin{aligned} x^2 + y(y-4) &= 0 \\ 2x(y-2) &= 0 \end{aligned}$$

From the second equation we get $x = 0 \vee y = 2$.

α) Let $x = 0$; then from the first equation we get $y = 0 \vee y = 4$.

β) Let $y = 2$; then from the first equation we get $x = -2 \vee x = 2$.

Thus, we get the points: $[0,0], [0,4], [2,-2], [2,2]$. However, only point [2,2] is an interior point of G, $([0,4],[2,-2] \notin G,\ [0,0] \in \partial G)$. We denote $A_0 \equiv [2,2]$, $f(A_0) = -\dfrac{16}{3} = -5.\overline{3}$.

B) Now we will investigate the boundary of G. The boundary of G can be divided into two parts:

$$\Gamma_1 = \left\{[x,y] \in \mathbf{E}_2 \,:\, y = \frac{x^2}{3}, x \in [0;9]\right\}, \qquad \Gamma_2 = \{[x,y] \in \mathbf{E}_2 \,:\, y = 3x, x \in [0;9]\}$$

Part Γ_1: The function value of f on Γ_1 depends only on one variable:

$$F_1(x) \equiv f(x, \frac{x^2}{3}) = \frac{x^3}{3} + x\frac{x^4}{9} - 4x\frac{x^2}{3} = \frac{x^5}{9} - x^3$$

Part Γ_2: The function value of f on Γ_2 also depends only on one variable:

$$F_2(x) \equiv f(x, 3x) = \frac{28}{3}x^3 - 12x^2$$

Ba) We will investigate these functions on the open interval $x \in (0; 9)$, (and we will evaluate the function values at $x = 0$ and $x = 9$ in part Bb)). We get the critical points of F_1 from the relation:

$$F_1'(x) = \frac{5}{9}x^4 - 3x^2 = 0$$

Thus, $x = 0,\ \sqrt{\frac{27}{5}},\ -\sqrt{\frac{27}{5}}$. However, only the point $x = \sqrt{\frac{27}{5}}$ is an interior point of the interval $[0; 9]$. After evaluation of the y-coordinate ($y = \frac{x^2}{3} = \frac{9}{5}$) we denote $A_1 \equiv [\sqrt{\frac{27}{5}}, \frac{9}{5}]$. We can compute the function value of f at $A_1 : f(A_1) = F_1(\sqrt{\frac{27}{5}}) \doteq$ $\doteq -5.019389$.

Part Γ_2:

$$F_2'(x) = 28x^2 - 24x = 0 \quad \Rightarrow \quad x = 0, \frac{6}{7}.$$

The interior point of $[0; 9]$ is $x = \frac{6}{7}$. After evaluation of the y-coordinate ($y = 3x = \frac{18}{7}$) we denote $A_2 \equiv [\frac{6}{7}, \frac{18}{7}]$. We can compute the function value of f at $A_2 : f(A_2) = F_2(\frac{6}{7}) \doteq -2.938775$.

Bb) Now we evaluate the function values $f(0, 0) = F_1(0) = F_2(0)$, and $f(9, 27) = F_1(9) = F_2(9)$:

$$x = 0 \quad \Rightarrow \quad y = \frac{x^2}{3} = 3x = 0 \quad \Rightarrow \quad A_3 \equiv [0, 0], \quad f(A_3) = 0$$

$$x = 9 \quad \Rightarrow \quad y = \frac{x^2}{3} = 3x = 27 \quad \Rightarrow \quad A_4 \equiv [9, 27], \quad f(A_4) = 5832.$$

If we compare the function values of f at points $A_0, A_1, ..., A_4$ we get: function f has the global minimum on G at the point $A_0 \equiv [2, 2]$ and the global maximum on G at the point $A_4 \equiv [9, 27]$. (Point A_0 is an interior point of G, point A_4 is a boundary point of G.)

II.8. Exercises.

1. Find the function's domain and range.

$f(x, y) = e^{16-x^2-y^2}$ $\qquad$ $f(x, y) = \dfrac{1}{x(y + 3)}$ $\qquad$ $f(x, y) = \ln(e^2 + x^2 + y^2)$

$f(x, y) = \sqrt[4]{y - x}$ $\qquad$ $f(x, y) = \sqrt{y - x^2}$ $\qquad$ $f(x, y) = \cos(3x^2 - 2y + 5)$

$f(x,y,z) = \sqrt{x^2+y^2+z^2-1}$ $\qquad f(x,y,z) = yz\ln x$

$f(x,y,z) = \dfrac{1}{x^2+y^2+z^2}$ $\qquad f(x,y,z) = \arctan(x+y+z)$

2. Do the following limits exist? If yes, evaluate them.

$$\lim_{[x,y]\to[0,0]} \frac{3x^2-y^2+5}{x^2+y^2+2} \qquad \lim_{[x,y]\to[0,0]} \frac{x^4}{x^4+y^2} \qquad \lim_{[x,y]\to[0,0]} \frac{e^y\sin x}{x}$$

$$\lim_{[x,y]\to[0,0]} \frac{xy}{|xy|} \qquad \lim_{[x,y]\to[2,2]} \frac{x+y-4}{\sqrt{x+y}-2} \qquad \lim_{[x,y]\to[0,0]} \frac{x+y}{x-y}$$

$$\lim_{[x,y]\to[0,0]} \frac{x-y+2\sqrt{x}-2\sqrt{y}}{\sqrt{x}-\sqrt{y}} \qquad \lim_{\substack{[x,y]\to[2,-4]\\ y\neq -4,\ x\neq x^2}} \frac{y+4}{x^2y-xy+4x^2-4x}$$

3. At what points $[x,y]$ in the plane are the functions continuous?

$$f(x,y) = \frac{x+y}{x-y} \qquad f(x,y) = \frac{x^2+y^4+1}{x^2+x-12} \qquad f(x,y) = \frac{1}{x^2-2y}$$

$$f(x,y) = \ln\frac{y}{x} \qquad f(x,y) = \cos(x^2+xy) \qquad f(x,y) = \mathrm{e}^{\frac{1}{x+y}}$$

4. At what points $[x,y,z]$ in space are the functions continuous?

$$f(x,y,z) = \frac{1}{x^2+z^2-4} \qquad f(x,y,z) = \ln xyz \qquad f(x,y,z) = e^z\sin(x+y)$$

$$f(x,y,z) = \frac{x+y}{x-y} \qquad f(x,y,z) = \ln\frac{1}{xyz} \qquad f(x,y,z) = \frac{1}{|xy|+|z|}$$

$$f(x,y,z) = \frac{1}{\ln\sqrt{x^2+y^2+z^2}} \qquad f(x,y,z) = \frac{y+4}{x^2y-xy+4x^2-4x}$$

5. Find $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$.

$$f(x,y) = x^2-7xy+13y^2 \qquad f(x,y) = (x+2)^2(y+3) \qquad f(x,y) = x^2(3y-5)^7$$

$$f(x,y) = x\sin(xy) \qquad f(x,y) = \ln(x^2y) \qquad f(x,y) = \frac{2x}{x-\sin y}$$

$$f(x,y) = \frac{x+y}{x-y} \qquad f(x,y) = \ln(x^2-2y) \qquad f(x,y) = \sqrt{x^2+y^2}$$

$$f(x,y) = e^x\ln y \qquad f(x,y) = \frac{1}{\tan(\frac{y}{x})} \qquad f(x,y) = ye^{x^2y}$$

6. Find $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ and $\frac{\partial f}{\partial z}$.

$$f(x,y,z) = \frac{x^5y^2}{z^3} \qquad f(x,y,z) = x-\sqrt{y^2+z^2} \qquad f(x,y,z) = \arctan(x+y+z)$$

$$f(x,y,z) = xy+yz+zx \qquad f(x,y,z) = \sqrt{x^2+y^2+z^2} \qquad f(x,y,z) = \frac{1}{\sqrt{(x^2+y^2+z^2)}}$$

$$f(x,y,z) = x^2\sin^2 y\cos z^2 \qquad f(x,y,z) = \frac{x^2}{\sqrt{y^2+z^2}} \qquad f(x,y,z) = \frac{e^z+\ln y^2}{\sqrt{x}}$$

7. Find the second order partial derivatives of the following functions.

$f(x,y) = x^2y + \cos y + y\sin x$ $\qquad f(x,y) = xe^y + y + x^5y^4 - 13$

$f(x,y) = e^{x+3y} + x\ln y + y\ln x + 3$ $\qquad f(x,y) = y + x^2y + 4y^3x - \ln(y^2+x)$

$f(x,y) = y^2 + y(\sin x - x^4)$ $\qquad f(x,y) = x^2 + 5xy + \sin(xy) + xe^{\frac{y^2}{2}}$

8. Evaluate $\operatorname{grad} f$ at point M and directional derivative $\frac{\partial f}{\partial \vec{s}}(M)$.

$f(x,y) = x^2 + 2xy - 3y^2,$ $\qquad M = [1,1],$ $\qquad \vec{s} = (3,4)$

$f(x,y,z) = x^2 + 2y^2 - 3z^3 - 17,$ $\qquad M = [1,1,1],$ $\qquad \vec{s} = (1,1,1)$

$f(x,y,z) = \cos(xy) + e^{yz} + \ln(zx),$ $\qquad M = [1,0,0.5],$ $\qquad \vec{s} = (1,2,2)$

9. Show that the following equations $F(x,y,z) = 0$ define implicit functions $f : z = f(x,y)$ in the neighbourhoods of the given points $M \equiv [M_1, M_2, M_3]$, and find its partial derivatives $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ at $[M_1, M_2]$.

$F(x,y,z) = z^3 - xy + yz + y^3 - 2 = 0,$ $\qquad M = [1,1,1]$

$F(x,y,z) = x^2 - 2y^2 + z^2 - 4x + 2z - 5 = 0,$ $\qquad M = \left[-1, \sqrt{\frac{3}{2}}, 1\right]$

$F(x,y,z) = xz^2 - x^2y + y^2z + 2x - y = 0,$ $\qquad M = [0,1,1]$

$F(x,y,z) = \sin(x+y) + \sin(y+z) + \sin(x+z) = 0,$ $\qquad M = [\pi,\pi,\pi]$

10. Find the equations for the tangent planes and normal lines at points M on given surfaces $F(x,y,z) = 0$.

$F(x,y,z) = x^2 + y^2 + z^2 - 3 = 0,$ $\qquad M = [1,1,1]$

$F(x,y,z) = \cos(\pi x) - x^2y + e^{xz} + yz - 4 = 0,$ $\qquad M = [0,1,2]$

11. Find all the local maxima and local minima of the following functions.

$f(x,y) = 2xy - 5x^2 - 2y^2 + 4x + 4y - 4$ $\qquad f(x,y) = x^2 + xy + 3x + 2y + 5$

$f(x,y) = 5xy - 7x^2 + 3x - 6y + 2$ $\qquad f(x,y) = x^2 - 4xy + y^2 + 6y + 2$

$f(x,y) = 2x^2 + 3xy + 4y^2 - 5x + 2y$ $\qquad f(x,y) = x^2 - y^2 - 2x + 4y + 6$

$f(x,y) = 9x^3 + 3\frac{y^3}{3} - 4xy$ $\qquad f(x,y) = 8x^3 + y^3 + 6xy$

$f(x,y) = x^3 + y^3 + 3x^2 - 3y^2 - 8$ $\qquad f(x,y) = 2x^3 + 2y^3 - 9x^2 + 3y^2 - 12y$

$f(x,y) = 4xy - x^4 - y^4 - 11$ $\qquad f(x,y) = x^4 + y^4 + 4xy + 7$

12. Find all the global maxima and global minima of the functions on the given subsets.

$f(x,y) = 2x^2 - 4x + y^2 - 4y + 2,$ $\qquad G = \{[x,y] : x \geq 0,\ y \leq 2,\ y \geq 2x\}$

$f(x,y) = x^2 - xy + y^2 + 7,$ $\qquad G = \{[x,y] : x \geq 0,\ y \leq 4,\ y \geq x\}$

$f(x,y) = x^2 + xy + y^2 - 6x + 2,$ $\qquad G = \{[x,y] : 0 \leq x \leq 5,\ -3 \leq y \leq 3\}$

$f(x,y) = x^2 + xy + y^2 - 6x,$ $\qquad G = \{[x,y] : 0 \leq x \leq 5,\ -3 \leq y \leq 0\}$

$f(x,y) = 48xy - 32x^3 - 24y^2,$ $\qquad G = \{[x,y] : 0 \leq x \leq 1,\ 0 \leq y \leq 1\}$

$f(x,y) = x^2 - y^2,$ $\qquad G = \{[x,y] : x \geq -1,\ y \geq -1,\ x + 2y \leq 2\}$

II. Riemann Integral of a Function of One Variable

II.1. Motivation and definition of the Riemann integral.

II.1.1. Physical motivation. 1. Suppose that we have a spring or a thin rod of a finite length which need not be homogeneous. This can be caused e.g. by a varying cross–section of the rod or by varying density of the material that the rod is made of. We can assume that the rod covers the interval $\langle a, b\rangle$ on the x–axis and its longitudinal density (i.e. amount of mass per unit of length) is a function $y = \rho(x)$. We wish to evaluate the mass M of the rod.

If function ρ is constant then the problem is very easy – we can simply put M equal to the product of the constant longitudinal density ρ and the length of the interval $\langle a, b\rangle$, i.e. $M = \rho \cdot (b - a)$.

If ρ is not constant then we can divide the rod (i.e. the interval $\langle a, b\rangle$) into many shorter parts (the subintervals $\langle x_0, x_1\rangle$, $\langle x_1, x_2\rangle$, $\langle x_2, x_3\rangle$, $\dots$, $\langle x_{n-1}, x_n\rangle$ where $x_0 = a$ and $x_n = b$) and we can approximate function ρ by a constant on each of the subintervals. A reasonable value of this constant is $\rho(\zeta_i)$ for some $\zeta_i \in \langle x_{i-1}, x_i\rangle$ $(i = 1, 2, \dots, n)$. Then the approximate masses of the the shorter parts of the rod are

$$\rho(\zeta_1) \cdot \Delta x_1, \quad \rho(\zeta_2) \cdot \Delta x_2, \quad \dots\ , \ \rho(\zeta_n) \cdot \Delta x_n$$

where $\Delta x_1 = x_1 - x_0$, $\Delta x_2 = x_2 - x_1$, $\dots$, $\Delta x_n = x_n - x_{n-1}$. The approximate mass of the whole rod is

$$\sum_{i=1}^{n} \rho(\zeta_i) \cdot \Delta x_i\,.$$

We can naturally expect that this sum will approach the exact value of the total mass M of the rod if $n \to +\infty$ and the numbers Δx_i $(i = 1, 2, \dots, n)$ tend to zero.

2. Suppose that a car moves in a time interval $\langle a, b\rangle$ and its velocity is given by the function $y = v(t)$. We wish to compute the distance d the car travels in the time interval $\langle a, b\rangle$.

If the velocity v is constant then the distance is obviously $d = v \cdot (b - a)$.

If the velocity is varying then we can divide the time interval $\langle a, b\rangle$ into many shorter subintervals $\langle t_0, t_1\rangle$, $\langle t_1, t_2\rangle$, $\dots$, $\langle t_{i-1}, t_i\rangle$ (where $t_0 = a$ and $t_n = b$) and we can approximate the velocity by a constant on each of these shorter subintervals. A natural value of this constant is $v(\zeta_i)$ for some $\zeta_i \in \langle t_{i-1}, t_i\rangle$ $(i = 1, 2, \dots, n)$. The approximate distances moved in the time intervals $\langle t_0, t_1\rangle$, $\langle t_1, t_2\rangle$, $\dots$, $\langle t_{n-1}, t_n\rangle$ are

$$v(\zeta_1) \cdot \Delta t_1, \quad v(\zeta_2) \cdot \Delta t_2, \quad \dots\ , \ v(\zeta_n) \cdot \Delta t_n$$

where $\Delta t_1 = t_1 - t_0$, $\Delta t_2 = t_2 - t_1$, $\dots$, $\Delta t_n = t_n - t_{n-1}$. The approximate distance travelled in the whole time interval $\langle a, b\rangle$ is

$$\sum_{i=1}^{n} v(\zeta_i) \cdot \Delta t_i \,.$$

One can expect that this sum will approach the real distance that the car travels in the time interval $\langle a, b \rangle$ if $n \to +\infty$ and the numbers Δt_i $(i = 1, 2, \dots, n)$ tend to zero.

II.1.2. Geometric motivation. Suppose that f is a nonnegative and bounded function on an interval $\langle a, b \rangle$ and D is the region between its graph and the x–axis. (See Fig 1.) An important question is how to define and evaluate the area of D.

If f is a constant function on $\langle a, b \rangle$ then D is a rectangle and its area is equal to the product $f \cdot (b - a)$.

If function f is not constant then we can again subdivide the interval $\langle a, b \rangle$ into many short subintervals $\langle x_0, x_1 \rangle$, $\langle x_1, x_2 \rangle$, $\dots$, $\langle x_{n-1}, x_n \rangle$ (with $x_0 = a$ and $x_n = b$) and we can approximate f by a constant on each of these subintervals. A possible value of this constant is $f(\zeta_i)$ for some $\zeta_i \in \langle x_{i-1}, x_i \rangle$ $(i = 1, 2, \dots, n)$. Thus, we can approximate the area of the region below the graph of f on the subinterval $\langle x_{i-1}, x_i \rangle$ by the area of the rectangle with the sides $f(\zeta_i)$ and Δx_i $(\equiv x_i - x_{i-1})$. The approximate value of the area of the whole region D is equal to the total area of all the rectangles:

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

(See Fig. 1.) We can now define the area of D as a limit of this sum for $n \to +\infty$ and the lengths Δx_i of the subintervals $\langle x_{i-1}, x_i \rangle$ $(i = 1, 2, \dots, n)$ tending to zero.

One can observe that all the situations described in paragraphs II.1.1 and II.1.2 lead to the limit of a certain sum and the sum is the same in all the considered situations. We explain in the next paragraphs what we exactly understand under the limit of this sum, what we call it, how we denote it and how we can evaluate it.

II.1.3. Partition of an interval. Let $\langle a, b\rangle$ be a bounded closed interval. A system of points $x_0, x_1, \ldots, x_n$ such that $a = x_0 < x_1 < \ldots < x_n = b$ is called a *partition* of the interval $\langle a, b\rangle$. If this partition is named P then we write

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

The *norm* of partition P is the number $\|P\| = \max_{i=1,\ldots,n}(x_i - x_{i-1})$. (Thus, $\|P\|$ is the length of the largest of the subintervals $\langle x_0, x_1\rangle, \langle x_1, x_2\rangle, \ldots, \langle x_{n-1}, x_n\rangle$ and it informs us how "fine" partition P is.)

II.1.4. Riemann sums and their limit. Let $y = f(x)$ be a bounded function on the interval $\langle a, b\rangle$ and let P be the partition of $\langle a, b\rangle$ given by (II.1). Denote by Δx_i the length of the i–th subinterval $\langle x_{i-1}, x_i\rangle$ (i.e. $\Delta x_i = x_i - x_{i-1}$). Let V be a system of points $\zeta_1 \in \langle x_0, x_1\rangle, \zeta_2 \in \langle x_1, x_2\rangle, \ldots, \zeta_n \in \langle x_{n-1}, x_n\rangle$. Then the *Riemann sum* of function f on the interval $\langle a, b\rangle$ corresponding to partition P and system V is

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

We say that number S is the *limit of the Riemann sums* $s(f, P, V)$ *as* $\|P\| \to 0+$ if to every given $\epsilon > 0$ there exists $\delta > 0$ such that for every partition P of $\langle a, b\rangle$ and for every choice of V, $\|P\| < \delta$ implies $|s(f, P, V) - S| < \epsilon$. We write:

$$\lim_{\|P\|\to 0+} s(f, P, V) = S. \tag{II.2}$$

II.1.5. Riemann integral. If the limit in (II.2) exists then function f is called *integrable* in the interval $\langle a, b\rangle$ and S is called the *Riemann integral* of function f on $\langle a, b\rangle$. The integral is usually denoted as

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

The numbers a and b in this integral are called the *limits of integration*, a being the *lower limit* and b being the *upper limit*. The integrated function is called the *integrand*.

The Riemann integral is also often called the *definite integral*.

II.1.6. The area of the region between the graph of a function and the x–axis. It follows from paragraph II.1.2 and the definition of the Riemann integral that if f is a nonnegative and integrable function on the interval $\langle a, b\rangle$ then the *area* of the region between the graph of f and the x–axis can be defined as the value of the integral $\int_a^b f\, dx$.

By analogy, if function f is nonpositive and integrable on the interval $\langle a, b\rangle$ then the area of the region bounded by the x–axis (from above) and the graph of f (from below) can be defined as $-\int_a^b f\, dx$.

In a general case, when f has both negative and positive values in the interval $\langle a, b\rangle$, the integral $\int_a^b f\, dx$ expresses the sum of the areas of all the parts of the region between the graph of f and the x–axis, but the contributions of the parts below the x–axis are taken negatively.

You will see in Chapter III that the area of more general sets than the regions below or above the graph of a function $y = f(x)$ can be defined by means of a so called two–dimensional measure m_2 and evaluated by means of a double integral.

II.1.7. Extension of the definition of the Riemann integral. If function f is integrable in the interval $\langle a, b\rangle$ then we put

$$\int_b^a f\ dx = -\int_a^b f\ dx.$$

Specially, we also put $\displaystyle\int_a^a f\ dx = 0.$

II.1.8. The mean value of function f on an interval. Let function f be integrable in the interval $\langle a, b\rangle$. The number

$$\mu\ =\ \frac{1}{b-a}\int_a^b f\ dx$$

is called the <u>*mean value*</u> (or the <u>*average value*</u>) of function f on the interval $\langle a, b\rangle$.

The mean value has the following geometric interpretation: Suppose for simplicity that function f is nonnegative on the interval $\langle a, b\rangle$. Then the mean value μ is such a number that the region between the graph of f and the x–axis has the same area as the rectangle with the sides $b - a$ and μ. It is clear that

$$\inf_{x\in\langle a,b\rangle} f(x)\ \le\ \mu\ \le\ \sup_{x\in\langle a,b\rangle} f(x). \tag{II.3}$$

II.2. Integrability (existence of the Riemann integral) – sufficient conditions.

The two statements "the Riemann integral $\int_a^b f(x)\ dx$ exists" and "function f is integrable in the interval $\langle a, b\rangle$" say exactly the same.

Most of the functions you will use in various applications will be integrable. Nevertheless, you should be aware that there also exist "bad" functions such that the limit of the Riemann sums (II.2) does not exist. Thus, the Riemann integral of these functions also does not exist. These functions are called *non–integrable.* The next theorem and Remark II.2.2 give sufficient conditions for the integrability of function f (i.e. for the existence of the Riemann integral of f).

II.2.1. Existence theorem for the Riemann Integral. *Let function f be continuous on the interval $\langle a, b\rangle$. Then f is integrable in $\langle a, b\rangle$.*

II.2.2. Remark. This theorem can be generalized:

Let function f be bounded and piecewise–continuous on the interval $\langle a, b\rangle$. Then it is integrable in $\langle a, b\rangle$.

(A function f is said to be *piecewise–continuous* in the interval $\langle a,b\rangle$ if $\langle a,b\rangle$ can be divided into a finite number of subintervals such that f is continuous in the interior of each of them.)

II.3. Important properties of the Riemann integral.

II.3.1. Theorem. (The domination inequality for the Riemann integral.) *If functions f and g are both integrable in the interval $\langle a,b\rangle$ and $g(x)\le f(x)$ for all $x\in\langle a,b\rangle$ then*

$$\int_a^b g\,dx \;\le\; \int_a^b f\,dx.$$

Specially, if $f(x)\ge 0$ for all $x\in\langle a,b\rangle$ then $\displaystyle\int_a^b f\,dx\ge 0.$

II.3.2. Theorem. (Boundedness of the Riemann integral.) *If function f is integrable in the interval $\langle a,b\rangle$ and $m\le f(x)\le M$ for all $x\in\langle a,b\rangle$ then*

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

Both theorems II.3.1 and II.3.2 easily follow from the definition of the Riemann integral. Theorem II.3.1 tells us that if function f dominates function g on $\langle a,b\rangle$ and the functions f and g are both integrable in $\langle a,b\rangle$ then also the integral of f dominates the integral of g on $\langle a,b\rangle$. The inequality in II.3.2 shows that the value of the Riemann integral can be estimated by means of the lower bound and the upper bound of function f.

II.3.3. Theorem. (Linearity of the Riemann integral.) *If functions f and g are integrable in $\langle a,b\rangle$ and $\alpha\in\mathbf{R}$ then*

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

(This property is already known from the theory of the indefinite integral.)

II.3.4. 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*

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

II.3.5. Theorem. a) *If function f is integrable in the interval $\langle a,b\rangle$ and if $\langle c,d\rangle\subset\langle a,b\rangle$ then f is also integrable in $\langle c,d\rangle$.*

b) *If functions f and g are both integrable in the interval $\langle a,b\rangle$ then the product $f\cdot g$ is also integrable in $\langle a,b\rangle$.*

c) *If function f is integrable in the interval $\langle a,b\rangle$ and function g differs from f in at most a finite number of points then function g is also integrable in $\langle a,b\rangle$ and*

$$\int_a^b g\,dx = \int_a^b f\,dx.$$

Item a) is an immediate consequence of the definition of the Riemann integral.

Item b) is a statement about the integrability of a function which is the product of two other functions. However, bear in mind that the fact that "the integrability of f and g implies the integrability of $f\cdot g$" does not mean that $\int_a^b f\cdot g\,dx = \left(\int_a^b f\,dx\right)\cdot\left(\int_a^b g\,dx\right)$!

Item c) tells us that the change of function values of f in at most a finite number of points does not affect the existence or the value of the integral $\int_a^b f\,dx$. In other words: The existence and the value of the integral $\int_a^b f\,dx$ do not depend on function values of f in a finite number of points. Thus, function f need not even be defined in a finite number of points of the interval $\langle a,b\rangle$ and this has no influence on the existence and the value of $\int_a^b f\,dx$. Specially, it plays no role whether the integral $\int_a^b f\,dx$ is considered on a closed or on an open interval!

II.3.6. Theorem. (The Riemann integral as a function of its upper limit.) *Suppose that function f is integrable in the interval $\langle a,b\rangle$. Then*

a) the function $F(x)=\int_a^x f(t)\,dt$ is continuous in $\langle a,b\rangle$,

b) the equality

$$\frac{d}{dx}\int_a^x f(t)\,dt \;=\; f(x) \tag{II.4}$$

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

Function $G(x)=\int_x^b f(t)\,dx$ (with the variable lower limit) is also continuous in $\langle a,b\rangle$. However, it satisfies the equality in b) with the change of the sign:

$$\frac{d}{dx}\int_x^b f(t)\,dt \;=\; -f(x). \tag{II.5}$$

(This is a consequence of the equation $G(x)=\int_a^b f(t)\,dt - F(x)$.)

Equalities (II.4) and (II.5) can also be modified for the boundary points of the interval $\langle a,b\rangle$ so that if function f is right–continuous at point a (respectively left–continuous at point b) then $F'_+(a)=f(a)$ (respectively $F'_-(b)=f(b)$).

The validity of statement a) follows (at least intuitively) from the geometric interpretation of the Riemann integral (see paragraphs II.1.2 and II.1.6). Formula (II.4) can be proved in this way:

$$F'(x)=\lim_{h\to 0}\frac{F(x+h)-F(x)}{h} \;=\; \lim_{h\to 0}\frac{1}{h}\Big[\int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt\Big] \;=$$

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

where $\mu(h)$ is the mean value of function f on the interval with the end points x and $x+h$. The continuity of f at point x and (II.3) imply that $\mu(h) \to f(x)$ if $h \to 0$. This proves (II.4).

II.3.7. Remark. It follows from Theorem II.3.6 that if function f is continuous in interval J and $c \in J$ then the function $F(x) = \int_c^x f(t)\,dt$ is an antiderivative to function f in J.

II.3.8. Remark. Formula (II.4) can be generalized. If function f is continuous in interval I and $a(x)$, $b(x)$ are differentiable functions of variable x in interval J with their values in I then

$$\frac{d}{dx}\int_{a(x)}^{b(x)} f(t)\,dt \;=\; f(b(x))\cdot b'(x) - f(a(x))\cdot a'(x)$$

for $x \in J$.

II.4. Evaluation of the Riemann integral.

We come to one of the fundamental topics of this chapter – to the question how to evaluate the integral $\int_a^b f\,dx$. Due to its importance, the next theorem is called the Fundamental Theorem of Integral Calculus:

II.4.1. Theorem. *If function f is continuous in the interval $\langle a, b\rangle$ and F is an antiderivative to f in $\langle a, b\rangle$ then*

$$\int_a^b f\,dx \;=\; F(b) - F(a). \tag{II.6}$$

The formula (II.6) is called the *Newton–Leibnitz formula*. The difference $F(b) - F(a)$ is often written in a shorter form: $F(b) - F(a) = [\,F\,]_a^b$

The proof of the Fundamental Theorem of Integral Calculus is easy: The function $G(x) = \int_a^x f(t)\,dt$ is also an antiderivative to f in $\langle a, b\rangle$. Thus, there exists a constant c such that $F = G + c$ on $\langle a, b\rangle$. This means that $F(a) = G(a) + c = c$ (because $G(a) = 0$) and $F(b) = G(b) + c = G(b) + F(a)$. This yields: $\int_a^b f(t)\,dt = G(b) = F(b) - F(a)$.

The Newton–Leibnitz formula connects the indefinite and the definite integral: When you know the indefinite integral of f on the interval $\langle a, b\rangle$ then you also know all antiderivatives to f on $\langle a, b\rangle$. You can choose any of them and use it in the Newton–Leibnitz formula to obtain the value of the definite integral of f on $\langle a, b\rangle$. The fact that the indefinite integral and the antiderivative are so important in the calculation of the definite integral was one of the main reasons why you have learned to compute indefinite integrals.

You already know that all antiderivatives to function f on the interval $\langle a, b\rangle$ differ at most in an additive constant. Thus, if you choose e.g. an antiderivative

$F + k$ (where k is a constant) instead of F and you use it in the Newton–Leibnitz formula, you get

$$\int_a^b f\,dx = [F+k]_a^b = \big(F(b)+k\big) - \big(F(a)+k\big) = F(b) - F(a).$$

The result is the same as in formula (II.6). Hence you can see that it is not important which one of the infinitely many antiderivatives to f on $\langle a, b\rangle$ you use.

II.4.2. Example. $\displaystyle\int_0^\pi \sin x\,dx = [-\cos x]_0^\pi = (-\cos\pi) - (-\cos 0) = 2.$

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

II.4.3. Theorem. (Integration by parts for the Riemann integral.) *Let the functions u and v have continuous derivatives in the interval $\langle a, b\rangle$. Then*

$$\int_a^b u' \cdot v\,dx = [u\cdot v]_a^b - \int_a^b u\cdot v'\,dx. \tag{II.7}$$

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

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

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

II.4.5. Theorem. (Integration by substitution for the Riemann integral.) *Let function g have a continuous derivative in the interval $\langle a, b\rangle$ and let g map $\langle a, b\rangle$ into interval J. Let function f be continuous in J. Then*

$$\int_a^b f(g(x))\cdot g'(x)\,dx = \int_{g(a)}^{g(b)} f(s)\,ds. \tag{II.8}$$

Formula (II.8) can be used in two situations: you wish to evaluate the integral on the left hand side and you transform it to the integral on the right hand side (if the integral on the right hand side is simpler) OR vice versa.

II.4.6. Example. Let us evaluate $\displaystyle\int_0^{\pi/2} \sin^2 x\cdot\cos x\,dx.$

If we put $\langle a, b\rangle = \langle 0, \pi/2\rangle$, $s = g(x) = \sin x$, $f(s) = s^2$, $J = (-\infty, +\infty)$, we can see that all the assumptions of Theorem II.4.5 are satisfied. Moreover, $g(0) = \sin 0 = 0$ and $g(\pi/2) = \sin(\pi/2) = 1$. Applying formula (II.8), we obtain:

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

II.4.7. Example. Let us evaluate $\displaystyle\int_0^2 \sqrt{4-x^2}\,dx$.

We can take this integral for the integral on the right hand side of (II.8) (with the variable denoted by x instead of s). The function $f(x)=\sqrt{4-x^2}$ is continuous on $\langle 0,2\rangle$ and so the integral exists. Put $x=g(t)=2\sin t$, $dx=g'(t)\,dt=2\cos t\,dt$. We have $g(a)=2\sin a=0$ and $g(b)=2\sin b=2$. Thus, we can choose $a=0$ and $b=\pi/2$. Then all the assumptions of Theorem II.4.5 are satisfied and we obtain:

$$\int_0^2 \sqrt{4-x^2}\,dx = \int_0^{\pi/2} \sqrt{4-4\sin^2 t}\cdot 2\cos t\,dt = \int_0^{\pi/2} 4\cos^2 t\,dt =$$
$$= \int_0^{\pi/2} 2\,(1+\cos 2t)\,dt = \big[2t+\sin 2t\big]_0^{\pi/2} = \pi.$$

II.4.8. Remark. Suppose that you have to evaluate a Riemann integral on the interval $\langle a,b\rangle$ and you wish to use integration by parts or a substitution. Then you have two possibilities:

1) You can use Theorem II.4.3 or Theorem II.4.5. You transform the integral to other (simpler) integrals and you deal with the upper and the lower limits of all the integrals during the computation. This approach is explained in examples II.4.5, II.4.6 and II.4.7.
2) You can first compute the integral as an indefinite integral on the interval $\langle a,b\rangle$ and then you apply the Newton–Leibnitz formula (II.6) on $\langle a,b\rangle$.

   To show what we exactly mean by this, let us compute the integral from example II.4.6 once again, this time by the method we are just explaining. Thus, let us start 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

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

   Formula (II.6) now gives: $\displaystyle\int_0^{\pi/2} \sin^2 x\cos x\,dx = \big[\tfrac{1}{3}\sin^3 x\big]_0^{\pi/2} = \tfrac{1}{3}.$

   As you will observe after having solved a larger number of examples, approach 1), based on direct application of integration by parts or integration by substitution to definite integrals, is usually technically simpler and less laborious.

II.5. Numerical integration.

