MINISTRY OF EDUCATION AND SCIENCE
OF THE RUSSIAN FEDERATION

BASHKIR STATE UNIVERSITY

RUSLAN A. SHARIPOV

COURSE OF ANALYTICAL GEOMETRY

The textbook

UFA 2011

УДК 514.123
ББК 22.151
Ш25

Referee: The division of Mathematical Analysis of Bashkir State Pedagogical University (БГПУ) named after Miftakhetdin Akmulla, Ufa.

In preparing Russian edition of this book I used the computer typesetting on the base of the $\mathcal{AMS}$-TEX package and I used Cyrillic fonts of the Lh-family distributed by the CyrTUG association of Cyrillic TEX users. English edition of this book is also typeset by means of the $\mathcal{AMS}$-TEX package.

Шарипов Р. А.
Ш25 Курс аналитической геометрии:
Учебное пособие / Р. А. Шарипов. — Уфа: РИЦ БашГУ, 2010. — 228 с.
ISBN 978-5-7477-2574-4

Учебное пособие по курсу аналитической геометрии адресовано студентам математикам, физикам, а также студентам инженерно-технических, технологических и иных специальностей, для которых государственные образовательные стандарты предусматривают изучение данного предмета.

УДК 514.123
ББК 22.151

ISBN 978-5-7477-2574-4
English Translation

CONTENTS.

PREFACE.

The elementary geometry, which is learned in school, deals with basic concepts such as *a point, a straight line, a segment.* They are used to compose more complicated concepts: *a polygonal line, a polygon, a polyhedron.* Some curvilinear forms are also considered: *a circle, a cylinder, a cone, a sphere, a ball.*

The analytical geometry basically deals with the same geometric objects as the elementary geometry does. The difference is in a method for studying these objects. The elementary geometry relies on visual impressions and formulate the properties of geometric objects in its axioms. From these axioms various theorems are derived, whose statements in most cases are also revealed in visual impressions. The analytical geometry is more inclined to a numeric description of geometric objects and their properties.

The transition from a geometric description to a numeric description becomes possible due to coordinate systems. Each coordinate system associate some groups of numbers with geometric points, these groups of numbers are called coordinates of points. The idea of using coordinates in geometry belongs French mathematician Rene Descartes. Simplest coordinate systems suggested by him at present time are called *Cartesian coordinate systems.*

The construction of Cartesian coordinates particularly and the analytical geometry in general are based on the concept of *a vector.* The branch of analytical geometry studying vectors is called *the vector algebra.* The vector algebra constitutes the first chapter of this book. The second chapter explains *the theory of straight lines and planes* and *the theory of curves of the second order.* In the third chapter *the theory of surfaces of the second order* is explained in brief.

The book is based on lectures given by the author during several years in Bashkir State University. It was planned as the first book in a series of three books. However, it happens that

the second and the third books in this series were written and published before the first book. These are

– «Course of linear algebra and multidimensional geometry» [1];
– «Course of differential geometry» [2].

Along with the above books, the following books were written:

– «Representations of finite group» [3];
– «Classical electrodynamics and theory of relativity» [4];
– «Quick introduction to tensor analysis» [5].
– «Foundations of geometry for university students and high school students» [6].

The book [3] can be considered as a continuation of the book [1] which illustrates the application of linear algebra to another branch of mathematics, namely to the theory of groups. The book [4] can be considered as a continuation of the book [2]. It illustrates the application of differential geometry to physics. The book [5] is a brief version of the book [2]. As for the book [6], by its subject it should precede this book. It could br recommended to the reader for deeper logical understanding of the elementary geometry.

I am grateful to Prof. R. R. Gadylshin and Prof. D. I. Borisov for reading and refereeing the manuscript of this book and for valuable advices.

December, 2010. R. A. Sharipov.

CHAPTER I

VECTOR ALGEBRA.

§ 1. Three-dimensional Euclidean space. Acsiomatics and visual evidence.

Like the elementary geometry explained in the book [6], the analytical geometry in this book is a geometry of three-dimensional space $\mathbb{E}$. We use the symbol $\mathbb{E}$ for to denote the space that we observe in our everyday life. Despite being seemingly simple, even the empty space $\mathbb{E}$ possesses a rich variety of properties. These properties reveal through the properties of various geometric forms which are comprised in this space or potentially can be comprised in it.

Systematic study of the geometric forms in the space $\mathbb{E}$ was initiated by ancient Greek mathematicians. It was Euclid who succeeded the most in this. He has formulated the basic properties of the space $\mathbb{E}$ in five postulates, from which he derived all other properties of $\mathbb{E}$. At the present time his postulates are called axioms. On the basis of modern requirements to the rigor of mathematical reasoning the list of Euclid's axioms was enlarged from five to twenty. These twenty axioms can be found in the book [6]. In favor of Euclid the space that we observe in our everyday life is denoted by the symbol $\mathbb{E}$ and is called the three-dimensional Euclidean space.

The three-dimensional Euclidean point space $\mathbb{E}$ consists of points. All geometric forms in it also consist of points. subsets of the space $\mathbb{E}$. Among subsets of the space $\mathbb{E}$ straight lines

and planes (see Fig. 1.2) play an especial role. They are used in the statements of the first eleven Euclid's axioms. On the base of these axioms the concept of a segment (see Fig. 1.2) is introduced. The concept of a segment is used in the statement of the twelfth axiom.

The first twelve of Euclid's axioms appear to be sufficient to define the concept of a ray and the concept of an angle between two rays outgoing from the same point. The concepts

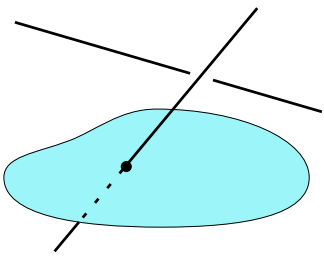

Fig. 1.1

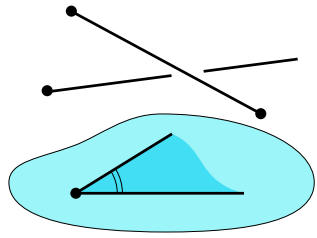

Fig. 1.2

of a segment and an angle along with the concepts of a straight line and a plane appear to be sufficient in order to formulate the remaining eight Euclid's axioms and to build the elementary geometry in whole.

Even the above survey of the book [6], which is very short, shows that building the elementary geometry in an axiomatic way on the basis of Euclid's axioms is a time-consuming and laborious work. However, the reader who is familiar with the elementary geometry from his school curriculum easily notes that proof of theorems in school textbooks are more simple than those in [6]. The matter is that proofs in school textbooks are not proofs in the strict mathematical sense. Analyzing them carefully, one can find in these proofs the usage of some non-proved propositions which are visually obvious from a supplied drawing since we have a rich experience of living within the space $\mathbb{E}$. Such proofs can be transformed to strict mathematical proofs by filling omissions,

i. e. by proving visually obvious propositions used in them.

Unlike [6], in this book I do not load the reader by absolutely strict proofs of geometric results. For geometric definitions, constructions, and theorems the strictness level is used which is close to that of school textbooks and which admits drawings and visually obvious facts as arguments. Whenever it is possible I refer the reader to strict statements and proofs in [6]. As far as the analytical content is concerned, i. e. in equations, in formulas, and in calculations the strictness level is applied which is habitual in mathematics without any deviations.

§ 2. Geometric vectors. Vectors bound to points.

DEFINITION 2.1. A geometric vectors $\overrightarrow{AB}$ is a straight line segment in which the direction from the point A to the point B is specified. The point A is called the *initial point* of the vector $\overrightarrow{AB}$, while the point B is called its *terminal point*.

The direction of the vector $\overrightarrow{AB}$ in drawing is marked by an arrow (see Fig. 2.1). For this reason vectors sometimes are called *directed segments*.

Fig. 2.1

Each segment $[AB]$ is associated with two different vectors: $\overrightarrow{AB}$ and $\overrightarrow{BA}$. The vector $\overrightarrow{BA}$ is usually called the *opposite vector* for the vector $\overrightarrow{AB}$.

Note that the arrow sign on the vector $\overrightarrow{AB}$ and bold dots at the ends of the segment $[AB]$ are merely symbolic signs used to make the drawing more clear. When considered as sets of points the vector $\overrightarrow{AB}$ and the segment $[AB]$ do coincide.