You will remember from the theory of the indefinite integral that an antiderivative to a given function f often exists, but it cannot be obtained by standard methods of integration and it cannot be expressed in a "closed form" (i.e. by a formula prescribing a finite number of operations). Analogously, it often happens that the the Riemann integral $\int_a^b f\,dx$ exists, but it cannot be evaluated by a standard integration based on the Newton–Leibnitz formula. However, there exist approximate methods (also called numerical methods) which enable us to evaluate the integral approximately, with an error as small as we wish. We shall explain two such methods in this section. Both these methods usually require the performance of a higher number of arithmetic operations in order to reach a higher accuracy (i.e. a smaller error). Therefore approximate methods are usually used on computers.

Both the methods are based on the partition

$$P: \quad a = x_0 \,<\, x_1 \,<\, x_2 \,\ldots\, <\, x_{n-1} \,<\, x_n = b \tag{II.9}$$

of the interval $\langle a, b\rangle$ to n subintervals $\langle x_{k-1}, x_k\rangle$ $(k = 1, 2, \ldots, n)$ of equal length h. Thus,

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

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

II.5.1. The trapezoidal method. Suppose that we approximate function f by a linear function on each of the subintervals $\langle x_{k-1}, x_k\rangle$. A linear function is uniquely determined by the requirement that its graph (a straight line) passes through two chosen points. Let these points be $[x_{k-1}, y_{k-1}]$ and $[x_k, y_k]$. Then the linear function has the equation $y = y_{k-1} + (y_k - y_{k-1})/h \cdot (x - x_{k-1})$. We can easily integrate it on the interval $\langle x_{k-1}, x_k\rangle$ and we obtain $I_k = h \cdot (y_{k-1} + y_k)/2$. I_k is the area of the trapezoid (see Fig. 2). When we sum all the numbers I_1, I_2, ... , I_n, we get

$$T_n \;=\; \frac{h}{2} \cdot \left[\, y_0 + 2y_1 + 2y_2 + \ldots + 2y_{n-1} + y_n \,\right]. \tag{II.10}$$

T_n is an approximate value of the Riemann integral $\int_a^b f \; dx$. The geometric sense of T_n is seen on Fig. 2 – it is the sum of the areas of n trapezoids constructed on the intervals $\langle x_0, x_1\rangle$, $\langle x_1, x_2\rangle$ ... , $\langle x_{n-1}, x_n\rangle$.

As to the accuracy of the approximation, it generally holds that the finer the partition of $\langle a, b\rangle$, the better are the results. In other words, the accuracy of the approximation increases with increasing n (i.e. decreasing h). It can be proved that if f'' is continuous on $\langle a, b\rangle$ and M is an upper bound for the values of $|f''|$ on $\langle a, b\rangle$ then the following error estimate holds:

$$\left| T_n - \int_a^b f\,dx \right| \le \frac{b-a}{12}\,h^2\,M. \qquad \text{(II.11)}$$

II.5.2. Simpson's method. Suppose now that n is an even number. We can approximate function f by a quadratic function on each of the subintervals $\langle x_0, x_2 \rangle$, $\langle x_2, x_4 \rangle$, ... , $\langle x_{n-2}, x_n \rangle$. A quadratic function on a subinterval $\langle x_{k-2}, x_k \rangle$ ($k = 2, 4, \dots, n$) is uniquely defined by the requirement that its graph (a parabola) passes through three chosen points – let it be the points $[x_{k-2}, y_{k-2}]$, $[x_{k-1}, y_{k-1}]$, $[x_k, y_k]$. The integral of this quadratic function on $\langle x_{k-2}, x_k \rangle$ can be relatively easily evaluated – you can check that it is $I_k = h \cdot (y_{k-2} + 4y_{k-1} + y_k)/3$. Summing all the numbers I_2, I_4, ... , I_n, we obtain

$$S_n = \frac{h}{3} \cdot \left[y_0 + 4y_1 + 2y_2 + 4y_3 + \ldots + 2y_{n-2} + 4y_{n-1} + y_n \right]. \qquad \text{(II.12)}$$

Provided that the fourth derivative $f^{(4)}$ of function f is continuous on $\langle a, b \rangle$ and M is an upper bound for the values of $|f^{(4)}|$ on $\langle a, b \rangle$, the following error estimate holds:

$$\left| S_n - \int_a^b f\,dx \right| \le \frac{b-a}{180}\,h^4\,M. \qquad \text{(II.13)}$$

II.6. Improper Riemann integral.

A fundamental assumption in the definition of the Riemann integral $\int_a^b f\,dx$ was the boundedness of the interval $\langle a, b \rangle$ and the boundedness of function f on $\langle a, b \rangle$. However, we often need to work with integrals whose domain of integration (the interval) or the integrand (the function) are unbounded. Such integrals, where either the interval or the integrand (or both) are unbounded, are called improper Riemann integrals. We will explain the definition of the improper Riemann integral in this section.

Suppose that function f is defined in the interval $\langle a, b)$ and that it is integrable on each interval $\langle a, t \rangle$ (for $a \le t < b$). If the limit

$$\lim_{t \to b-} \int_a^t f(x)\,dx$$

exists, then its value is called an *improper Riemann integral* with a *singular upper limit*.

The improper Riemann integral of function f is denoted in the same way as the "usual" Riemann integral, i.e. $\int_a^b f\,dx$. Thus, we can write:

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

The improper Riemann integral with a *singular lower limit* can be defined quite analogously. The definition can even be extended to the case when both the limits are singular: If the two integrals $\int_a^c f\,dx$ and $\int_c^b f\,dx$ exist (the first one as an improper integral with a singular lower limit and the second one as an improper integral with a singular upper limit) and their sum is defined (i.e. it is not for example $-\infty+\infty$) then we put $\int_a^b f\,dx = \int_a^c f\,dx + \int_c^b f\,dx$.

If function f is integrable (in the sense of paragraph II.1.5) on the interval $\langle a, b\rangle$ then the improper Riemann integral of f on $\langle a, b\rangle$ coincides with the "usual" Riemann integral of f on $\langle a, b\rangle$. Thus, the improper Riemann integral represents an extension of the definition of the "usual" Riemann integral.

The value of the improper Riemann integral $\int_a^b f\,dx$ can either be finite (we say that the integral $\int_a^b f\,dx$ *converges*) or it can be infinite (the integral $\int_a^b f\,dx$ *diverges*).

II.6.1. Example. Let us evaluate the improper Riemann integral $\displaystyle\int_1^{+\infty} \frac{1}{x}\,dx$.
The function $f(x) = 1/x$ is continuous on the interval $\langle 1, +\infty)$ and it has an antiderivative $F(x) = \ln x$. Thus,

$$\int_1^t \frac{1}{x}\,dx = F(t) - F(1) = \ln t - \ln 1 = \ln t.$$

Since $\displaystyle\lim_{t\to+\infty} \int_1^t \frac{1}{x}\,dx = \lim_{t\to+\infty} \ln t = +\infty$, we get: $\displaystyle\int_1^{+\infty} \frac{1}{x}\,dx = +\infty$.

II.7. Historical remark.

Both differential calculus (i.e. limits, derivatives, their applications, etc.) and integral calculus (i.e. integrals) are together called *calculus*. Many aspects of finding and analytically describing tangent lines were worked out by René Descartes (1596–1650), Bonaventura Cavalieri (1598–1647), Pierre de Fermat (1601–1665) and others. However, we usually consider Sir Isaac Newton (1642–1727) and Baron Gottfried Wilhelm Leibnitz (1646–1716) to be the inventors of calculus. They were the first to understand that the process of finding tangents and the process of finding areas are mutually inverse. Since they lived in the same time, the question of priority over the invention of calculus has lead to the bitter controversies. Leibnitz was accused of copying Newton's work and the Royal Society of London did not exonerate him from this charge after investigating the matter. Present–day historians and mathematicians consider that Leibnitz's and Newton's inventions were simultaneous, but independent. Nevertheless, the disputation caused a split in the mathematical world for one and half centuries. The followers of Newton, mostly British, pursued his methods while Leibnitz's pupils, mostly French, Germans and Swiss, followed his approach. Due to Leibnitz's superior notation and his simpler mathematical language, his followers were able to be more successful than their British counterparts in the further development of calculus.

The original historical definition of the definite integral was different from the definitions you can find in present–day literature. This is especially due to the fact that the concept accepted at the time of Newton and Leibnitz is not quite correct from the present–day point of view. However, since this concept is very simple, let us explain it.

Suppose that f is a function defined and bounded on the interval $\langle a, b\rangle$. We divide the interval $\langle a, b\rangle$ into infinitely many "infinitely small" parts. A typical "infinitely small" part is an interval $\langle x, x+dx\rangle$, where dx is an "infinitely small" positive number. The product $f(x)\cdot dx$ has the following geometric interpretation: The region between the graph of f and the interval $\langle x, x+dx\rangle$ on the x–axis can be taken as an "infinitely narrow" rectangle and if $f(x)>0$ then $f(x)\cdot dx$ is the area of this rectangle. The sum of all "infinitely small" numbers $f(x)\cdot dx$ (for all $x\in\langle a, b\rangle$) was called the *definite integral* of function f on the interval $\langle a, b\rangle$.

The incorrectness of this approach can immediately be seen – the notion of an "infinitely small" positive real number dx is wrong: such a number does not exist! If you do not believe this, then imagine that you have such a number. Is it e.g. 10^{-6}? No, because you can find many positive numbers less than 10^{-6}. And what about $dx=10^{-20}$? Even this dx is not "infinitely small" because there exist many other positive numbers, less than 10^{-20}. You can see that the concept of an "infinitely small positive number" logically leads to the contradiction. Mathematics cannot allow itself to work with notions which are not defined precisely. (Overlooking this rule has often in the past lead to surprising contradictions or confusions in mathematical theories and models.) This motivated Georg F.B.Riemann (1826–1866) to study the definite integral in detail and put it on solid logical foundations.

Nevertheless, in spite of the logical incorrectness of the concept of an infinitely small positive number dx, the idea often still appears in various applications, and we have not completely abandoned it. We will use it again in paragraphs IV.2.1, IV.4.1 and V.2.1 which have a motivating character and whose main purpose it to show that the following definitions of various types of integrals are reasonable and that the integrals have some physical sense.

II.8. Exercises.

1. Do the following Riemann integrals exist?

$$\int_{-2}^{1}\frac{x+1}{x^2-x-6}\,dx \qquad \int_{1}^{2}\frac{\ln x}{x}\,dx \qquad \int_{0}^{1}\frac{\sin x}{x}\,dx$$

$$\int_{-1}^{5}e^{-x}\,dx \qquad \int_{-2.5}^{3}\frac{x}{\ln(x+3)}\,dx \qquad \int_{-2}^{-1}\frac{x^2+1}{x^3-2x^2+x}\,dx$$

2. Evaluate the following integrals.

$$\int_{-1}^{1}(3x^2-4x+7)\,dx \qquad \int_{0}^{1}(8t^3-12t^2+5)\,dt \qquad \int_{1}^{2}\frac{4}{s^2}\,ds$$

$$\int_{1}^{27}x^{-4/3}\,dx \qquad \int_{0}^{2}s\sqrt{4x+1}\,dx \qquad \int_{0}^{1}\frac{36\,du}{(2u+1)^3}$$

$$\int_{1}^{2}\left(w+\frac{1}{w^2}\right)dw \qquad \int_{0}^{1/2}x^3\,(1+9x^4)^{-3/2}\,dx \qquad \int_{0}^{\pi}\sin 5r\,dr$$

$$\int_{0}^{\pi}\cos 3\varphi\,d\varphi \qquad \int_{0}^{3\pi}\cos^2\left(\frac{x}{6}\right)dx \qquad \int_{0}^{\pi}\tan^2\left(\frac{\theta}{3}\right)d\theta$$

$\int_0^{\pi/2} 5\,(\sin\, x)^{3/2}\, \cos\, x\; dx$ $\int_{\pi/2}^{\pi/2} 15\,(\sin\, 3x)^4\, \cos\, 3x\; dx$ $\int_{-1}^{1} 2x\, \sin\,(1-x^2)\; dx$

$\int_2^3 \frac{dv}{v^3-2v^2+v}$ $\int_1^4 \frac{1+\sqrt{z}}{z^2}\; dz$ $\int_{-1}^{1} \frac{a^2}{a^2+y^2}\; dy \quad (a \neq 0)$

$\int_0^{\pi/2} x\, \cos\, x\; dx$ $\int_{-3\pi/2}^{-2\pi} \cos^6 x\; dx$ $\int_0^1 \ln\,(a+x)\; dx \quad (a>0)$

$\int_0^{\pi/2} \frac{dx}{2+\cos\, x}\; dx$ $\int_0^2 x^2\sqrt{4-x^2}\; dx$ $\int_0^1 \frac{dx}{1+\sqrt{x}}$

$\int_0^{\ln\, 2} \sqrt{e^x-1}\; dx$ $\int_e^{e^2} \frac{dx}{x\, \ln\, x}$ $\int_0^{\pi/8} \sin^3(4x)\; dx$

3. Find the area of the region between the graph of f and the x–axis.

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

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

4. Find the average value of

$f(x) = \sqrt{3x}$ over $\langle 0, 3\rangle$ $f(x) = \sqrt{ax}$ over $\langle 0, a\rangle$

$f(x) = mx + b$ over $\langle -1, 1\rangle$ $f(x) = mx + b$ over $\langle -k, k\rangle$

5. Evaluate the following improper integrals.

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

$\int_2^{+\infty} \frac{dx}{x^2-1}$ $\int_0^{+\infty} y^2\, e^{-x}\; dx$ $\int_0^5 \frac{1}{\sqrt{x}}\; dx$

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

6. Evaluate $F'(x)$ if function F is defined by the following integrals.

$F(x) = \int_{1/x}^{\sqrt{x}} \cos\,(t^2)\; dt\,; \quad x > 0$ $F(x) = \int_0^{2x} \frac{\sin\, x}{x}\; dx$

$F(x) = \int_x^0 \sqrt{1+t^4}\; dt$ $F(x) = \int_{x^2}^{x^3} \ln\, t\; dt\,; \quad x > 0$

III. Riemann Integral of a Function of Two and Three Variables

III.1. The double integral – motivation and definition. The two–dimensional Jordan measure and measurable sets in $\mathbf{E}_2$.

III.1.1. Physical motivation. Suppose that we have a thin plate covering the rectangle $R = \langle a, b\rangle \times \langle c, d\rangle$ in the xy–plane. The plate need not be homogeneous and so its planar density (i.e. the amount of mass per unit area – let us denote it $\rho(x, y)$) varies with the position of the point $[x, y]$. We wish to evaluate the mass M of the plate.

If density ρ is constant then $M = \rho \cdot (b - a) \cdot (d - c)$. (Why?)

In a general case when the density is not constant, we can subdivide rectangle R into many smaller pieces $R_1, \dots, R_n$ by a network of lines parallel to the x– and y–axes. If rectangles $R_1, \dots, R_n$ are "small enough" then ρ can be approximated by a constant on each of them. A reasonable value of this constant is $\rho(Z_i)$ where Z_i is some point from R_i. Then the approximate mass of the part of the plate covering rectangle R_i is $\rho(Z_i \cdot \Delta x_i\, \Delta y_i$ where Δx_i and Δy_i are the lengths of sides of R_i. The approximate mass of the whole plate is

$$\sum_{i=1}^{n} \rho(Z_i) \cdot \Delta x_i\, \Delta y_i\, .$$

It is now natural to expect that the exact value of the total mass M of the plate will be equal to the limit of this sum as $n \to +\infty$, and the numbers Δx_i and Δy_i $(i = 1, 2, \dots, n)$ tend to zero.

III.1.2. Geometric motivation. Suppose that $z = f(x, y)$ is a nonnegative function on set $R \in \mathbf{E}_2$ and we wish to define and evaluate the volume V of the region between the graph of function f and the xy–plane. Suppose for simplicity that R is the rectangle $\langle a, b\rangle \times \langle c, d\rangle$.

If f is a constant function on R then the volume is $V = f \cdot m_2(R) = f \cdot (b - a) \cdot (d - c)$.

If f is not a constant function then we can use the same partition of R into n smaller pieces $R_1, \dots, R_n$ as in paragraph III.1.1 and we can approximate the volume of the region between the graph of function f on rectangle R_i and the xy–plane by the number $f(Z_i) \cdot \Delta x_i\, \Delta y_i$ for some $Z_i \in R_i$ $(i = 1, 2, \dots, n)$. The volume of the whole region between the graph of f on set R and the xy–plane can be approximated by the sum

$$\sum_{i=1}^{n} f(Z_i) \cdot \Delta x_i\, \Delta y_i\, .$$

The volume of the region between the graph of function f on R and the xy–plane can now be naturally defined as the limit for $n \to +\infty$ and the numbers Δx_i, Δy_i $(i = 1, 2, \dots, n)$ tending to zero.

III.1.3. Rectangular region in $\mathbf{E}_2$ and its partition. If $\langle a, b\rangle$ is a closed interval on the x–axis and $\langle c, d\rangle$ is a closed interval on the y–axis then the Cartesian product $R = \langle a, b\rangle \times \langle c, d\rangle$ forms a rectangle in $\mathbf{E}_2$. We can subdivide this rectangle by a network of lines parallel to the x– and y–axes into n smaller rectangles $R_1, \dots, R_n$. The system of these smaller rectangles is called the *partition* of rectangle R.

If this partition is named P and if the lengths of sides of smaller rectangles $R_1, \dots, R_n$ are $\Delta x_1, \Delta y_1, \dots, \Delta x_n, \Delta y_n$ then the number which is equal to the maximum of $\Delta x_1, \Delta y_1, \dots, \Delta x_n, \Delta y_n$ is denoted by $\|P\|$ and it is called the *norm* of partition P.

III.1.4. Riemann sums and their limit. Let $z = f(x, y)$ be a bounded function on a bounded set $D \subset \mathbf{E}_2$. Let R be the smallest rectangle in $\mathbf{E}_2$ whose sides are parallel to the x– and y–axes and which contains D. Let P be a partition of R to smaller rectangles $R_1, \dots R_n$ whose lengths of sides are $\Delta x_1, \Delta y_1, \dots, \Delta x_n, \Delta y_n$. The smaller rectangles can be numbered so that those of them which are inside D are $R_1, \dots, R_m$. (See Fig. 3.) Let V be a system of points $Z_i \in R_i$ $(i = 1, 2, \dots, m)$. Then the *Riemann sum* of function f on set D corresponding to partition P and system V is

$$s(f, P, V) = \sum_{i=1}^{m} f(Z_i) \cdot \Delta x_i \, \Delta y_i \, .$$

We say that number S is the *limit of the Riemann sums* $s(f, P, V)$ *as* $\|P\| \to 0+$ if to every given $\epsilon > 0$ there exists $\delta > 0$ such that for every partition P of R and for every choice of V, $\|P\| < \delta$ implies $|s(f, P, V) - S| < \epsilon$. We write:

$$\lim_{\|P\| \to 0+} s(f, P, V) = S. \qquad \text{(III.1)}$$

III.1.5. The double integral. If the limit in (III.1) exists, then function f is called *integrable* in set D and S is called the *double integral* of function f on D. The integral is usually denoted as

$$\iint_D f(x, y) \, dx \, dy \qquad \text{or} \qquad \iint_D f \, dx \, dy.$$

III.1.6. Remark. It follows from paragraph III.1.2 and from the definition of the double integral that if the function $y = f(x, y)$ is nonnegative and integrable on set $D \in \mathbf{E}_2$ then the integral $\iint_D f\, dx\, dy$ defines and evaluates the volume of the region between the graph of f on D and the xy–plane. However, you will see in Sections III.5 – III.7 that the volumes of even more general regions in $\mathbf{E}_3$ can also be defined and evaluated by means of volume integrals.

The notion of a bounded set in $\mathbf{E}_2$ is too general for practical purposes. For example, it can be shown that there exist bounded sets $D \in \mathbf{E}_2$ such that even the constant function is not integrable on D. In order to distinguish between these "bad" sets and other "reasonable" sets, we introduce the notion of a so called measurable set.

III.1.7. A measurable set in $\mathbf{E}_2$ and its Jordan measure. Suppose that D is a bounded set in $\mathbf{E}_2$. We say that this set is *measurable* (in the sense of Jordan) if the constant function $f(x, y) = 1$ is integrable on D. In this case, we call the number

$$m_2(D) \;=\; \iint_D dx\, dy$$

the *two–dimensional Jordan measure* of set D.

$m_2(D)$ has a very simple geometric interpretation – it defines and evaluates the *area of set D*.

It is important to have a criterion which enables us easily to recognize some simple measurable sets. We will give such a criterion in paragraph III.1.10. However, we first list some sets whose measure is zero.

III.1.8. Some sets whose two–dimensional Jordan measure is zero. It can be proved for example that the following sets in $\mathbf{E}_2$ have the measure equal to zero:

a) Sets consisting of a finite number of points.

b) Graphs of continuous functions $y = \varphi(x)$ or $x = \psi(y)$ on closed bounded intervals.

c) So called simple smooth curves, respectively simple piecewise–smooth curves (see Section IV.1).

The next theorem is quite obvious, and it also concerns sets of measure zero.

III.1.9. Theorem. a) *If $N_1, N_2, \ldots, N_n$ are sets whose measure is zero then* $m_2\left(\cup_{i=1}^{n} N_i\right) = 0$.

b) *If $M \subset N$ and $m_2(N) = 0$ then $m_2(M) = 0$.*

III.1.10. Theorem. (Sufficient and necessary condition for measurability of a set in $\mathbf{E}_2$.) *A bounded set $D \subset \mathbf{E}_2$ is measurable if and only if $m_2(\partial D) = 0$ (where ∂D is the boundary of D).*

III.2. Existence and important properties of the double integral.

The two statements "f is integrable on set D" and "the double integral $\iint_D f\,dx\,dy$ exists" say exactly the same.

III.2.1. Existence theorem for the double Integral. *Let D be a measurable set in $\mathbf{E}_2$ and let f be a bounded function on D whose set of discontinuities has measure m_2 equal to zero. Then f is integrable on D.*

In particular, if D is a measurable set and f is a bounded continuous function on D then f is integrable on D.

III.2.2. Important properties of the double integral. The double integral has many properties which are exactly the same as the properties of the one–dimensional Riemann integral explained in paragraphs III.3.1 – III.3.5. Let us mention only several of them:

a) **(Linearity of the double integral.)** *If functions f and g are integrable on set $D \subset \mathbf{E}_2$ and $\alpha \in \mathbf{R}$ then*

$$\iint_D (f+g)\,dx\,dy = \iint_D f\,dx\,dy + \iint_D g\,dx\,dy,$$

$$\iint_D \alpha \cdot f\,dx\,dy = \alpha \cdot \iint_D f\,dx\,dy.$$

b) **(Additivity of the double integral with respect to the set.)** *If D_1 and D_2 are measurable sets such that $m_2\Big(D_1 \cap D_2\Big) = 0$ (i.e. D_1 and D_2 are not overlapping) and if f is integrable on D_1 and on D_2 then*

$$\iint_{D_1} f\,dx\,dy + \iint_{D_2} f\,dx\,dy = \iint_{D_1 \cup D_2} f\,dx\,dy.$$

c) *If function f is integrable on set $D \in \mathbf{E}_2$ and function g differs from f at most on a set whose measure is zero then g is also integrable on D and*

$$\iint_D g\,dx\,dy = \iint_D f\,dx\,dy.$$

d) *If $D \subset \mathbf{E}_2$ and $m_2(D) = 0$ then* $\iint_D f\,dx\,dy = 0$ *for every function f.*

Proposition c) shows that the behaviour of the integrated function on a set of measure zero does not affect the existence and the value of the double integral. Thus, from the point of view of integration on set D, whose boundary has measure zero, it is not important whether D is considered open (i.e. without its boundary) or closed (i.e. with its boundary).

III.3. Evaluation of the double integral – Fubini's theorem and transformation to the polar coordinates.

Fubini's theorem transforms the evaluation of a double integral to the computation of two single (= one–dimensional) integrals. It can be applied if the domain of integration is a so called elementary region.

III.3.1. Elementary region in $\mathbf{E}_2$. a) Let $y = \phi_1(x)$ and $y = \phi_2(x)$ be continuous functions on the interval $\langle a, b\rangle$ and let $\phi_1(x) \leq \phi_2(x)$ for all $x \in \langle a, b\rangle$. Then the set

$$D = \big\{[x,y] \in \mathbf{E}_2;\ a \leq x \leq b,\ \phi_1(x) \leq y \leq \phi_2(x)\big\}$$

is called the *elementary region relative to the x–axis*.
b) Let $x = \psi_1(y)$ and $x = \psi_2(y)$ be continuous functions on the interval $\langle c, d\rangle$ and let $\psi_1(y) \leq \psi_2(y)$ for all $y \in \langle c, d\rangle$. Then the set

$$D = \big\{[x,y] \in \mathbf{E}_2;\ c \leq y \leq d,\ \psi_1(y) \leq x \leq \psi_2(y)\big\}$$

is called the *elementary region relative to the y–axis*.

Elementary regions are measurable sets in $\mathbf{E}_2$. Let us now explain the idea of integrating of function $z = f(x,y)$ on the elementary region relative to the x–axis (see Fig. 4a). Imagine that we can cut the region into infinitely many infinitely narrow vertical stripes. One such stripe is the line segment PQ on Fig. 4a. We first integrate f on each such segment as a function of one variable y – we obtain $F(x) = \int_{\phi_1(x)}^{\phi_2(x)} f(x,y)\,dy$. Certainly, this depends on x because the position of the line segment PQ depends on x. Then we integrate $F(x)$ as a function of x from a to b. Thus, we obtain formula (III.2) (see the next paragraph III.3.2).

The next theorem precisely formulates the assumptions under which we can apply the described method, and it also treats the case when D is an elementary region relative to the y–axis.

III.3.2. Fubini's theorem for the double integral. a) *Let D be the elementary region relative to the x–axis from paragraph III.3.1. Let the function $z = f(x,y)$ be continuous on D. Then*

$$\iint_D f(x,y)\,dx\,dy \;=\; \int_a^b \Big(\int_{\phi_1(x)}^{\phi_2(x)} f(x,y)\,dy\Big)\,dx. \tag{III.2}$$

b) *Let D be the elementary region relative to the y–axis from paragraph III.3.1. Let function $z = f(x,y)$ be continuous on D. Then*

$$\iint_D f(x,y)\,dx\,dy \;=\; \int_c^d \Big(\int_{\psi_1(y)}^{\psi_2(y)} f(x,y)\,dx\Big)\,dy. \tag{III.3}$$

III.3.3. Example Evaluate the integral $\iint_D (2x+3y+5)\,dx\,dy$ where D is the region bounded by the curves $y = \frac{1}{2}\,x$, $y = 1/x$ and $x = \sqrt{2}$.

The given curves divide the xy–plane into various regions (Sketch a figure!) but only one of them is bounded and this is D. It can be described as the set of all points $[x,y] \in \mathbf{E}_2$ such that $\sqrt{2}/2 \le x \le \sqrt{2}$ and $1/x \le y \le 2x$.

D is measurable (because it is bounded and its boundary has the measure equal to zero – see Theorem III.1.11). Function f is continuous on D. Thus, using the Fubini theorem III.3.2, we obtain:

$$\iint_D (2x+3y+5)\,dx\,dy \;=\; \int_{\sqrt{2}/2}^{\sqrt{2}} \Big(\int_{1/x}^{2x} (2x+3y+5)\,dy\Big)\,dx \;=$$

$$= \int_{\sqrt{2}/2}^{\sqrt{2}} \big[2xy + \tfrac{3}{2}y^2 + 5y\big]_{y=1/x}^{y=2x}\,dx \;=\; \int_{\sqrt{2}/2}^{\sqrt{2}} \Big(4x^2 + 6x^2 + 10x - 2 - \frac{3}{2x^2} - \frac{5}{x}\Big)\,dx \;=$$

$$= \Big[\frac{10x^3}{3} + 5x^2 - 2x + \frac{3}{2x} - 5\,\ln\,x\Big]_{\sqrt{2}/2}^{\sqrt{2}} \;=\; \tfrac{79}{12}\sqrt{2} + 7.5 - 5\,\ln\,2.$$

You remember that a powerful method for computation of a one–dimensional integral is the method of substitution. This method can also be used when we evaluate a double integral. When applying it, we usually say that we transform the integral to new coordinates. The most–used new coordinates in $\mathbf{E}_2$ are so called polar (or generalized polar) coordinates.

III.3.4. Polar coordinates in $\mathbf{E}_2$. The position of a point $X \in \mathbf{E}_2$ is uniquely given by its *polar coordinates* r, φ whose geometric meaning is the following: r is the distance of X from the origin O and φ is the angle between the positive part of the x–axis and the line segment OX. φ is measured from the x–axis towards the line segment OX. (Sketch a figure!) The relation between the Cartesian coordinates x, y and the polar coordinates r, φ is given by the equations

$$x \;=\; r\,\cos\,\varphi, \quad y \;=\; r\,\sin\,\varphi. \tag{III.4}$$

III.3.5. Transformation of the double integral to the polar coordinates. Suppose that we have to evaluate the integral $\iint_D f(x,y)\,dx\,dy$. We can use the equations (III.4) and replace x, y by $r\,\cos\,\varphi$, respectively $r\,\sin\,\varphi$. However, we must also

a) change D (analogously to the change of the limits in the one–dimensional integral if we apply the method of substitution – see Theorem II.4.6),

b) substitute for the term $dx\,dy$ in the integral (analogously to the equation $dx = g'(t)\,dt$ if we use the substitution $x = g(t)$ in a one–dimensional integral).

Set D corresponds to some other set D' in the polar coordinates. Optimally, every point $[x, y] \in D$ should have just one opposite point $[r, \varphi] \in D'$ (such that $x = r\cos\varphi$ and $y = r\sin\varphi$). However, since the sets of measure zero play no role, the one–to–one correspondence between the points $[x, y] \in D$ and the points $[r, \varphi] \in D'$ can be disturbed on a set of measure zero, both on the side of D and on the side of D'.

It can be proved that $dx\,dy$ must be substituted in this way:

$$dx\,dy = r\,dr\,d\varphi. \tag{III.5}$$

The factor r on the right hand side is a so called "Jacobian" (the abbreviation of the "Jacobi determinant") and you will find more about it in Section III.9.

The transformation of the double integral to the polar coordinates has a sense if it leads either to the simplification of the integrand (see example III.3.6) or to the simplification of the domain of integration (see example III.3.7). It follows from the geometric sense of the polar coordinates that $r \geq 0$ and φ can be taken from an interval whose length does not exceed 2π (i.e. the interval $(0, 2\pi)$).

III.3.6. Example. Evaluate the integral $\iint_D (x+y)\,dx\,dy$ where $D = \{[x, y] \in \mathbf{E}_2;\ x > 0,\ y > 0,\ x^2 + y^2 < 4\}$.

D is the intersection of the disk (with the center at the origin and radius 2) and the first quadrant. It corresponds to the domain $D' = \{[r, \varphi] \in \mathbf{E}_2;\ 0 < r < 2,\ 0 < \varphi < \pi/2\}$ in the polar coordinates. Thus, using the transformation (III.4), (III.5), we obtain:

$$\iint_D (x+y)\,dx\,dy = \iint_{D'} (r\cos\varphi + r\sin\varphi)\,r\,dr\,d\varphi =$$

$$= {}^{1)} \int_0^2 \Big(\int_0^{\pi/2} r^2 (\cos\varphi + \sin\varphi)\,d\varphi \Big)\,dr = \int_0^2 r^2 \big[\sin\varphi - \cos\varphi\big]_0^{\pi/2}\,dr =$$

$$= \int_0^2 2r^2\,dr = \tfrac{16}{3}.$$

[1]) We have applied Fubini's theorem.

III.3.7. Example. Evaluate the integral $\iint_D (x^2+y^2)^{-1/2}\,dx\,dy$ where D is a triangle with the vertices $[0, 0]$, $[1, 0]$, $[1, 1]$.

D can also be described as the set of all points $[x, y]$ such that $0 < x < 1$ and $0 < y < x$. Transforming these inequalities to the polar coordinates, we obtain

$$0 < r\cos\varphi < 1, \quad 0 < r\sin\varphi < r\cos\varphi. \tag{III.6}$$

The second inequality implies: $0 < \sin\varphi < \cos\varphi$ which means that $0 < \varphi < \pi/4$. The first inequality in (III.6) implies: $0 < r < 1/\cos\varphi$. Hence D corresponds to the set $D' = \{[r, \varphi] \in \mathbf{E}_2; 0 < \varphi < \pi/4,\ 0 \leq r \leq 1/\cos\varphi\}$ in the polar coordinates. Thus, using the transformation (III.4), (III.5) and afterwards applying Fubini's theorem, we obtain:

$$\iint_D \frac{1}{\sqrt{x^2+y^2}}\,dx\,dy \;=\; \iint_{D'} \frac{1}{r}\,r\,dr\,d\varphi \;=\; \int_0^{\pi/4}\Big(\int_0^{1/\cos\varphi} dr\Big)\,d\varphi \;=$$

$$= \int_0^{\pi/4} \frac{1}{\cos\varphi}\,d\varphi \;=\; \int_0^{\pi/4} \frac{\cos\varphi}{1-\sin^2\varphi}\,d\varphi \;=\; ^{2)} \int_0^{\sqrt{2}/2} \frac{dt}{1-t^2} \;=$$

$$= \tfrac{1}{2}\int_0^{\sqrt{2}/2}\Big(\frac{1}{1-t}+\frac{1}{1+t}\Big)\,dt \;=\; \tfrac{1}{2}\big[-\ln(1-t)+\ln(1+t)\big]_0^{\sqrt{2}/2} \;=\; \tfrac{1}{2}\ln\frac{2+\sqrt{2}}{2-\sqrt{2}}.$$

[2]) We have used the substitution $\sin\varphi = t$, $\cos\varphi\,d\varphi = dt$.

III.3.8. Generalized polar coordinates in $\mathbf{E}_2$. We shall denote these coordinates again by r, φ. They are analogous to the polar coordinates, though their origin need not be the same as the origin of the Cartesian coordinates and they are not "isotropic", i.e. the rate of change of r is different in the x–direction and in the y–direction. The relation between the Cartesian coordinates x, y and the generalized polar coordinates r, φ is

$$x \;=\; x_0 + ar\cos\varphi, \quad y \;=\; y_0 + br\sin\varphi \tag{III.7}$$

where $[x_0, y_0]$ is a given point in $\mathbf{E}_2$ and a, b are positive constants.