A direction on a segment, which makes it a vector, can mean different things in different situations. For instance, drawing a vector $\overrightarrow{AB}$ on a geographic map, we usually mark the displacement of some object from the point A to the point B. However,

if it is a weather map, the same vector $\overrightarrow{AB}$ can mean the wind direction and its speed at the point A. In the first case the length of the vector $\overrightarrow{AB}$ is proportional to the distance between the points A and B. In the second case the length of $\overrightarrow{AB}$ is proportional to the wind speed at the point A.

There is one more difference in the above two examples. In the first case the vector $\overrightarrow{AB}$ is bound to the points A and B by its meaning. In the second case the vector $\overrightarrow{AB}$ is bound to the point A only. The fact that its arrowhead end is at the point B is a pure coincidence depending on the scale we used for translating the wind speed into the length units on the map. According to what was said, geometric vectors are subdivided into two types:

1) purely geometric;
2) conditionally geometric.

Only displacement vectors belong to the first type; they actually bind some two points of the space $\mathbb{E}$. The lengths of these vectors are always measured in length units: centimeters, meters, inches, feets etc.

Vectors of the second type are more various. These are velocity vectors, acceleration vectors, and force vectors in mechanics; intensity vectors for electric and magnetic fields, magnetization vectors in magnetic materials and media, temperature gradients in non-homogeneously heated objects et al. Vectors of the second type have a geometric direction and they are bound to some point of the space $\mathbb{E}$, but they have not a geometric length. Their lengths can be translated to geometric lengths only upon choosing some scaling factor.

Zero vectors or null vectors hold a special position among geometric vectors. They are defined as follows.

Definition 2.2. A geometric vector of the space $\mathbb{E}$ whose initial and terminal points do coincide with each other is called a *zero vector* or a *null vector*.

A geometric null vector can be either a purely geometric vector

or a conditionally geometric vector depending on its nature.

§ 3. Equality of vectors.

Definition 3.1. Two geometric vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are called *equal* if they are equal in length and if they are *codirected*, i. e. $|AB| = |CD|$ and $\overrightarrow{AB} \uparrow\uparrow \overrightarrow{CD}$.

The vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are said to be codirected if they lie on a same line as shown in Fig. 3.1 of if they lie on parallel lines as shown in Fig. 3.2. In both cases they should be pointing in the

Fig. 3.1

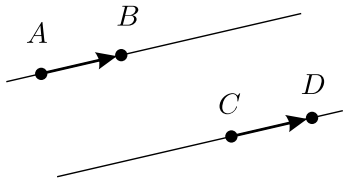

Fig. 3.2

same direction. Codirectedness of geometric vectors and their equality are that very visually obvious properties which require substantial efforts in order to derive them from Euclid's axioms (see [6]). Here I urge the reader not to focus on the lack of rigor in statements, but believe his own geometric intuition.

Zero geometric vectors constitute a special case since they do not fix any direction in the space.

Definition 3.2. All null vectors are assumed to be codirected to each other and each null vector is assumed to be codirected to any nonzero vector.

The length of all null vectors is zero. However, depending on the physical nature of a vector, its zero length is complemented with a measure unit. In the case of zero force it is zero newtons,

in the case of zero velocity it is zero meters per second. For this reason, testing the equality of any two zero vectors, one should take into account their physical nature.

DEFINITION 3.3. All null vectors of the same physical nature are assumed to be equal to each other and any nonzero vector is assumed to be not equal to any null vector.

Testing the equality of nonzero vectors by means of the definition 3.1, one should take into account its physical nature. The equality $|AB| = |CD|$ in this definition assumes not only the equality of numeric values of the lengths of $\overrightarrow{AB}$ and $\overrightarrow{CD}$, but assumes the coincidence of their measure units as well.

A remark. Vectors are geometric forms, i. e. they are sets of points in the space $\mathbb{E}$. However, the equality of two vectors introduced in the definition 3.1 differs from the equality of sets.

§ 4. The concept of a free vector.

Defining the equality of vectors, it is convenient to use parallel translations. Each parallel translation is a special transformation of the space $p\colon \mathbb{E} \to \mathbb{E}$ under which any straight line is mapped onto itself or onto a parallel line and any plane is mapped onto itself or onto a parallel plane. When applied to vectors, parallel translation preserve their length and their direction, i. e. they map each vector onto a vector equal to it, but usually being in a different place in the space. The number of parallel translations is infinite. As appears, the parallel translations are so numerous that they can be used for testing the equality of vectors.

DEFINITION 4.1. A geometric vector $\overrightarrow{CD}$ is called equal to a geometric vector $\overrightarrow{AB}$ if there is a parallel translation $p\colon \mathbb{E} \to \mathbb{E}$ that maps the vector $\overrightarrow{AB}$ onto the vector $\overrightarrow{CD}$, i. e. such that $p(A) = C$ and $p(B) = D$.

The definition 4.1 is equivalent to the definition 3.1. I do not prove this equivalence, relying on its visual evidence and

assuming the reader to be familiar with parallel translations from the school curriculum. A more meticulous reader can see the theorems 8.4 and 9.1 in Chapter VI of the book [6].

THEOREM 4.1. *For any two points A and C in the space $\mathbb{E}$ there is exactly one parallel translation $p\colon \mathbb{E} \to \mathbb{E}$ mapping the point A onto the point C, i. e. such that $p(A) = C$.*

The theorem 4.1 is a visually obvious fact. On the other hand it coincides with the theorem 9.3 from Chapter VI in the book [6], where it is proved. For these two reasons we exploit the theorem 4.1 without proving it in this book.

Lei's apply the theorem 4.1 to some geometric vector $\overrightarrow{AB}$. Let C be an arbitrary point of the space $\mathbb{E}$ and let p be a parallel translation taking the point A to the point C. The existence and uniqueness of such a parallel translation are asserted by the theorem 4.1. Let's define the point D by means of the formula $D = p(B)$. Then, according to the definition 4.1, we have

$$\overrightarrow{AB} = \overrightarrow{CD}.$$

These considerations show that each geometric vector $\overrightarrow{AB}$ has a copy equal to it and attached to an arbitrary point $C \in \mathbb{E}$. In the other words, by means of parallel translations each geometric vector $\overrightarrow{AB}$ can be replicated up to an infinite set of vectors equal to each other and attached to all points of the space $\mathbb{E}$.

DEFINITION 4.2. A *free vector* is an infinite collection of geometric vectors which are equal to each other and whose initial points are at all points of the space $\mathbb{E}$. Each geometric vector in this infinite collection is called a *geometric realization* of a given free vector.

Free vectors can be composed of purely geometric vectors or of conditionally geometric vectors as well. For this reason one can consider free vectors of various physical nature.

In drawings free vectors are usually presented by a single geometric realization or by several geometric realizations if needed.

Geometric vectors are usually denoted by two capital letters: $\overrightarrow{AB}$, $\overrightarrow{CD}$, $\overrightarrow{EF}$ etc. Free vectors are denoted by single lowercase letters: $\vec{\mathbf{a}}$, $\vec{\mathbf{b}}$, $\vec{\mathbf{c}}$ etc. Arrows over these letters are often omitted since it is usually clear from the context that vectors are considered. Below in this book I will not use arrows in denoting free vectors. However, I will use boldface letters for them. In many other books, but not in my book [1], this restriction is also removed.

§ 5. Vector addition.

Assume that two free vectors $\mathbf{a}$ and $\mathbf{b}$ are given. Let's choose some arbitrary point A and consider the geometric realization of the vector $\mathbf{a}$ with the initial point A. Let's denote through B the terminal point of this geometric realization. As a result we get $\mathbf{a} = \overrightarrow{AB}$. Then we consider the geometric realization of the vector $\mathbf{b}$ with initial point B and denote through C its terminal point. This yields $\mathbf{b} = \overrightarrow{BC}$.

Definition 5.1. The geometric vector $\overrightarrow{AC}$ connecting the initial point of the vector $\overrightarrow{AB}$ with the terminal point of the vector $\overrightarrow{BC}$ is called the *sum* of the vectors $\overrightarrow{AB}$ and $\overrightarrow{BC}$:

$$\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}. \tag{5.1}$$

The vector $\overrightarrow{AC}$ constructed by means of the vectors $\mathbf{a} = \overrightarrow{AB}$ and $\mathbf{b} = \overrightarrow{BC}$ can be replicated up to a free vector $\mathbf{c}$ by parallel translations to all points of the space $\mathbb{E}$. Such a vector $\mathbf{c}$ is naturally called the sum of the free vectors $\mathbf{a}$ and $\mathbf{b}$. For this vector we write $\mathbf{c} = \mathbf{a} + \mathbf{b}$. The correctness of such a definition is guaranteed by the following lemma.