By analogy with (III.5), it can be proved that if we transform a double integral to the generalized polar coordinates then $dx\,dy$ must be substituted in this way:

$$dx\,dy \;=\; rab\;dr\,d\varphi. \tag{III.8}$$

The factor rab on the right hand side is again the "Jacobian", and it will be explained in Section III.9.

The transformation of a double integral to the generalized polar coordinates usually simplifies the integral if the integrand depends on x and y through the expression $(x-x_0)^2/a^2 + (y-y_0)^2/b^2$ or if the domain of integration is the interior of an ellipse $(x-x_0)^2/a^2 + (y-y_0)^2/b^2 = 1$ or a sector of an ellipse.

III.3.9. Example. Evaluate the integral $\iint_D x\,dx\,dy$ where $D = \{[x,y] \in \mathbf{E}_2;\ (x-2)^2 + (y-1)/4^2 \le 1\}$.

We can observe that if we use the transformation

$$x = 2 + r\cos\varphi, \quad y = 1 + 2r\sin\varphi \tag{III.9}$$

then the points $[x,y]$ fill up D if and only if the points $[r,\varphi]$ fill up the set $D' = \{[r,\varphi] \in \mathbf{E}_2;\ 0 \le r \le 1,\ 0 \le \varphi < 2\pi\}$. Using transformation (III.9), equality $dx\,dy = 2r\,dr\,d\varphi$ (following from (III.8)) and also applying Fubini's theorem, we get:

$$\iint_D x\,dx\,dy \;=\; \iint_{D'} (2 + r\cos\varphi)\,dr\,d\varphi \;=\; \int_0^{2\pi}\Big(\int_0^1 (2 + r\cos\varphi)\,dr\Big)\,d\varphi \;=$$

$$= \int_0^{2\pi} \big[2r + \tfrac{1}{2}r^2\cos\varphi\big]_0^1\,d\varphi \;=\; \int_0^{2\pi} (2 + \tfrac{1}{2}\cos\varphi)\,d\varphi \;=\; 4\pi.$$

III.3.10. Remark. In fact, transformation (III.9) is not a one–to–one mapping of set D' onto set D in example III.3.8. The one–to–one correspondence is disturbed on the subset $D'_0 = \{[r, \varphi] \in \mathbf{E}_2;\ r = 0,\ 0 \le \varphi < 2\pi\}$ of D'. (You can observe that D'_0 is a subset of the boundary of D'.) This is clear, because transformation (III.9) maps all points of D'_0 onto the point $[2, 1]$ in D. Thus, the point $[2, 1]$ in D has infinitely many opposite points in D' – i.e. all points of D'_0. However, since $m_2(D'_0) = 0$, this does not affect the existence and the value of the integral.

III.4. Some physical applications of the double integral.

Suppose that a two–dimensional thin plate coincides with a measurable set D in the xy–plane. The plane need not be homogeneous. It means that its planar density (amount of mass per unit of area) need not be constant. Let the planar density be given by function $\rho(x, y)$. The double integral enables us to define and evaluate some fundamental mechanical characteristics of the plate. Suppose that ρ is expressed in $[\mathrm{kg} \cdot \mathrm{m}^{-2}]$. Then we have:

Mass $\quad M = \iint_D \rho(x, y)\, dx\, dy \quad [\mathrm{kg}]$,

Static moment about the x–axis $\quad M_x = \iint_D y \cdot \rho(x, y)\, dx\, dy \quad [\mathrm{kg} \cdot \mathrm{m}]$,

Static moment about the y–axis $\quad M_y = \iint_D x \cdot \rho(x, y)\, dx\, dy \quad [\mathrm{kg} \cdot \mathrm{m}]$,

Center of mass $[x_m, y_m] \quad x_m = \dfrac{M_y}{M}, \quad y_m = \dfrac{M_x}{M} \quad [\mathrm{m}]$,

Moment of inertia about the x–axis $\quad J_x = \iint_D y^2 \cdot \rho(x, y)\, dx\, dy \quad [\mathrm{kg} \cdot \mathrm{m}^2]$,

Moment of inertia about the y–axis $\quad J_y = \iint_D x^2 \cdot \rho(x, y)\, dx\, dy \quad [\mathrm{kg} \cdot \mathrm{m}^2]$,

Moment of inertia about the origin $\quad J_0 = \iint_D (x^2 + y^2) \cdot \rho(x, y)\, dx\, dy \quad [\mathrm{kg} \cdot \mathrm{m}^2]$.

Suggest a formula for the moment of inertia about a general straight line in $\mathbf{E}_2$ whose equation is $ax + by + c = 0$!

III.5. The volume integral – motivation and definition. The three–dimensional Jordan measure and measurable sets in $\mathbf{E}_3$.

The theory of the volume integral is almost identical with the theory of the double integral. The main difference lies in the simple fact that we have one more dimension. Thus, we can repeat almost everything that was written about the double integral. The same holds for the three–dimensional Jordan measure. This is why we present the theory of the volume integral very briefly and we do not explain the details.

On the other hand, since one more dimension causes higher variety of possible domains of integration as well as integrated functions, you will see that the methods of evaluation of the volume integral, though their techniques are again based on Fubini's theorem and on the transformation to other coordinates, are usually technically more complicated than in the case of the double integral.

III.5.1. Physical motivation. Suppose that we have a body whose density is $\rho(x, y, z)$. We wish to evaluate the mass M of the body. Suppose for simplicity that the body has the form of the block $B = \langle a, b\rangle \times \langle c, d\rangle \times \langle r, s\rangle$.

If density ρ is constant then $M = \rho \cdot (b - a) \cdot (d - c) \times (s - r)$. (Why?)

However, in a general situation when the density is not constant, we can subdivide B into n smaller rectangular cells $B_1, \ldots, B_n$ by planes parallel to the coordinate planes xy, xz and yz. If the cells B_1 are "small enough" then we can approximate ρ by a constant on each of them. A possible value of this constant is $\rho(Z_i)$ where Z_i is some point from the cell R_i. Then the approximate mass of cell R_i is $\rho(Z_i) \cdot \Delta x_i\, \Delta y_i\, \Delta z_i$ where Δx_i, Δy_i and Δz_i are the lengths of sides of R_i. The approximate mass of the whole body is

$$\sum_{i=1}^{n} \rho(Z_i) \cdot \Delta x_i\, \Delta y_i\, \Delta z_i.$$

The exact value of the mass M of the body is equal to the limit of this sum as $n \to +\infty$ and the numbers Δx_i, Δy_i, Δz_i $(i = 1, 2, \ldots, n)$ tend to zero.

III.5.2. A block in $\mathbf{E}_3$ and its partition. If $\langle a, b\rangle$ is a bounded closed interval on the x–axis, $\langle c, d\rangle$ is a bounded closed interval on the y–axis and $\langle r, s\rangle$ is a closed bounded interval on the z–axis then the set $B = \langle a, b\rangle \times \langle c, d\rangle \times \langle r, s\rangle$ forms a block in $\mathbf{E}_3$. We can subdivide this block to n rectangular cells $B_1, \ldots, B_n$ by planes parallel to the xy–plane, xz–plane and yz–plane. The system of these cells is called the *partition* of B.

If this partition is named P and if the lengths of sides of smaller cells $B_1, \ldots,$ B_n are Δx_1, Δy_1, $\Delta z_1, \ldots, \Delta x_n$, Δy_n, Δz_n then the maximum of all these lengths is denoted by $\|P\|$ and it is called the *norm* of partition P.

III.5.3. Riemann sums and their limit. Let $u = f(x, y, z)$ be a bounded function on a bounded set $D \subset \mathbf{E}_3$. Let B be the smallest block in $\mathbf{E}_2$ whose sides are parallel to the xy–, xz– and yz–planes and which contains D. Let P be a partition of B into smaller rectangular cells $B_1, \ldots B_n$ whose lengths of sides are Δx_1, Δy_1, $\Delta z_1, \ldots, \Delta x_n$, Δy_n, Δz_n. These smaller cells can be numbered so that those of them which are inside D are $B_1, \ldots, B_m$. Let V be a system of points $Z_i \in B_i$ $(i = 1, 2, \ldots, m)$. Then the *Riemann sum* of function f on set D corresponding to partition P and system V is

$$s(f, P, V) = \sum_{i=1}^{m} f(Z_i) \cdot \Delta x_i\, \Delta y_i\, \Delta z_i.$$

We say that number S is the *limit of the Riemann sums* $s(f,P,V)$ *as* $\|P\| \to 0+$ if to every given $\epsilon > 0$ there exists $\delta > 0$ such that for every partition P of B and for every choice of V, $\|P\| < \delta$ implies $|s(f,P,V) - S| < \epsilon$. We write:

$$\lim_{\|P\| \to 0+} s(f,P,V) \;=\; S. \tag{III.10}$$

III.5.4. The volume integral. If the limit in (III.10) exists, then function f is called *integrable* in set D and S is called the *volume integral* of function f on D. The integral is usually denoted as

$$\iiint_D f(x,y)\, dx\, dy\, dz \qquad \text{or} \qquad \iiint_D f\; dx\, dy\, dz.$$

You may also find another name for the volume integral in literature: the *triple integral*.

III.5.5. A measurable set in $\mathbf{E}_3$ and its Jordan measure. Suppose that D is a bounded set in $\mathbf{E}_3$. We say that D is *measurable* (in the sense of Jordan) if the constant function $f(x,y,z) = 1$ is integrable on D. In this case, we call the number

$$m_3(D) \;=\; \iiint_D dx\, dy\, dz$$

the *three–dimensional Jordan measure* of set D.

$m_3(D)$ has an important geometric meaning – it defines and evaluates the *volume of set D*.

III.5.6. Some sets whose three–dimensional Jordan measure is zero. The following sets in $\mathbf{E}_3$ have the measure equal to zero:

a) Sets consisting of a finite number of points or bounded curves.

b) Graphs of continuous functions $z = \varphi(x,y)$ or $y = \psi(x,z)$ or $x = \eta(y,z)$ defined on bounded measurable sets in $\mathbf{E}_2$.

c) So called simple smooth surfaces, respectively simple piecewise–smooth surfaces (see Section V.1).

Usually, if M is a set in $\mathbf{E}_k$ (with $k = 1$, 2 or 3) and we say that M has measure zero, we mean that the k–dimensional measure of M is zero, i.e. $m_k(M) = 0$.

III.5.7. Theorem. a) *If N_1, N_2, …, N_n are sets in $\mathbf{E}_3$ whose measure is zero then* $m_3\Big(\cup_{i=1}^n N_i\Big) = 0$.

b) *If $M \subset N$ and $m_3(N) = 0$ then $m_3(M) = 0$.*

III.5.8. Theorem. (A sufficient and necessary condition for measurability of a set in $\mathbf{E}_3$.) *A bounded set $D \subset \mathbf{E}_3$ is measurable if and only if $m_3(\partial D) = 0$ (where ∂D is the boundary of D).*

III.6. Existence and important properties of the volume integral.

The two statements "f is integrable on set D" and "the volume integral $\iint_D f\, dx\, dy$ exists" say exactly the same.

III.6.1. Existence theorem for the volume Integral. *Let D be a measurable set in $\mathbf{E}_3$ and let f be a bounded function on D whose set of discontinuities has measure m_3 equal to zero. Then f is integrable on D.*

Specifically, if D is a measurable set in $\mathbf{E}_3$ and f is a bounded continuous function on D then f is integrable on D.

III.6.2. Important properties of the volume integral.

a) **(Linearity of the volume integral.)** *If functions f and g are integrable on set $D \subset \mathbf{E}_3$ and $\alpha \in \mathbf{R}$ then*

$$\iiint_D (f+g)\, dx\, dy\, dz \;=\; \iiint_D f\, dx\, dy\, dz + \iiint_D g\, dx\, dy\, dz,$$

$$\iiint_D \alpha \cdot f\, dx\, dy\, dz \;=\; \alpha \cdot \iiint_D f\, dx\, dy\, dz.$$

b) **(Additivity of the volume integral with respect to the set.)** *If D_1 and D_2 are measurable non–overlapping sets in $\mathbf{E}_3$ (i.e. $m_3\big(D_1 \cap D_2\big) = 0$) and f is integrable on D_1 and D_2 then*

$$\iiint_{D_1} f\, dx\, dy\, dz + \iiint_{D_2} f\, dx\, dy\, dz \;=\; \iiint_{D_1 \cup D_2} f\, dx\, dy\, dz.$$

c) *If function f is integrable on set $D \in \mathbf{E}_3$ and function g differs from f at most on a set whose measure is zero, then g is also integrable on D and*

$$\iiint_D g\, dx\, dy\, dz \;=\; \iiint_D f\, dx\, dy\, dz.$$

d) *If $D \subset \mathbf{E}_3$ and $m_3(D) = 0$ then* $\displaystyle\iiint_D f\, dx\, dy\, dz = 0$ *for every function f.*

Thus, the behaviour of the integrated function on a set of measure zero does not affect the existence and the value of the volume integral.

III.7. Evaluation of the volume integral – Fubini's theorem and transformation to the cylindrical and to the spherical coordinates.

Fubini's theorem for the volume integral transforms the evaluation of the integral to the computation of one single and one double integral. It can be applied if the domain of integration D is a so called elementary region:

III.7.1. Elementary region in $\mathbf{E}_3$. a) Let D_{xy} be a measurable closed set in $\mathbf{E}_2$ and $z = \phi_1(x, y)$ and $z = \phi_2(x, y)$ be continuous functions on D_{xy} such that $\phi_1(x, y) \le \phi_2(x, y)$ for all $[x, y] \in D_{xy}$. Then the set

$$D = \big\{[x, y, z] \in \mathbf{E}_3;\ [x, y] \in D_{xy},\ \phi_1(x, y) \le z \le \phi_2(x, y)\big\}$$

is called the <u>*elementary region relative to the xy–plane*</u>. (See Fig. 5.)

We can also define quite analogously an elementary region relative to the xz–plane and an elementary region relative to the yz–plane. Try to write down these definitions for yourself!

Elementary regions are measurable sets in $\mathbf{E}_3$. The idea of integrating function $f(x, y, z)$ on the elementary region relative to the xy–plane is the following: Imagine that we cut the region into infinitely many vertical line segments. One of them is the line segment PQ in Fig. 5. We first integrate f on each such segment as a function of one variable z – we obtain $F(x, y) = \int_{\phi_1(x,y)}^{\phi_2(x,y)} f(x, y, z)\, dz$. This depends on x and y because the position of the line segment PQ depends on x and y. Then we integrate $F(x, y)$ as a function of x and y on set D_{xy}. Thus, we obtain formula (III.11) (see the next paragraph III.7.2).

III.7.2. Fubini's theorem for the volume integral. *Let D be the elementary region relative to the xy–plane from paragraph III.7.1. Let function $u = f(x, y, z)$ be continuous on D. Then*

$$\iiint_D f(x, y, z)\, dx\, dy\, dz = \iint_{D_{xy}} \Big(\int_{\phi_1(x,y)}^{\phi_2(x,y)} f(x, y, z)\, dz \Big)\, dx\, dy. \qquad \text{(III.11)}$$

Formulate for yourself analogous theorems for the integration on the elementary region relative to the xz–plane and the elementary region relative to the yz–plane!

III.7.3. Example Evaluate the integral $\iiint_D (z + 2)\ dx\, dy\, dz$ where D is the region in $\mathbf{E}_3$ bounded by the surfaces $x^2 + y^2 = 2$, $z = -2 - x$, $z = 2 + y$.

The given surfaces divide $\mathbf{E}_3$ into more regions, but only one of them is bounded and this is D. It is a part of the cylinder $x^2 + y^2 \le 2$ bounded by the plane $z = -2 - x$

from below and by the plane $z = 2+y$ from above. Thus, D is the elementary region relative to the xy–plane with $D_{xy} = \{[x, y] \in \mathbf{E}_2;\ x^2+y^2 \leq 2\}$ and $\phi_1(x, y) = -2-x$, $\phi_2(x, y) = 2 + y$. Applying Theorem III.7.2, we obtain

$$\iiint_D (z+2)\ dx\, dy\, dz \ = \ \iint_{D_{xy}} \Big(\int_{-2-x}^{2+y} (z+2)\ dz \Big)\ dx\, dy \ =$$

$$= \ \iint_{x^2+y^2<2} \big[z^2/2 + 2z \big]_{-2-x}^{2+y}\ dx\, dy \ =$$

$$= {}^{1)} \int_0^{\sqrt{2}} \Big(\int_0^{2\pi} \big(\tfrac{1}{2}\, r^2 \sin^2\varphi - \tfrac{1}{2}\, r^2 \cos^2\varphi + 4r \sin\varphi - 4r \cos\varphi + 8 \big)\, r\, d\varphi \Big)\ dr \ =$$

$$= \int_0^{\sqrt{2}} 16\pi\, r\ dr \ = \ 16\pi.$$

[1]) We transform the double integral on D_{xy} to the polar coordinates.

III.7.4. Remark. Fubini's theorem III.7.2 transforms the volume integral to the composition of the two integrals – the outside double integral and the inside single integral. It is sometimes quite useful to do this conversely, i.e. to transform the volume integral to the outside single integral and the inside double integral. We allow ourselves to omit the corresponding theory (because it is quite analogous to the contents of paragraphs III.7.1 and III.7.2) and we show this procedure in the following example.

III.7.5. Example. Evaluate the volume V of the oblique cone $C = \{[x, y, z] \in \mathbf{E}_3;\ 0 < z < 5,\ (x-2z)^2 + y^2 < z^2\}$.

The volume of C is

$$V \ = \ \iiint_C dx\, dy\, dz \ = \ \int_0^5 \Big(\iint_{(x-2z)^2+y^2<z^2} dx\, dy \Big)\ dz \ =$$

$$= {}^{2)} \int_0^5 \Big(\int_0^z \Big(\int_0^{2\pi} rz\ d\varphi \Big)\ dr \Big)\ dz \ = \ \int_0^5 \pi\, z^3\, dz \ = \ \frac{625\,\pi}{4}.$$

[2]) The inside double integral is transformed from the Cartesian coordinates x, y to the generalized polar coordinates r, φ by the equations $x = 2z + r\cos\varphi$, $y = r\sin\varphi$.

III.7.6. Cylindrical coordinates in $\mathbf{E}_3$. The cylindrical coordinates of the point $X = [x, y, z] \in \mathbf{E}_3$ are r, φ, w whose geometric meaning is the following: r, φ are the polar coordinates of the point $[x, y]$ in the xy–plane and $w = z$. Thus, the relation between the Cartesian coordinates x, y, z of point X and its cylindrical coordinates are:

$$x = r\cos\varphi, \quad y = r\sin\varphi, \quad z = w. \tag{III.12}$$

When transforming a volume integral to the cylindrical coordinates, we must also substitute for $dx\,dy\,dz$. It can be proved that the right substitution is

$$dx\,dy\,dz = r\;dr\,d\varphi\,dw. \tag{III.13}$$

(See Section III.9 for more details.)

It follows from the geometric sense of r, φ and w that $r \geq 0$, φ can be taken from any interval whose length is at most 2π and $w \in \mathbf{R}$. The transformation of a volume integral to the cylindrical coordinates usually simplifies the integral if the domain of integration is a cylinder (or a sector of a cylinder) or if the integrand depends on x and y mainly through the expression $x^2 + y^2$.

The transformation of the volume integral $\iiint_D f\;dx\,dy\,dz$ leads to another integral, in variables r, φ, w, on a set D'. Optimally, equations (III.12) should define a one–to–one mapping of D' onto D. Nevertheless, since the behaviour of the integrand on a set of measure zero (the three–dimensional measure m_3 because we are dealing with the volume integral) does not affect the integral, the one–to–one correspondence between the points of D' and D can be disturbed on a set of measure zero both on the side of D' and on the side of D. This is also valid for transformations to other coordinates (spherical, generalized cylindrical, generalized spherical, etc.) and so we will not deal with it in detail in this section.

III.7.7. Example. Find the volume of the region D bounded below by the plane $z = 0$, laterally by the circular cylinder $x^2 + (y-1)^2 = 1$, and above by the paraboloid $z = x^2 + y^2$.

The volume of D is identical with the three–dimensional Jordan measure of D (see paragraph III.5.5). So we denote it by $m_3(D)$. It is defined by the integral $\iiint_D dx\,dy\,dz$. Let us evaluate this integral by the transformation to the cylindrical coordinates.

Region D is a set of points $[x,y,z] \in \mathbf{E}_3$ such that $0 \leq z \leq x^2 + y^2$ and $x^2 + (y-1)^2 \leq 1$. The last inequality is equivalent to $x^2 + y^2 - 2y \leq 0$. Transforming it to the cylindrical coordinates r, φ and w, we obtain

$$r^2 - 2r\sin\varphi \leq 0,$$
$$r \leq 2\sin\varphi.$$

The inequalities $0 \leq z \leq x^2 + y^2$ are equivalent to

$$0 \leq w \leq r^2.$$

These inequalities give the limits of integration with respect to r and w. The φ–limits of integration can be found by means of the orthogonal projection of D on the xy–plane. The projection is the disk D_{xy} with the center $[0,1]$ and radius 1. The angle made by all possible straight lines passing through the origin and entering D_{xy}, measured from the positive part of the x–axis, runs from $\varphi = 0$ to $\varphi = \pi$. Hence the volume is

$$m_3(D) = \iiint_D dx\,dy\,dz = \int_0^{\pi} \int_0^{2\sin\varphi} \int_0^{r^2} r\;dw\;dr\;d\varphi = \int_0^{\pi} 4\sin^4\varphi\;d\varphi = \tfrac{3}{2}\pi.$$

III.7.8. Spherical coordinates in $\mathbf{E}_3$. The spherical coordinates of the point $X = [x, y, z]$ in $\mathbf{E}_3$ are r, φ and ϑ. They have this geometric sense: r is the distance of point X from the origin O. φ is the angle between the line segment OX' (where X' is the orthogonal projection of point X to the xy–plane) and the positive part of the x–axis (measured from the x–axis). ϑ is the angle between the line segment OX' and the line segment OX (measured from the line segment OX'. This geometric interpretation of the spherical coordinates easily leads to the following equations:

$$x = r\,\cos\vartheta\,\cos\varphi, \quad y = r\,\cos\vartheta\,\sin\varphi, \quad z = r\,\sin\vartheta. \qquad \text{(III.14)}$$

When transforming a volume integral from the Cartesian coordinates x, y, z to the spherical coordinates r, φ, ϑ, it is also necessary to transform $dx\,dy\,dz$. It can be proved that

$$dx\,dy\,dz \;=\; r^2\,\cos\vartheta\;dr\,d\varphi\,d\vartheta. \qquad \text{(III.15)}$$

(See Section III.9 for more details.)

It follows from the geometric sense of r, φ and ϑ that $r \geq 0$, φ can be taken from any interval whose length is at most 2π and ϑ can be taken from an interval whose length does not exceed π (usually $(-\pi/2, \pi/2)$). The transformation of a volume integral to the spherical coordinates usually simplifies the integral if the domain of integration is a ball or a sector of a ball (a ball is the interior of a sphere) or if the integrand depends on x, y and z mainly through the expression $x^2 + y^2 = z^2$.

III.7.9. Example. Find the volume of the upper region D cut from the ball $x^2 + y^2 + z^2 \leq 1$ by the cone $z^2 = \frac{1}{3}\,(x^2 + y^2)$.

The inequality defining the ball in the spherical coordinates r, φ and ϑ is simple: $r \leq 1$. Substituting from (III.14) to the equation of the cone, we get:

$$\begin{aligned} r^2\,\sin^2\vartheta \;&=\; \tfrac{1}{3}\,r^2\,\cos^2\vartheta\,\bigl(\cos^2\varphi + \sin^2\varphi\bigr), \\ \sin^2\vartheta \;&=\; \tfrac{1}{3}\,\cos^2\vartheta, \\ \tan\vartheta \;&=\; \pm\sqrt{3}/3 \end{aligned}$$

which means that $\vartheta = \pm\pi/6$. Since D is the region above the cone, the ϑ–coordinates of its points satisfy: $\vartheta \in \langle \pi/6,\, \pi/2 \rangle$. Finally, all possible straight lines passing through the origin sweep over D as the angle φ runs from 0 to 2π. Thus,

$$m_3(D) \;=\; \iiint_D dx\,dy\,dz \;=\; \int_0^1 \int_0^{2\pi} \int_{\pi/6}^{\pi/2} r^2 \,\cos\vartheta \; d\vartheta \; d\varphi \; dr \;=\; \tfrac{1}{3}\pi.$$

III.7.10. Generalized cylindrical coordinates in $\mathbf{E}_3$. We will again denote these coordinates of the point $[x,y,z] \in \mathbf{E}_3$ by r, φ, w. They are analogous to the cylindrical coordinates, though their origin need not coincide with the origin O of the Cartesian coordinates. r, φ represent the generalized polar coordinates of the point $[x,y]$ in the xy–plane and w is a linear function of z (and vice versa). Thus, the relations between the Cartesian coordinates and the generalized cylindrical coordinates are:

$$x = x_0 + ar\,\cos\varphi, \quad y = y_0 + br\,\sin\varphi, \quad z = z_0 + cw, \tag{III.16}$$

where $[x_0, y_0, z_0]$ is a chosen point in $\mathbf{E}_3$ (the origin of the generalized cylindrical coordinates) and a, b, c are positive parameters.

Analogously to (III.8) and (III.13), when we transform a volume integral to the generalized cylindrical coordinates, we must substitute for $dx\,dy\,dz$ in accordance with the following equation:

$$dx\,dy\,dz \;=\; abc\,r\;dr\,d\varphi\,dw. \tag{III.17}$$

(See Section III.9 for more details.)

The transformation of a volume integral to the generalized cylindrical coordinates can simplify the integral either if the domain of integration is a part of the elliptic cylinder $(x-x_0)^2/a^2 + (y-y_0)^2/b^2 < R^2$ or if the integrand depends on x and y mainly through the expression $(x-x_0)^2/a^2 + (y-y_0)^2/b^2$.

III.7.11. Generalized spherical coordinates in $\mathbf{E}_3$. We will denote these coordinates of the point $[x,y,z] \in \mathbf{E}_3$ in the same way as the spherical coordinates, i.e. by r, φ, w. The difference between the spherical coordinates and the generalized spherical coordinates is that the generalized spherical coordinates need not have their origin at the same point as the Cartesian coordinates and they are not "isotropic". This means that r can increase with the distance from the origin with the different rate in the x–direction, y–direction and z–direction. The relations between the Cartesian coordinates and the generalized spherical coordinates are:

$$x = x_0 + ar\,\cos\vartheta\,\cos\varphi, \quad y = y_0 + br\,\cos\vartheta\,\sin\varphi, \quad z = z_0 + cr\,\sin\vartheta, \tag{III.18}$$

where $[x_0, y_0, z_0]$ is a chosen point in $\mathbf{E}_3$ (the origin of the generalized spherical coordinates) and a, b, c are positive parameters.

Analogously to (III.8) and (III.15), when we transform a volume integral to the generalized cylindrical coordinates, we must substitute for $dx\,dy\,dz$ in accordance with the following equation:

$$dx\,dy\,dz \;=\; abc\,r^2\,\cos\vartheta\;dr\,d\varphi\,dw. \tag{III.19}$$

(See Section III.9 for more details.)

The transformation of a volume integral to the generalized spherical coordinates usually simplifies the integral either if the domain of integration is the ellipsoid $(x-$

$x_0)^2/a^2 + (y - y_0)^2/b^2 + (z - z_0)^2/c^2 < R^2$ (or its sector) or if the integrand depends on x, y and z mainly through the expression $(x-x_0)^2/a^2+(y-y_0)^2/b^2+(z-z_0)/c^2$.

III.8. Some physical applications of the volume integral.

Suppose that a three–dimensional body has the form of a measurable set D in $\mathbf{E}_3$. The body need not be homogeneous and so its density (amount of mass per unit of volume) need not be constant. Let the density be given by function $\rho(x, y, z)$. The volume integral enables us to define and evaluate some fundamental mechanical characteristics of the body. Suppose that ρ is expressed in $[\mathrm{kg} \cdot \mathrm{m}^{-3}]$. Then we have:

Mass $\quad M = \iiint_D \rho(x, y, z)\, dx\, dy\, dz \quad [\mathrm{kg}]$,

Static moment about the xy–plane $\quad M_{xy} = \iiint_D z \cdot \rho(x, y, z)\, dx\, dy\, dz \quad [\mathrm{kg} \cdot \mathrm{m}]$,

Static moment about the xz–plane $\quad M_{xz} = \iiint_D y \cdot \rho(x, y, z)\, dx\, dy\, dz \quad [\mathrm{kg} \cdot \mathrm{m}]$,

Static moment about the yz–plane $\quad M_{yz} = \iiint_D x \cdot \rho(x, y, z)\, dx\, dy\, dz \quad [\mathrm{kg} \cdot \mathrm{m}]$,

Center of mass $[x_m, y_m, z_m]$ $\quad x_m = \dfrac{M_{yz}}{M}, \quad y_m = \dfrac{M_{xz}}{M}, \quad z_m = \dfrac{M_{xy}}{M} \quad [\mathrm{m}]$,

Moment of inertia about the x–axis $\; J_x = \iiint_D (y^2+z^2)\, \rho(x, y, z)\, dx\, dy\, dz \quad [\mathrm{kg} \cdot \mathrm{m}^2]$,

Moment of inertia about the y–axis $\; J_y = \iiint_D (x^2+z^2)\, \rho(x, y, z)\, dx\, dy\, dz \quad [\mathrm{kg} \cdot \mathrm{m}^2]$,

Moment of inertia about the z–axis $\; J_z = \iiint_D (x^2+y^2)\, \rho(x, y, z)\, dx\, dy\, dz \quad [\mathrm{kg} \cdot \mathrm{m}^2]$,

Moment of inertia about the origin $\; J_0 = \iint_D (x^2+y^2+z^2)\, \rho(x, y, z)\, dx\, dy\, dz \; [\mathrm{kg} \cdot \mathrm{m}^2]$.

Try to suggest the formula for the moment of inertia about a general straight line in $\mathbf{E}_3$ whose parametric equations are $x = x_0 + u_1 t$, $y = y_0 + u_2 t$, $z = z_0 + u_3 t$; $t \in \mathbf{R}$!

III.9. A remark to the method of substitution in the double and volume integral.

The idea of the method of substitution is the same in the double and in the volume integral. This is why we shall treat it together for both types of integrals in this section. Thus,

– $\mathbf{E}_k$ will mean either $\mathbf{E}_2$ (if $k = 2$) or $\mathbf{E}_3$ (if $k = 3$),

– the integral $\int$ will mean either $\iint$ (if $k = 2$) or $\iiint$ (if $k = 3$),

– point $X \in \mathbf{E}_k$ will denote either $[x_1, x_2]$ (if $k = 2$) or $[x_1, x_2, x_3]$, (if $k = 3$) and

– dX will denote either $dx_1\, dx_2$ (if $k = 2$) or $dx_1\, dx_2\, dx_3$ (if $k = 3$).

We already know some widely used substitutions – they are given by equations (III.4), (III.7), (III.12), (III.14), (III.16) and (III.18) and they are called the transformation to the polar (or generalized polar) coordinates, transformation to the cylindrical coordinates, etc. We will explain the method of substitution on a general level in this section.

Suppose that $D \subset \mathbf{E}_k$ ($k = 2$ or $k = 3$) and we have to evaluate the integral $\int_D f(X)\, dX$. If every point $X \in D$ can be expressed as $X = \mathcal{F}(Y)$ where points Y are taken from some other set $D' \subset \mathbf{E}_k$ then the integral can be transformed to the integral in the variables Y on the domain of integration D'. However, it is clear that this can be done only if both the integrals on D and D' exist and mapping $\mathcal{F}$ has certain properties. We will discuss them in the following.

III.9.1. Regular mapping and its Jacobian. Let D and D' be domains in $\mathbf{E}_k$. Suppose that $\mathcal{F}$ is a mapping of D' to D. Equation $X = \mathcal{F}(Y)$ means

$$x_1 = \phi_1(y_1, y_2), \quad x_2 = \phi_2(y_1, y_2) \qquad \text{for } k = 2,$$
$$x_1 = \phi_1(y_1, y_2, y_3), \quad x_2 = \phi_2(y_1, y_2, y_3), \quad x_3 = \phi_3(y_1, y_2, y_3) \qquad \text{for } k = 3.$$

The determinant

$$J_{\mathcal{F}}(Y) = \left|\frac{\partial \phi_i}{\partial y_j}(Y)\right|_{i,j=1,2} \quad (\text{for } k = 2) \qquad \text{or} \qquad J_{\mathcal{F}}(Y) = \left|\frac{\partial \phi_i}{\partial y_j}(Y)\right|_{i,j=1,2,3} \quad (\text{for } k = 3)$$

is called the *Jacobian* or the *Jacobi determinant* of mapping $\mathcal{F}$.

Mapping $\mathcal{F}$ is called *regular* if functions ϕ_i ($i = 1, 2$ or $i = 1, 2, 3$) have continuous partial derivatives in set D' and $J_{\mathcal{F}}(Y) \neq 0$ in all points $Y \in D'$.

III.9.2. A one–to–one mapping. You already know the notion of a one–to–one mapping. Mapping $\mathcal{F}$ is called *one–to–one* if

$$Y,\ Z \in D',\ \ Y \neq Z \implies \mathcal{F}(Y) \neq \mathcal{F}(Z).$$

III.9.3. Example. Verify that the mapping given by equations (III.4) is a one–to–one regular mapping of the open rectangle $D' = \{[r, \varphi] \in \mathbf{E}_2;\ r \in (0, 2),\ \varphi \in (0, 2\pi)\}$ onto the domain $D = \{[x_1, x_2] \in \mathbf{E}_2;\ x_1^2 + x_2^2 < 4\} - \{[x_1, x_2] \in \mathbf{E}_2;\ x_1 \in \langle 0, 1),\ x_2 = 0\}$ (D is an open disk (with its center at the origin and radius 2) minus the line segment connecting the points $O = [0, 0]$ and $P = [2, 0]$. Sketch a figure!

The one–to–one correspondence between the points $[x_1, x_2] \in D$ and $[r, \varphi] \in D'$ is obvious. To verify the regularity of mapping (III.4), let us evaluate the Jacobian of this mapping. Equations (III.4) can also be written as

$$x_1 \;=\; \phi_1(r, \varphi) \;=\; r\, \cos\varphi, \quad x_2 \;=\; \phi_2(r, \varphi) \;=\; r\, \sin\varphi.$$

The Jacobian is:

$$J(r,\varphi) \;=\; \begin{vmatrix} \dfrac{\partial\phi_1}{\partial r}, & \dfrac{\partial\phi_1}{\partial\varphi} \\ \dfrac{\partial\phi_2}{\partial r}, & \dfrac{\partial\phi_2}{\partial\varphi} \end{vmatrix} \;=\; \begin{vmatrix} \cos\varphi, & -r\,\sin\varphi \\ \sin\varphi, & r\,\cos\varphi \end{vmatrix} \;=\; r.$$

Now it is seen that functions ϕ_1 and ϕ_2 have continuous partial derivatives in domain D' and $r \neq 0$ in D'. Thus, mapping (III.4) is regular in D'.