Lemma 5.1. *The sum* $\mathbf{c} = \mathbf{a} + \mathbf{b} = \overrightarrow{AC}$ *of two free vectors* $\mathbf{a} = \overrightarrow{AB}$ *and* $\mathbf{b} = \overrightarrow{BC}$ *expressed by the formula* (5.1) *does not*

depend on the choice of a point A at which the geometric realization $\overrightarrow{AB}$ of the vector $\mathbf{a}$ *begins.*

Proof. In addition to A, let's choose another initial point E. Then in the above construction of the sum $\mathbf{a}+\mathbf{b}$ the vector $\mathbf{a}$ has

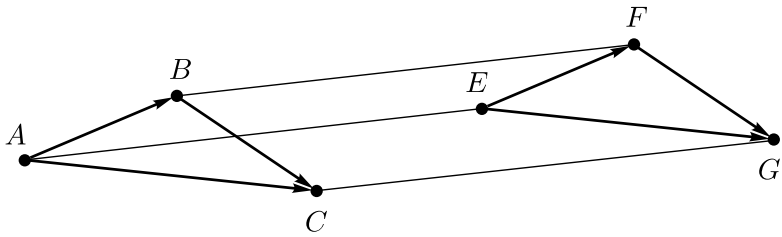

Fig. 5.1

two geometric realizations $\overrightarrow{AB}$ and $\overrightarrow{EF}$. The vector $\mathbf{b}$ also has two geometric realizations $\overrightarrow{BC}$ and $\overrightarrow{FG}$ (see Fig. 5.1). Then

$$\overrightarrow{AB} = \overrightarrow{EF}, \qquad \overrightarrow{BC} = \overrightarrow{FG}. \tag{5.2}$$

Instead of (5.1) now we have two equalities

$$\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}, \qquad \overrightarrow{EG} = \overrightarrow{EF} + \overrightarrow{FG}. \tag{5.3}$$

Let's denote through p a parallel translation that maps the point A to the point E, i. e. such that $p(A) = E$. Due to the theorem 4.1 such a parallel translation does exist and it is unique. From $p(A) = E$ and from the first equality (5.2), applying the definition 4.1, we derive $p(B) = F$. Then from $p(B) = F$ and from the second equality (5.2), again applying the definition 4.1, we get $p(C) = G$. As a result we have

$$p(A) = E, \qquad p(C) = G. \tag{5.4}$$

The relationships (5.4) mean that the parallel translation p maps the vector $\overrightarrow{AC}$ to the vector $\overrightarrow{EG}$. Due to the definition 4.1 this fact yields $\overrightarrow{AC} = \overrightarrow{EG}$. Now from the equalities (5.3) we derive

$$\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{EF} + \overrightarrow{FG}. \tag{5.5}$$

The equalities (5.5) complete the proof of the lemma 5.1. □

The addition rule given by the formula (5.1) is called the *triangle rule*. It is associated with the triangle ABC in Fig. 5.1.

§ 6. Multiplication of a vector by a number.

Let **a** be some free vector. Let's choose some arbitrary point A and consider the geometric realization of the vector **a** with initial point A. Then we denote through B the terminal point of this geometric realization of **a**. Let α be some number. It can be

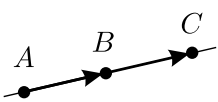

Fig. 6.1

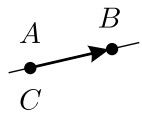

Fig. 6.2

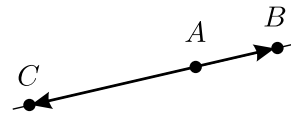

Fig. 6.3

either positive, negative, or zero.

Let $\alpha > 0$. In this case we lay a point C onto the line AB so that the following conditions are fulfilled:

$$\overrightarrow{AC} \uparrow\uparrow \overrightarrow{AB}, \qquad |AC| = |\alpha| \cdot |AB|. \tag{6.1}$$

As a result we obtain the drawing which is shown in Fig. 6.1.

If $\alpha = 0$, we lay the point C so that it coincides with the point A. In this case the vector $\overrightarrow{AC}$ appears to be zero as shown in Fig. 6.2 and we have the relationship

$$|AC| = |\alpha| \cdot |AB|. \tag{6.2}$$

In the case $\alpha < 0$ we lay the point C onto the line AB so that the following two conditions are fulfilled:

$$\overrightarrow{AC} \uparrow\downarrow \overrightarrow{AB}, \qquad |AC| = |\alpha| \cdot |AB|. \tag{6.3}$$

This arrangement of points is shown in Fig. 6.3.

Definition 6.1. In each of the three cases $\alpha > 0$, $\alpha = 0$, and $\alpha < 0$ the geometric vector $\overrightarrow{AC}$ defined through the vector $\overrightarrow{AB}$ according to the drawings in Fig. 6.1, in Fig. 6.2, and in Fig. 6.3 and according to the formulas (6.1), (6.2), and (6.3) is called the product of the vector $\overrightarrow{AB}$ by the number α. This fact is expressed by the following formula:

$$\overrightarrow{AC} = \alpha \cdot \overrightarrow{AB}. \tag{6.4}$$

The case $\mathbf{a} = \mathbf{0}$ is not covered by the above drawings in Fig. 6.1, in Fig. 6.2, and in Fig. 6.3. In this case the point B coincides with the points A and we have $|AB| = 0$. In order to provide the equality $|AC| = |\alpha| \cdot |AB|$ the point C is chosen coinciding with the point A. Therefore the product of a null vector by an arbitrary number is again a null vector.

The geometric vector $\overrightarrow{AC}$ constructed with the use of the vector $\mathbf{a} = \overrightarrow{AB}$ and the number α can be replicated up to a free vector $\mathbf{c}$ by means of the parallel translations to all points of the space $\mathbb{E}$. Such a free vector $\mathbf{c}$ is called the product of the free vector $\mathbf{a}$ by the number α. For this vector we write $\mathbf{c} = \alpha \cdot \mathbf{a}$. The correctness of this definition of $\mathbf{c} = \alpha \cdot \mathbf{a}$ is guaranteed by the following lemma.

Lemma 6.1. *The product* $\mathbf{c} = \alpha \cdot \mathbf{a} = \overrightarrow{AC}$ *of a free vector* $\mathbf{a} = \overrightarrow{AB}$ *by a number* α *expressed by the formula* (6.4) *does not depend on the choice of a point* A *at which the geometric realization of the vector* $\mathbf{a}$ *is built.*

Proof. Let's prove the lemma for the case $\mathbf{a} \neq \mathbf{0}$ and $\alpha > 0$. In addition to A we choose another initial point E. Then in the

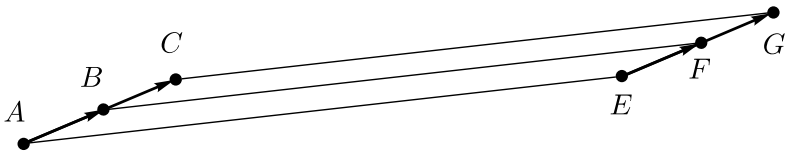

Fig. 6.4

construction of the product $\alpha \cdot \mathbf{a}$ the vector $\mathbf{a}$ gets two geometric realizations $\overrightarrow{AB}$ and $\overrightarrow{EF}$ (see Fig. 6.4). Hence we have

$$\overrightarrow{AB} = \overrightarrow{EF}. \tag{6.5}$$

Let's denote through p the parallel translation that maps the point A to the point E, i. e. such that $p(A) = E$. Then from the equality (6.5), applying the definition 4.1, we derive $p(B) = F$. The point C is placed on the line AB at the distance $|AC| = |\alpha| \cdot |AB|$ from the point A in the direction of the vector $\overrightarrow{AB}$. Similarly, the point G is placed on the line EF at thew distance $|EG| = |\alpha| \cdot |EF|$ from the point E in the direction of the vector $\overrightarrow{EF}$. From the equality (6.5) we derive $|AB| = |EF|$. Therefore $|AC| = |\alpha| \cdot |AB|$ and $|EG| = |\alpha| \cdot |EF|$ mean that $|AC| = |EF|$. Due to $p(A) = E$ and $p(B) = F$ the parallel translation p maps the line AB onto the line EF. It preserves lengths of segments and maps codirected vectors to codirected ones. Hence $p(C) = G$. Along with $p(A) = E$ due to the definition 4.1 the equality $p(C) = G$ yields $\overrightarrow{AC} = \overrightarrow{EG}$, i. e.

$$\alpha \cdot \overrightarrow{AB} = \alpha \cdot \overrightarrow{EF}.$$

The lemma 6.1 is proved for the case $\mathbf{a} \neq \mathbf{0}$ and $\alpha > 0$. Its proof for the other cases is left to the reader as an exercise. □

Exercise 6.1. *Consider the cases $\alpha = 0$ and $\alpha < 0$ for $\mathbf{a} \neq \mathbf{0}$*

and consider the case $\mathbf{a} = \mathbf{0}$. *Prove the lemma 6.1 for these cases and provide your proof with drawings analogous to that of Fig 6.4.*

§ 7. Properties of the algebraic operations with vectors.