III.9.4. Theorem. *Let $X = \mathcal{F}(Y)$ be a one–to–one regular mapping of domain $D' \in \mathbf{E}_k$ onto domain $D \in \mathbf{E}_k$. Then*

$$\int_D f(X)\,dX \;=\; \int_{D'} f(\mathcal{F}(Y))\cdot|J_{\mathcal{F}}(Y)|\,dY, \tag{III.20}$$

if both integrals exist.

III.9.5. Remark. We already know that adding (or subtracting) a set of measure zero to the domain of integration does not influence the existence or the value of the integral. This enables us to generalize Theorem III.9.4:

If the assumptions of Theorem III.9.4 are satisfied and A, respectively A', are sets in $\mathbf{E}_k$ which differ from D, respectively D', only in sets of measure zero (i.e. $m_k\big((A-D)\cup(D-A)\big) = m_k\big((A'-D')\cup(D'-A')\big) = 0$) then

$$\int_A f(X)\,dX \;=\; \int_{A'} f(\mathcal{F}(Y))\cdot|J_{\mathcal{F}}(Y)|\,dY, \tag{III.21}$$

if both integrals exist.

It is seen from equations (III.20) and (III.21) that dX in the integral on the left–hand side changes to $|J_{\mathcal{F}}(Y)|\;dY$ in the integral on the right–hand side. We have already shown that if $k = 2$ and $Y = [r,\varphi]$ represents the polar coordinates then the Jacobian is equal to r (see example III.9.3). Thus, the equation

$$dX \;=\; |J_{\mathcal{F}}(Y)|\;dY \tag{III.22}$$

implies, as a special case, equation (III.5). Computing the Jacobians of mappings (III.7), (III.12), (III.14), (III.16) and (III.18), we can see that general equation (III.22) also implies special equations (III.8), (III.13), (III.15), (III.17) and (III.19).

III.10. Exercises.

1. Do the following integrals exist?

$$\iint_D \frac{dx\,dy}{x+y+1}\,; \quad D = \langle 0,1\rangle \times \langle 0,1\rangle$$

$$\iint_D \frac{\sin(x^2+y^2)}{x^2+y^2}\,dx\,dy\,; \quad D = \{[x,y]\in\mathbf{E}_2;\; x^2+y^2 \le 9\}$$

$$\iint \frac{x}{x^2+y^2}\,dx\,dy\,; \quad D = \{[x,y]\in\mathbf{E}_2;\; x^2+y^2 \le 9\}$$

$\displaystyle\iint_D \frac{dx\,dy}{(1-xy)^2}$; D is the square $PQRS$ where $P=[1,2]$, $Q=[3,2]$, $R=[3,4]$, $S=[1,4]$

$\displaystyle\iiint_D \sqrt{1-\frac{x^2}{4}-\frac{y^2}{9}-\frac{z^2}{16}}\,dx\,dy\,dz$; $D=\left\{[x,y,z]\in\mathbf{E}_3;\ \frac{x^2}{4}+\frac{y^2}{9}+\frac{z^2}{16}<1\right\}$

$\displaystyle\iiint_D \sqrt{x^2+y^2+z^2}\,dx\,dy\,dz$; $D=\{[x,y,z]\in\mathbf{E}_3;\ x^2+y^2+z^2\le 0\}$

2. Find the area of the region R in the xy–plane enclosed by the curves

a) $y=2x+4,\ \ y=4-x^2$

b) $xy=1,\ \ y=x,\ \ x=4$

c) $y=x^2,\ \ 4y=x^2,\ \ y=4$

d) $y^2=4+x,\ \ x+3y=0$

e) $y=\ln x,\ \ x-y=1,\ \ y=-1$

f) $x^2+y^2=2x,\ \ x^2+y^2=4x,\ \ y=x,\ \ y=0$

3. Find the area of the region in xy–plane bounded on the right by the parabola $y=x^2$, on the left by the line $x+y=2$, and above by the line $y=4$.

4. Find the volume of set R in $\mathbf{E}_3$ if

a) R is the region under the paraboloid $z=x^2+y^2$, above the triangle enclosed by the lines $y=x,\ x=0,\ x+y=2$, in the xy–plane

b) R is the region under the parabolic cylinder $z=x^2$, above the domain enclosed by the parabola $y=6-x^2$ and the line $y=x$, in the xy–plane

c) R is the region in the half–space $z\ge 0$ bounded by the surfaces $x^2+y^2-z^2=0$, $z=6-x^2-y^2$

d) R is the region in the half–space $z\ge 0$ bounded by the surfaces $az=x^2+y^2$, $x^2+y^2+x^2=2a^2$, $(a>0)$

e) R is the region bounded by the surfaces $y^2=4a^2-3ax$, $y^2=ax$, $z=h$, $z=-h$, $(a>0,\ h>0)$

f) R is the region bounded by the surfaces $x^2+y^2=z$, $x^2+y^2=2z$, $z=0$

5. Evaluate the following integrals.

a) $\displaystyle\iint_D (1+x)\,dx\,dy$; D is the region in $\mathbf{E}_2$ enclosed by the lines $y=x^2-4$, $y=-3x$

b) $\displaystyle\iint_D \frac{dx\,dy}{(x+y)^2}$; $D=\langle 3,4\rangle\times\langle 1,2\rangle$

c) $\displaystyle\iint_D xy\,dx\,dy$; D is the region in $\mathbf{E}_2$ enclosed by the line $y=x-4$ and by the parabola $y^2=2x$

d) $\displaystyle\iiint_V (x+y+z)\,dx\,dy\,dz$; $V=\langle 0,1\rangle\times\langle 0,2\rangle\times\langle 0,3\rangle$

e) $\displaystyle\iiint_V x\,dx\,dy\,dz$; V is the region in $\mathbf{E}_3$ bounded by the surfaces $x=0,\ y=0,\ z=0,\ y=2,\ x+z=1$

f) $\displaystyle\iiint_V xy^2z^3\,dx\,dy\,dz\,;$ V is the region in $\mathbf{E}_3$ bounded by the surfaces $z = xy$, $y = x,\ \ x = 1,\ \ z = 0$

g) $\displaystyle\iiint_V x^3y^2z\,dx\,dy\,dz\,;$ $V = \big\{[x, y, z] \in \mathbf{E}_3;\ 0 \le x \le 1,\ 0 \le y \le x,\ 0 \le z \le xy\big\}$

h) $\displaystyle\iiint_V y\,\cos(x+z)\,dx\,dy\,dz\,;$ V is the region in $\mathbf{E}_3$ bounded by the surfaces $y = \sqrt{x},\ \ y = 0,\ \ z = 0,\ \ x + z = \pi/2$

i) $\displaystyle\iint_D \sqrt{1 - x^2 - y^2}\,dx\,dy\,;$ $D = \big\{[x, y] \in \mathbf{E}_2;\ x^2 + y^2 \le 1\big\}$

j) $\displaystyle\iint_D \frac{dx\,dy}{(x^2+y^2)^2}\,;$ D is the region in the xy–plane enclosed by the lines $y = x$, $y = 2x$ and by the circles $x^2 + y^2 = 4x,\ \ x^2 + y^2 = 8x$

k) $\displaystyle\iint_D y\,dx\,dy\,;$ D is the upper half of the disk $(x-a)^2 + y^2 \le a^2$ $(a > 0)$

l) $\displaystyle\iint_D x\,dx\,dy\,;$ D is the sector of the disk $x^2 + y^2 \le a^2$ consisting of the points $[x, y]$ such that $x \ge 0$ and $-x \le \sqrt{3}\,y \le 1$

m) $\displaystyle\iiint_V \sqrt{x^2 + y^2 + z^2}\,dx\,dy\,dz\,;$ V is the ball $x^2 + y^2 + z^2 \le a^2,$ $(a > 0)$

n) $\displaystyle\iiint_D (x + y + z)^2\,dx\,dy\,dz\,;$ D is the region in the half–space $z \ge 0$ bounded by the paraboloid $z = \frac{1}{2}\,(x^2 + y^2)$ and by the sphere $x^2 + y^2 + z^2 = 3$

o) $\displaystyle\iiint_V z\,dx\,dy\,dz\,;$ V is the region in $\mathbf{E}_3$ bounded by the surfaces $z = \sqrt{x^2 + y^2}$ and $z = 1$

p) $\displaystyle\iiint_D (x^2 + y^2)\,dx\,dy\,dz\,;$ D is the region in $\mathbf{E}_3$ bounded by the surfaces $x^2 + y^2 = 2z$ and $z = 2$

q) $\displaystyle\iiint_V \Big(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\Big)\,dx\,dy\,dz\,;$ V is the interior of the ellipsoid $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1,$ $(a > 0,$ $b > 0,\ c > 0)$

6. Find the center of mass of the homogeneous regions in $\mathbf{E}_2$ bounded by the curves
a) $y = \sin\,x,\ \ y = 0\,;\ \ x \in \langle 0, \pi\rangle$, b) $x^2 + y^2 = a^2,\ \ y = 0\,;\ \ (y \ge 0,\ a > 0)$,
c) $y^2 = ax,\ \ x = 0,\ \ y = a\,;\ \ (y > 0,\ a > 0)$, d) $y^2 = 4x + 4,\ \ y^2 = -2x + 4$.

7. Evaluate the moment of inertia with respect to the x–axis of the homogeneous region in $\mathbf{E}_2$, bounded by the lines $y = x/2,\ \ y = a,\ \ x = a$ $(a > 0)$. (The density is $\rho = 1$.)

8. Evaluate the mass of the body in $\mathbf{E}_3$ bounded by the surfaces
a) $x = 0,\ x = a,\ y = 0,\ y = b,\ z = 0,\ z = c$ $(a > 0,\ b > 0,\ c > 0)$ if the density is $\rho(x, y, z) = x + y + z$,
b) $2x + z = 2a,\ x + z = a,\ y^2 = ax,\ y = 0$ (for $y > 0$) if $a > 0$ and the density is $\rho(x, y, z) = y$,

c) $x^2+y^2+z^2=a^2$, $x^2+y^2+z^2=4a^2$ $(a>0)$ if the density is $\rho(x,y,z)=2/\sqrt{x^2+y^2+z^2}$.

IV. Line Integrals

IV.1. Simple curves.

We need to specify exactly what we understand under the notion of a curve (i.e. a line) before we begin to deal with line integrals. A variety of definitions of many types of curves appear in literature. We will restrict ourselves to two types of curves: so called simple smooth curves and simple piecewise–smooth curves. They are also called simple regular or simple piecewise–regular curves.

The idea of the definition of a simple smooth curve is the following: Imagine that a mass point moves in $\mathbf{E}_2$ or in $\mathbf{E}_3$ in a time interval $\langle a,b\rangle$ and its position at time t is $P(t)$. Then P is a mapping of the interval $\langle a,b\rangle$ to $\mathbf{E}_k$ (with $k=2$ or $k=3$). Suppose that the velocity of the moving mass point is continuous, bounded and different from zero in all times $t\in(a,b)$. (This means that mapping P has a continuous, bounded and non–zero derivative at all points $t\in(a,b)$). Suppose further that the mass point cannot be at the same place at two different times, with a possible exception when the motion starts and finishes at the same point. (This means that mapping P is one–to–one in the interval $\langle a,b\rangle$ with a possible exception when $P(a)=P(b)$.) Then the path travelled by the mass point is called a simple smooth curve. The position function P is called the parametrization of the curve. You will find the precise definition of a simple smooth curve in paragraph IV.1.1.

Since $P(t)$ (for fixed t) is a point in $\mathbf{E}_k$ (where $k=2$ or $k=3$), it has two or three coordinates. Let us denote them $\phi(t)$, $\psi(t)$, respectively $\phi(t)$, $\psi(t)$, $\vartheta(t)$. Then ϕ, ψ, respectively ϕ, ψ, ϑ, are the functions of one variable t defined in the interval $\langle a,b\rangle$. They will be called the *coordinate functions* of mapping P and we will write

$$P(t)=\big[\phi(t),\psi(t)\big] \qquad \text{if } k=2,$$

$$P(t)=\big[\phi(t),\psi(t),\vartheta(t)\big] \qquad \text{if } k=3.$$

The derivative of P with respect to the parameter will be denoted by the dot, in accordance with the customs in physics. We will take the derivative for a vector and we will therefore enclose its components in parentheses:

$$\dot P(t)=\big(\dot\phi(t),\dot\psi(t)\big) \qquad \text{if } k=2,$$

$$\dot P(t)=\big(\dot\phi(t),\dot\psi(t),\dot\vartheta(t)\big) \qquad \text{if } k=3.$$

The coordinate functions of parametrization P are also often denoted by $x(t)$, $y(t)$, respectively by $x(t)$, $y(t)$, $z(t)$.

IV.1.1. Simple smooth curve. Let P be a continuous mapping of a closed bounded interval $\langle a,b\rangle$ to $\mathbf{E}_k$ (where $k=2$ or $k=3$). Suppose that

a) mapping P is one–to–one in the interval $\langle a,b\rangle$ with a possible exception when $P(a)=P(b)$,

b) P has a bounded, continuous and non–zero derivative $\dot{P}$ in the interval (a,b).

Then the set $C=\{X=P(t)\in\mathbf{E}_k;\ t\in\langle a,b\rangle\}$ is called a <u>*simple smooth curve*</u> in $\mathbf{E}_k$. Mapping P is called the <u>*parametrization*</u> of curve C.

The simple smooth curve C is called <u>*closed*</u> if $P(a)=P(b)$.

The vector $\dot{P}(t)$ is tangent to the simple smooth curve C at every "interior" point $P(t)$ of curve C (i.e. point $P(t)$ corresponding to $t\in(a,b)$). The vector $\dot{P}(t)/|\dot{P}(t)|$ is also tangent to C at point $P(t)$ and moreover, its length is equal to one. We can choose the <u>*orientation*</u> of curve C so that we put the <u>*unit tangent vector*</u> $\vec{\tau}$ to C at point $P(t)$

either
$$\vec{\tau} = \frac{\dot{P}(t)}{|\dot{P}(t)|} \quad \text{(for all } t\in(a,b)\text{)} \qquad \text{(IV.1)}$$

or
$$\vec{\tau} = -\frac{\dot{P}(t)}{|\dot{P}(t)|} \quad \text{(for all } t\in(a,b)\text{).} \qquad \text{(IV.2)}$$

We say that curve C is <u>*oriented in accordance with its parametrization P*</u> if the unit tangent vector $\vec{\tau}$ to C is given by formula (IV.1).

In other words, we say that simple smooth curve C is oriented in accordance with its parametrization P if the parametrization defines the motion along C in the direction corresponding to the orientation of C.

If the simple smooth curve C is oriented in accordance with the parametrization P then the point $P(a)$ is called the <u>*initial point of C*</u> (we denote it $i.p.\,C$) and the point $P(b)$ is called the <u>*terminal point of C*</u> (we denote it $t.p.\,C$).

If the orientation of C is opposite to parametrization P then the position of the initial and the terminal point of C is also opposite: $i.p.\,C=P(b)$ and $t.p.\,C=P(a)$.

A simple smooth curve C which is not closed can also be oriented so that one of the points $P(a)$, $P(b)$ is chosen to be the initial point of C and the second one to be the terminal point of C.

Every simple smooth curve has infinitely many parametrizations. This is clear if you take into account that every path can be travelled by infinitely many possible motions.

IV.1.2. Example. Every line segment in $\mathbf{E}_k$ is a simple smooth curve. For instance, the line segment AB in $\mathbf{E}_3$ with $A=[1,2,4]$ and $B=[3,-1,7]$ can be parametrized by the mapping

$$P(t)=A+t\cdot(B-A);\quad t\in\langle 0,1\rangle.$$

This means that the coordinate functions ϕ, ψ and ϑ of parametrization P are:

$$x = \phi(t) = 1 + 2t, \quad y = \psi(t) = 2 - 3t, \quad z = \vartheta(t) = 4 + 3t; \quad t \in \langle 0, 1 \rangle.$$

The simple smooth curve C, identical with the line segment AB, is oriented in accordance with the above parametrization if $A = i.p.\, C$ and $B = t.p.\, C$.

IV.1.3. Example. The part of the parabola $y = x^2 + 1$ between the points $[1, 2]$ and $[3, 10]$ (oriented from $[1, 2]$ to $[3, 10]$) is a simple smooth curve in $\mathbf{E}_2$. Its possible parametrization is e.g. the mapping

$$P \ : \quad x = \phi(t) = t, \quad y = \psi(t) = t^2 + 1; \quad t \in \langle 1, 3 \rangle.$$

Since this parametrization defines the motion on the curve from its initial to its terminal point, the curve is oriented in accordance with parametrization P.

IV.1.4. Example. The arc $x^2 + y^2 = 9,\ x \geq 0,\ y \geq 0,$ oriented from the point $[3, 0]$ to the point $[0, 3]$, is a simple smooth curve in $\mathbf{E}_2$. Its possible parametrization, generating the same orientation, is

$$x = \phi(t) = 3 \cos t, \quad y = \psi(t) = 3 \sin t; \quad t \in \langle 0, \pi/2 \rangle.$$

IV.1.5. Example. The circle $C : \ (x-3)^2 + y^2 = 4$ in $\mathbf{E}_2$ (oriented counter–clockwise) is a closed simple smooth curve. Its parametrization is for instance the mapping

$$x = \phi(t) = 3 + 2 \cos t, \quad y = \psi(t) = 2 \sin t; \quad t \in \langle 0, 2\pi \rangle.$$

You can easily verify that C is oriented in accordance with this parametrization.

IV.1.6. Simple piecewise–smooth curve. Let $C_1, \ldots, C_m$ be simple smooth curves in $\mathbf{E}_k$ such that

a) $t.p.\, C_1 = i.p.\, C_2,\ t.p.\, C_2 = i.p.\, C_3,\ \ldots,\ t.p.\, C_{m-1} = i.p.\, C_m,$

b) except for the points named in a) and except for the possibility when $i.p.\, C_1 = t.p.\, C_m$, any two of the curves $C_1, \ldots, C_m$ have no more common points.

Then the set $C = \cup_{i=1}^m C_i$ is called the *simple piecewise–smooth curve* in $\mathbf{E}_k$.

The orientation of the simple piecewise–smooth curve C is given by the orientation of its smooth parts $C_1, \ldots, C_m$. We put $i.p.\, C$ (the initial point of C) $= i.p.\, C_1$ and $t.p.\, C$ (the terminal point of C) $= t.p.\, C_m$.

The curve which differs from a simple piecewise–smooth curve C only by its orientation will be denoted by $-C$.

A simple piecewise–smooth curve C is called *closed* if $i.p.\, C = t.p.\, C$.

A simple piecewise–smooth curve in $\mathbf{E}_k$ (where $k = 2$ or $k = 3$) is a bounded set in $\mathbf{E}_k$ whose k–dimensional measure m_k equals zero.

It is obvious that the notion of a simple piecewise–smooth curve is a generalization of the notion of a simple smooth curve. Always when we will use the word "curve" in the coming sections, we will have in mind a simple piecewise–smooth curve. Similarly, a "closed curve" will mean a closed simple piecewise–smooth curve. We will give more details about the curve if they are important.

IV.2. The line integral of a scalar function.

A scalar function is a function whose values are scalars, i.e. in our case real numbers. We use the name "scalar function" only in order to distinguish it from a "vector function" which will be treated in the next sections.

IV.2.1. Physical motivation. Suppose that a spring or a wire has the form of a simple smooth curve C in $\mathbf{E}_k$ (with $k = 2$ or $k = 3$). Suppose further that the longitudinal density of the wire is ρ. ρ need not be a constant, and it is generally a function of two variables x, y (if $k = 2$) or three variables x, y, z (if $k = 3$). We wish to evaluate the mass M of the wire.

The idea of evaluation of the total mass of the wire is exactly the same as the idea of computation of the mass of a one–dimensional rod, used in paragraph II.1.1. We could explain it by means of a partition of the wire into many "short" pieces, similarly as we divided the interval $\langle a, b\rangle$ into many "short" subintervals in paragraph II.1.1. However, let us use another approach – an approach based on the idea of the partition of an interval into infinitely many "infinitely short" subintervals. This idea was explained in Section II.7 (and we promised to come back to it).

Thus, suppose that P is a parametrization of curve C which is defined in the interval $\langle a, b\rangle$. A typical "infinitely short" subinterval of $\langle a, b\rangle$ has the form $\langle t, t+dt\rangle$. Parametrization P maps this interval to the "infinitely short" part of C with the end points $P(t)$ and $P(t + dt)$. We can take this part for a line segment whose length is $ds = |P(t+dt) - P(t)| = |\dot{P}(t)|\, dt$. The mass of this segment is $dM = \rho(P(t)) \cdot ds = \rho(P(t)) \cdot |\dot{P}(t)|\, dt$. The total mass of the whole wire is

$$M = \int_a^b \rho(P(t)) \cdot |\dot{P}(t)|\, dt.$$

IV.2.2. The line integral of a scalar function on a simple smooth curve. Let C be a simple smooth curve in $\mathbf{E}_2$ or $\mathbf{E}_3$ and P be its parametrization defined in the interval $\langle a, b\rangle$. Let f be a scalar function defined on C. We say that f is *integrable* on curve C if the Riemann integral $\int_a^b f(P(t)) \cdot |\dot{P}(t)|\, dt$ exists. We denote this integral by $\int_C f\, ds$ and we call it the *line integral of a scalar function* f on the simple smooth curve C.

IV.2.3. Remark. The integrability of function f on a simple smooth curve C and the line integral $\int_C f\, ds$ are defined by means of a parametrization of curve C. Since C can be parametrized in infinitely many ways, there arises the question whether the integrability of f on curve C as well as the value of the integral $\int_C f\, ds$ can depend on a concrete choice of parametrization of C. The answer is NO. It can be proved that neither the existence, nor the value of the line integral $\int_C f\, ds$ depends on the concrete choice of parametrization of curve C.

IV.2.4. The line integral of a scalar function on a simple piecewise–smooth curve. Let C be a simple piecewise–smooth curve in $\mathbf{E}_2$ or $\mathbf{E}_3$ which is a union of simple smooth curves C_1, C_2, …, C_m (see paragraph IV.1.3). Let f be a scalar function defined on C. If function f is integrable on each of curves C_1, C_2, ..., C_m then we say that it is integrable on curve C and we put

$$\int_C f\ ds = \sum_{i=1}^{m} \int_{C_i} f\ ds. \tag{IV.5}$$

The integral on the left hand side is called the *line integral of the scalar function* f on the simple piecewise–smooth curve C.

The line integral of a scalar function is also often called the *line integral of the 1st kind*.

Instead of $\int_C f\ ds$, we can also write for example $\int_C f(x,y)\ ds$ (if $C \subset \mathbf{E}_2$ and f is a function of two variables) or $\int_C f(x,y,z)\ ds$ (if $C \subset \mathbf{E}_3$ and f is a function of three variables). The symbol ds at the end of the integral can also be replaced e.g. by dl, dr, etc.

IV.2.5. The length of a curve. If C is a simple piecewise–smooth curve then the value of the integral $\int_C ds$ is called the *length* of curve C.

Specially, if C is a simple smooth curve and P is its parametrization defined in the interval $\langle a, b\rangle$ then the length of curve C is

$$l(C) = \int_C ds = \int_a^b |\dot{P}(t)|\ dt. \tag{IV.6}$$

IV.2.6. Some important properties of the line integral of a scalar function. Since the line integral of a scalar function is defined by means of the Riemann integral, most of its properties are quite analogous. Let us mention only some of them:

a) **(Existence of the line integral.)** *If function f is continuous on curve C then it is integrable on C (i.e. the integral $\int_C f\ ds$ exists).*

b) **(Linearity of the line integral.)** *If functions f and g are integrable on curve C and $\alpha \in \mathbf{R}$ then*

$$\int_C (f+g)\ ds = \int_C f\ ds + \int_C g\ ds,$$

$$\int_C \alpha \cdot f\ ds = \alpha \cdot \int_C f\ ds.$$

c) *If function f is integrable on curve C and function g differs from f at most in a finite number of points then g is also integrable on C and*

$$\int_C g\ ds = \int_C f\ ds.$$

d) *If function f is integrable on curve C then it is also integrable on curve $-C$ and*

$$\int_{-C} f\ ds = \int_C f\ ds.$$

Assertion a) can be generalized in this way: *If C is a simple piecewise–smooth curve and f is continuous on each of its smooth parts then f is integrable on C.*

Assertion d) says that <u>neither the existence, nor the value of the line integral of a scalar function depends on the orientation of the curve</u>!

IV.2.7. Evaluation of the line integral of a scalar function. The line integral of function f on a simple smooth curve C can be evaluated by means of a parametrization of C. Thus, if P is such a parametrization, defined in the interval $\langle a, b\rangle$, and function f is integrable on C then we can use the formula

$$\int_C f\, ds \;=\; \int_a^b f(P(t))\cdot|\dot{P}(t)|\, dt. \tag{IV.7}$$

This formula follows immediately from the definition of the line integral on a simple smooth curve – see paragraph IV.2.2.

If C is a simple piecewise–smooth curve which is a union of simple smooth curves $C_1, C_2, \dots, C_m$ (see paragraph IV.1.6) then the line integral of function f on curve C can be computed by means of formula (IV.5).

IV.2.8. Example. C is the union of two line segments $C_1 = OP$ and $C_2 = PQ$ where $O = [0,0,0]$, $P = [1,1,0]$ and $Q = [1,1,1]$. Integrate the function $f(x,y,z) = x - 3y^2 + z$ over C.

The simplest parametrizations of C_1 and C_2 we can think of are:

$$\begin{aligned} C_1 :\quad & P_1(t) = O + (P-O)\,t = [t,t,0]; \quad t \in \langle 0,1\rangle, \\ C_2 :\quad & P_2(t) = P + (Q-P)\,t = [1,1,t]; \quad t \in \langle 0,1\rangle. \end{aligned}$$

We can easily find that $\dot{P}_1(t) = (1,1,0)$, $\dot{P}_2(t) = (0,0,1)$ and so $|\dot{P}_1(t)| = \sqrt{2}$ and $|\dot{P}_2(t)| = 1$. Using formulas (IV.7) and (IV.5), we obtain

$$\begin{aligned} \int_C (x - 3y^2 + z)\, ds \;&=\; \int_{C_1} (x - 3y^2 + z)\, ds \;+\; \int_{C_2} (x - 3y^2 + z)\, ds \;= \\ &=\; \int_0^1 (t - 3t^2)\,\sqrt{2}\, dt \;+\; \int_0^1 (t-2)\, dt \;=\; -\,\frac{\sqrt{2}+3}{2}. \end{aligned}$$

IV.2.9. Example. Evaluate the line integral $\int_C (x^2 + y)\, ds$ where C is the circle $x^2 + (y-5)^2 = 4$. We do not specify the orientation of C because it does not affect the line integral of a scalar function on C.

C can be parametrized e.g. by the mapping

$$x = \phi(t) = 2\cos t, \quad y = \psi(t) = 5 + 2\sin t; \quad t \in \langle 0,\, 2\pi\rangle.$$

In order to use formula (IV.6), we also need to express $|\dot{P}(t)|$:

$$|\dot{P}(t)| \;=\; \big|(-2\sin t,\, 2\cos t)\big| \;=\; 2.$$

Thus, we obtain

$$\int_C (x^2 + y)\, ds \;=\; \int_0^{2\pi} (4\cos^2 t + 5 + 2\sin t)\cdot 2\, dt \;=\; 20\,\pi.$$

IV.3. Some physical applications of the line integral of a scalar function.

Suppose that a wire or a spring has the form of a curve C in $\mathbf{E}_k$ ($k = 2$ or $k = 3$). The wire need not be homogeneous and so its longitudinal density (amount of mass per unit of length) need not be constant. Let the density be given by function $\rho(x, y)$ (if $k = 2$) or $\rho(x, y, z)$ (if $k = 3$). The line integral of a scalar function enables us to define and evaluate some fundamental mechanical characteristics of curve C. Suppose that ρ is expressed in $[\text{kg} \cdot \text{m}^{-1}]$. Then we have:

I. $\underline{k = 2}$ Mass $M = \int_C \rho(x, y)\, ds \quad [\text{kg}]$,

Static moment about the x–axis $M_x = \int_C y \cdot \rho(x, y)\, ds \quad [\text{kg} \cdot \text{m}]$,

Static moment about the y–axis $M_y = \int_C x \cdot \rho(x, y)\, ds \quad [\text{kg} \cdot \text{m}]$,

Center of mass $[x_m, y_m]$ $x_m = \dfrac{M_y}{M}, \quad y_m = \dfrac{M_x}{M} \quad [\text{m}]$,

Moment of inertia about the x–axis $J_x = \int_C y^2 \cdot \rho(x, y)\, ds \quad [\text{kg} \cdot \text{m}^2]$,

Moment of inertia about the y–axis $J_y = \int_C x^2 \cdot \rho(x, y)\, ds \quad [\text{kg} \cdot \text{m}^2]$,

Moment of inertia about the origin $J_0 = \int_C (x^2 + y^2) \cdot \rho(x, y)\, ds \quad [\text{kg} \cdot \text{m}^2]$.

II. $\underline{k = 3}$ Mass $M = \int_C \rho(x, y, z)\, ds \quad [\text{kg}]$,

Static moment about the xy–plane $M_{xy} = \int_C z \cdot \rho(x, y, z)\, ds \quad [\text{kg} \cdot \text{m}]$,

Static moment about the xz–plane $M_{xz} = \int_C y \cdot \rho(x, y, z)\, ds \quad [\text{kg} \cdot \text{m}]$,

Static moment about the yz–plane $M_{yz} = \int_C x \cdot \rho(x, y, z)\, ds \quad [\text{kg} \cdot \text{m}]$,

Center of mass $[x_m, y_m, z_m]$ $x_m = \dfrac{M_{yz}}{M}, \quad y_m = \dfrac{M_{xz}}{M}, \quad z_m = \dfrac{M_{xy}}{M} \quad [\text{m}]$,

Moment of inertia about the x–axis $J_x = \int_C (y^2 + z^2) \cdot \rho(x, y, z)\, ds \quad [\text{kg} \cdot \text{m}^2]$,

Moment of inertia about the y–axis $J_y = \int_C (x^2 + z^2) \cdot \rho(x, y, z)\, ds \quad [\text{kg} \cdot \text{m}^2]$,

Moment of inertia about the z–axis $J_z = \int_C (x^2 + y^2) \cdot \rho(x, y, z)\, ds \quad [\text{kg} \cdot \text{m}^2]$, Moment of inertia about the origin $J_0 = \int_C (x^2 + y^2 + z^2) \cdot \rho(x, y, z)\, ds$ [kg · m²].

Try to suggest the formula for the moment of inertia about a general straight line in $\mathbf{E}_3$ whose parametric equations are $x = x_0 + u_1 t$, $y = y_0 + u_2 t$, $z = z_0 + u_3 t$; $t \in \mathbf{R}$!

IV.4. The line integral of a vector function.

A vector function in a domain $D \subset \mathbf{E}_3$ is a mapping which assigns to every point $[x, y, z] \in D$ a vector. We shall denote vectors and vector functions by boldface letters, like for example $\mathbf{u}$, $\mathbf{f}$, etc. However, you can also use $\vec{u}$, $\vec{f}$, etc.

If $\mathbf{f}$ is a vector function in domain D then $\mathbf{f}(x, y, z)$ has three components. We shall denote them by $U(x, y, z)$, $V(x, y, z)$ and $W(x, y, z)$. U, V and W are scalar functions in domain D. We shall write

$$\mathbf{f}(x, y, z) = \big(U(x, y, z),\, V(x, y, z),\, W(x, y, z)\big) \quad \text{or simply} \quad \mathbf{f} = (U, V, W).$$

If we denote by $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ the unit vectors oriented successively in accordance with the x–axis, the y–axis and the z–axis, we can also write:

$$\mathbf{f}(x, y, z) = U(x, y, z)\, \mathbf{i} + V(x, y, z)\, \mathbf{j} + W(x, y, z)\, \mathbf{k} \quad \text{or} \quad \mathbf{f} = U\, \mathbf{i} + V\, \mathbf{j} + W\, \mathbf{k}.$$

A vector function in domain $D \subset \mathbf{E}_3$ is also often called a *vector field* in D. We shall say that the vector function (or the vector field) $\mathbf{f} = (U, V, W)$ is continuous (respectively differentiable) in D if all its components U, V and W are continuous (respectively differentiable) functions in domain D.

The denotation and the used terminology in the case of two–dimensional vector functions is analogous, the only difference being that we have one variable and one component less.

IV.4.1. Physical motivation. Suppose that a body moves along a curve C due to the action of a force $\mathbf{f}$. We wish to evaluate the work A done by force $\mathbf{f}$ along curve C. The force is generally the function of three variables x, y and z. Let us apply the idea of an "infinitely small" positive number explained in Section II.7 and let us imagine that curve C can be decomposed to infinitely many "infinitely short" parts. A position of a typical part is $[x, y, z]$, its length is ds and the unit tangent vector to C at point $[x, y, z]$ is $\vec{\tau}(x, y, z)$. The work dA of force $\mathbf{f}$ done by its action on the considered "infinitely short" part is $dA = \mathbf{f}(x, y, z) \cdot \vec{\tau}(x, y, z)\, ds$. Hence the total work of force f along the whole curve C is

$$A = \int_C \mathbf{f}(x, y, z) \cdot \vec{\tau}(x, y, z)\, ds.$$

Since the product $\mathbf{f}(x, y, z) \cdot \vec{\tau}(x, y, z)$ is a scalar, the integral is the integral of a scalar function which is already known.

IV.4.2. The line integral of a vector function. Let C be a simple piecewise–smooth curve in $\mathbf{E}_k$ (where $k = 2$ or $k = 3$) and let $\mathbf{f}$ be a vector function (with k components) defined on C. We say that the vector function $\mathbf{f}$ is integrable on curve C if the scalar function $\mathbf{f}\cdot\vec{\tau}$ is integrable on C (in the sense explained in paragraphs IV.2.2 and IV.2.4). The integral $\int_C \mathbf{f}\cdot\vec{\tau}\,ds$ is called the *line integral of a vector function* $\mathbf{f}$ on curve C and it is usually denoted by $\int_C \mathbf{f}\cdot d\mathbf{s}$.

The line integral of a vector function is also often called the *line integral of the 2nd kind*.

IV.4.3. Remark. The fact that the unit tangent vector $\vec{\tau}$ need not exist in all points of a simple piecewise–smooth curve C does not matter. The points where $\vec{\tau}$ need not be defined are the points where the smooth parts of C are connected and the number of these points is at most finite. The line integral of a vector function is defined by means of the line integral of a scalar function and we already know that this integral does not depend on the behaviour of the integrand at points whose number is finite. (See paragraph IV.2.6, part c).)

IV.4.4. Remark. The line integral of a vector function can be denoted and written down in various ways. It is very important to understand them and to recognize correctly what they mean. We will explain one of the other possible ways of writing the line integral of a vector function in this paragraph.

Suppose that a vector function $\mathbf{f}$ has components U, V and W. Thus, $\mathbf{f} = (U, V, W) = U\cdot\mathbf{i}+V\cdot\mathbf{j}+W\cdot\mathbf{k}$. Comparing the two integrals $\int_C \mathbf{f}\cdot\vec{\tau}\,ds$ and $\int_C \mathbf{f}\cdot d\mathbf{s}$ which mean the same, we obtain the formal equality $\vec{\tau}\,ds = d\mathbf{s}$. The term $d\mathbf{s}$ is considered as an "infinitely short" tangent vector to curve C and its components are often denoted by dx, dy and dz. Thus, we have

$$\vec{\tau}\,ds \;=\; d\mathbf{s} \;=\; (dx, dy, dz) \;=\; \mathbf{i}\,dx + \mathbf{j}\,dy + \mathbf{k}\,dz.$$

The scalar product $\mathbf{f}\cdot d\mathbf{s}$ can now be expressed

$$\mathbf{f}\cdot d\mathbf{s} \;=\; (U, V, W)\cdot(dx, dy, dz) \;=\; U\,dx + V\,dy + W\,dz$$

and the line integral of the vector function $\mathbf{f}$ can be written as

$$\int_C \mathbf{f}\cdot\vec{\tau}\,ds \;=\; \int_C \mathbf{f}\cdot d\mathbf{s} \;=\; \int_C \big(U\,dx + V\,dy + W\,dz\big). \tag{IV.8}$$

It is clear that if C is a curve in $\mathbf{E}_2$ and $\mathbf{f}$ is a two–dimensional vector function with the components U and V then

$$\int_C \mathbf{f}\cdot\vec{\tau}\,ds \;=\; \int_C \mathbf{f}\cdot d\mathbf{s} \;=\; \int_C \big(U\,dx + V\,dy\big). \tag{IV.9}$$

IV.4.5. Example. Using the notation explained in the previous paragraph, you can observe that the integral $\int_C (2x^2+3y)\,dz$ is in fact the line integral of a vector function $(0,\,0,\,2x^2+3y)$:

$$\int_C (2x^2+3y)\,dz \;=\; \int_C 0\cdot dx + 0\cdot dy + (2x^2+3y)\,dz \;=\; \int_C \big(0,\,0,\,2x^2+3y\big)\cdot d\mathbf{s}.$$

IV.4.6. Remark. The line integral of vector function $\mathbf{f}$ is defined by means of the line integral of the scalar function $\mathbf{f} \cdot \vec{\tau}$ and so the main properties of the line integral of a vector function are the same as the properties of the line integral of a scalar function. Thus, we can simply rewrite items a), b) and c) of paragraph IV.2.6 with the function $\mathbf{f} \cdot \vec{\tau}$ and we obtain valid assertions for the line integral of a vector function. (Do it for yourself!)

The main and very important difference between the line integral of a scalar function and the line integral of a vector function is that <u>the line integral of a vector function depends on the orientation of the curve</u>. More precisely:

IV.4.7. Theorem. *If a vector function* $\mathbf{f}$ *is integrable on curve* C *then it is also integrable on curve* $-C$ *and*

$$\int_{-C} \mathbf{f} \cdot d\mathbf{s} = -\int_C \mathbf{f} \cdot d\mathbf{s}.$$

This theorem follows immediately from the definition of the line integral of a vector function. The integral $\int_C \mathbf{f} \cdot d\mathbf{s}$ is equal to the integral $\int_C \mathbf{f} \cdot \vec{\tau}\, ds$ where $\vec{\tau}$ is the unit tangent vector. It shows the direction given by the orientation of the curve. If we change the orientation then the unit tangent vector changes its sign and this leads to the change of sign of the integral.

IV.4.8. Evaluation of the line integral of a vector function. The line integral of a vector function $\mathbf{f}$ on a simple smooth curve C can be evaluated by means of a parametrization of C. Suppose that P is such a parametrization, defined in the interval $\langle a, b\rangle$. Suppose further that curve C is oriented in accordance with parametrization P. Then the unit tangent vector in every "interior" point of curve C can be expressed as $\vec{\tau} = \dot{P}(t)/|\dot{P}(t)|$. Now using the definition of the line integral of the vector function $\mathbf{f}$ and formula (IV.7), we obtain

$$\int_C \mathbf{f} \cdot d\mathbf{s} = \int_C \mathbf{f} \cdot \vec{\tau}\, ds = \int_a^b \mathbf{f}(P(t)) \cdot \frac{\dot{P}(t)}{|\dot{P}(t)|}\, |\dot{P}(t)|\, dt,$$

$$\int_C \mathbf{f} \cdot d\mathbf{s} = \int_a^b \mathbf{f}(P(t)) \cdot \dot{P}(t)\, dt. \tag{IV.10}$$

Substituting here $\mathbf{f} = (U, V, W)$, $P(t) = \big[\phi(t),\, \psi(t),\, \vartheta(t)\big]$ and $\dot{P}(t) = \big(\dot{\phi}(t),\, \dot{\psi}(t),\, \dot{\vartheta}(t)\big)$, we get:

$$\int_C \mathbf{f} \cdot d\mathbf{s} = \int_a^b \big[U\, \dot{\phi}(t) + V\, \dot{\psi}(t) + W\, \dot{\vartheta}(t)\big]\, dt \tag{IV.11}$$

where $U = U\big(\phi(t),\, \psi(t),\, \vartheta(t)\big)$, $V = V\big(\phi(t),\, \psi(t),\, \vartheta(t)\big)$ and $W = W\big(\phi(t),\, \psi(t),\, \vartheta(t)\big)$. Formula (IV.11) can also be formally obtained from (IV.8) if we use the substitution $x = \phi(t)$, $y = \psi(t)$, $z = \vartheta(t)$ and $dx = \dot{\phi}(t)\, dt$, $dy = \dot{\psi}(t)\, dt$, $dz = \dot{\vartheta}(t)\, dt$.

If curve C is not oriented in accordance with parametrization P (i.e. P generates the opposite orientation of C) then formula (IV.11) holds with the sign "$-$" in front of the integral on the right hand side.

The line integral of a vector function on a simple piecewise–smooth curve C which is a union of the simple smooth curves $C_1, C_2, \ldots, C_m$ (see paragraph IV.1.6 for details) can be evaluated in such a way that we first compute the integral on each smooth part $C_1, C_2, \ldots, C_m$ of curve C (e.g. by means of the parametrization of these parts) and then we use the fact that the integral on C is equal to the sum of the integrals on $C_1, C_2, \ldots, C_m$.

Finally, the line integral of a vector function can also be sometimes evaluated by means of the Green theorem, the Stokes theorem or formula (VI.1). (You will find the details in paragraphs IV.5.5., V.6.6 and VI.1.5.)

IV.4.9. Example. Find the work done by the force $\mathbf{f}(x,y,z) = (y-x^2)\,\mathbf{i} + (z-y^2)\,\mathbf{j} + (x-z^2)\,\mathbf{k}$ over the curve $C: \;\; P(t) = [t, t^2, t^3]; \;\; t \in \langle 0, 1\rangle$ from $[0,0,0]$ to $[1,1,1]$.

Curve C is defined by means of its parametrization P. Since $[0,0,0] = i.p.\,C = P(0)$ and $[1,1,1] = t.p.\,C = P(1)$, C is oriented in accordance with parametrization P. We can easily find that $\dot{P}(t) = (1, 2t, 3t^2)$. Using formula (IV.10), we obtain

$$\int_C \mathbf{f}\cdot d\mathbf{s} = \int_C \big(y-x^2,\, z-y^2,\, x-z^2\big)\cdot d\mathbf{s} = \int_0^1 \big(0,\, t^3-t^4,\, t-t^6\big)\cdot\big(1,\, 2t,\, 3t^2\big)\, dt =$$

$$= \int_0^1 \big[\, 2t^4 - 2t^5 + 3t^3 - 3t^8 \big]\, dt \; = \; \big[\, \tfrac{2}{5}t^5 - \tfrac{2}{6}t^6 + \tfrac{3}{4}t^4 - \tfrac{3}{9}t^9 \,\big]_0^1 \; = \; \tfrac{29}{60}.$$

IV.4.10. Circulation of a vector field around a closed curve. Let C be a closed curve in $\mathbf{E}_2$ or in $\mathbf{E}_3$ and let $\mathbf{f}$ be a vector field (= a vector function) defined on C. The line integral $\int_C \mathbf{f}\cdot d\mathbf{s}$ is called the <u>*circulation*</u> of $\mathbf{f}$ around curve C. In order to stress the fact that C is a closed curve, the integral is also often denoted as $\oint_C \mathbf{f}\cdot d\mathbf{s}$.

IV.5. Green's theorem.

This section deals with the line integral of a vector function on a closed curve in $\mathbf{E}_2$. The vector function is also supposed to be two–dimensional (i.e. to have two components).

The next theorem says something that is very obvious at first sight. We do not give the proof of the theorem. Nevertheless, if you were to see the proof, you would be surprised that it is not easy.

Bear in mind the convention that if nothing else is specified then "curve" denotes a simple piecewise–smooth curve and "closed curve" denotes a closed simple piecewise–smooth curve. (See paragraph IV.1.6.)

IV.5.1. Jordan's theorem. *Let C be a closed curve in $\mathbf{E}_2$. Then there exist two disjoint domains G_1 and G_2 in $\mathbf{E}_2$ such that C is their common boundary and*

a) $\mathbf{E}_2 = G_1 \cup C \cup G_2$,

b) one of the domains G_1, G_2 is bounded and the second one is unbounded.

IV.5.2. Interior and exterior of a closed curve in $\mathbf{E}_2$. Let C be a closed curve in $\mathbf{E}_2$ and G_1, G_2 be the domains whose existence is given by Jordan's theorem. That domain of G_1, G_2 which is bounded is called the *interior* of curve C and it is denoted by $Int\, C$. The second domain, which is unbounded, is called the *exterior* of C and it is denoted by $Ext\, C$.

IV.5.3. Positive and negative orientation of a closed curve in $\mathbf{E}_2$. Let C be a closed curve in $\mathbf{E}_2$. We say that C is oriented *positively* if, when moving on C in accordance with its orientation, we have its interior on our left–hand side. (See Fig. 9.) In the opposite case, i.e. in the case when the interior of C is on our right–hand side when moving along C in accordance with its orientation, we say that C is oriented *negatively*.

IV.5.4. Remark. Definition IV.5.3 is very simple and you can easily imagine what it says, because you know where you have your left hand and your right hand. However, you can also observe that this definition is not correct from a purely logical point of view. Why not? – It is clear: Mathematical notions must be defined precisely and must not depend on our knowledge of where we have our left hand and our right hand. In other words: How would you explain the above definition to an intelligent being (for example an extra–terrestrial) who does not have two hands and is not used to distinguishing between "left" and "right"?

Since the logically correct definition of the positive (respectively negative) orientation of a closed curve in $\mathbf{E}_2$ is not so easy and the above (not quite correct) definition is satisfactory for our purposes, we do not show the correct definition in this text.

IV.5.5. Green's theorem. *Suppose that*

a) a vector function $\mathbf{f} = (U, V)$ has continuous partial derivatives in domain $D \subset \mathbf{E}_2$,

b) C is a positively oriented closed curve in D such that $Int\ C \subset D$.

Then

$$\oint_C \mathbf{f} \cdot d\mathbf{s} = \iint_{Int\ C} \left(\frac{\partial V}{\partial x} - \frac{\partial U}{\partial y} \right) dx\, dy. \tag{IV.11}$$

IV.5.6. Remark. Using the formal equality $\mathbf{f} \cdot d\mathbf{s} = U\, dx + V\, dy$ as in (IV.8), we can write formula (IV.11) in the form

$$\oint_C U\, dx + V\, dy = \iint_{Int\ C} \left(\frac{\partial V}{\partial x} - \frac{\partial U}{\partial y} \right) dx\, dy. \tag{IV.12}$$

If all the assumptions of Green's theorem are satisfied with the exception that curve C is oriented negatively then formulas (IV.11) and (IV.12) hold with the sign "$-$" in front of the integrals on the right hand sides.

IV.5.7. Example. Evaluate the circulation of the vector field $\mathbf{f} = (-x^2y, xy^2)$ around the positively oriented circle $x^2 + y^2 = \alpha^2$ (where $\alpha > 0$).

If we denote the components of $\mathbf{f}$ by U and V then we get:

$$\frac{\partial V}{\partial x} - \frac{\partial U}{\partial y} = \frac{\partial}{\partial x}(xy^2) - \frac{\partial}{\partial y}(-x^2y) = x^2 + y^2.$$

It can easily be verified that all the assumptions of Green's theorem are satisfied and so we obtain:

$$\oint_C -x^2y\,dx + xy^2\,dy = \iint_{x^2+y^2\le\alpha^2}(x^2+y^2)\,dx\,dy \overset{1)}{=} \int_0^{2\pi}\Big(\int_0^{\alpha} r^3\,dr\Big)\,d\varphi = \frac{\pi\alpha^4}{2}.$$

[1]) We have transformed the double integral to the polar coordinates.

IV.5.8. Remark. If the components U and V of vector function $\mathbf{f}$ are such that $\partial V/\partial x - \partial U/\partial y = 1$ then Green's theorem can be used to evaluate the area of $Int\,C$. For example, if we choose $U = -\frac{1}{2}y$, $V = \frac{1}{2}x$ and C is a closed curve in $\mathbf{E}_2$ then

$$\tfrac{1}{2}\oint_C -y\,dx + x\,dy = \tfrac{1}{2}\iint_{Int\ C}\Big(\frac{\partial x}{\partial x} - \frac{\partial(-y)}{\partial y}\Big)\,dx\,dy = \iint_{Int\,C} dx\,dy = m_2(Int\ C).$$

IV.6. Exercises.

1. Decide about the existence of the integral $\int_C ds/(x^2+y^2)$ where C is the circle with its center at point S and radius 1.

a) $S = [0,0]$ b) $S = [1,0]$ $S = [0,-2]$

2. Evaluate the length of curve C which is defined by its parametrization.

a) $P(t) = [3t,\, 3t^2,\, 2t^3]$, $t \in \langle 0,1\rangle$

b) $P(t) = [a\cos t,\, a\sin t,\, bt]$, $t \in \langle 0,2\pi\rangle$ $(a>0,\ b>0)$

3. Evaluate the following integrals. (Which of them are are the line integrals of a scalar function and which of them are the line integrals of a vector function?)

a) $\displaystyle\int_C \frac{ds}{x-y}$; C is the part of the straight line $y = \frac{1}{2}x - 2$ between the points $[0,-2]$ and $[4,0]$

b) $\displaystyle\int_C y\ ds$; C is the part of the parabola $y^2 = 2px$ between the points $[0,0]$ and $[2p,2p]$ $(p>0)$

c) $\displaystyle\int_C xy\ ds$; C is the part of the ellipse $x^2/a^2 + y^2/b^2 = 1$ in the second quadrant

d) $\displaystyle\int_C \sqrt{2y}\ ds$; C is the part of the cycloid $x = a(t-\sin t)$, $y = a(1-\cos t)$ corresponding to $t \in \langle 0,2\pi\rangle$

e) $\displaystyle\int_C (x-y)\ ds$; C is the circle $x^2+y^2 = 2x$

f) $\displaystyle\int_C \frac{z^2}{x^2+y^2}\ ds$; C: $x = a\cos t$, $y = a\sin t$, $z = at$, $t\in\langle 0,2\pi\rangle$ $(a>0)$

g) $\int_C xyz\, ds$; C is the intersection of the surfaces $x^2 + y^2 + z^2 = R^2$ and $x^2 + y^2 = R^2/4$ in the first octant $(R > 0)$

h) $\int_C (x + y)\, ds$; C is the quarter of the circle $x^2 + y^2 + z^2 = R^2$, $y = x$ in the first octant

i) $\int_C x\, dy$; C is the triangle with its sides on the coordinate x– and y–axes and on the line $x/2 + y/3 = 1$, oriented positively

j) $\int_C (x + y)\, dx$; C is the line segment from $[a, 0]$ to $[0, b]$

k) $\int_C (x^2 - y^2)\, dx$; C is the part of the parabola $y = x^2$ from $[0, 0]$ to $[2, 4]$

l) $\int_C [-x \cos y\, dx + y \sin x\, dy]$; C is the line segment from $[0, 0]$ to $[\pi, 2\pi]$

m) $\int_C (y, -x) \cdot d\mathbf{s}$; C is the ellipse $x^2/a^2 + y^2/b^2 = 1$, oriented positively

n) $\int_C \frac{y^2\, dx - x^2\, dy}{x^2 + y^2}$; C is the part of the circle $x^2 + y^2 = a^2$ $(a > 0)$ in the first and in the second quadrant, oriented from $[a, 0]$ to $[-a, 0]$

o) $\int_C (2a - y, -a + y) \cdot d\mathbf{s}$; C: $x = a(t - \sin t)$, $y = a(1 - \cos t)$, $t \in \langle 0, 2\pi \rangle$; C is oriented from $[2\pi a, 0]$ to $[0, 0]$

p) $\int_C [y^2\, dx + z^2\, dy + x^2\, dz]$; C is the intersection of the sphere $x^2 + y^2 + z^2 = R^2$ with the cylindrical surface $x^2 + y^2 = Rx$ $(R > 0)$, $z \geq 0$; C is oriented positively as viewed from the origin $O = [0, 0, 0]$

q) $\int_C [2xy\, \mathbf{i} - x^2 \mathbf{j}] \cdot d\mathbf{s}$; C is the union of the line segments leading from $[0, 0]$ to $[2, 0]$ and from $[2, 0]$ to $[2, 1]$

r) $\int_C [yz\, \mathbf{i} + z\sqrt{R^2 - y^2}\, \mathbf{j} + xy\, \mathbf{k}] \cdot d\mathbf{s}$; C: $x = R \cos t$, $y = R \sin t$, $z = at/(2\pi)$ $(a > 0)$, C is oriented from its intersection with the plane $z = 0$ to its intersection with the plane $z = a$

s) $\int_C (1 - x^2)y\, dx + x(1 + y^2)\, dy$; C is the boundary of the square $\langle 0, 2 \rangle \times \langle 0, 2 \rangle$ oriented positively

t) $\int_C (e^{xy} + 2x \cos y)\, dx + (e^{xy} - x^2 \sin y)\, dy$; C is the circle $x^2 + y^2 = 8$ oriented positively

u) $\int_C (xy + x + y)\, dx + (xy + x - y)\, dy$; C is the ellipse $9x^2 + 36x + 4y^2 = 0$ oriented negatively

v) $\int_C (x + y)\, dx - 2x\, dy$; C is the boundary of the triangle with its sides on the lines $x = 0$, $y = 0$, $x + y = 5$, oriented negatively

w) $\int_C (dx/y - dy/x)$; C is the boundary of the triangle with the vertices $[1, 1]$, $[2, 1]$, $[2, 2]$, oriented positively

x) $\displaystyle\int_C [2(x^2+y^2)\,\mathbf{i}+(x+y)^2\,\mathbf{j}]\;d\mathbf{s}$; C is the boundary of the triangle with the vertices $[1,1]$, $[2,2]$, $[1,3]$, oriented positively

4. Evaluate the work done by force $\mathbf{f}$ over curve C.

a) $\mathbf{f}=(x-y,x)$, C is the boundary of the square with the vertices $[-2,-2]$, $[1,-2]$, $[1,1]$, $[-2,1]$, oriented clockwise

b) $\mathbf{f}=(x+y,2x)$, C is the circle $x=a\cos t,\ y=a\sin t,\ t\in\langle 0,2\pi\rangle$

c) $\mathbf{f}=(y,2)$, C is the closed curve which consists of the semi–axes and the quarter of the ellipse $x=2\cos t,\ y=\sin t,\ t\in\langle 0,2\pi\rangle$ in the first quadrant

5. C_1 is the line segment from $[0,0]$ to $[1,1]$, C_2 is the part of the parabola $y=x^2$ from $[0,0]$ to $[1,1]$, $I_1=\int_{C_1}(x+y)^2\,dx-(x-y)^2\,dy$, $I_2=\int_{C_2}(x+y)^2\,dx-(x-y)^2\,dy$. Applying the Green theorem, evaluate the difference I_1-I_2.

6. Using the line integral, evaluate the area of the interior of the closed curve which consists of the arc of the cycloid $x=a(t-\sin t),\ y=a(1-\cos t),\ t\in\langle 0,2\pi\rangle$ and the line segment connecting the points $[0,0]$ and $[2\pi a,0]$.

7. Using the line integral, derive the formula for the area of the interior of the ellipse $x^2/a^2+y^2/b^2=1$.

8. Using the line integral, evaluate the area of the interior of the closed curve whose equation in the polar coordinates is $r=a(1-\cos\varphi)$ where $a>0$ (a so called "cardioid").

9. Using the line integral, evaluate the area of the interior of a so called asteroid, whose equation is $x^{2/3}+y^{2/3}=a^{2/3}$ $(a>0)$. (You can use its parametric equations $x=a\cos^3 t,\ y=a\sin^3 t,\ t\in\langle 0,2\pi\rangle$.)

10. Evaluate the circulation of vector field $\mathbf{f}$ around the closed curve C. If it is possible, apply the Green theorem.

a) $\mathbf{f}(x,y)=\big(e^x\sin y-y^2,\ e^x\cos y-1\big)$, $C=C_1\cup C_2$, $C_1=\big\{[x,y]\in\mathbf{E}_2;\ x^2+y^2+2x=0,\ y\le 0\big\}$, $C_2=\big\{[x,y]\in\mathbf{E}_2;\ -1\le x\le 0,\ y=0\big\}$, C is oriented positively.

b) $\mathbf{f}(x,y)=(x+y)\,\mathbf{i}+(y-x)\,\mathbf{j}$, $C=\big\{[x,y]\in\mathbf{E}_2;\ x^2/a^2+y^2/b^2=1\big\}$, C is oriented negatively.

c) $\mathbf{f}(x,y)=(x^2,y^2)$, C is the perimeter of a triangle with the vertices $A=[1,1]$, $B=[2,1]$, $D=[2,3]$, oriented positively.

d) $\mathbf{f}(x,y)=(y^2,-x)$, C is the perimeter of a triangle with the vertices $A=[1,1]$, $B=[1,3]$, $D=[2,2]$, oriented negatively.

V. Surface Integrals

V.1. Simple surfaces.

We will work with two types of surfaces in $\mathbf{E}_3$: so called simple smooth surfaces and so called simple piecewise–smooth surfaces.

The idea of the definition of a simple smooth surface is the following. Imagine that you have a subset B of $\mathbf{E}_2$, bounded by a closed curve Γ. You cut set B from $\mathbf{E}_2$, you move it somewhere to $\mathbf{E}_3$ and you deform it elastically so that you do not disturb its smoothness. This means that you can stretch it in various directions, but you cannot break it and you cannot paste two different points together. Thus, you get a simple smooth surface in $\mathbf{E}_3$.

This can be easily expressed mathematically. The described procedure moves every point $[u,v] \in B$ to some other point $P(u,v) = \big(\phi(u,v),\, \psi(u,v),\, \vartheta(u,v)\big)$ in $\mathbf{E}_3$. Thus, P is a mapping of B to $\mathbf{E}_3$. The requirement that the deformation of B is elastic and smooth leads to the condition that P (i.e. the functions ϕ, ψ and ϑ) is continuous and has continuous partial derivatives in a sufficiently large subset of B. The requirement that two different points belonging originally to B cannot be pasted together leads to the condition that mapping P is one–to–one in B.

The functions

$$x = \phi(u,v), \quad y = \psi(u,v), \quad z = \vartheta(u,v)$$

are called the *coordinate functions* of mapping P.

We shall denote the partial derivatives of mapping P with respect to the variables u, v by P_u and P_v and we shall work with them as with vectors. Hence

$$P_u(u,v) = \Big(\frac{\partial\phi(u,v)}{\partial u}, \frac{\partial\psi(u,v)}{\partial u}, \frac{\partial\vartheta(u,v)}{\partial u}\Big) \quad \text{or shortly} \quad P_u = \Big(\frac{\partial\phi}{\partial u}, \frac{\partial\psi}{\partial u}, \frac{\partial\vartheta}{\partial u}\Big),$$

$$P_v(u,v) = \Big(\frac{\partial\phi(u,v)}{\partial v}, \frac{\partial\psi(u,v)}{\partial v}, \frac{\partial\vartheta(u,v)}{\partial v}\Big) \quad \text{or shortly} \quad P_v = \Big(\frac{\partial\phi}{\partial v}, \frac{\partial\psi}{\partial v}, \frac{\partial\vartheta}{\partial v}\Big).$$

The vector product of vectors P_u and P_v will be denoted by $P_u \times P_v$. Have in mind that

$$P_u \times P_v = \begin{vmatrix} \mathbf{i}, & \mathbf{j}, & \mathbf{k} \\ \frac{\partial\phi}{\partial u}, & \frac{\partial\psi}{\partial u}, & \frac{\partial\vartheta}{\partial u} \\ \frac{\partial\phi}{\partial v}, & \frac{\partial\psi}{\partial v}, & \frac{\partial\vartheta}{\partial v} \end{vmatrix} = \Big(\frac{\partial\psi}{\partial u}\frac{\partial\vartheta}{\partial v} - \frac{\partial\vartheta}{\partial u}\frac{\partial\psi}{\partial v},\ \frac{\partial\vartheta}{\partial u}\frac{\partial\phi}{\partial v} - \frac{\partial\phi}{\partial u}\frac{\partial\vartheta}{\partial v},\ \frac{\partial\phi}{\partial u}\frac{\partial\psi}{\partial v} - \frac{\partial\psi}{\partial u}\frac{\partial\phi}{\partial v}\Big).$$

We describe the notion of a simple smooth surface once again, this time precisely, in the following definition.

V.1.1. Simple smooth surface. Let $\Omega \subset \mathbf{E}_2$, $P = (\phi, \psi, \vartheta)$ be a mapping of Ω to $\mathbf{E}_3$, Γ be a closed simple piecewise–smooth curve in Ω, and let $B = \Gamma \cup Int\,\Gamma$. Suppose that

a) mapping P is continuous and one–to–one in B,

b) P has continuous and bounded partial derivatives P_u and P_v in $B - K$ where $K = \emptyset$ or K is a finite set of points on the boundary Γ of set B,

c) $P_u \times P_v \neq \vec{0}$ in $B - K$.

Then the set $\sigma = \{X = P(u,v) \in \mathbf{E}_3;\ [u,v] \in B\}$ is called the *simple smooth surface* in $\mathbf{E}_3$. Mapping P is called the *parametrization* of surface σ. The set $C = \{X = P(u,v) \in \mathbf{E}_3;\ [u,v] \in \Gamma\}$ is called the *boundary* of surface σ.

The boundary of a simple smooth surface is a closed simple piecewise–smooth curve in $\mathbf{E}_3$. Instead of the word boundary, you may also find the denotation "contour" or "margin" in literature.

Every simple smooth surface has infinitely many parametrizations. (Compare with the analogous statement about the simple smooth curve in paragraph IV.1.1.)

Our definition of a simple smooth surface is relatively straightforward. However, this is paid for by the fact that, for instance, a "nice" surface like a sphere is not a simple smooth surface. All attempts to modify the definition of a simple smooth surface so that it will also include the sphere always lead to such great complications that they do not pay off. This is a consequence of the geometrical structure of the three–dimensional space $\mathbf{E}_3$ – it provides such a variety of possible forms of surfaces that we must be very careful in order to avoid confusion in our definitions and theorems. However, you will see that we do not exclude spheres (and other similar surfaces) from the class of surfaces that we will deal with – they can be treated as so called simple piecewise–smooth surfaces, whose definition is given in paragraphs V.1.5 and V.1.6.

V.1.2. Orientation of a simple smooth surface. Normal vector. Let P be a parametrization of a simple smooth surface σ, defined in set $B \subset \mathbf{E}_2$, and let $X = P(u,v)$ for $[u,v] \in B - K$ (see definition V.1.1). Then the vectors $P_u(u,v)$ and $P_v(u,v)$ are tangent to σ at point X and due to condition c) from definition V.1.1, they are linearly independent. Their vector product is perpendicular to both of them, and so it is also perpendicular to surface σ. If we divide the vector product by its length, we obtain a unit vector, perpendicular to σ at point X.

We can choose the *orientation* of surface σ in such a way that we put the *normal vector* $\mathbf{n}$ (i.e. the vector which is perpendicular to σ, its length is one and its direction shows the orientation of σ)

either
$$\mathbf{n} = \frac{P_u(u,v) \times P_v(u,v)}{|P_u(u,v) \times P_v(u,v)|} \quad \text{for all } [u,v] \in B - K \tag{V.1}$$

or
$$\mathbf{n} = -\frac{P_u(u,v) \times P_v(u,v)}{|P_u(u,v) \times P_v(u,v)|} \quad \text{for all } [u,v] \in B - K. \tag{V.2}$$

If $\mathbf{n}$ is given by formula (V.1) then we say that the simple smooth surface σ is oriented in accordance with its parametrization P.

Thus, the simple smooth surface σ is oriented by the choice of a normal vector $\mathbf{n}$ (i.e. a unit vector perpendicular to σ) at any point where the perpendicular direction to σ is defined. The normal vector is oriented to the same side of σ at every point where it exists. This means that it changes continuously if you move on σ.

V.1.3. The relation between the orientation of a simple smooth surface and its boundary. We say that boundary C of a simple smooth surface σ is *oriented in accordance with* σ if your left hand shows the direction of the normal vector $\mathbf{n}$ on σ when you move on C in the sense of its orientation. (See Fig. 11.)

It is obvious that this definition is not logically quite correct (for the same reasons as in the case of definition IV.5.3). Nevertheless, it is instructive, simple and it cannot lead to confusion.

V.1.4. Example. Surface σ is a part of the cone $z = \sqrt{x^2 + y^2}$, corresponding to $x \geq 0$ and $x^2 + y^2 \leq 4$. It is oriented "upwards", i.e. the third component of the normal vector is positive. Show that σ is a simple smooth surface, find its parametrization, decide whether σ is or is not oriented in accordance with your parametrization, and define the orientation of the boundary of σ so that it is oriented in accordance with the orientation of σ.

In order to parametrize σ, we can use the equation $z = \sqrt{x^2+y^2}$: We can put

$$P \;:\quad x = \phi(u,v) = u, \quad y = \psi(u,v) = v, \quad z = \vartheta(u,v) = \sqrt{u^2+v^2}$$

for $[u,v] \in B$ where set B is defined by means of the conditions $u \geq 0$ and $u^2+v^2 \leq 4$:

$$B \;=\; \{[u,v] \in \mathbf{E}_2;\; u \geq 0,\; u^2+v^2 \leq 4\}.$$

It can be verified that mapping P has all the properties which are required in definition V.1.1, and so σ is a simple smooth surface and P is the parametrization of σ. The partial derivatives P_u and P_v are

$$P_u(u,v) \;=\; \Big(1,\; 0,\; \frac{u}{\sqrt{u^2+v^2}}\Big), \qquad P_v(u,v) \;=\; \Big(0,\; 1,\; \frac{v}{\sqrt{u^2+v^2}}\Big),$$

their vector product is

$$P_u \times P_v \;=\; \Big(-\frac{u}{\sqrt{u^2+v^2}},\; -\frac{v}{\sqrt{u^2+v^2}},\; 1\Big)$$

and the length of this vector product is $\sqrt{2}$. Thus, the unit vector $P_u \times P_v/|P_u \times P_v|$ perpendicular to σ equals $P_u \times P_v/\sqrt{2}$ and it is seen that the third component of this vector is positive. Hence it coincides with the given normal vector $\mathbf{n}$ and so we can say that surface σ is oriented in accordance with our parametrization P.

The orientation of the boundary C of σ which corresponds to the orientation of σ is marked in Fig. 12. For example, the unit tangent vector to C at the point $X = [2,0,2]$ is $\vec{\tau} = (0,1,0)$.

V.1.5. A simple piecewise–smooth surface consisting of two simple smooth surfaces. Suppose that σ_1 and σ_2 are two oriented simple smooth surfaces whose boundaries C_1 and C_2 are either both oriented in accordance with σ_1 and σ_2 or they are both oriented opposite to σ_1 and σ_2. Suppose that

a) $\sigma_1 \cap \sigma_2 = C_1 \cap C_2$ and this set forms a simple piecewise–smooth curve or more (a finite number of) such curves,

b) the orientation of C_1 and C_2 is opposite in all points of $C_1 \cap C_2$. (See Fig. 13.)

Then the union $\sigma = \sigma_1 \cup \sigma_2$ is called a *simple piecewise–smooth surface* in $\mathbf{E}_3$, consisting of two simple smooth surfaces σ_1 and σ_2.

The orientation of σ is given by the orientation of σ_1 and σ_2.

The boundary of σ is the closure of the set $(C_1 \cup C_2) - (C_1 \cap C_2)$. (See Fig. 13.) It is either empty, or it is one simple piecewise–smooth curve (see Fig. 13), or it consists of more (a finite number of) simple piecewise–smooth curves.

V.1.6. A simple piecewise–smooth surface consisting of more simple smooth surfaces. Let σ_1 and σ_2 be the simple smooth surfaces from the previous paragraph. If we successively, respecting the same rules, connect other simple smooth surfaces $\sigma_3, \sigma_4, \ldots, \sigma_m$ to the union $\sigma \cup \sigma_2$, we obtain a simple piecewise–smooth surface in $\mathbf{E}_3$ which consists of m simple smooth surfaces $\sigma_1, \sigma_2, \ldots, \sigma_m$. (See Fig. 14.)