The addition of vectors and their multiplication by numbers are two basic algebraic operations with vectors in the three-dimensional Euclidean point space $\mathbb{E}$. Eight basic properties of these two algebraic operations with vectors are usually considered. The first four of these eight properties characterize the operation of addition. The other four characterize the operation of multiplication of vectors by numbers and its correlation with the operation of vector addition.

THEOREM 7.1. *The operation of addition of free vectors and the operation of their multiplication by numbers possess the following properties:*

1) *commutativity of addition:* $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$;
2) *associativity of addition:* $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$;
3) *the feature of the null vector:* $\mathbf{a} + \mathbf{0} = \mathbf{a}$;
4) *the existence of an opposite vector: for any vector* $\mathbf{a}$ *there is an opposite vector* $\mathbf{a}'$ *such that* $\mathbf{a} + \mathbf{a}' = \mathbf{0}$;
5) *distributivity of multiplication over the addition of vectors:* $k \cdot (\mathbf{a} + \mathbf{b}) = k \cdot \mathbf{a} + k \cdot \mathbf{b}$;
6) *distributivity of multiplication over the addition of numbers:* $(k + q) \cdot \mathbf{a} = k \cdot \mathbf{a} + q \cdot \mathbf{a}$;
7) *associativity of multiplication by numbers:* $(k\,q) \cdot \mathbf{a} = k \cdot (q \cdot \mathbf{a})$;
8) *the feature of the numeric unity:* $1 \cdot \mathbf{a} = \mathbf{a}$.

Let's consider the properties listed in the theorem 7.1 one by one. Let's begin with the commutativity of addition. The sum $\mathbf{a} + \mathbf{b}$ in the left hand side of the equality $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ is calculated by means of the triangle rule upon choosing some geometric realizations $\mathbf{a} = \overrightarrow{AB}$ and $\mathbf{b} = \overrightarrow{BC}$ as shown in Fig. 7.1.

Let's draw the line parallel to the line BC and passing through the point A. Then we draw the line parallel to the

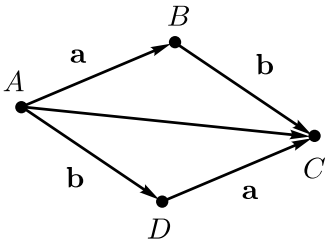

Fig. 7.1

line AB and passing through the point C. Both of these lines are in the plane of the triangle ABC. For this reason they intersect at some point D. The segments $[AB]$, $[BC]$, $[CD]$, and $[DA]$ form a parallelogram.

Let's mark the vectors $\overrightarrow{DC}$ and $\overrightarrow{AD}$ on the segments $[CD]$ and $[DA]$. It is easy to see that the vector $\overrightarrow{DC}$ is produced from the vector $\overrightarrow{AB}$ by applying the parallel translation from the point A to the point D. Similarly the vector $\overrightarrow{AD}$ is produced from the vector $\overrightarrow{BC}$ by applying the parallel translation from the point B to the point A. Therefore $\overrightarrow{DC} = \overrightarrow{AB} = \mathbf{a}$ and $\overrightarrow{BC} = \overrightarrow{AD} = \mathbf{b}$. Now the triangles ABC and ADC yield

$$\begin{aligned} \overrightarrow{AC} &= \overrightarrow{AB} + \overrightarrow{BC} = \mathbf{a} + \mathbf{b}, \\ \overrightarrow{AC} &= \overrightarrow{AD} + \overrightarrow{DC} = \mathbf{b} + \mathbf{a}. \end{aligned} \tag{7.1}$$

From (7.1) we derive the required equality $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$.

The relationship $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ and Fig. 7.1 yield another method for adding vectors. It is called the *parallelogram rule*. In order to add two vectors $\mathbf{a}$ and $\mathbf{b}$ their geometric realizations $\overrightarrow{AB}$ and $\overrightarrow{AD}$ with the common initial point A are used. They are completed up to the parallelogram $ABCD$. Then the diagonal of this parallelogram is taken for the geometric realization of the sum: $\mathbf{a} + \mathbf{b} = \overrightarrow{AB} + \overrightarrow{AD} = \overrightarrow{AC}$.

EXERCISE 7.1. *Prove the equality* $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ *for the case where* $\mathbf{a} \parallel \mathbf{b}$. *For this purpose consider the subcases*

1) $\mathbf{a} \uparrow\uparrow \mathbf{b}$;
2) $\mathbf{a} \uparrow\downarrow \mathbf{b}$ *and* $|\mathbf{a}| > |\mathbf{b}|$;
3) $\mathbf{a} \uparrow\downarrow \mathbf{b}$ *and* $|\mathbf{a}| = |\mathbf{b}|$;
4) $\mathbf{a} \uparrow\downarrow \mathbf{b}$ *and* $|\mathbf{a}| < |\mathbf{b}|$.

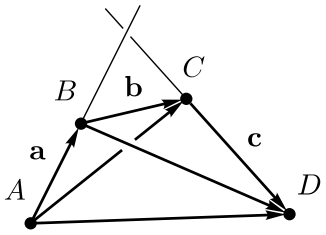

Fig. 7.2

The next property in the theorem 7.1 is the associativity of the operation of vector addition. In order to prove this property we choose some arbitrary initial point A and construct the following geometric realizations of the vectors: $\mathbf{a} = \overrightarrow{AB}$, $\mathbf{b} = \overrightarrow{BC}$, and $\mathbf{c} = \overrightarrow{CD}$. Applying the triangle rule of vector addition to the triangles ABC and ACD, we get the relationships

$$\begin{aligned} \mathbf{a} + \mathbf{b} &= \overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}, \\ (\mathbf{a} + \mathbf{b}) + \mathbf{c} &= \overrightarrow{AC} + \overrightarrow{CD} = \overrightarrow{AD} \end{aligned} \tag{7.2}$$

(see Fig. 7.2). Applying the same rule to the triangles BCD and ABD, we get the analogous relationships

$$\begin{aligned} \mathbf{b} + \mathbf{c} &= \overrightarrow{BC} + \overrightarrow{CD} = \overrightarrow{BD}, \\ \mathbf{a} + (\mathbf{b} + \mathbf{c}) &= \overrightarrow{AB} + \overrightarrow{BD} = \overrightarrow{AD}. \end{aligned} \tag{7.3}$$

The required relationship $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$ now is immediate from the formulas (7.2) and (7.3).

A remark. The tetragon $ABCD$ in Fig. 7.2 is not necessarily planar. For this reason the line CD is shown as if it goes under the line AB, while the line BD is shown going over the line AC.

The feature of the null vector $\mathbf{a} + \mathbf{0} = \mathbf{a}$ is immediate from the triangle rule for vector addition. Indeed, if an initial point A for the vector $\mathbf{a}$ is chosen and if its geometric realization $\overrightarrow{AB}$ is built, then the null vector $\mathbf{0}$ is presented by its geometric realization $\overrightarrow{BB}$. From the definition 5.1 we derive $\overrightarrow{AB} + \overrightarrow{BB} = \overrightarrow{AB}$ which yields $\mathbf{a} + \mathbf{0} = \mathbf{a}$.

The existence of an opposite vector is also easy to prove. Assume that the vector $\mathbf{a}$ is presented by its geometric realization

$\overrightarrow{AB}$. Let's consider the opposite geometric vector $\overrightarrow{BA}$ and let's denote through $\mathbf{a}'$ the corresponding free vector. Then

$$\mathbf{a} + \mathbf{a}' = \overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = \mathbf{0}.$$

The distributivity of multiplication over the vector addition follows from the properties of the *homothety transformation* in the Euclidean space $\mathbb{E}$ (see § 11 of Chapter VI in [6]). It is sometimes called the *similarity transformation*, which is not quite exact. Similarity transformations constitute a larger class of transformations that comprises homothety transformations as a subclass within it.