The surface which differs from a simple piecewise–smooth surface σ only by its orientation will be denoted by $-\sigma$.

V.1.7. A closed simple piecewise– smooth surface. A simple piecewise – smooth surface σ whose boundary is the empty set is called *closed*.

V.1.8. Example. A surface which consists of two simple smooth surfaces $\sigma_1 : x^2 + (y+1)^2 = 2;\ y \geq 0$ and $\sigma_2 : x^2 + (y-1)^2 = 2;\ y \leq 0$ is a closed simple piecewise–smooth surface.

Other examples of closed simple piecewise–smooth surfaces are: the surface of a cube, a sphere, an ellipsoid, etc.

Similarly as in the case of curves, we can make an agreement that whenever we will use the word "surface", we will mean a simple piecewise–smooth surface. A "closed surface" will mean a closed simple piecewise–smooth surface. More details about the surfaces will be specified if they are important and necessary.

V.2. The surface integral of a scalar function.

V.2.1. Physical motivation. Suppose that a desk has the form of a simple smooth surface σ in $\mathbf{E}_3$ and its surface density (i.e. amount of mass per unit area) is ρ. ρ is generally a function of three variables x, y, z. We wish to evaluate the total mass M of the desk.

Suppose that P is a parametrization of σ which is defined in set $B \subset \mathbf{E}_2$. Imagine that B can be decomposed to infinitely many "infinitely small" squares of the form $\langle u, u+du \rangle \times \langle v, v+dv \rangle$. P maps each of these squares to the part of σ. Since the square $\langle u, u+du \rangle \times \langle v, v+dv \rangle$ is supposed to be "infinitely small", its image on σ can be taken for an "infinitely small" parallelogram with the vertices $A_1 = P(u,v)$, $A_2 = P(u+du, v) = P(u,v) + P_u(u,v) \cdot du$, $A_3 = P(u+du, v+dv) = P(u,v) + P_u(u,v) \cdot du + P_v(u,v) \cdot dv$ and $A_4 = P(u, v+dv) = P(u,v) + P_v(u,v) \cdot dv$. (See Fig. 15.) Its area is $dp = |A_2 - A_1| \cdot |A_3 - A_1| \cdot \sin\alpha$ and this can be expressed as $dp = |(A_2 - A_1) \times (A_3 - A_1)|$. Substituting here for points A_1, A_2, A_3, we obtain: $dp = |P_u(u,v) \times P_v(u,v)|\, du\, dv$. The mass of the parallelogram $A_1A_2A_3A_4$ is $dM = \rho(A) \cdot dp = \rho(P(u,v)\, |P_u(u,v) \times P_v(u,v)|\, du\, dv$ and the total mass of the whole desk (surface) σ is

$$M = \iint_B \rho(P(u,v)) \cdot |P_u(u,v) \times P_v(u,v)|\, du\, dv.$$

V.2.2. The surface integral of a scalar function on a simple smooth surface. Let σ be a simple smooth surface in $\mathbf{E}_3$ and P be its parametrization defined in set $B \subset \mathbf{E}_2$. Let f be a scalar function defined on σ. We say that f is *integrable* on surface σ if the double integral $\iint_B f(P(u,v)) \cdot |P_u(u,v) \times P_v(u,v)|\, du\, dv$ exists. We denote this integral by $\iint_\sigma f\ dp$ and we call it the *surface integral of a scalar function* f on the simple smooth surface σ.

V.2.3. Remark. The integrability of function f on a simple smooth surface σ and the surface integral $\iint_\sigma f\ dp$ are defined by means of a parametrization of surface σ. However, analogously to the line integral of a scalar function (see remark IV.2.3), it can be proved that neither the existence nor the value of the surface integral $\iint_\sigma f\ dp$ depends on the concrete choice of parametrization of surface σ.

V.2.4. The surface integral of a scalar function on a simple piecewise–smooth surface. Let σ be a simple piecewise–smooth surface in $\mathbf{E}_3$ which is a union of simple smooth surfaces $\sigma_1, \sigma_2, \dots, \sigma_m$ (see paragraphs V.1.5 and V.1.6). Let f be a scalar function defined on σ. If function f is integrable on each of surfaces $\sigma_1, \sigma_2, \dots, \sigma_m$ then we say that it is integrable on surface σ and we put

$$\iint_\sigma f\, dp = \sum_{i=1}^{m} \iint_{\sigma_i} f\, dp. \tag{V.3}$$

The integral on the left hand side is called the *surface integral of the scalar function* f on the simple piecewise–smooth surface σ.

The surface integral of a scalar function is also often called the *surface integral of the 1st kind*.

Instead of $\iint_\sigma f\, dp$, we can also write for example $\iint_\sigma f(x, y, z)\, dp$.

V.2.5. The area of a surface. If σ is a simple piecewise–smooth surface, then the value of the integral $\iint_\sigma dp$ is called the *area* of surface σ.

Specifically, if σ is a simple smooth surface and P is its parametrization defined in set $B \subset \mathbf{E}_2$ then the area of surface σ is

$$p(\sigma) = \iint_\sigma dp = \iint_B |P_u(u,v) \times P_v(u,v)|\, du\, dv. \tag{V.4}$$

V.2.6. Some important properties of the surface integral of a scalar function. Since the surface integral of a scalar function is defined by means of the double integral, its basic properties are the same as the corresponding properties of the double integral. Let us mention only some of them:

a) **(Existence of the surface integral.)** *If function f is continuous on surface σ then it is integrable on σ (i.e. the integral $\iint_\sigma f\, dp$ exists).*

b) **(Linearity of the surface integral.)** *If functions f and g are integrable on surface σ and $\alpha \in \mathbf{R}$ then*

$$\iint_\sigma (f+g)\, dp = \iint_\sigma f\, dp + \iint_\sigma g\, dp,$$

$$\iint_\sigma \alpha \cdot f\, dp = \alpha \cdot \iint_\sigma f\, dp.$$

c) *If function f is integrable on surface σ and function g differs from f at most in a finite number of points or curves then g is also integrable on σ and*

$$\iint_\sigma g\, dp = \iint_\sigma f\, dp.$$

d) *If function f is integrable on surface σ then it is also integrable on surface $-\sigma$ and*

$$\iint_{-\sigma} f\, dp = \iint_\sigma f\, dp.$$

Assertion a) can be generalized in this way: *If σ is a simple piecewise–smooth surface and f is continuous on each of its smooth parts then f is integrable on σ.*

Assertion d) says that neither the existence nor the value of the surface integral of a scalar function depends on the orientation of the surface.

V.2.7. Evaluation of the surface integral of a scalar function. The surface integral of function f on a simple smooth surface σ can be evaluated by means of a parametrization of σ. Thus, if P is such a parametrization, defined in set $B \subset \mathbf{E}_2$, and function f is integrable on σ then we can use the formula

$$\iint_\sigma f\, dp = \iint_B f(P(u,v)) \cdot |P_u(u,v) \times P_v(u,v)|\, du\, dv. \tag{V.5}$$

This formula follows immediately from the definition of the surface integral on a simple smooth surface – see paragraph V.2.2.

If σ is a simple piecewise–smooth surface which is a union of simple smooth surfaces σ_1, σ_2, ..., σ_m (see paragraphs V.1.5 and V.1.6) then the surface integral of function f on surface σ can be computed by means of formula (V.3).

V.2.8. Example. Integrate the function $f(x,y,z) = x + 2y$ over the the surface $\sigma:\ x+y+z=1,\ x \geq 0,\ y \geq 0,\ z \geq 0$.

Surface σ can be parametrized by the mapping

$$P(u,v): \quad x = u, \quad y = v, \quad z = 1 - u - v\,; \quad [u,v] \in B$$

where $B = \{[u,v] \in \mathbf{E}_2;\ 0 \leq u \leq 1,\ 0 \leq v \leq 1-u\}$. We can find that $P_u = (1,\ 0,\ -1)$, $P_v = (0,\ 1,\ -1)$, $P_u \times P_v = (1,1,1)$ and $|P_u \times P_v| = \sqrt{3}$. Using formula (V.5), we get

$$\iint_\sigma (x+2y)\, dp = \iint_B (u+2v)\sqrt{3}\, du\, dv = \sqrt{3}\int_0^1 \int_0^{1-u} (u+2v)\, dv\, du =$$
$$= \sqrt{3}\int_0^1 (1-u)\, du = \sqrt{3}/2.$$

V.2.9. Example. Integrate the function $g(x,y,z) = xyz$ over the surface of the cube cut from the first octant by the planes $x = 1$, $y = 1$ and $z = 1$.

The cube can also be expressed as the Cartesian product $\langle 0,1\rangle \times \langle 0,1\rangle \times \langle 0,1\rangle$. Its surface has six sides. Since $xyz = 0$ on the three sides that lie in the coordinate planes, the integral over the surface of the cube is equal to

$$\iint_{\sigma_1} xyz\, dp + \iint_{\sigma_2} xyz\, dp + \iint_{\sigma_3} xyz\, dp$$

where σ_1 is the square region $x = 1$, $0 \leq y \leq 1$, $0 \leq z \leq 1$, σ_2 is the square region $y = 1$, $0 \leq x \leq 1$, $0 \leq z \leq 1$ and σ_3 is the square region $z = 1$, $0 \leq x \leq 1$, $0 \leq y \leq 1$. σ_1 can be naturally parametrized by the mapping

$$P(u,v): \quad x = 1, \quad y = u, \quad z = v\,; \quad [u,v] \in B = \langle 0,1\rangle \times \langle 0,1\rangle.$$

It is obvious that $P_u = (0,1,0)$, $P_v = (0,0,1)$, $P_u \times P_v = (1,0,0)$ and $|P_u \times P_v| = 1$. Using formula (V.5), we obtain

$$\iint_{\sigma_1} xyz \; dp \;=\; \int_0^1 \int_0^1 uv \; du\, dv \;=\; \tfrac{1}{4}.$$

Due to the symmetry, the integrals over σ_2 and σ_3 are also $\frac{1}{4}$. Hence, the integral over the surface of the cube is equal to $\frac{3}{4}$.

V.2.10. Example. Although the sphere $\sigma_R: \; x^2+y^2+z^2 = R^2$ is not a simple smooth surface, there exists a mapping P of a closed set $B \subset \mathbf{E}_2$ onto the sphere which has all the properties of a parametrization (see paragraph V.1.1) with one exception: it is not one–to–one on the whole set B. (It is one–to–one in the interior of B, but not on the boundary of B.) Mapping P is defined by the equations

$$\begin{aligned} x &= \phi(u,v) = R \cos u \cos v, \\ y &= \psi(u,v) = R \sin u \cos v, \\ z &= \vartheta(u,v) = R \sin v \end{aligned}$$

for $u \in \langle 0, 2\pi \rangle$, $v \in \langle -\pi/2, \pi/2 \rangle$. (You can observe that the background of P is the expression of the coordinates of the points of sphere σ_R in the spherical coordinates.) Since P fail to satisfy all the requirements on the parametrization only on the set of two–dimensional measure zero (the boundary of B) and it is already known that the behaviour of integrands on sets of two–dimensional measure zero does not affect double or surface integrals, P can be used in the evaluation of the surface integral on sphere σ_R in the same way as if it were a parametrization. (In fact, mappings whose properties differ from the required properties of parametrizations only on sets of two–dimensional measure zero are also often, not quite correctly, called the parametrizations.)

Thus, if for example $f(x,y,z) = x^2 + y^2$ then, using formula (V.5), we obtain

$$\iint_{\sigma_R} f(x,y,z) \; dp \;=\; \iint_B R^2 \cos^2 v \, |P_u(u,v) \times P_v(u,v)| \; du\, dv.$$

Vectors P_u, P_v, $P_u \times P_v$ and the number $|P_u \times P_v|$ are:

$$\begin{aligned} P_u(u,v) &= \big(-R \sin u \cos v, \; R \cos u \cos v, \; 0\big), \\ P_v(u,v) &= \big(-R \cos u \sin v, \; -R \sin u \sin v, \; R \cos v\big), \\ P_u(u,v) \times P_v(u,v) &= \big(R^2 \cos u \cos^2 v, \; R^2 \sin u \cos^2 v, \; R^2 \sin v \cos v\big), \\ |P_u(u,v) \times P_v(u,v)| &= R^2 \cos v. \end{aligned}$$

Substituting this to the above integral and applying Fubini's theorem III.3.2, we get

$$\iint_B R^2 \cos^2 v \, |P_u(u,v) \times P_v(u,v)| \; du\, dv \;=\; \int_0^{2\pi} \Big(\int_{-\pi/2}^{\pi/2} R^4 \cos^3 v \, dv \Big) \, du \;=\; \tfrac{8}{3} \, \pi \, R^4.$$

V.2.11. Remark. The approach explained in example V.2.10 can also be used in connection with other simple piecewise–smooth surfaces, such as ellipsoids and conic surfaces. If for instance σ is the conic surface $x^2+y^2=z^2$ corresponding to $z \in \langle 0,4\rangle$ then the mapping

$$P: \quad x=\phi(u,v)=u\cos v, \quad y=\psi(u,v)=u\sin v, \quad z=\vartheta(u,v)=u$$

(defined for $u \in \langle 0,2\rangle$, $v \in \langle 0,2\pi\rangle$) has similar properties as mapping P from example V.2.10: It satisfies all the conditions of the parametrization (see paragraph V.1.1) with the exception that it is one–to–one only in the interior of its domain, i.e. in $(0,2\pi)\times(0,2)$ and not in $\langle 0,2\pi\rangle\times\langle 0,2\rangle$. Nevertheless, mapping P can be used in the evaluation of the surface integral on σ in the same way as if it were a parametrization of σ.

V.3. Some physical applications of the surface integral of a scalar function.

Suppose that a desk has the form of surface σ in $\mathbf{E}_3$. The desk need not be homogeneous, and so its surface density (amount of mass per unit of area) need not be constant. Let the density be given by function $\rho(x,y,z)$. The surface integral of a scalar function can be used to define and evaluate some mechanical characteristics of surface σ. Suppose that ρ is expressed in $[\mathrm{kg}\cdot\mathrm{m}^{-2}]$. Then we have:

Mass $\quad M = \iint_\sigma \rho(x,y,z)\,dp \quad [\mathrm{kg}]$,

Static moment about the xy–plane $\quad M_{xy} = \iint_\sigma z\cdot\rho(x,y,z)\,dp \quad [\mathrm{kg}\cdot\mathrm{m}]$,

Static moment about the xz–plane $\quad M_{xz} = \iint_\sigma y\cdot\rho(x,y,z)\,dp \quad [\mathrm{kg}\cdot\mathrm{m}]$,

Static moment about the yz–plane $\quad M_{yz} = \iint_\sigma x\cdot\rho(x,y,z)\,dp \quad [\mathrm{kg}\cdot\mathrm{m}]$,

Center of mass $[x_m,y_m,z_m]$ $\quad x_m=\dfrac{M_{yz}}{M}, \quad y_m=\dfrac{M_{xz}}{M}, \quad z_m=\dfrac{M_{xy}}{M} \quad [\mathrm{m}]$,

Moment of inertia about the x–axis $\quad J_x = \iint_\sigma (y^2+z^2)\cdot\rho(x,y,z)\,dp \quad [\mathrm{kg}\cdot\mathrm{m}^2]$,

Moment of inertia about the y–axis $\quad J_y = \iint_\sigma (x^2+z^2)\cdot\rho(x,y,z)\,dp \quad [\mathrm{kg}\cdot\mathrm{m}^2]$,

Moment of inertia about the z–axis $\quad J_z = \iint_\sigma (x^2+y^2)\cdot\rho(x,y,z)\,dp \quad [\mathrm{kg}\cdot\mathrm{m}^2]$,

Moment of inertia about the origin $\quad J_0 = \iint_\sigma (x^2+y^2+z^2)\cdot\rho(x,y,z)\,dp \quad [\mathrm{kg}\cdot\mathrm{m}^2]$.

Derive the formula for the moment of inertia about a general straight line in $\mathbf{E}_3$ whose parametric equations are $x=x_0+u_1t$, $y=y_0+u_2t$, $z=z_0+u_3t$; $t\in\mathbf{R}$!

V.4. The surface integral of a vector function.

V.4.1. Physical motivation. Suppose that σ is a surface in the flow of an incompressible fluid and we wish to express the flux of the fluid through surface σ per unit time. By "flux", we understand the volume of the fluid that flows through the surface. Suppose that the fluid moves with a steady velocity $\mathbf{v}(x, y, z)$, and $\mathbf{n}(x, y, z)$ is the normal vector to σ at the point $[x, y, z]$. The flux of the fluid through an "infinitely small" part of surface σ which finds itself at $[x, y, z]$ and its area is dp is $\mathbf{v}(x, y, z) \cdot \mathbf{n}(x, y, z)\ ds$. Thus, the total flux through the whole surface σ is $\iint_\sigma \mathbf{v}(x, y, z) \cdot \mathbf{n}(x, y, z)\ ds$.

The same approach can also be used if we wish to evaluate e.g. the flux of a magnetic field through a given surface.

Let us recall that the idea of an "infinitely small" part of σ is not logically quite precise (see also Section II.7 for further details). However, if we apply the idea carefully, it can be useful especially in situations when we need to derive formulas expressing various geometrical and physical quantities.

V.4.2. The surface integral of a vector function. Let σ be a simple piecewise–smooth surface in $\mathbf{E}_3$ and let $\mathbf{f}$ be a vector function (with three components) defined on σ. We say that the vector function $\mathbf{f}$ is integrable on surface σ if the scalar function $\mathbf{f} \cdot \mathbf{n}$ is integrable on σ (in the sense explained in paragraphs V.2.2 and V.2.4). The integral $\iint_\sigma \mathbf{f} \cdot \mathbf{n}\ dp$ is called the <u>*surface integral of a vector function*</u> $\mathbf{f}$ on surface σ, and it is usually denoted by $\iint_\sigma \mathbf{f} \cdot d\mathbf{p}$.

The surface integral of a vector function is also often called the <u>*surface integral of the 2nd kind*</u>. It defines the <u>*flux*</u> of vector field $\mathbf{f}$ through surface σ.

V.4.3. Remark. The fact that the normal vector $\mathbf{n}$ need not exist in all points of a simple piecewise–smooth surface σ does not matter. $\mathbf{n}$ need not be defined at points where the smooth parts of σ are connected and they form at most a finite number of lines. The surface integral of a vector function is defined by means of the surface integral of a scalar function and we already know that this integral does not depend on the behaviour of the integrand in a finite number of points or curves. (See paragraph V.2.6, part c).)

V.4.4. Remark. It is very important to understand various ways in which the surface integral of a vector function can be written down, and to recognize correctly what they mean.

If the vector function $\mathbf{f}$ has components U, V and W then the integrals

$$\iint_\sigma \mathbf{f} \cdot d\mathbf{p}, \quad \iint_\sigma \mathbf{f} \cdot \mathbf{n}\ dp, \quad \iint_\sigma (U, V, W) \cdot d\mathbf{p},$$

$$\iint_\sigma (U, V, W) \cdot \mathbf{n}\ dp, \quad \iint_\sigma (U\mathbf{i} + V\mathbf{j} + W\mathbf{k}) \cdot d\mathbf{p}, \quad \iint_\sigma (U\mathbf{i} + V\mathbf{j} + W\mathbf{k}) \cdot \mathbf{n}\ dp$$

have the same meaning.

Another denotation of the surface integral of a vector function also sometimes appears in literature. It is based on the idea of expressing vector $d\mathbf{p}$ in the form

$$d\mathbf{p} = \big(dy\,dz,\; dx\,dz,\; dx\,dy\big) = \mathbf{i}\; dy\,dz + \mathbf{j}\; dx\,dz + \mathbf{k}\; dx\,dy.$$

Substituting this to the integral $\iint_\sigma (U,V,W)\cdot d\mathbf{p}$ and computing the scalar product, we obtain

$$\iint_\sigma \mathbf{f}\cdot d\mathbf{p} = \iint_\sigma (U,V,W)\cdot d\mathbf{p} = \iint_\sigma U\,dy\,dz + V\,dx\,dz + W\,dx\,dy.$$

Nevertheless, we think that this last notation of the surface integral of a vector function can lead to confusion, so we will not use it.

V.4.5. Remark. The surface integral of vector function $\mathbf{f}$ is defined by means of the surface integral of the scalar function $\mathbf{f}\cdot\mathbf{n}$ and so the main properties of the surface integral of a vector function are the same as the properties of the surface integral of a scalar function. Thus, we can rewrite items a), b) and c) of paragraph V.2.6 with the function $\mathbf{f}\cdot\mathbf{n}$ instead of f and we obtain valid statements for the surface integral of a vector function. (Do it for yourself!)

The main difference between the surface integral of a scalar function and the surface integral of a vector function is that the surface integral of a vector function depends on the orientation of the surface. This is the content of the following theorem:

V.4.6. Theorem. *If a vector function* $\mathbf{f}$ *is integrable on surface* σ *then it is also integrable on surface* $-\sigma$ *and*

$$\iint_{-\sigma} \mathbf{f}\cdot d\mathbf{p} = -\iint_\sigma \mathbf{f}\cdot d\mathbf{p}.$$

This theorem is an immediate consequence of the definition of the surface integral of a vector function. The integral $\iint_\sigma \mathbf{f}\cdot d\mathbf{p}$ equals $\iint_\sigma \mathbf{f}\cdot\mathbf{n}\,dp$ where $\mathbf{n}$ is the normal vector to σ. $\mathbf{n}$ defines the orientation of surface σ. If we change the orientation then vector $\mathbf{n}$ changes its sign and hence also the surface integral $\iint_\sigma \mathbf{f}\cdot\mathbf{n}\,dp$ changes its sign.

V.4.7. Evaluation of the surface integral of a vector function. The surface integral of a vector function $\mathbf{f}$ on a simple smooth surface σ can be evaluated by means of a parametrization of σ. Let P be such a parametrization, defined in set $B\subset\mathbf{E}_2$. Let σ be oriented in accordance with parametrization P. Then the normal vector $\mathbf{n}$ to σ can be expressed in all "interior points" of σ as $\mathbf{n} = P_u\times P_v/|P_u\times P_v|$. (See paragraph V.1.2, formula (V.1).) Using the formula $\iint_\sigma \mathbf{f}\cdot d\mathbf{p} = \iint_\sigma \mathbf{f}\cdot\mathbf{n}\,dp$ and formula (V.5), we obtain

$$\iint_\sigma \mathbf{f}\cdot d\mathbf{p} = \iint_\sigma \mathbf{f}\cdot\mathbf{n}\,dp =$$
$$= \iint_B \mathbf{f}(P(u,v))\cdot\frac{P_u(u,v)\times P_v(u,v)}{|P_u(u,v)\times P_v(u,v)|}\,|P_u(u,v)\times P_v(u,v)|\,du\,dv,$$

$$\iint_\sigma \mathbf{f}\cdot d\mathbf{p} \;=\; \iint_B \mathbf{f}(P(u,v))\cdot \big(P_u(u,v)\times P_v(u,v)\big)\; du\, dv. \tag{V.6}$$

If surface σ is not oriented in accordance with parametrization P (i.e. P generates the opposite orientation of σ) then formula (V.6) holds with the sign "$-$" in front of the integral on the right hand side.

The surface integral of a vector function on a simple piecewise–smooth surface σ which is a union of simple smooth surfaces $\sigma_1,\ \sigma_2,\ \dots,\ \sigma_m$ (see paragraph V.1.6 for details) can be computed in such a way that we first evaluate the integral on each smooth part $\sigma_1,\ \sigma_2,\ \dots,\ \sigma_m$ of surface σ (e.g. by means of the parametrization of these parts) and then we use the formula

$$\iint_\sigma \mathbf{f}\cdot d\mathbf{p} \;=\; \sum_{i=1}^{m} \iint_{\sigma_i} \mathbf{f}\cdot d\mathbf{p}.$$

Some simple piecewise–smooth surfaces, e.g. spheres, ellipsoids and parts of cones, can be described by means of a mapping whose properties differ from the required properties of parametrizations (see paragraph V.1.1) only on a set of two–dimensional measure zero. Examples of such mappings are given in paragraphs V.2.10 and V.2.11. These mappings (let us recall that they are also often, not quite correctly, called the parametrizations) can be used in formula (V.6) in the same way as parametrizations.

The other possible way of evaluating the surface integral of a vector function is to apply the Gauss–Ostrogradsky or Stokes theorem. These theorems will be explained in Section V.6.

V.4.8. Example. Find the flux of the vector field $\mathbf{f}(x,y,z) = yz\,\mathbf{j} + z^2\,\mathbf{k}$ through the surface σ cut from the semicircular cylinder $y^2+z^2=4,\ z\geq 0$ by the planes $x=-1$ and $x=1$. Surface σ is oriented by its outward normal vector.

We can parametrize surface σ by the mapping

$$P(u,v):\quad x=u,\quad y=2\cos v,\quad z=2\sin v\,;\quad [u,v]\in B=\langle -1,1\rangle\times\langle 0,\pi\rangle.$$

We can find that $P_u = (1,\,0,\,0)$, $P_v = (0,\,-2\sin v,\,2\cos v)$ and $P_u\times P_v = (0,\,-2\cos v,\,-2\sin v)$. The unit vector perpendicular to σ for example at the point $[\,0,\,0,\,1\,]$ (which corresponds to $u=0$ and $v=\pi/2$), expressed by means of parametrization P is

$$\frac{P_u\times P_v}{|P_u\times P_v|}\;\Big|_{[u,v]=[0,\,\pi/2]} \;=\; (0,\,0,\,-1).$$

Since surface σ is oriented outward, the above vector is equal to $-\mathbf{n}$ (where $\mathbf{n}$ is the normal vector to σ at the point $[\,0,\,0,\,1\,]$). Hence, parametrization P generates the opposite orientation of surface σ. This means that if we use formula (V.6), we must write the "$-$" sign in front of the integral of the right hand side:

$$\iint_\sigma (yz\,\mathbf{j}+z^2\,\mathbf{k})\cdot d\mathbf{p} \;=\; -\iint_B \big[\,4\sin v\,\cos v\,\mathbf{j} + 4\sin^2 v\,\mathbf{k}\,\big]\cdot(P_u\times P_v)\;du\,dv \;=$$

$$= -\int_{-1}^{1}\int_{0}^{\pi} \big(0,\, 4\sin v \cos v,\, 4\sin^2 v\big)\cdot\big(0,\, -2\cos v,\, -2\sin v\big)\, dv\, du =$$
$$= -\int_{-1}^{1}\int_{0}^{\pi} \big[-8\sin v \cos^2 v - 8\sin^3 v\big]\, dv\, du = 32.$$

V.5. Operators div and curl.

V.5.1. Divergence of a vector field. Let $\mathbf{f} = (U, V.W)$ be a differentiable vector field in domain $D \subset \mathbf{E}_3$. The *divergence* of $\mathbf{f}$ is a scalar field in D which is denoted by div $\mathbf{f}$ and it is defined by the equation

$$\operatorname{div} \mathbf{f} = \frac{\partial U}{\partial x} + \frac{\partial V}{\partial y} + \frac{\partial W}{\partial z}.$$

V.5.2. Curling of a vector field. Let $\mathbf{f} = (U, V.W)$ be a differentiable vector field in domain $D \subset \mathbf{E}_3$. The *curling* of $\mathbf{f}$ is a vector field in D which is denoted by curl $\mathbf{f}$ and it is defined by the equation

$$\operatorname{curl} \mathbf{f} = \begin{vmatrix} \mathbf{i}, & \mathbf{j}, & \mathbf{k} \\ \dfrac{\partial}{\partial x}, & \dfrac{\partial}{\partial y}, & \dfrac{\partial}{\partial z} \\ U, & V, & W \end{vmatrix} = \Big(\frac{\partial W}{\partial y} - \frac{\partial V}{\partial z},\ \frac{\partial U}{\partial z} - \frac{\partial W}{\partial x},\ \frac{\partial V}{\partial x} - \frac{\partial U}{\partial y}\Big).$$

Instead of the curling of a vector field, denoted by curl $\mathbf{f}$, we often speak about the *rotation* of a vector field, and we denote it by rot $\mathbf{f}$.

V.5.3. The operator nabla. We denote by ∇ and refer to as the operator nabla the vector whose components are operators of partial differentiation with respect to x, y and z. Thus

$$\nabla = \Big(\frac{\partial}{\partial x},\ \frac{\partial}{\partial y},\ \frac{\partial}{\partial z}\Big).$$

(We use the word "operator" because it prescribes performance of some operations – in our case performance of partial derivatives with respect to x, y and z.)

The operator nabla is often used in the denotation of various scalar or vector fields. You already know that the gradient of a scalar field ϕ is a vector field whose components are the partial derivatives of ϕ with respect to x, y and z. This can be expressed by means of the operator ∇ in this way:

$$\operatorname{grad} \phi = \nabla\phi = \Big(\frac{\partial \phi}{\partial x},\ \frac{\partial \phi}{\partial y},\ \frac{\partial \phi}{\partial z}\Big).$$

On the other hand, divergence of a vector field $\mathbf{f} = (U, V, W)$, which is a scalar field, can be written down by means of the scalar product of ∇ and $\mathbf{f}$:

$$\operatorname{div} \mathbf{f} = \nabla\cdot\mathbf{f} = \nabla\cdot(U, V, W) = \frac{\partial U}{\partial x} + \frac{\partial V}{\partial y} + \frac{\partial W}{\partial z}.$$

Finally, curling of a vector field $\mathbf{f} = (U, V, W)$ can be written down as the vector product of ∇ with $\mathbf{f}$:

$$\operatorname{curl} \mathbf{f} \;=\; \nabla \times \mathbf{f} \;=\; \begin{vmatrix} \mathbf{i}, & \mathbf{j}, & \mathbf{k} \\ \dfrac{\partial}{\partial x}, & \dfrac{\partial}{\partial y}, & \dfrac{\partial}{\partial z} \\ U, & V, & W \end{vmatrix} = \Big(\frac{\partial W}{\partial y} - \frac{\partial V}{\partial z},\; \frac{\partial U}{\partial z} - \frac{\partial W}{\partial x},\; \frac{\partial V}{\partial x} - \frac{\partial U}{\partial y}\Big).$$

V.5.4. Remark. Operators grad, div and curl play an important role in the theory of all possible types of fields (flow fields of various fluids, gravity field, electrostatic field, electromagnetic field, etc.). Their mathematical properties and relations are therefore very interesting not only from the point of view of applied mathematics itself, but also from the point of view of many other disciplines. More detailed study of these properties would go beyond the scope of this text. Nevertheless, let us mention two formulas whose validity follows immediately from the definitions of grad, div and curl and it can be easily verified:

If ϕ is a twice–differentiable scalar field in a domain $D \subset \mathbf{E}_3$ and $\mathbf{f}$ is a twice–differentiable vector field in D then

$$\operatorname{curl} \operatorname{grad} \phi \;=\; \vec{0} \;=\; (0,\, 0,\, 0), \tag{V.7}$$

$$\operatorname{div} \operatorname{curl} \mathbf{f} \;=\; 0. \tag{V.8}$$

You already know the geometrical meaning of the gradient of a scalar function ϕ from Chapter I – $\operatorname{grad} \phi$ is the vector which shows the direction of the greatest growth of function ϕ. The physical sense of the other two operators, div and curl, will be explained in paragraphs V.6.5 and V.6.8.

V.6. The Gauss–Ostrogradsky theorem and the Stokes theorem.

We already know the "two–dimensional" Jordan theorem – see paragraph IV.5.1. The next paragraph contains a "three–dimensional" version of the same theorem. It says again something that is very clear at first sight. However, you would be surprised that it is quite complicated to prove. (We do not show the proof in this text.)

Note that if no other details are given then "surface" refers to a simple piecewise–smooth surface and "closed surface" means a closed simple piecewise–smooth surface. (See paragraph V.1.8.)

V.6.1. Jordan's theorem. *Let σ be a closed surface in $\mathbf{E}_3$. Then there exist two disjoint domains G_1 and G_2 in $\mathbf{E}_3$ such that σ is their common boundary and*

a) $\mathbf{E}_3 = G_1 \cup \sigma \cup G_2$,

b) one of the domains G_1, G_2 is bounded and the second is unbounded.

V.6.2. Interior and exterior of a closed surface in $\mathbf{E}_3$. Let σ be a closed surface in $\mathbf{E}_3$ and G_1, G_2 be the domains whose existence is given by Jordan's theorem. That domain of G_1, G_2 which is bounded is called the <u>*interior*</u> of surface σ and it is denoted by $Int\,\sigma$. The second domain, which is unbounded, is called the <u>*exterior*</u> of σ and it is denoted by $Ext\,\sigma$.

We say that the closed surface σ is oriented to its exterior (respectively to its interior) if its normal vector (at all points of σ where it exists) is oriented to the exterior of σ (respectively to the interior of σ).

V.6.3. The Gauss–Ostrogradsky theorem. *Suppose that*

a) vector function $\mathbf{f}$ *has continuous partial derivatives in domain* $D \subset \mathbf{E}_3$*,*

b) σ *is a closed surface in* D*, oriented to its exterior and such that* $Int\,\sigma \subset D$*.*

Then
$$\iint_\sigma \mathbf{f}\cdot d\mathbf{p} \;=\; \iiint_{Int\ \sigma} \operatorname{div}\,\mathbf{f}\;\; dx\,dy\,dz. \tag{V.9}$$

V.6.4. Example. Calculate the flux of the field $\mathbf{f}(x,y,z) = xy\,\mathbf{i} + yz\,\mathbf{j} + xz\,\mathbf{k}$ outward through the surface σ of the cube $\langle 0,1\rangle \times \langle 0,1\rangle \times \langle 0,1\rangle$.

All the components of $\mathbf{f}$ are continuously differentiable in the whole space $\mathbf{E}_3$ and the considered surface is a closed surface in $\mathbf{E}_3$, oriented outward. Thus, the Gauss–Ostrogradsky theorem yields

$$\iint_\sigma \mathbf{f}\cdot d\mathbf{p} \;=\; \iiint_{Int\ \sigma} \operatorname{div}\,\mathbf{f}\;\, dx\,dy\,dz \;=$$
$$= \int_0^1\int_0^1\int_0^1 \Big(\frac{\partial(xy)}{\partial x} + \frac{\partial(yz)}{\partial y} + \frac{\partial(xz)}{\partial z}\Big)\, dx\,dy\,dz = \int_0^1\int_0^1\int_0^1 (y+z+x)\,dx\,dy\,dz = \tfrac{3}{2}.$$

V.6.5. Physical sense of divergence. Suppose that the vector function $\mathbf{f}$ has continuous partial derivatives in domain $D \subset \mathbf{E}_3$ and $A \in D$. Denote by σ_r the sphere with center A and radius r, oriented to its exterior. Then

$$\operatorname{div}\,\mathbf{f}(A) \;=\; \lim_{r\to 0+} \frac{\operatorname{div}\,\mathbf{f}(A)}{\frac{4}{3}\,\pi r^3} \iiint_{Int\ \sigma_r} dx\,dy\,dz \;=$$
$$=\; \lim_{r\to 0+} \frac{1}{\frac{4}{3}\,\pi r^3} \iiint_{Int\ \sigma_r} \operatorname{div}\,\mathbf{f}(x,y,z)\;\, dx\,dy\,dz \;=\; \lim_{r\to 0+} \frac{1}{\frac{4}{3}\,\pi r^3} \iint_{\sigma_r} \mathbf{f}(x,y,z)\cdot d\mathbf{p}.$$

If the vector field $\mathbf{f}$ has a source at point A then $\iint_{\sigma_r} \mathbf{f}\cdot d\mathbf{p}$ is positive for $r > 0$ sufficiently small and the limit of this integral divided by the volume of $Int\ \sigma_r$ for $r \to 0+$ gives the intensity of the source. Thus, $\operatorname{div}\,\mathbf{f}(A)$ expresses the intensity of the source of $\mathbf{f}$ at point A.

For example, if $\mathbf{v}$ is the velocity of a moving incompressible fluid then $\operatorname{div}\,\mathbf{v} = 0$ in all points of the flow field. This follows from the fact that the conservation of mass, together with the incompressibility of the fluid, guarantees that the fluid cannot arise or disappear at any point A and so the velocity field has no sources (positive or negative). (The equation $\operatorname{div}\,\mathbf{v} = 0$ is the very well known *equation of continuity* for incompressible fluids – you will hear more about it later, in mechanics of fluids.)

V.6.6. Stokes' theorem. *Suppose that*

a) *vector function* $\mathbf{f}$ *has continuous partial derivatives in domain* $D \subset \mathbf{E}_3$,

b) σ *is a surface in* D *whose boundary* C *is oriented in accordance with* σ.

Then
$$\oint_C \mathbf{f} \cdot d\mathbf{s} \;=\; \iint_\sigma \operatorname{curl} \mathbf{f} \cdot d\mathbf{p}. \tag{V.10}$$

V.6.7. Example. Evaluate the line integral $\int_C \mathbf{f} \cdot d\mathbf{s}$ where $\mathbf{f}(x,y,z) = xz\,\mathbf{i} + xy\,\mathbf{j} + 3xz\,\mathbf{k}$ and C is the boundary of surface σ which is the portion of the plane $2x + y + z = 2$ in the first octant. C is oriented counter–clockwise as viewed from above.

A vector perpendicular to σ is given by the coefficients from the equation of σ: $(2,1,1)$. If you sketch a figure, you can observe that surface σ is oriented in accordance with its boundary C if its normal vector differs from the vector $(2,1,1)$ only by the length: $\mathbf{n} = (2,1,1)/\sqrt{6}$.

The components of vector field $\mathbf{f}$ are continuously differentiable in $\mathbf{E}_3$ and $\operatorname{curl} \mathbf{f} = (0,\, x - 3z,\, y)$. Thus, the Stokes theorem yields

$$\int_C \mathbf{f} \cdot d\mathbf{s} \;=\; \iint_\sigma \operatorname{curl} \mathbf{f} \cdot d\mathbf{p} \;=\; \iint_\sigma \big(0,\, x - 3z,\, y\big) \cdot d\mathbf{p}.$$

Surface σ can be parametrized by the mapping

$$P(u,v) : \quad x = u, \quad y = v, \quad z = 2 - 2u - v\,; \quad [u,v] \in B$$

where $B = \{[u,v] \in \mathbf{E}_2;\ 0 \le u \le 1,\ 0 \le v \le 2 - 2u\}$. We can easily find that $P_u = (1,\,0,\,-2)$, $P_v = (0,\,1,\,-1)$ and $P_u \times P_v = (2,1,1)$. Since the orientation of the last vector is the same as the orientation of the normal vector $\mathbf{n}$, surface σ is oriented in accordance with parametrization P. Using formula (V.6), we obtain

$$\iint_\sigma \big(0,\, x - 3z,\, y\big) \cdot d\mathbf{p} \;=$$
$$= \iint_B \big(0,\, u - 6 + 6u + 3v,\, v\big) \cdot (2,1,1)\; du\, dv = \int_0^1 \int_0^{2-2u} [\, 7u + 4v - 6\,]\; dv\, du \;=\; -1.$$

V.6.8. Physical sense of curling. Suppose that the vector function $\mathbf{f}$ has continuous partial derivatives in the domain $D \subset \mathbf{E}_3$, $A \in D$ and $\mathbf{a}$ is a vector whose length is one. Denote by σ_r the disk with center A, radius r and normal vector $\mathbf{a}$. Denote further by C_r the circle which is the boundary of the disk σ_r and is oriented in accordance with σ_r. Then we have

$$\operatorname{curl} \mathbf{f}(A) \cdot \mathbf{a} \;=\; \lim_{r \to 0+} \frac{\operatorname{curl} \mathbf{f}(A) \cdot \mathbf{a}}{\pi r^2} \iint_{\sigma_r} dp \;=\; \lim_{r \to 0+} \frac{1}{\pi r^2} \iint_{\sigma_r} \operatorname{curl} \mathbf{f}(x,y,z) \cdot \mathbf{a}\; dp \;=$$

$$= \lim_{r\to 0+} \frac{1}{\pi r^2} \iint_{\sigma_r} \text{curl}\, \mathbf{f}(x,y,z) \cdot d\mathbf{p} = \lim_{r\to 0+} \frac{1}{\pi r^2} \oint_{C_r} \mathbf{f}(x,y,z) \cdot d\mathbf{s}.$$

Thus, curl $\mathbf{f}(A)$ is the vector whose scalar product with any unit vector $\mathbf{a}$ expresses the intensity of circulation of $\mathbf{f}$ around circles perpendicular to $\mathbf{a}$, oriented in accordance with $\mathbf{a}$.

V.7. Exercises.

1. $\sigma = \{[x,y,z] \in \mathbf{E}_3;\ x^2+y^2+z=4,\ x \ge 0,\ y \ge 0,\ z \ge 0\}$, σ is oriented so that its normal vector $\mathbf{n}$ at every point of σ satisfies $\mathbf{n}\cdot\mathbf{i} \ge 0$.

a) Verify that the mapping $P(u,v) = [2\cos u,\ 2\sin u,\ v]$ for $[u,v] \in B = \langle -\pi/2, \pi/2\rangle \times \langle 1,4\rangle$ is a parametrization of σ (i.e. that is has all the properties named in paragraph V.1.1). Decide whether σ is oriented in accordance with this parametrization.

b) Show that the mapping $Q(u,v) = [\sqrt{4-u^2},\ u,\ v]$ for $[u,v] \in B = \langle -2,2\rangle \times \langle 1,4\rangle$ is not a parametrization of σ.

2. σ is the half–sphere $\{[x,y,z] \in \mathbf{E}_3;\ x^2+y^2+z^2=a^2,\ z \ge 0\}$ $(a>0)$, oriented by the normal vector $\mathbf{n} = (n_1,n_2,n_3)$ such that $n_3 \ge 0$. Set B is $B = \{[u,v] \in \mathbf{E}_2;\ u^2+v^2 \le a^2\}$.

a) Show that the mapping

$$P(u,v) = \left[\frac{2a^2u}{a^2+u^2+v^2},\ \frac{2a^2v}{a^2+u^2+v^2},\ \frac{2a^3}{a^2+u^2+v^2} - a\right]; \quad [u,v] \in B$$

is a parametrization of surface σ (i.e. it has all the properties named in paragraph V.1.1). Decide whether σ is oriented in accordance with this parametrization.

b) Show that the mapping $P(u,v) = [u,\ v,\ \sqrt{a^2-u^2-v^2}]$, $[u,v] \in B$, is not a parametrization of surface σ.

3. σ is a simple smooth surface, oriented by the normal vector $\mathbf{n}$. Find its parametrization, show that it has all the properties named in paragraph V.1.1. and determine whether σ is oriented in accordance with the chosen parametrization.

a) σ is the triangle with the vertices $A = [1,-1,2]$, $B = [2,1,3]$, $C = [-1,2,4]$, $\mathbf{n}\cdot\mathbf{j} < 0$

b) $\sigma = \{[x,y,z] \in \mathbf{E}_3;\ x^2+y^2=4,\ x \ge 0,\ 0 \le z \le 4\}$, $\mathbf{n} = (1,0,0)$ at the point $P = [2,0,2]$

c) $\sigma = \{[x,y,z] \in \mathbf{E}_3;\ x^2+y^2=z,\ y \ge 0,\ z \le 1\}$, $P = [0,0,0]$, $\mathbf{n} = (0,0,-1)$

d) σ is the parallelogram with the vertices $A = [1,1,1]$, $B = [1,4,4]$, $C = [0,5,6]$, $D = [0,2,3]$, $\mathbf{n}\times\mathbf{k} > 0$ at every point of σ

e) σ is the disk in the plane $x = 2$ with its center at the point $[2,-1,3]$ and radius $r = 4$, $\mathbf{n} = (-1,0,0)$ at every point of σ

f) $\sigma = \{[x,y,z] \in \mathbf{E}_3;\ x^2+y^2+z^2=4;\ z \ge \sqrt{2}\}$, $P = [0,0,2]$, $\mathbf{n} = (0,0,1)$

g) $\sigma = \{[x,y,z] \in \mathbf{E}_3;\ xy-z=0,\ x^2+y^2 \le a^2\}$ $(a>0)$, $\mathbf{n}\cdot\mathbf{k} > 0$

4. Verify that the set $\sigma = \{X \in \mathbf{E}_3;\ X = P(u,v),\ [u,v] \in B\}$ is a simple smooth surface in $\mathbf{E}_3$ and P is its parametrization.

a) $P(u,v) = [u,\ 4u^2 + 9v^2,\ v],\ \ B = \{[u.v] \in \mathbf{E}_2;\ u^9/9 + v^2/4 \le 1\}$

b) $P(u,v) = [u,\ v,\ 4 - u - v],\ \ B = \{[u,v] \in \mathbf{E}_2;\ u \ge 0,\ v \ge 0,\ u + v \le 4\}$

c) $P(u,v) = [3\cos u \cos v,\ 3 \sin u \cos v,\ 3 \sin v],\ \ B = \langle 0, \pi\rangle \times \langle 0, \pi/4\rangle$

5. Verify that σ is a simple piecewise–smooth surface. Find parametrizations of the simple smooth parts of σ.

a) $\sigma = \{[x,y,z] \in \mathbf{E}_3;\ z = 4 - \sqrt{x^2+y^2},\ z \ge 0\}$

b) $\sigma = \{[x,y,z] \in \mathbf{E}_3;\ z = 4 - \sqrt{x^2+y^2},\ 0 \le z \le 2\}$

c) $\sigma_1 \cup \sigma_2$ where $\sigma_1 = \{[x,y,z] \in \mathbf{E}_3;\ x^2+y^2 \le 16,\ z = 0\}$, $\sigma_2 = \{[x,y,z] \in \mathbf{E}_3;\ z = 4 - \sqrt{x^2+y^2},\ z \ge 0\}$

d) σ is the boundary of $D = \{[x,y,z] \in \mathbf{E}_3;\ x^2+y^2+z^2 \le 4,\ 4 - x^2 - y^2 \le 4z\}$

6. Decide about the existence of the integral $\iint_\sigma f\ dp$.

a) $f(x,y,z) = (xy \ln|x|)/z,\ \ \sigma = \{[x,y,z] \in \mathbf{E}_3;\ (x-2)^2 + y^2 + z^2 = 1,\ z \ge 0\}$

b) $f(x,y,z) = (xy \ln|x|)/z,\ \ \sigma = \{[x,y,z] \in \mathbf{E}_3;\ z = 1 + x^2 + y^2,\ z \le 2\}$

c) $f(x,y,z) = (x^2+y^2+z^2-1)^{-1}$, σ is the sphere with its center at the point $S = [0,0,3]$ and radius $r = a$

7. Evaluate the area of the surfaces from examples 4c, 5b, 5c, 3g

8. Evaluate the surface integrals

a) $\displaystyle\iint_\sigma xyz\ dp,\ \ \sigma = \{[x,y,z] \in \mathbf{E}_3;\ y^2 + 9z^2 = 9,\ 1 \le x \le 3,\ y \ge 0,\ z \ge 0\}$

b) $\displaystyle\iint_\sigma xz\ dp$, σ is the triangle with the vertices $A = [1,0,0]$, $[0,1,0]$ and $C = [0,0,1]$

c) $\displaystyle\iint_\sigma x(^2 + y^2)\ dp,\ \ \sigma = \{[x,y,z] \in \mathbf{E}_3;\ x^2+y^2+z^2 = a^2\}$, $(a > 0)$

d) $\displaystyle\iint_\sigma (xy + yz + xz)\ dp,\ \ \sigma = \{[x,y,z] \in \mathbf{E}_3;\ y = \sqrt{x^2+z^2},\ x^2 + z^2 \le 2x\}$

e) $\displaystyle\iint_\sigma (x + y + z)\ dp,\ \ \sigma = \{[x,y,z] \in \mathbf{E}_3;\ x^2+y^2+z^2 = a^2,\ x \le 0\}$ $(a > 0)$

f) $\displaystyle\iint_\sigma \frac{dp}{x^2+y^2+z^2},\ \ \sigma = \{[x,y,z] \in \mathbf{E}_3;\ x^2+y^2 = 9,\ 0 \le z \le 3\}$

g) $\displaystyle\iint_\sigma (x^2 + y^2)\ dp$, σ is the surface from example 5d

9. Find the center of mass of surface σ if mass is distributed on σ with the density ρ.

a) $\rho(x,y,z) = x,\ \ \sigma = \{[x,y,z] \in \mathbf{E}_3;\ x = \sqrt{y^2+z^2},\ y \ge 0,\ 0 \le x \le 2\}$

b) $\rho(x,y,z) = xyz,\ \ \sigma = \{[x,y,z] \in \mathbf{E}_3;\ x^2+z^2 = 4,\ x \ge 0,\ z \ge 0,\ 0 \le y \le 3\}$

10. C is the circular cone with the radius of its basis $r = a > 0$ and the height $h > 0$. Surface σ is the boundary of C. The planar density of the mass distribution on σ is $\rho = const$. Evaluate the moment of inertia of σ with respect to the axis of cone C.

11. Evaluate the flux of vector field $\mathbf{f}$ through surface σ.

a) $\mathbf{f}(x,y,z) = y\,\mathbf{i} - x\,\mathbf{j} + z\mathbf{k}$, $\sigma = \{\,[x,y,z] \in \mathbf{E}_3;\ z = 4 - x^2 - y^2/9,\ y \geq 0,\ z \geq 0\,\}$, σ is oriented so that the angle between its normal vector (at any point) and the vector $\mathbf{k} = (0,0,1)$ is acute (i.e. less than $\pi/2$).

b) $\mathbf{f}(x,y,z) = (0,0,y)$, σ is the triangle with the vertices $A = [0,0,0]$, $B = [5,0,1]$, $C = [1,4,1]$, oriented so that the angle between its normal vector (at any point) and the vector $\mathbf{k} = (0,0,1)$ is acute.

c) $\mathbf{f}(x,y,z) = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$, σ is the cylindrical surface $x^2 + y^2 = 9$, $0 \leq z \leq 4$, oriented to the interior of the cylinder.

d) $\mathbf{f}(x,y,z) = (x^3, y^3, z^3)$, σ is the part (corresponding to $z \geq 0$) of the sphere with its center $S = [0,0,0]$ and radius $r = 2$, oriented so that its normal vector at the point $[0,0,2]$ is $\mathbf{n} = (0,0,1)$.

e) $\mathbf{f}(x,y,z) = x\,\mathbf{i} - x\,\mathbf{j} + y\,\mathbf{k}$, σ is the parallelogram $A = [0,0,0]$, $B = [0,3,3]$, $C = [-1,4,5]$, $D = [-1,1,2]$, oriented by the normal vector $\mathbf{n} = (1,-1,1)$.

f) $\mathbf{f}(x,y,z) = (x^2-y^2,\ y^2-z^2,\ z^2-x^2)$, $\sigma = \sigma_1 \cup \sigma_2$, $\sigma_1 = \big\{[x,y,0] \in \mathbf{E}_3;\ x^2+y^2 \leq 1,\ x \geq 0,\ y \geq 0\big\}$, $\sigma_2 = \big\{[x,0,z] \in \mathbf{E}_3;\ x^2+z^2 \leq 1,\ z \geq 0,\ z \geq 0\big\}$, the normal vector to σ_2 is $\mathbf{n} = (0,-1,0)$.

g) $\mathbf{f}(x,y,z) = x\,\mathbf{i} + y\,\mathbf{j} - 2z\,\mathbf{k}$, $\sigma = \big\{[x,y,z] \in \mathbf{E}_3;\ y = 9 - \sqrt{x^2+z^2},\ y \geq 3\big\}$, $\mathbf{n} \cdot \mathbf{j} \leq 0$.

h) $\mathbf{f}(x,y,z) = (x^2, y^2, z^2)$, $\sigma = \big\{[x,y,z] \in \mathbf{E}_3;\ y^2/16 + z^2/4 = 1,\ z \geq 0,\ 0 \leq x \leq 3\big\}$, the normal vector $\mathbf{n}$ at the point $P = [1,0,2]$ is $\mathbf{n} = (0,0,-1)$.

i) $\mathbf{f}(x,y,z) = x\,\mathbf{i} + y\,\mathbf{j} - z\,\mathbf{k}$, $\sigma = \big\{[x,y,z] \in \mathbf{E}_3;\ x^2+y^2+z^2 = 4,\ x \geq 0\big\}$, the normal vector $\mathbf{n}$ at the point $P = [2,0,0]$ is $\mathbf{n} = \mathbf{i} = (1,0,0)$.

j) $\mathbf{f}(x,y,z) = (z,x,y)$, $\sigma = \big\{[x,y,z] \in \mathbf{E}_3;\ x+z = 2,\ x^2+y^2 \leq 4\big\}$, the normal vector is $\mathbf{n} = (1,0,1)/\sqrt{2}$.

12. Evaluate the flux of vector field $\mathbf{f}$ through the closed surface σ. If it is possible, apply the Gauss-Ostrogradsky theorem.

a) $\mathbf{f}(x,y,z) = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$, σ is the sphere with the center at the point $S_0 = [x_0, y_0, z_0]$ and radius $r = a > 0$, oriented outward.

b) $\mathbf{f}(x,y,z) = y\,\mathbf{i} + z\,\mathbf{j} + x\,\mathbf{k}$, σ is the sphere with the center at the point $S_0 = [x_0, y_0, z_0]$ and radius $r = a > 0$, oriented outward.

c) $\mathbf{f}(x,y,z) = (x^2, y^2, z^2)$, σ is the sphere with the center at the point $S_0 = [x_0, y_0, z_0]$ and radius $r = a > 0$, oriented outward.

d) $\mathbf{f}(x,y,z) = (y, 2x, -z)$, σ is the boundary of the set $D = \big\{[x,y,z] \in \mathbf{E}_3;\ x^2 + y^2 \leq a^2,\ -a \leq z \leq a\big\}$ $(a > 0)$, σ is oriented to the interior.

e) $\mathbf{f}(x,y,z) = (x^2, y^2, z^2)$, σ is the boundary of the set $D = \big\{[x,y,z] \in \mathbf{E}_3;\ -2 \leq z \leq 4 - x^2 - y^2,\ x^2 + y^2 \leq 4\big\}$, σ is oriented to the interior.

f) $\mathbf{f}(x,y,z) = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$, σ is the boundary of the set $D = \{[x,y,z] \in \mathbf{E}_3;\ y^2+z^2 < x^2,\ 0 < x < 3\}$, σ is oriented to the exterior.

g) $\mathbf{f}(x,y,z) = (x^3, z, y)$, σ is the boundary of the set $D = \{[x,y,z] \in \mathbf{E}_3;\ x^2+y^2 < z < 4\}$, σ is oriented to the exterior.

h) $\mathbf{f}(x,y,z) = 2xy\,\mathbf{i} - y^2\mathbf{j} + 2z\,\mathbf{k}$, $\sigma = \{[x,y,z] \in \mathbf{E}_3;\ x^2/4 + y^2/4 + z^2/9 = 1\}$, σ is oriented to the interior.

i) $\mathbf{f}(x,y,z) = x^3\mathbf{i} + y^3\mathbf{j} + z^3\mathbf{k}$, σ is the boundary of the set $D = \{[x,y,z] \in \mathbf{E}_3;\ x^2+y^2+z^2 \leq a^2,\ y > 0\}$ $(a > 0)$, σ is oriented to the exterior.

j) $\mathbf{f} = (x-1, y+2, 2)$, σ is a closed surface in $\mathbf{E}_3$, oriented outwards, such that $m_3(Int\ \sigma) = 5$.

13. Evaluate the circulation of vector field $\mathbf{f}$ around the closed curve C. If it is possible, apply the Stokes theorem.

a) $\mathbf{f}(x,y,z) = (y, z, x)$, C is the intersection of the cylindrical surface $x^2+y^2 = 4$ and the plane $x+z = 0$. C is oriented so that its unit tangent vector at the point $[2,0,-2]$ is $\vec{\tau} = (0,1,0)$.

b) $\mathbf{f}(x,y,z) = (xy, yz, zx)$, $C = \{[x,y,z] \in \mathbf{E}_3;\ x+z = 1,\ y^2+z^2 = 1\}$. C is oriented in accordance with the surface $\lambda = \{[x,y,z] \in \mathbf{E}_3;\ x+y = 1,\ y^2+z^2 \leq 1\}$ and the orientation of λ is given by the normal vector $\mathbf{n} = (1,0,1)/\sqrt{2}$.

c) $\mathbf{f}(x,y,z) = (z+1)\,\mathbf{i} + (x-y)\,\mathbf{j} + y\,\mathbf{k}$, C is a circle which is the intersection of the sphere $x^2+y^2+z^2 = 2$ with the plane $x+y+z = 0$. C is oriented clockwise as viewed from the point $[0,0,10]$.

d) $\mathbf{f}(x,y,z) = \big(-y/(x^2+y^2),\ x/(x^2+y^2),\ 2z\big)$, C is a circle $x^2+y^2 = a^2$ $(a > 0)$, $z = h$ $(h > 0)$, oriented counterclockwise if viewed from the point $[0,0,2h]$.

VI. Potential and Solenoidal Vector Field

VI.1. Independence of the line integral of a vector function on the path. Potential vector field.

VI.1.1. Independence of the line integral of a vector function on the path. Let $\mathbf{f}$ be a k–dimensional vector field in domain $D \subset \mathbf{E}_k$ (for $k = 2$ or $k = 3$). Suppose that for any two curves C_1 and C_2 in D, such that $i.p.\,C_1 = i.p.\,C_2$ and $t.p.\,C_1 = t.p.\,C_2$, the line integrals $\int_{C_1} \mathbf{f} \cdot d\mathbf{s}$ and $\int_{C_2} \mathbf{f} \cdot d\mathbf{s}$ exist and are equal. Then we say that the <u>*line integral of* $\mathbf{f}$ *does not depend on the path in* D</u>.

VI.1.2. Theorem. *The line integral of vector function* $\mathbf{f}$ *does not depend on the path in domain* $D \subset \mathbf{E}_k$ *(for* $k = 2$ *or* $k = 3$*) if and only if the circulation of* $\mathbf{f}$ *around every closed curve* C *in* D *is equal to zero.*

P r o o f : a) Suppose at first that the line integral of $\mathbf{f}$ does not depend on the path in D, and C is a closed curve in D. Then C can be decomposed to the union of two curves K_1 and K_2 such that $t.p.\,K_1 = i.p.\,K_2$ and $t.p.\,K_2 = i.p.\,K_1$. Putting $C_1 = K_1$ and $C_2 = -K_2$, we get two curves in D with the same initial and terminal points. The independence of the line integral of $\mathbf{f}$ on the path implies that $\int_{C_1} \mathbf{f} \cdot d\mathbf{s} = \int_{C_2} \mathbf{f} \cdot d\mathbf{s}$. This implies

$$\oint_C \mathbf{f} \cdot d\mathbf{s} \;=\; \int_{K_1} \mathbf{f} \cdot d\mathbf{s} + \int_{K_2} \mathbf{f} \cdot d\mathbf{s} \;=\; \int_{C_1} \mathbf{f} \cdot d\mathbf{s} - \int_{C_2} \mathbf{f} \cdot d\mathbf{s} \;=\; 0.$$

b) Suppose now that the circulation of $\mathbf{f}$ along any closed curve in D is zero. Let C_1 and C_2 be two curves in D such that $i.p.\,C_1 = i.p.\,C_2$ and $t.p.\,C_1 = t.p.\,C_2$. Suppose for simplicity that curves C_1 and C_2 do not have any other common points, i.e. they do not intersect or touch at any other points. Then the union $C = C_1 \cup (-C_2)$ is a closed curve in D and so the circulation of $\mathbf{f}$ around C is zero. This implies:

$$\int_{C_1} \mathbf{f} \cdot d\mathbf{s} \;=\; \int_{C_1 \cup (-C_2)} \mathbf{f} \cdot d\mathbf{s} + \int_{C_2} \mathbf{f} \cdot d\mathbf{s} \;=\; \int_{C_2} \mathbf{f} \cdot d\mathbf{s}$$

A similar approach can be used in the case that the curves C_1 and C_2 have more common points than their initial and terminal points, and it can also be proved in this case that the line integrals of $\mathbf{f}$ on C_1 and C_2 are equal.

VI.1.3. Potential vector field. We say that the vector field $\mathbf{f}$ in domain $D \subset \mathbf{E}_k$ (for $k = 2$ or $k = 3$) is *potential field* in D if there exists a scalar field φ in D such that

$$\mathbf{f} \;=\; \operatorname{grad} \varphi$$

in D. Scalar function φ is called the *potential* of $\mathbf{f}$ in D.

VI.1.4. Remark. Recall that

$$\operatorname{grad} \varphi = \left(\frac{\partial \varphi}{\partial x}, \frac{\partial \varphi}{\partial y}\right) \;\; (\text{if } k = 2), \qquad \operatorname{grad} \varphi = \left(\frac{\partial \varphi}{\partial x}, \frac{\partial \varphi}{\partial y}, \frac{\partial \varphi}{\partial z}\right) \;\; (\text{if } k = 3).$$

It is quite obvious that if $\mathbf{f}$ is a potential field in domain D and D' is a domain such that $D' \subset D$ then $\mathbf{f}$ is also a potential field in domain D'.

You will see later (in paragraph VI.2.4) that e.g. gravitational and electric fields are examples of potential fields.

A scalar function φ which is a potential of a potential vector field $\mathbf{f}$ has some properties which are in some sense analogous to the properties of an antiderivative. For example:

If $\mathbf{f}$ is a potential vector field in domain $D \subset \mathbf{E}_k$ then its potential φ is unique up to an additive constant.

This means that:

a) $\varphi + c$ *is also a potential of* $\mathbf{f}$ *in* D *for every real constant* c.

b) *Any other potential* ζ *of* $\mathbf{f}$ *in* D *differs from* φ *at most in an additive constant. In other words: If* ζ *is another potential of* $\mathbf{f}$ *in* D *then there exists a constant* c *such that* $\zeta = \varphi + c$ *in* D.

Assertion a) is very simple – the equation $\operatorname{grad}\varphi = \mathbf{f}$ in D immediately implies that $\operatorname{grad}(\varphi + c) = \mathbf{f}$ and so $\varphi + c$ is also a potential of $\mathbf{f}$ in D.

Assertion b) is also quite obvious: If φ and ζ are two potentials of $\mathbf{f}$ in D then $\mathbf{f} = \operatorname{grad}\varphi$ and $\mathbf{f} = \operatorname{grad}\zeta$ in D and so $\operatorname{grad}\zeta - \operatorname{grad}\varphi = \operatorname{grad}(\zeta - \varphi) = \vec{0}$. However the only function whose gradient is zero in domain D is a constant function. This implies the existence of a constant c such that $\zeta - \varphi = c$ and so $\zeta = \varphi + c$ in D.

Another analogy between the potential φ of a potential vector field and the antiderivative to a function of one variable is the similarity of formula (IV.1) (see the next theorem) and the Newton–Leibnitz formula (II.6).

VI.1.5. Theorem. *If* $\mathbf{f}$ *is a continuous and potential vector field in domain* D, φ *is a potential of* $\mathbf{f}$ *in* D *and* C *is a curve in* D *then*

$$\int_C \mathbf{f} \cdot d\mathbf{s} \;=\; \varphi(t.p.\,C) - \varphi(i.p.\,C). \tag{VI.1}$$

P r o o f : Since φ is a potential of $\mathbf{f}$ in D, $\mathbf{f}$ is equal to $\operatorname{grad}\varphi$ in D. Suppose that C is a simple smooth curve in D and P is its parametrization defined in the interval $\langle a, b\rangle$ such that curve C is oriented in accordance with P. Let us denote by $x(t)$, $y(t)$ and $z(t)$ the coordinate functions of parametrization P. Then

$$\int_C \mathbf{f} \cdot d\mathbf{s} \;=\; \int \operatorname{grad}\varphi \cdot d\mathbf{s} \;=\; \int_a^b \operatorname{grad}\varphi(P(t)) \cdot \dot{P}(t)\; dt \;=\; \int_a^b \Big(\frac{\partial\varphi}{\partial x}\big(x(t), y(t), z(t)\big),$$

$$\frac{\partial\varphi}{\partial y}\big(x(t), y(t), z(t)\big),\; \frac{\partial\varphi}{\partial z}\big(x(t), y(t), z(t)\big)\Big) \cdot \big(\dot{x}(t),\; \dot{y}(t),\; \dot{z}(t)\big)\; dt \;=$$

$$= \int_a^b \Big[\, \frac{\partial\varphi}{\partial x}\big(x(t), y(t), z(t)\big)\; \dot{x}(t) \;+\; \frac{\partial\varphi}{\partial y}\big(x(t), y(t), z(t)\big)\; \dot{y}(t) +$$

$$+ \frac{\partial\varphi}{\partial z}\big(x(t), y(t), z(t)\big)\; \dot{z}(t) \Big]\; dt \;=\; \int_a^b \frac{d}{dt}\; \varphi\big(x(t), y(t), z(t)\big)\; dt \;=$$

$$= \varphi\big(x(b), y(b), z(b)\big) - \varphi\big(x(a), y(a), z(a)\big) \;=\; \varphi(t.p.\,C) - \varphi(i.p.\,C).$$

The same equality can also be proved in the case when C is a simple piecewise–smooth curve.

The next theorem is perhaps the most important theorem in this section.

VI.1.6. Theorem. *Suppose that* $\mathbf{f}$ *is a continuous vector field in domain* $D \subset \mathbf{E}_k$ *(for* $k = 2$ *or* $k = 3$*). Then the next two conditions are equivalent:*

a) $\mathbf{f}$ *is a potential vector field in* D.

b) *The line integral of* $\mathbf{f}$ *does not depend on the path in* D.

P r o o f : The implication a) $\Longrightarrow$ b) is the consequence of formula (VI.1).

Let us now prove the opposite implication, i.e. b) $\Longrightarrow$ a). Suppose that condition b) is fulfilled. Let us denote by U, V and W the components of $\mathbf{f}$. Choose a point $A \in D$. The point A was chosen arbitrarily, but we take it as a fixed point from now. Let $X = [x, y, z]$ be any other point of D. Let us define

$$\varphi(x, y, z) \;=\; \int_C \mathbf{f} \cdot d\mathbf{s} \tag{VI.2}$$

where C is a curve in D such that $i.p.\, C = A$ and $t.p.\, C = X$. (It follows from the independence of the line integral of $\mathbf{f}$ on the path in D that the value of $\varphi(x, y, z)$ does not depend on the concrete choice of the curve C connecting the points A and $X = [x, y, z]$.) We claim that $\operatorname{grad} \varphi = \mathbf{f}$ in D. To prove this, it is sufficient to show that

$$\frac{\partial \varphi}{\partial x}(X) = U(X), \quad \frac{\partial \varphi}{\partial y}(X) = V(X) \quad \text{and} \quad \frac{\partial \varphi}{\partial z}(X) = W(X). \tag{VI.3}$$

Let us prove for example the first of these equalities. Using the definition of the partial derivative of φ with respect to x at point $X = [x, y, z]$, we obtain

$$\frac{\partial \varphi}{\partial x}(X) \;=\; \frac{\partial \varphi}{\partial x}(x, y, z) \;=\; \lim_{h \to 0} \frac{\varphi(x + h, y, z) - \varphi(x, y, z)}{h}.$$

$\varphi(x + h, y, z)$ can be expressed as the line integral of $\mathbf{f}$ on the simple piecewise – smooth curve which is the union of C and the line segment XY leading from point X to the point $Y = [x + h, y, z]$. The unit tangent vector on XY is $\vec{\tau} = (1, 0, 0)$. Thus, we get

$$\frac{\partial \varphi}{\partial x}(X) \;=\; \lim_{h \to 0} \frac{1}{h} \Big[\int_{C \cup XY} \mathbf{f} \cdot d\mathbf{s} - \int_C \mathbf{f} \cdot d\mathbf{s} \Big] \;=\; \lim_{h \to 0} \frac{1}{h} \int_{XY} \mathbf{f} \cdot d\mathbf{s} \;=$$

$$=\; \lim_{h \to 0} \frac{1}{h} \int_{XY} (U, V, W) \cdot (1, 0, 0)\, ds \;=\; \lim_{h \to 0} \frac{1}{h} \int_{XY} U\, ds \;=\; U(X).$$

The equalities in (VI.3) show that $\mathbf{f} = \operatorname{grad} \varphi$ in D and so the vector field $\mathbf{f}$ is potential in D.

VI.1.7. Remark. If $\mathbf{f}$ is a potential vector field in domain $D \subset \mathbf{E}_k$ (for $k = 2$ or $k = 3$) then the line integral of $\mathbf{f}$ is independent on the path in D (by Theorem VI.1.6) and this means that the circulation of $\mathbf{f}$ on every closed curve in D is zero (by Theorem VI.1.2). If $\mathbf{f}$ has a physical meaning of a force then we can say that the work done by force $\mathbf{f}$ over every closed curve C is zero. Due to this fact, potential vector fields are also often called *conservative fields*.