Let $\mathbf{a} \nparallel \mathbf{b}$ and let the sum of vectors $\mathbf{a} + \mathbf{b}$ is calculated according to the triangle rule as shown in Fig. 7.3. Assume that

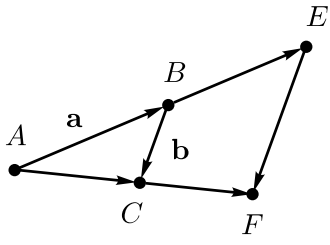

Fig. 7.3

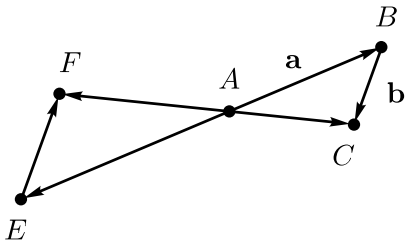

Fig. 7.4

$k > 0$. Let's construct the homothety transformation $h_{kA} \colon \mathbb{E} \to \mathbb{E}$ with the center at the point A and with the homothety factor k. Let's denote through E the image of the point B under the transformation h_{kA} and let's denote through F the image of the point C under this transformation:

$$E = h_{kA}(B), \qquad F = h_{kA}(C).$$

Due to the properties of the homothety the line EF is parallel to

the line BC and we have the following relationships:

$$\begin{aligned} &\overrightarrow{EF} \uparrow\uparrow \overrightarrow{BC}, && |EF| = |k| \cdot |BC|, \\ &\overrightarrow{AE} \uparrow\uparrow \overrightarrow{AB}, && |AE| = |k| \cdot |AB|, \\ &\overrightarrow{AF} \uparrow\uparrow \overrightarrow{AC}, && |AF| = |k| \cdot |AC|. \end{aligned} \tag{7.4}$$

Comparing (7.4) with (6.1) and taking into account that we consider the case $k > 0$, from (7.4) we derive

$$\overrightarrow{AE} = k \cdot \overrightarrow{AB}, \qquad \overrightarrow{EF} = k \cdot \overrightarrow{BC}, \qquad \overrightarrow{AF} = k \cdot \overrightarrow{AC}. \tag{7.5}$$

The relationships (7.5) are sufficient for to prove the distributivity of the mutiplication of vectors by numbers over the operation of vector addition. Indeed, from (7.5) we obtain:

$$\begin{aligned} &k \cdot (\mathbf{a} + \mathbf{b}) = k \cdot (\overrightarrow{AB} + \overrightarrow{BC}) = k \cdot \overrightarrow{AC} = \overrightarrow{AF} = \\ &= \overrightarrow{AE} + \overrightarrow{EF} = k \cdot \overrightarrow{AB} + k \cdot \overrightarrow{BC} = k \cdot \mathbf{a} + k \cdot \mathbf{b}. \end{aligned} \tag{7.6}$$

The case where $\mathbf{a} \nparallel \mathbf{b}$ and $k < 0$ is very similar to the case just above. In this case Fig. 7.3 is replaced by Fig. 7.4. Instead of the relationships (7.4) we have the relationships

$$\begin{aligned} &\overrightarrow{EF} \uparrow\downarrow \overrightarrow{BC}, && |EF| = |k| \cdot |BC|, \\ &\overrightarrow{AE} \uparrow\downarrow \overrightarrow{AB}, && |AE| = |k| \cdot |AB|, \\ &\overrightarrow{AF} \uparrow\downarrow \overrightarrow{AC}, && |AF| = |k| \cdot |AC|. \end{aligned} \tag{7.7}$$

Taking into account $k < 0$ from (7.7) we derive (7.5) and (7.6).

In the case $k = 0$ the relationship $k \cdot (\mathbf{a} + \mathbf{b}) = k \cdot \mathbf{a} + k \cdot \mathbf{b}$ reduces to the equality $\mathbf{0} = \mathbf{0} + \mathbf{0}$. This equality is trivially valid.

EXERCISE 7.2. *Prove that* $k\cdot(\mathbf{a}+\mathbf{b})=k\cdot\mathbf{a}+k\cdot\mathbf{b}$ *for the case* $\mathbf{a}\parallel\mathbf{b}$. *For this purpose consider the subcases*

1) $\mathbf{a}\uparrow\uparrow\mathbf{b}$;
2) $\mathbf{a}\uparrow\downarrow\mathbf{b}$ *and* $|\mathbf{a}|>|\mathbf{b}|$;
3) $\mathbf{a}\uparrow\downarrow\mathbf{b}$ *and* $|\mathbf{a}|=|\mathbf{b}|$;
4) $\mathbf{a}\uparrow\downarrow\mathbf{b}$ *and* $|\mathbf{a}|<|\mathbf{b}|$.

In each of these subcases consider two options: $k>0$ *and* $k<0$.

Let's proceed to proving the distributivity of multiplication of vectors by numbers over the addition of numbers. The case $\mathbf{a}=\mathbf{0}$ in the equality $(k+q)\cdot\mathbf{a}=k\cdot\mathbf{a}+q\cdot\mathbf{a}$ is trivial. In this case the equality $(k+q)\cdot\mathbf{a}=k\cdot\mathbf{a}+q\cdot\mathbf{a}$ reduces to $\mathbf{0}=\mathbf{0}+\mathbf{0}$.

The cases $k=0$ and $q=0$ are also trivial. In these cases the equality $(k+q)\cdot\mathbf{a}=k\cdot\mathbf{a}+q\cdot\mathbf{a}$ reduces to the equalities $q\cdot\mathbf{a}=\mathbf{0}+q\cdot\mathbf{a}$ and $k\cdot\mathbf{a}=k\cdot\mathbf{a}+\mathbf{0}$ respectively.

Let's consider the case $\mathbf{a}\neq\mathbf{0}$ and for the sake of certainty let's assume that $k>0$ and $q>0$. Let's choose some arbitrary point A and let's build the geometric realizations of the vector $k\cdot\mathbf{a}$ with the initial point A. Let B be the terminal point of this geometric realization. Then $\overrightarrow{AB}=k\cdot\mathbf{a}$. Similarly, we construct the geometric realization $\overrightarrow{BC}=q\cdot\mathbf{a}$. Due to $k>0$ and $q>0$ the vectors $\overrightarrow{AB}$ and $\overrightarrow{BC}$ are codirected to the vector $\mathbf{a}$. These vectors lie on the line AB (see Fig. 7.5). The sum of these two vectors

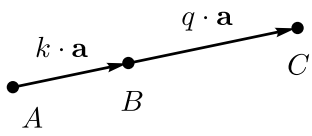

Fig. 7.5

$$\overrightarrow{AC}=\overrightarrow{AB}+\overrightarrow{BC} \tag{7.8}$$

lie on the same line and it is codirected to the vector $\mathbf{a}$. The length of the vector $\overrightarrow{AC}$ is given by the formula

$$|AC|=|AB|+|BC|=k\,|\mathbf{a}|+q\,|\mathbf{a}|=(k+q)\,|\mathbf{a}|. \tag{7.9}$$

Due to $\overrightarrow{AC} \uparrow\uparrow \mathbf{a}$ and due to $k+q>0$ from (7.9) we derive

$$\overrightarrow{AC} = (k+q)\cdot\mathbf{a}. \tag{7.10}$$

Let's substitute (7.10) and (7.8) and take into account the relationships $\overrightarrow{AB} = k\cdot\mathbf{a}$ and $\overrightarrow{BC} = q\cdot\mathbf{a}$ which follow from our constructions. As a result we get the required equality $(k+q)\cdot\mathbf{a} = k\cdot\mathbf{a}+q\cdot\mathbf{a}$.

EXERCISE 7.3. *Prove that* $(k+q)\cdot\mathbf{a} = k\cdot\mathbf{a}+q\cdot\mathbf{a}$ *for the case where* $\mathbf{a}\neq\mathbf{0}$, *while* k *and* q *are two numbers of mutually opposite signs. For the case consider the subcases*

1) $|k|>|q|$; 2) $|k|=|q|$; 3) $|k|<|q|$.

The associativity of the multiplication of vectors by numbers is expressed by the equality $(k\,q)\cdot\mathbf{a} = k\cdot(q\cdot\mathbf{a})$. If $\mathbf{a}=\mathbf{0}$, this equality is trivially fulfilled. It reduces to $\mathbf{0}=\mathbf{0}$. If $k=0$ or if $q=0$, it is also trivial. In this case it reduces to $\mathbf{0}=\mathbf{0}$.

Let's consider the case $\mathbf{a}\neq\mathbf{0}$, $k>0$, and $q>0$. Let's choose some arbitrary point A in the space $\mathbb{E}$ and build the geometric realization of the vector $q\cdot\mathbf{a}$ with the initial point A. Let B be the terminal point of this geometric realization. Then $\overrightarrow{AB} = q\cdot\mathbf{a}$ (see Fig. 7.6). Due to $q>0$ the vector $\overrightarrow{AB}$ is codirected with the vector $\mathbf{a}$. Let's build the vector $\overrightarrow{AC}$ as the product $\overrightarrow{AC} = k\cdot\overrightarrow{AB} = k\cdot(q\cdot\mathbf{a})$ relying upon the definition 6.1. Due to $k>0$ the vector $\overrightarrow{AC}$ is also codirected with $\mathbf{a}$. The lengths of $\overrightarrow{AB}$ and $\overrightarrow{AC}$ are given by the formulas

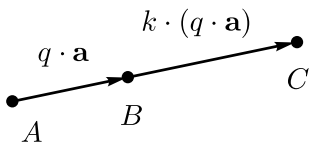

Fig. 7.6

$$|AB| = q\,|\mathbf{a}|, \qquad |AC| = k\,|AB|. \tag{7.11}$$

From the relationships (7.11) we derive the equality

$$|AC| = k\,(q\,|\mathbf{a}|) = (k\,q)\,|\mathbf{a}|. \tag{7.12}$$

The equality (7.12) combined with $\overrightarrow{AC} \uparrow\uparrow \mathbf{a}$ and $k\,q > 0$ yields $\overrightarrow{AC} = (k\,q)\cdot\mathbf{a}$. By our construction $\overrightarrow{AC} = k\cdot\overrightarrow{AB} = k\cdot(q\cdot\mathbf{a})$. As a result now we immediately derive the required equality $(k\,q)\cdot\mathbf{a} = k\cdot(q\cdot\mathbf{a})$.

EXERCISE 7.4. *Prove the equality* $(k\,q)\cdot\mathbf{a} = k\cdot(q\cdot\mathbf{a})$ *in the case where* $\mathbf{a}\neq\mathbf{0}$*, while the numbers* k *and* q *are of opposite signs. For this case consider the following two subcases:*

$$1)\ \ k>0\ \ \textit{and}\ \ q<0; \qquad\qquad 2)\ \ k<0\ \ \textit{and}\ \ q>0.$$

The last item 8 in the theorem 7.1 is trivial. It is immediate from the definition 6.1.

§ 8. Vectorial expressions and their transformations.

The properties of the algebraic operations with vectors listed in the theorem 7.1 are used in transforming vectorial expressions. Saying a vectorial expression one usually assumes a formula such that it yields a vector upon performing calculations according to this formula. In this section we consider some examples of vectorial expressions and learn some methods of transforming these expressions.

Assume that a list of several vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is given. Then one can write their sum setting brackets in various ways:

$$\begin{aligned}&(\mathbf{a}_1+\mathbf{a}_2)+(\mathbf{a}_3+(\mathbf{a}_4+\ldots+(\mathbf{a}_{n-1}+\mathbf{a}_n)\ldots)),\\ &(\ldots((((\mathbf{a}_1+\mathbf{a}_2)+\mathbf{a}_3)+\mathbf{a}_4)+\ldots+\mathbf{a}_{n-1})+\mathbf{a}_n).\end{aligned} \tag{8.1}$$

There are many ways of setting brackets. The formulas (8.1) show only two of them. However, despite the abundance of the

ways for brackets setting, due to the associativity of the vector addition (see item 2 in the theorem 7.1) all of the expressions like (8.1) yield the same result. For this reason the sum of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ can be written without brackets at all:

$$\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3 + \mathbf{a}_4 + \ldots + \mathbf{a}_{n-1} + \mathbf{a}_n. \tag{8.2}$$

In order to make the formula (8.2) more concise the summation sign is used. Then the formula (8.2) looks like

$$\mathbf{a}_1 + \mathbf{a}_2 + \ldots + \mathbf{a}_n = \sum_{i=1}^{n} \mathbf{a}_i. \tag{8.3}$$

The variable i in the formula (8.3) plays the role of the cycling variable in the summation cycle. It is called a *summation index*. This variable takes all integer values ranging from $i = 1$ to $i = n$. The sum (8.3) itself does not depend on the variable i. The symbol i in the formula (8.3) can be replaced by any other symbol, e. g. by the symbol j or by the symbol k:

$$\sum_{i=1}^{n} \mathbf{a}_i = \sum_{j=1}^{n} \mathbf{a}_j = \sum_{k=1}^{n} \mathbf{a}_k. \tag{8.4}$$

The trick with changing (redesignating) a summation index used in (8.4) is often applied for transforming expressions with sums.

The commutativity of the vector addition (see item 1 in the theorem 7.1) means that we can change the order of summands in sums of vectors. For instance, in the sum (8.2) we can write

$$\mathbf{a}_1 + \mathbf{a}_2 + \ldots + \mathbf{a}_n = \mathbf{a}_n + \mathbf{a}_{n-1} + \ldots + \mathbf{a}_1.$$

The most often application for the commutativity of the vector addition is changing the summation order in multiple sums. Assume that a collection of vectors $\mathbf{a}_{ij}$ is given which is indexed

by two indices i and j, where $i = 1, \ldots, m$ and $j = 1, \ldots, n$. Then we have the equality

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbf{a}_{ij} = \sum_{j=1}^{n}\sum_{i=1}^{m}\mathbf{a}_{ij} \tag{8.5}$$

that follows from the commutativity of the vector addition. In the same time we have the equality

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbf{a}_{ij} = \sum_{j=1}^{m}\sum_{i=1}^{n}\mathbf{a}_{ji} \tag{8.6}$$

which is obtained by redesignating indices. Both methods of transforming multiple sums (8.5) and (8.6) are used in dealing with vectors.

The third item in the theorem 7.1 describes the property of the null vector. This property is often used in calculations. If the sum of a part of summands in (8.3) is zero, e. g. if the equality

$$\mathbf{a}_{k+1} + \ldots + \mathbf{a}_n = \sum_{i=k+1}^{n}\mathbf{a}_i = \mathbf{0}$$

is fulfilled, then the sum (8.3) can be transformed as follows:

$$\mathbf{a}_1 + \ldots + \mathbf{a}_n = \sum_{i=1}^{n}\mathbf{a}_i = \sum_{i=1}^{k}\mathbf{a}_i = \mathbf{a}_1 + \ldots + \mathbf{a}_k.$$

The fourth item in the theorem 7.1 declares the existence of an opposite vector $\mathbf{a}'$ for each vector $\mathbf{a}$. Due to this item we can define the subtraction of vectors.

Definition 8.1. The difference of two vectors $\mathbf{a} - \mathbf{b}$ is the sum of the vector $\mathbf{a}$ with the vector $\mathbf{b}'$ opposite to the vector $\mathbf{b}$. This fact is written as the equality

$$\mathbf{a} - \mathbf{b} = \mathbf{a} + \mathbf{b}'. \tag{8.7}$$

EXERCISE 8.1. *Using the definitions 6.1 and 8.1, show that the opposite vector* $\mathbf{a}'$ *is produced from the vector* $\mathbf{a}$ *by multiplying it by the number* -1, *i. e.*

$$\mathbf{a}' = (-1) \cdot \mathbf{a}. \tag{8.8}$$

Due to (8.8) the vector $\mathbf{a}'$ *opposite to* $\mathbf{a}$ *is denoted through* $-\mathbf{a}$ *and we write* $\mathbf{a}' = -\mathbf{a} = (-1) \cdot \mathbf{a}$.

The distributivity properties of the multiplication of vectors by numbers (see items 5 and 6 in the theorem 7.1) are used for expanding expressions and for collecting similar terms in them:

$$\alpha \cdot \left(\sum_{i=1}^{n} \mathbf{a}_i \right) = \sum_{i=1}^{n} \alpha \cdot \mathbf{a}_i, \tag{8.9}$$

$$\left(\sum_{i=1}^{n} \alpha_i \right) \cdot \mathbf{a} = \sum_{i=1}^{n} \alpha_i \cdot \mathbf{a}. \tag{8.10}$$

Transformations like (8.9) and (8.10) can be found in various calculations with vectors.