The path independence of the line integral in potential vector field $\mathbf{f}$ has an important physical meaning: It says that the amount of work done by force $\mathbf{f}$ in moving from point $A \in D$ to point $B \in D$ is the same for all paths in D, leading from A to B. This is known to hold for example in a gravitational or an electric field, where the amount of work it takes to move a mass particle or a charge from point A to point B depends only on the position of A and B and not on the path taken between A and B.

Since the value of the integral $\int_C \mathbf{f} \cdot d\mathbf{s}$ of potential vector field $\mathbf{f}$ depends only on the position of the initial point and the terminal point of curve C, this integral is often written as

$$\int_C \mathbf{f} \cdot d\mathbf{s} = \int_A^B \mathbf{f} \cdot d\mathbf{s}$$

where $A = i.p.\, C$ and $B = t.p.\, C$.

We have shown the proofs of the last two theorems because they are very instructive and their ideas can also be used in other situations. Formula (VI.1) provides a very simple way of evaluating the line integral of a potential vector field when the potential φ of $\mathbf{f}$ is known. On the other hand, the idea of the proof of Theorem VI.1.6 shows a method of finding a potential φ of a vector field $\mathbf{f}$ provided that it is known that $\mathbf{f}$ is a potential vector field. This approach will be applied to concrete example (see paragraph VI.2.1). We will also show another method of finding a potential φ of $\mathbf{f}$ – see examples VI.2.1., VI.2.2 and VI.2.4.

We have seen that the potential vector field $\mathbf{f}$ in domain D has interesting and useful properties – especially that its line integral does not depend on the path in D (see Theorem VI.1.6) and moreover, its line integral can be evaluated by means of formula (VI.1) (see Theorem VI.1.5). It is therefore very important to be able to recognize whether a given vector field in domain D is or is not a potential vector field in D. The next paragraphs will deal with this question. We will distinguish between the two–dimensional case (see paragraphs VI.I.8–VI.1.13) and the three–dimensional case (see paragraphs VI.1.14–VI.1.18).

VI.1.8. Theorem. (Potential field in $\mathbf{E}_2$ – the necessary condition.) *Suppose that* $\mathbf{f} = (U, V)$ *is a potential vector field in domain* $D \subset \mathbf{E}_2$ *Suppose that the components* U *and* V *of* $\mathbf{f}$ *are continuously differentiable functions in* D. *Then*

$$\frac{\partial V}{\partial x} - \frac{\partial U}{\partial y} = 0 \quad \text{in } D. \tag{VI.4}$$

P r o o f : If φ is a potential of $\mathbf{f}$ in D then $\mathbf{f} = \big(\partial\varphi/\partial x,\ \partial\varphi/\partial y\big)$. Hence we have

$$U = \frac{\partial \varphi}{\partial x}, \qquad V = \frac{\partial \varphi}{\partial y}$$

in D. This form of U and V, together with the information about the continuity of the partial derivatives of U and V in D (which means the continuity of the second partial derivatives of φ in D), easily implies condition (VI.4).

VI.1.9. Remark. Condition (VI.4) is a necessary but not a sufficient condition. This means that the two–dimensional vector field $\mathbf{f}$ in domain $D \subset \mathbf{E}_2$ whose components are continuously differentiable functions in D can be a potential vector field in D only if condition (VI.4) is fulfilled, but on the other hand the validity of condition (VI.4) itself does not guarantee that vector field $\mathbf{f}$ really is a potential field in D. We can demonstrate this through the next example.

VI.1.10. Example. The vector function $\mathbf{f} = (U, V) = \left(-\dfrac{y}{x^2+y^2},\ \dfrac{x}{x^2+y^2}\right)$ satisfies condition (VI.4) in the domain

$$D=\{[x,y]\in\mathbf{E}_2;\ x^2+y^2>0\}.$$

(Verify it for yourself!)

Let us now evaluate the circulation of $\mathbf{f}$ around the circle $C_r:\ x^2+y^2=r^2$ (where $r>0$), whose orientation is positive (see paragraph IV.5.3). It can easily be checked that the mapping $P:\ x=\phi(t)=r\cos t,\ y=\psi(t)=r\sin t$ (for $t\in\langle 0,2\pi\rangle$) is a parametrization of C_r which also generates the positive orientation of C_r. Thus, we have

$$\oint_{C_r}\mathbf{f}\cdot d\mathbf{s}=\oint_{C_r}-\frac{y}{x^2+y^2}\,dx+\frac{x}{x^2+y^2}\,dy=$$

$$=\int_0^{2\pi}\Big[-\frac{r\sin t}{r^2}(-r\sin t)+\frac{r\cos t}{r^2}\,r\cos t\Big]\,dt=\int_0^{2\pi}dt=2\pi.$$

Thus, although the components of $\mathbf{f}$ satisfy condition (VI.4), $\mathbf{f}$ is not a potential vector field in D because we have shown that the circulation of $\mathbf{f}$ around at least one closed curve in D is different from zero. (See Theorem VI.1.2.)

Our next goal is to formulate sufficient conditions which will guarantee that a given two–dimensional vector field will be potential in domain $D\subset\mathbf{E}_2$. We will therefore need the notion of a so called simply connected domain in $\mathbf{E}_2$:

VI.1.11. A simply connected domain in $\mathbf{E}_2$. Domain $D\subset\mathbf{E}_2$ is said to be *simply connected* if each closed curve C in D can be contracted to a point in D without ever leaving D.

A simply connected domain in $\mathbf{E}_2$ can also be defined as such domain $D\subset\mathbf{E}_2$ that the interior of every closed curve C in D is a subset of D.

Roughly speaking, domains which have bounded "holes" (and look like a Swiss cheese) are not simply connected, while domains which do not have such holes are simply connected. The examples of domains which are simply connected are: the whole plane $\mathbf{E}_2$, the half–plane, $\mathbf{E}_2$ minus a half–line, interiors of closed curves, etc. Examples of domains that are not simply connected include: $\mathbf{E}_2$ minus one point (e.g. domain D from example VI.1.10), $\mathbf{E}_2$ minus a bounded subset and an open disk minus one point.

VI.1.12. Theorem. (Potential field in $\mathbf{E}_2$ – sufficient conditions.) *Let*

a) D be a simply connected domain in $\mathbf{E}_2$ and

b) $\mathbf{f}=(U,V)$ be a vector field in D, whose components U, V are continuously differentiable functions in D and they satisfy the condition

$$\frac{\partial V}{\partial x}-\frac{\partial U}{\partial y}=0\quad\text{in } D. \tag{VI.4}$$

Then $\mathbf{f}$ is a potential vector field in D.

P r o o f : We will prove that $\mathbf{f}$ is a potential vector field in D if we show that the circulation of $\mathbf{f}$ around every closed curve in D equals zero. (See Theorem VI.1.2.)

Thus, let C be a closed curve in D. Since D is simply connected, $Int\ C \subset D$. Applying Green's theorem (see paragraph IV.5.5), we obtain

$$\oint_C \mathbf{f} \cdot d\mathbf{s} \;=\; \pm \iint_{Int\ C} \Big(\frac{\partial V}{\partial x} - \frac{\partial U}{\partial y} \Big)\, dx\, dy \;=\; 0.$$

(The "+" sign is valid if C is positively oriented and "−" holds if C is negatively oriented. However, since the integral on the right hand side equals zero, the signs are not important.)

VI.1.13. Example. We have seen in example VI.1.10 that the vector field

$$\mathbf{f} \;=\; (U, V) \;=\; \Big(-\frac{y}{x^2+y^2},\; \frac{x}{x^2+y^2} \Big)$$

is not potential in the domain $D = \big\{[x,y] \in \mathbf{E}_2;\ x^2+y^2 > 0\big\}$. However, it is potential in any sub–domain $D' \subset D$ which is simply connected! For instance, $\mathbf{f}$ is potential in the upper half–plane $D' = \big\{[x,y] \in \mathbf{E}_2;\ y > 0\big\}$. This follows immediately from Theorem VI.1.12. (The validity of condition (VI.4) was already mentioned in example VI.1.10.)

Thus, we know that vector field $\mathbf{f}$ is potential in domain D', but we do not know the potential φ of $\mathbf{f}$ in D'. We will deal with methods of finding the potential in Section VI.2, and we will also return to this example. (See example VI.2.2.)

The following paragraphs deal with three–dimensional potential vector fields. However, you can find many analogies with the contents of paragraphs VI.1.8–VI.1.13, which deal with two–dimensional potential vector fields.

VI.1.14. Theorem. (Potential field in $\mathbf{E}_3$ − the necessary condition.) *Suppose that* $\mathbf{f}$ *is a potential vector field in domain* $D \subset \mathbf{E}_3$. *Suppose that the components of* $\mathbf{f}$ *are continuously differentiable functions in* D. *Then*

$$\operatorname{curl} \mathbf{f} \;=\; \vec{0} \quad \textit{in } D. \tag{VI.5}$$

P r o o f : Condition (VI.5) is an immediate consequence of formula (V.7) and the fact that $\mathbf{f}$ can be expressed in the form $\mathbf{f} = \operatorname{grad} \varphi$ for some scalar function φ (the potential of $\mathbf{f}$ in D).

VI.1.15. Remark. Suppose that vector field $\mathbf{f}$ has the components U, V and W. Writing $\operatorname{curl} \mathbf{f}$ in components (see paragraph V.5.2), we can observe that condition (VI.5) says the same as the three equations

$$\frac{\partial W}{\partial y} - \frac{\partial V}{\partial z} = 0, \qquad \frac{\partial U}{\partial z} - \frac{\partial W}{\partial x} = 0, \qquad \frac{\partial V}{\partial x} - \frac{\partial U}{\partial y} = 0. \tag{VI.6}$$

It can be observed that Theorem VI.1.8 (dealing with the two–dimensional case) is a consequence of the more general Theorem VI.1.14. Indeed, if we have a two–dimensional potential vector field (U, V) in domain $D \subset \mathbf{E}_2$ then $\mathbf{f} = (U, V, 0)$ is a three dimensional potential vector field in the three–dimensional domain $D \times \mathbf{R}$. Applying Theorem VI.1.14 now to this vector field, writing condition (VI.5) in the form of equations (VI.6) and using the fact that U and V do not depend on z, we can see that the first two equations in (VI.6) are automatically satisfied and the third equation in (VI.6) is identical with (VI.4).

VI.1.16. Remark. Analogously to condition (VI.4), condition (VI.5) is the necessary condition, but it is not a sufficient condition! This may be seen from the three–dimensional version of example VI.1.10: The vector function

$$\mathbf{f} = (U, V, W) = \Big(-\frac{y}{x^2+y^2},\ \frac{x}{x^2+y^2},\ 0 \Big)$$

satisfies condition (VI.5) in the domain $D = \big\{ [x,y,z] \in \mathbf{E}_3;\ x^2+y^2 > 0 \big\}$, but vector field $\mathbf{f}$ is still not potential in D, because the circulation of $\mathbf{f}$ around the circle $C_r:\ x^2+y^2 = r^2,\ z = 0$ (where $r > 0$) is either 2π or -2π, in dependence on the chosen orientation of C_r. (This can be computed similarly as in example VI.1.10.)

However, condition (VI.5) becomes a sufficient condition if it is completed by the assumption about the form of domain D.

VI.1.17. A simply connected domain in $\mathbf{E}_3$. Domain $D \subset \mathbf{E}_3$ is said to be *simply connected* if each closed curve C in D can be contracted to a point in D without ever leaving D.

VI.1.18. Theorem. (Potential field in $\mathbf{E}_3$ – sufficient conditions.) *Let*

a) *D be a simply connected domain in $\mathbf{E}_3$ and*

b) *$\mathbf{f}$ be a vector field in D, whose components are continuously differentiable functions in D and they satisfy the condition*

$$\text{curl}\ \mathbf{f} = \vec{0} \quad \text{in } D. \tag{VI.5}$$

Then $\mathbf{f}$ is a potential vector field in D.

VI.1.19. Remark. A vector field $\mathbf{f}$ can also be potential in a domain D which is not simply connected. However, this cannot be verified by means of Theorem VI.1.12 (in the two–dimensional case) or Theorem VI.2.18 (in the three–dimensional case).

Theorems VI.1.8, VI.1.12, VI.1.14 and VI.1.18 will be applied to concrete examples in the next section.

VI.2. How to find a potential.

In this section, we will deal with two methods of finding a potential φ of a potential vector field. We will explain these methods with concrete examples.

VI.2.1. Example. $\mathbf{f} = \big(y^2 + y \cos x + 6x,\ 2xy + \sin x + 5\big)$. a) Is $\mathbf{f}$ a potential field in $\mathbf{E}_2$? b) If yes, find its potential. c) Compute the integral $\int_C \mathbf{f} \cdot d\mathbf{s}$ on curve C which is the part of the parabola $y = x^2 + 2$ from point $[0,2]$ to point $[2,6]$.

a) The components of vector field $\mathbf{f}$ are continuously differentiable functions and you can easily check that they satisfy equation (VI.4) in $\mathbf{E}_2$. The whole plane $\mathbf{E}_2$ is a simply connected domain. Thus, by Theorem VI.1.12, $\mathbf{f}$ is a potential field in $\mathbf{E}_2$. Let us denote by φ its potential.

1st method of finding a potential. It follows from the definition of the potential (see paragraph VI.1.3) that $\mathbf{f} = \operatorname{grad} \varphi$ in $\mathbf{E}_2$. This means that

$$\frac{\partial \varphi}{\partial x} = y^2 + y \cos x + 6x, \quad \frac{\partial \varphi}{\partial y} = 2xy + \sin x + 5. \tag{VI.7}$$

Integrating the first equality in (VI.7) with respect to x, we obtain

$$\varphi(x, y) = xy^2 + y \sin x + 3x^2 + C_1(y). \tag{VI.8}$$

(C_1 is the constant of integration which arose by the integration with respect to x. So it is a constant with respect to x. However, it can generally depend on y.) Integrating now the second equality in (VI.7) with respect to y, we get

$$\varphi(x, y) = xy^2 + y \sin x + 5y + C_2(x). \tag{VI.9}$$

(C_2 is the constant of integration which appeared after the integration with respect to y. So it cannot depend on y. However, it can depend on x.) Comparing (VI.8) and (VI.9), we get

$$\begin{aligned} xy^2 + y \sin x + 3x^2 + C_1(y) &= xy^2 + y \sin x + 5y + C_2(x), \\ 3x^2 + C_1(y) &= 5y + C_2(x). \end{aligned}$$

This is satisfied if we put e.g. $C_1(y) = 5y$ and $C_2(x) = 3x^2$. Substituting this either to (VI.8) or to (VI.9) and using the uniqueness of the potential up to an additive constant (see paragraph VI.1.4), we get

$$\varphi(x, y) = xy^2 + y \sin x + 3x^2 + 5y + const. \tag{VI.10}$$

2nd method of finding a potential. This method follows the proof of Theorem VI.1.6, and the potential is constructed by means of formula (VI.2). Choose $O = [0, 0]$ as a fixed point and $X = [x_0, y_0]$ as a "variable" point and put $\varphi(x_0, y_0) = \int_C \mathbf{f} \cdot d\mathbf{s}$ where C is an arbitrary curve with the initial point O and the terminal point X. Let us choose curve C so that the computation of the line integral is as simple as possible. For instance: Put $C = OX' \cup X'X$ where OX' is the line segment leading from point O to the point $X' = [x_0, 0]$ and $X'X$ is the line segment leading from point X' to point X. Then we have:

$$\varphi(x_0, y_0) = \left(\int_{OX'} + \int_{X'X} \right) (y^2 + y \cos x + 6x)\, dx + (2xy + \sin x + 5)\, dy.$$

Since $y = 0$ and x varies from 0 to x_0 on the line segment OX', we have $dy = 0$ on OX' and

$$\int_{OX'} (y^2 + y \cos x + 6x)\, dx + (2xy + \sin x + 5)\, dy = \int_0^{x_0} 6x\, dx = 3x_0^2.$$

Further, $x = x_0$ and y varies from 0 to y_0 on $X'X$. Hence $dx = 0$ on $X'X$ and

$$\int_{X'X} (y^2 + y \cos x + 6x)\, dx + (2xy + \sin x + 5)\, dy =$$
$$= \int_0^{y_0} (2x_0 y + \sin x_0 + 5)\, dy = x_0 y_0^2 + y_0 \sin x_0 + 5y_0.$$

Thus, the value of the potential φ at point X is

$$\varphi(x_0, y_0) = \Big(\int_{OX'} + \int_{X'X} \Big) (y^2 + y \cos x + 6x)\, dx + (2xy + \sin x + 5)\, dy =$$
$$= 3x_0^2 + x_0 y_0^2 + y_0 \sin x_0 + 5y_0.$$

Writing $[x, y]$ instead of $[x_0, y_0]$ and taking into account that the potential is determined uniquely up to an additive constant, we obtain formula (VI.10).

c) Using now Theorem VI.1.5, we can evaluate the given line integral:

$$\int_C \mathbf{f} \cdot d\mathbf{s} = \int_{[0,2]}^{[2,6]} \mathbf{f} \cdot d\mathbf{s} = \varphi(2, 6) - \varphi(0, 2) = 104 + 6 \sin 2.$$

VI.2.2. Example. We already know from example VI.1.13 that the vector field

$$\mathbf{f} = (U, V) = \Big(-\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \Big)$$

is potential in the domain $D' = \big\{[x, y] \in \mathbf{E}_2;\ y > 0\big\}$. Using the first or the second method for computation of the potential, we can find that the potential of vector field $\mathbf{f}$ in D' is $\varphi(x, y) = -\arctan(x/y) + const.$

VI.2.3. Example. $\mathbf{f} = \big(y^2 + x^2,\ x - y\big)$. a) Is $\mathbf{f}$ a potential field in $\mathbf{E}_2$? b) If yes, find its potential.

a) If we denote by U and V the components of $\mathbf{f}$ then we can easily compute that

$$\frac{\partial V}{\partial x} - \frac{\partial U}{\partial y} = 1 - 2y \quad \text{in } \mathbf{E}_2.$$

Thus, condition (VI.4) is not satisfied and so applying Theorem VI.1.8, we can see that vector field $\mathbf{f}$ is not a potential field in $\mathbf{E}_2$.

VI.2.4. Example. It is known from physics that a particle with the mass M at the point $X_0 = [x_0, y_0, z_0]$ generates the gravitational field

$$\mathbf{g} = -\kappa M \frac{(x - x_0)\,\mathbf{i} + (y - y_0)\,\mathbf{j} + (z - z_0)\,\mathbf{k}}{[\,(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2\,]^{3/2}}$$

in $D = \mathbf{E}_3 - \{X_0\}$. Since $\mathbf{g}$ satisfies condition (VI.5) in $\mathbf{E}_3 - \{X_0\}$ (Verify this for yourself!) and the domain $\mathbf{E}_3 - \{X_0\}$ is simply connected (Why?), $\mathbf{g}$ is a potential vector field in $\mathbf{E}_3 - \{X_0\}$.

Potential φ of $\mathbf{g}$ satisfies

$$\frac{\partial \varphi(x, y, z)}{\partial x} = -\kappa M \frac{x - x_0}{[\,(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2\,]^{3/2}}.$$

Integrating this equation with respect to x, we obtain

$$\varphi(x,y,z) \;=\; \frac{\kappa M}{[\,(x-x_0)^2+(y-y_0)^2+(z-z_0)^2\,]^{1/2}} \;+\; C_1(y,z)$$

where $C_1(y,z)$ is the constant of integration. Putting the partial derivatives of φ with respect to y and z equal to the second and the third components of $\mathbf{g}$ and integrating with respect to y and z, we can obtain the same formulas for φ, only with $C_2(x,z)$ or $C_3(x,y)$ instead of $C_1(y,z)$. Comparing all three expressions of φ, we can see that we can put $C_1(y,z) = C_2(x,z) = C_3(x,y) = 0$ and so we get the potential of the gravitational field $\mathbf{g}$ in $\mathbf{E}_3 - \{X_0\}$:

$$\varphi(x,y,z) \;=\; \frac{\kappa M}{[\,(x-x_0)^2+(y-y_0)^2+(z-z_0)^2\,]^{1/2}} \;+\; const.$$

You can verify that the electric field generated by a charge Q at point X_0 is also a potential field in $\mathbf{E}_3 - \{X_0\}$ and its potential φ has a similar form.

VI.3. Solenoidal vector field.

VI.3.1 Solenoidal vector field. A vector field $\mathbf{f}$ in domain D is called *solenoidal* if its flux through any closed surface σ in D is zero. (Note that the flux of $\mathbf{f}$ through surface σ was defined in paragraph V.4.2 as $\iint_\sigma \mathbf{f}\cdot d\mathbf{p}$.)

VI.3.2. Theorem. (Solenoidal field in $\mathbf{E}_3$ – the necessary condition.) *Suppose that $\mathbf{f}$ is a solenoidal vector field in domain $D \subset \mathbf{E}_3$. Suppose that the components of $\mathbf{f}$ are continuously differentiable functions in D. Then*

$$\operatorname{div}\mathbf{f} \;=\; 0 \quad \textit{in } D. \tag{VI.11}$$

P r o o f : By contradiction. Suppose that there exists point $X_0 \in D$ such that $\operatorname{div}\mathbf{f}(X_0) \neq 0$. We can suppose that $\operatorname{div} f(X_0) > 0$ without loss of generality. It follows from the continuity of partial derivatives of the components of $\mathbf{f}$ that there exists a neighbourhood $U(X_0) \subset D$ such that $\operatorname{div}\mathbf{f} > 0$ in all points of $U(X_0)$. Let σ be a sphere with the center X_0 and with such a small radius that $\sigma \subset U(X_0)$. Using the Gauss–Ostrogradsky theorem (see paragraph V.6.3), we obtain

$$\iint_\sigma \mathbf{f}\cdot d\mathbf{p} \;=\; \pm \iiint_{Int\ \sigma} \operatorname{div}\mathbf{f}\;\; dx\,dy\,dz$$

where the "+" sign holds if σ is oriented to its exterior and the "−" sign holds in the opposite case. The integral on the right hand side is positive because $Int\ \sigma \subset U(X_0)$ and $\operatorname{div}\mathbf{f} > 0$ in $U(X_0)$. Thus, the flux of $\mathbf{f}$ through the closed surface σ is different from zero and so vector field $\mathbf{f}$ is not solenoidal in D. This is the desired contradiction.

VI.3.3. Remark. Analogously to conditions (VI.4) and (VI.5), condition (VI.11) is the necessary condition, but it is not a sufficient condition! This may be shown through the following example: The vector function

$$\mathbf{f} \;=\; \frac{x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}}{[\,x^2 + y^2 + z^2\,]^{3/2}}$$

satisfies condition (VI.11) in the domain $D = \mathbf{E}_3 - O$ where $O = [0,0,0]$. (You can check this for yourself.) However, $\mathbf{f}$ is not a solenoidal field in D. We can prove it so that we show that its flux through some closed surface σ in D is different from zero. Thus, let σ be for instance a sphere with the center O and radius R, oriented to its exterior. The flux of $\mathbf{f}$ through σ cannot be evaluated by means of the Gauss–Ostrogradsky theorem (see paragraph V.6.3) because the components of $\mathbf{f}$ do not satisfy the assumption of this theorem. (They are not continuous at point O which belongs to $Int\ \sigma$.) However, the surface integral $\iint_\sigma \mathbf{f}\cdot d\mathbf{p}$ can be computed by means of parametrization P discussed in paragraph V.2.10:

$$\begin{aligned} x &= \phi(u,v) = R\,\cos\,u\,\cos\,v,\\ y &= \psi(u,v) = R\,\sin\,u\,\cos\,v,\\ z &= \vartheta(u,v) = R\,\sin\,v \end{aligned}$$

for $u \in \langle 0, 2\pi\rangle$, $v \in \langle -\pi/2, \pi/2\rangle$. The vector $P_u \times P_v$ is

$$P_u(u,v) \times P_v(u,v) = \big(R^2\,\cos\,u\,\cos^2 v,\ R^2\,\sin\,u\,\cos^2 v,\ R^2\,\sin\,v\,\cos\,v\big).$$

Using now formula (V.6) and applying Fubini's theorem III.3.2, we get

$$\iint_\sigma \mathbf{f}\cdot d\mathbf{p} \;=\; \int_0^{2\pi}\Big(\int_{-\pi/2}^{\pi/2}\big(\cos^2 u\,\cos^3\,v + \sin^2 u\,\cos^3 v + \sin^2 v\,\cos\,v\big)\,dv\Big)\,du \;=$$

$$=\; \int_0^{2\pi}\Big(\int_{-\pi/2}^{\pi/2}\cos\,v\,dv\Big)\,du \;=\; 4\pi.$$

Condition (VI.11) becomes a sufficient condition if it is completed by an assumption about the form of domain D:

VI.3.4. Theorem. (Solenoidal field in $\mathbf{E}_3$ – sufficient conditions.) *Let*

a) *D be a domain in $\mathbf{E}_3$ such that if σ is any closed surface in D then $Int\ \sigma \subset D$,*

b) *$\mathbf{f}$ be a vector field in D, whose components are continuously differentiable functions in D and they satisfy the condition*

$$\text{div}\,\mathbf{f} \;=\; 0 \quad \textit{in } D. \tag{VI.11}$$

Then $\mathbf{f}$ is a solenoidal vector field in D.

P r o o f : Let σ be a closed surface in D. The flux of $\mathbf{f}$ through σ can be evaluated by means of the Gauss–Ostrogradsky theorem and if we also use condition (VI.11), we obtain

$$\iint_\sigma \mathbf{f}\cdot d\mathbf{p} \;=\; \pm\iiint_{Int\ \sigma} \text{div}\,\mathbf{f}\;\; dx\,dy\,dz \;=\; 0.$$

(The sign in front of the triple integral depends on the orientation of σ. However, it is not important because the integral is equal to zero.)

VI.3.5. Example. Vector field $\mathbf{f}$ from paragraph VI.3.3 is solenoidal in the domain $G = \{[x,y,z] \in \mathbf{E}_3;\ z > 0\}$. It can be verified that it satisfies condition (VI.11) in G and moreover, domain G has property a) formulated in Theorem VI.3.4.

VI.4. Exercises.

1. Find maximum domains in $\mathbf{E}_2$ in which the given vector field $\mathbf{f}$ is defined. Verify whether $\mathbf{f}$ is a potential field in these domains. If yes, find the potential φ of $\mathbf{f}$ and evaluate the integral $\int_A^B \mathbf{f} \cdot d\mathbf{s}$.

a) $\mathbf{f}(x, y) = y^2, 2xy), \ \ A = [1, 3], \ B = [3, 2]$

b) $\mathbf{f}(x, y) = \dfrac{-\mathbf{i} + \mathbf{j}}{(x-y)^2}, \ \ A = [1, 2], \ B = [4, 1]$

c) $\mathbf{f}(x, y) = (x^2, y^2), \ \ A = [0, 0], \ B = [3, 5]$

d) $\mathbf{f}(x, y) = \Big(y^2 - \dfrac{x}{\sqrt{y - x^2}} - 1, \ 2xy + \dfrac{1}{2\sqrt{y - x^2}}\Big), \ \ A = [0, 1], \ B = [1, 2]$

e) $\mathbf{f}(x, y) = \dfrac{y^2}{\sqrt{x}}\,\mathbf{i} + 4y\sqrt{x}\,\mathbf{j}, \ \ A = [1, 2], \ B = [4, -2]$

f) $\mathbf{f}(x, y) = \Big(1 - y^2 + \dfrac{1}{2\sqrt{y^2 + x}}, \ \dfrac{y}{\sqrt{y^2 + x}} - 2xy\Big), \ \ A = [-3, 2], \ B = [3, 1]$

g) $\mathbf{f}(x, y) = (xy, x + y), \ \ A = [0, 0], \ B = [1, 1]$

h) $\mathbf{f}(x, y) = \dfrac{-y\,\mathbf{i} + x\,\mathbf{j}}{(x-y)^2}, \ \ A = [2, 1], \ B = [6, 2]$

i) $\mathbf{f}(x, y) = \big(x^3y^2 + x, \ y^2 + yx^4\big), \ \ A = [3, -1], \ B = [1, 5]$

j) $\mathbf{f}(x, y) = \big(1 + y^2 \sin 2x, \ -2y \cos^2 x\big), \ \ A = [\pi, 1], \ B = [\pi/2, 2]$

k) $\mathbf{f}(x, y) = \Big(\ln y - \dfrac{e^y}{x^2}, \ \dfrac{e^y}{x} + \dfrac{x}{y}\Big), \ \ A = [1, 1], \ B = [1, 2]$

l) $\mathbf{f}(x, y) = \Big(\dfrac{x - 2y}{(y-x)^2} + x, \ \dfrac{y}{(y-x)^2} - y^2\Big), \ \ A = [0, 1], \ B = [1, 4]$

m) $\mathbf{f}(x, y) = \dfrac{4x\,\mathbf{i} + y\,\mathbf{j}}{4x^2 + y^2 - 4}, \ \ A = [0, 0], \ B = [0, 2]$

n) $\mathbf{f}(x, y) = (y \sin x, \ y - \cos x), \ \ A = [0, 1], \ B = [5, 2]$

o) $\mathbf{f}(x, y) = (\cos(2y) + y + x, \ y - 2x \sin(2y) + x), \ \ A = [0, 0], \ B = [-2, 2]$

p) $\mathbf{f}(x, y) = (y^2, \ 2xy), \ \ A = [2, 1], \ B = [0, 0]$

2. Function φ is the potential of vector field $\mathbf{f}$ in domain $D \subset \mathbf{E}_3$. Find D (a maximum possible), $\mathbf{f}$ and evaluate the work done by vector field f on curve C leading from point A to point B.

a) $\varphi(x, y, z) = xy + xz + yz, \ \ A = [-1, 2, -1], \ B = [3, 4, 1]$

b) $\varphi(x, y, z) = \ln|x^2 + y^2 + z^2 - 1|, \ \ A = -1, 1, 2], \ B = [-3, 4, -1]$

3. Find maximum domains in $\mathbf{E}_3$ in which the given vector field $\mathbf{f}$ is defined. Verify whether $\mathbf{f}$ is a potential field in these domains. If yes, find the potential φ.

a) $\mathbf{f}(x,y,z) = \left(\frac{y^2}{z}, \frac{2xy}{z}, -\frac{xy^2}{z^2}\right)$ b) $\mathbf{f}(x,y,z) = \left(\frac{x}{x^2+y^2}, \frac{y}{x^2+y^2}, 2z\right)$

c) $\mathbf{f}(x,y,z) = \frac{-x\,\mathbf{i} - y\,\mathbf{j} + \mathbf{k}/2}{\sqrt{z-x^2-y^2}}$ d) $\mathbf{f}(x,y,z) = \left(\frac{2x-y}{x^2+y^2}, \frac{x+2y}{x^2+y^2}, \ln z\right)$

e) $\mathbf{f}(x,y,z) = \left(\frac{1}{\sqrt{x}}\sin z - y^2,\ -2xy,\ 2\sqrt{x}\cos z\right)$ f) $\mathbf{f}(x,y,z) = y^2\mathbf{i} + z^2\mathbf{j} + x^2\mathbf{k}$

g) $\mathbf{f}(x,y,z) = \left(\frac{z}{x-y}, \frac{z}{y-x}, \ln(x-y) + \frac{1}{\sqrt{z}}\right)$ h) $\mathbf{f}(x,y,z) = \frac{x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}}{\sqrt{x^2+y^2+z^2}}$

i) $\mathbf{f}(x,y,z) = \left(e^x(y+z^2),\ e^x,\ ze^x\right)$ j) $\mathbf{f}(x,y,z) = \left(x-y,\ y^2,\ x+z\right)$

k) $\mathbf{f}(x,y,z) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}, 2z\right)$ l) $\mathbf{f}(x,y,z) = \left(\frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}}, \frac{1}{z}\right)$

4. Find maximum domains in $\mathbf{E}_3$ in which the given vector field $\mathbf{f}$ is defined. Verify whether $\mathbf{f}$ is a solenoidal field in these domains.

a) $\mathbf{f}(x,y,z) = \left(y^2, z^2, x^2\right)$ b) $\mathbf{f}(x,y,z) = \left(z-y,\ x-z,\ y-z\right)$

c) $\mathbf{f}(x,y,z) = \left(x,\ y,\ -2z\right)$ d) $\mathbf{f}(x,y,z) = \left(\frac{x}{y}, \sqrt{z-x^2}, -\frac{z}{y}\right)$

e) $\mathbf{f}(x,y,z) = \left(y,\ z,\ x^2\right)$ f) $\mathbf{f}(x,y,z) = \left(xy,\ 1-y^2,\ yz\right)$

g) $\mathbf{f}(x,y,z) = \frac{(yz,\ xz,\ -xy)}{y^2+z^2-1}$ h) $\mathbf{f}(x,y,z) = \left(5x^2,\ y-z,\ \ln z\right)$

5. $\mathbf{f}(x,y,z) = \left(\frac{x-y+z}{(x^2+y^2+z^2)^{3/2}}, \frac{x+y-z}{(x^2+y^2+z^2)^{3/2}}, \frac{-x+y+z}{(x^2+y^2+z^2)^{3/2}}\right)$

a) Show that $\operatorname{div} \mathbf{f} = 0$ in $\mathbf{E}_3 - \{[0,0,0]\}$.

b) Evaluate the flux of $\mathbf{f}$ through the sphere with the center at the origin and radius $r = 1$, oriented outward.

c) Decide whether $\mathbf{f}$ is a solenoidal vector field in $\mathbf{E}_3 - \{[0,0,0]\}$. (Why?)

d) Find a domain in $\mathbf{E}_3$ where vector field $\mathbf{f}$ is solenoidal.

Bibliography

1. Finney R.L. & Thomas G.B. Jr.: *Calculus.* Addison–Wesley Publishing Company, New York 1994
2. Čipera S. & Machalický M.: *Tematické celky pro přednášky z předmětu Matematika II.* Vysokoškolské skriptum. Vydavatelství ČVUT, Praha 1992
3. Čipera S. & Machalický M.: *Tematické celky pro cvičení z předmětu Matematika II.* Vysokoškolské skriptum. Vydavatelství ČVUT, Praha 1992
4. Neustupa J.: *Mathematics I.* Vysokoškolské skriptum. Vydavatelství ČVUT, Praha 1996