EXERCISE 8.2. *Using the relationships (8.7) and (8.8), prove the following properties of the operation of subtraction:*

$$\mathbf{a} - \mathbf{a} = \mathbf{0}; \qquad (\mathbf{a} + \mathbf{b}) - \mathbf{c} = \mathbf{a} + (\mathbf{b} - \mathbf{c});$$
$$(\mathbf{a} - \mathbf{b}) + \mathbf{c} = \mathbf{a} - (\mathbf{b} - \mathbf{c}); \qquad (\mathbf{a} - \mathbf{b}) - \mathbf{c} = \mathbf{a} - (\mathbf{b} + \mathbf{c});$$
$$\alpha \cdot (\mathbf{a} - \mathbf{b}) = \alpha \cdot \mathbf{a} - \alpha \cdot \mathbf{b}; \qquad (\alpha - \beta) \cdot \mathbf{a} = \alpha \cdot \mathbf{a} - \beta \cdot \mathbf{a}.$$

Here $\mathbf{a}$, $\mathbf{b}$, *and* $\mathbf{c}$ *are vectors, while* α *and* β *are numbers.*

The associativity of the multiplication of vectors by numbers (see item 7 in the theorem 7.1) is used expanding vectorial expressions. Here is an example of such usage:

$$\beta \cdot \left(\sum_{i=1}^{n} \alpha_i \cdot \mathbf{a}_i \right) = \sum_{i=1}^{n} \beta \cdot (\alpha_i \cdot \mathbf{a}_i) = \sum_{i=1}^{n} (\beta\, \alpha_i) \cdot \mathbf{a}_i. \tag{8.11}$$

In multiple sums this property is combined with the commutativity of the regular multiplication of numbers by numbers:

$$\sum_{i=1}^{m} \alpha_i \cdot \left(\sum_{j=1}^{n} \beta_j \cdot \mathbf{a}_{ij} \right) = \sum_{j=1}^{n} \beta_j \cdot \left(\sum_{i=1}^{m} \alpha_i \cdot \mathbf{a}_{ij} \right).$$

A remark. It is important to note that the associativity of the multiplication of vectors by numbers is that very property because of which one can omit the dot sign in writing a product of a number and a vector:

$$\alpha\, \mathbf{a} = \alpha \cdot \mathbf{a}.$$

Below I use both forms of writing for products of vectors by numbers intending to more clarity, conciseness, and aesthetic beauty of formulas.

The last item 8 of the theorem 7.1 expresses the property of the numeric unity in the form of the relationship $1 \cdot \mathbf{a} = \mathbf{a}$. This property is used in collecting similar terms and in finding common factors. Let's consider an example:

$$\mathbf{a} + 3 \cdot \mathbf{b} + 2 \cdot \mathbf{a} + \mathbf{b} = \mathbf{a} + 2 \cdot \mathbf{a} + 3 \cdot \mathbf{b} + \mathbf{b} = 1 \cdot \mathbf{a} + 2 \cdot \mathbf{a} +$$
$$+ 3 \cdot \mathbf{b} + 1 \cdot \mathbf{b} = (1 + 2) \cdot \mathbf{a} + (3 + 1) \cdot \mathbf{b} = 3 \cdot \mathbf{a} + 4 \cdot \mathbf{b}.$$

EXERCISE 8.3. *Using the relationship* $1 \cdot \mathbf{a} = \mathbf{a}$, *prove that the conditions* $\alpha \cdot \mathbf{a} = \mathbf{0}$ *and* $\alpha \neq 0$ *imply* $\mathbf{a} = \mathbf{0}$.

§ 9. Linear combinations. Triviality, non-triviality, and vanishing.

Assume that some set of n free vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is given. One can call it a collection of n vectors, a system of n vectors, or a family of n vectors either.

Using the operation of vector addition and the operation of multiplication of vectors by numbers, one can compose some

vectorial expression of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$. It is quite likely that this expression will comprise sums of vectors taken with some numeric coefficients.

DEFINITION 9.1. An expression of the form $\alpha_1 \mathbf{a}_1 + \ldots + \alpha_n \mathbf{a}_n$ composed of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is called a *linear combination* of these vectors. The numbers $\alpha_1, \ldots, \alpha_n$ are called the *coefficients of a linear combination*. If

$$\alpha_1 \mathbf{a}_1 + \ldots + \alpha_n \mathbf{a}_n = \mathbf{b}, \tag{9.1}$$

then the vector $\mathbf{b}$ is called the *value of a linear combination*.

In complicated vectorial expressions linear combinations of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ can be multiplied by numbers and can be added to other linear combinations which can also be multiplied by some numbers. Then these sums can be multiplied by numbers and again can be added to other subexpressions of the same sort. This process can be repeated several times. However, upon expanding, upon applying the formula (8.11), and upon collecting similar terms with the use of the formula (8.10) all such complicated vectorial expressions reduce to linear combinations of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$. Let's formulate this fact as a theorem.

THEOREM 9.1. *Each vectorial expression composed of vectors* $\mathbf{a}_1, \ldots, \mathbf{a}_n$ *by means of the operations of addition and multiplication by numbers can be transformed to some linear combination of these vectors* $\mathbf{a}_1, \ldots, \mathbf{a}_n$.

The value of a linear combination does not depend on the order of summands in it. For this reason linear combinations differing only in order of summands are assumed to be coinciding. For example, the expressions $\alpha_1 \mathbf{a}_1 + \ldots + \alpha_n \mathbf{a}_n$ and $\alpha_n \mathbf{a}_n + \ldots + \alpha_1 \mathbf{a}_1$ are assumed to define the same linear combination.

DEFINITION 9.2. A linear combination $\alpha_1 \mathbf{a}_1 + \ldots + \alpha_n \mathbf{a}_n$ composed of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is called *trivial* if all of its coefficients are zero, i. e. if $\alpha_1 = \ldots = \alpha_n = 0$.

DEFINITION 9.3. A linear combination $\alpha_1 \mathbf{a}_1 + \ldots + \alpha_n \mathbf{a}_n$ composed of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is called *vanishing* or *being equal to zero* if its value is equal to the null vector, i. e. if the vector $\mathbf{b}$ in (9.1) is equal to zero.

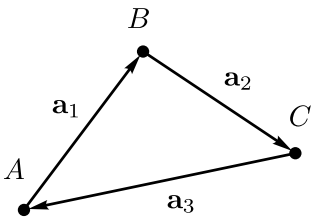

Fig. 9.1

Each trivial linear combination is equal to zero. However, the converse is not valid. Noe each vanishing linear combination is trivial. In Fig. 9.1 we have a triangle ABC. Its sides are marked as vectors $\mathbf{a}_1 = \overrightarrow{AB}$, $\mathbf{a}_2 = \overrightarrow{BC}$, and $\mathbf{a}_3 = \overrightarrow{CA}$. By construction the sum of these three vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ in Fig. 9.1 is zero:

$$\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3 = \overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{CA} = \mathbf{0}. \tag{9.2}$$

The equality (9.2) can be written as follows:

$$1 \cdot \mathbf{a}_1 + 1 \cdot \mathbf{a}_2 + 1 \cdot \mathbf{a}_3 = \mathbf{0}. \tag{9.3}$$

It is easy to see that the linear combination in the left hand side of the equality (9.3) is not trivial (see Definition 9.2), however, it is equal to zero according to the definition 9.3.

DEFINITION 9.4. A linear combination $\alpha_1 \mathbf{a}_1 + \ldots + \alpha_n \mathbf{a}_n$ composed of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is called *non-trivial* if it is not trivial, i. e. at least one of its coefficients $\alpha_1, \ldots, \alpha_n$ is not equal to zero.

§ 10. Linear dependence and linear independence.

DEFINITION 10.1. A system of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is called *linearly dependent* if there is a non-trivial linear combination of these vectors which is equal to zero.

The vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ shown in Fig. 9.1 is an example of a linearly dependent set of vectors.

It is important to note that the linear dependence is a property of systems of vectors, it is not a property of linear combinations. Linear combinations in the definition 10.1 are only tools for revealing the linear dependence.

It is also important to note that the linear dependence, once it is present in a collection of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$, does not depend on the order of vectors in this collection. This follows from the fact that the value of any linear combination and its triviality or non-triviality are not destroyed if we transpose its summands.

DEFINITION 10.2. A system of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is called *linearly independent*, if it is not linearly dependent in the sense of the definition 10.1, i. e. if there is no linear combination of these vectors being non-trivial and being equal to zero simultaneously.

One can prove the existence of a linear combination with the required properties in the definition 10.1 by finding an example of such a linear combination. However, proving the non-existence in the definition 10.2 is more difficult. For this reason the following theorem is used.

THEOREM 10.1 (**linear independence criterion**). *A system of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is linearly independent if and only if vanishing of a linear combination of these vectors implies its triviality.*

PROOF. The proof is based on a simple logical reasoning. Indeed, the non-existence of a linear combination of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$, being non-trivial and vanishing simultaneously means that a linear combination of these vectors is inevitably trivial whenever it is equal to zero. In other words vanishing of a linear combination of these vectors implies triviality of this linear combination. The theorem 10.1 is proved. □

THEOREM 10.2. *A system of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is linearly independent if and only if non-triviality of a linear combination of these vectors implies that it is not equal to zero.*

The theorem 10.2 is very similar to the theorem 10.1. However, it is less popular and is less often used.

Exercise 10.1. *Prove the theorem 10.2 using the analogy with the theorem 10.1.*

§ 11. Properties of the linear dependence.

Definition 11.1. The vector $\mathbf{b}$ is said *to be expressed as a linear combination* of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ if it is the value of some linear combination composed of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ (see (9.1)). In this situation for the sake of brevity the vector $\mathbf{b}$ is sometimes said *to be linearly expressed* through the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ or *to be expressed in a linear way* through $\mathbf{a}_1, \ldots, \mathbf{a}_n$.

There are five basic properties of the linear dependence of vectors. We formulate them as a theorem.

Theorem 11.1. *The relation of the linear dependence for a system of vectors possesses the following basic properties:*

1) *a system of vectors comprising the null vector is linearly dependent;*
2) *a system of vectors comprising a linearly dependent subsystem is linearly dependent itself;*
3) *if a system of vectors is linearly dependent, then at least one of these vectors is expressed in a linear way through other vectors of this system;*
4) *if a system of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is linearly independent, while complementing it with one more vector $\mathbf{a}_{n+1}$ makes the system linearly dependent, then the vector $\mathbf{a}_{n+1}$ is linearly expressed through the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$;*
5) *if a vector $\mathbf{b}$ is linearly expressed through some m vectors $\mathbf{a}_1, \ldots, \mathbf{a}_m$ and if each of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_m$ is linearly expressed through some other n vectors $\mathbf{c}_1, \ldots, \mathbf{c}_n$, then the vector $\mathbf{b}$ is linearly expressed through the vectors $\mathbf{c}_1, \ldots, \mathbf{c}_n$.*

The properties 1)–5) in the theorem 11.1 are relatively simple. Their proofs are purely algebraic, they do not require drawings. I do not prove them in this book since the reader can find their proofs in § 3 of Chapter I in the book [1].

Apart from the properties 1)–5) listed in the theorem 11.1, there is one more property which is formulated separately.

THEOREM 11.2 (**Steinitz**). *If the vectors* $\mathbf{a}_1, \ldots, \mathbf{a}_n$ *are linearly independent and if each of them is linearly expressed through some other vectors* $\mathbf{b}_1, \ldots, \mathbf{b}_m$, *then* $m \geqslant n$.

The Steinitz theorem 11.2 is very important in studying multidimensional spaces. We do not use it in studying the three-dimensional space $\mathbb{E}$ in this book.

§ 12. Linear dependence for $n = 1$.

Let's consider the case of a system composed of a single vector $\mathbf{a}_1$ and apply the definition of the linear dependence 10.1 to this system. The linear dependence of such a system of one vector $\mathbf{a}_1$ means that there is a linear combination of this single vector which is non-trivial and equal to zero at the same time:

$$\alpha_1 \mathbf{a}_1 = \mathbf{0}. \tag{12.1}$$

Non-triviality of the linear combination in the left hand side of (12.1) means that $\alpha_1 \neq 0$. Due to $\alpha_1 \neq 0$ from (12.1) we derive

$$\mathbf{a}_1 = \mathbf{0} \tag{12.2}$$

(see Exercise 8.3). Thus, the linear dependence of a system composed of one vector $\mathbf{a}_1$ yields $\mathbf{a}_1 = \mathbf{0}$.

The converse proposition is also valid. Indeed, assume that the equality (12.2) is fulfilled. Let's write it as follows:

$$1 \cdot \mathbf{a}_1 = \mathbf{0}. \tag{12.3}$$

The left hand side of the equality (12.3) is a non-trivial linear combination for the system of one vector $\mathbf{a}_1$ which is equal to zero. Its existence means that such a system of one vector is linearly dependent. We write this result as a theorem.

THEOREM 12.1. *A system composed of a single vector* $\mathbf{a}_1$ *is linearly dependent if and only if this vector is zero, i. e.* $\mathbf{a}_1 = \mathbf{0}$.

§ 13. Linear dependence for $n = 2$. Collinearity of vectors.

Let's consider a system composed of two vectors $\mathbf{a}_1$ and $\mathbf{a}_2$. Applying the definition of the linear dependence 10.1 to it, we get the existence of a linear combination of these vectors which is non-trivial and equal to zero simultaneously:

$$\alpha_1 \mathbf{a}_1 + \alpha_2 \mathbf{a}_2 = \mathbf{0}. \tag{13.1}$$

The non-triviality of the linear combination in the left hand side of the equality (13.1) means that $\alpha_1 \neq 0$ or $\alpha_2 \neq 0$. Since the linear dependence is not sensitive to the order of vectors in a system, without loss of generality we can assume that $\alpha_1 \neq 0$. Then the equality (13.1) can be written as

$$\mathbf{a}_1 = -\frac{\alpha_2}{\alpha_1}\,\mathbf{a}_2. \tag{13.2}$$

Let's denote $\beta_2 = -\alpha_2/\alpha_1$ and write the equality (13.2) as

$$\mathbf{a}_1 = \beta_2\,\mathbf{a}_2. \tag{13.3}$$

Note that the relationship (13.3) could also be derived by means of the item 3 of the theorem 11.1.

According to (13.3), the vector $\mathbf{a}_1$ is produced from the vector $\mathbf{a}_2$ by multiplying it by the number β_2. In multiplying a vector by a number it length is usually changed (see Formulas (6.1), (6.2), (6.3), and Figs. 6.1, 6.2, and 6.3). As for its direction, it

either is preserved or is changed for the opposite one. In both of these cases the vector $\beta_2\,\mathbf{a}_2$ is parallel to the vector $\mathbf{a}_2$. If $\beta_2 = 0\text{m}$ the vector $\beta_2\,\mathbf{a}_2$ appears to be the null vector. Such a vector does not have its own direction, the null vector is assumed to be parallel to any other vector by definition. As a result of the above considerations the equality (13.3) yields

$$\mathbf{a}_1 \parallel \mathbf{a}_2. \tag{13.4}$$

In the case of vectors for to denote their parallelism a special term *collinearity* is used.

Definition 13.1. Two free vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ are called *collinear*, if their geometric realizations are parallel to some straight line common to both of them.

As we have seen above, in the case of two vectors their linear dependence implies the collinearity of these vectors. The converse proposition is also valid. Assume that the relationship (13.4) is fulfilled. If both vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ are zero, then the equality (13.3) is fulfilled where we can choose $\beta_2 = 1$. If at least one the two vectors is nonzero, then up to a possible renumerating these vectors we can assume that $\mathbf{a}_2 \neq \mathbf{0}$. Having built geometric realizations $\mathbf{a}_2 = \overrightarrow{AB}$ and $\mathbf{a}_1 = \overrightarrow{AC}$, one can choose the coefficient β_2 on the base of the Figs. 6.1, 6.2, or 6.3 and on the base of the formulas (6.1), (6.2), (6.3) so that the equality (13.3) is fulfilled in this case either. As for the equality (13.3) itself, we write it as follows:

$$1 \cdot \mathbf{a}_1 + (-\beta_2) \cdot \mathbf{a}_2 = \mathbf{0}. \tag{13.5}$$

Since $1 \neq 0$, the left hand side of the equality (13.5) is a non-trivial linear combination of the vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ which is equal to zero. The existence of such a linear combination means that the vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ are linearly dependent. Thus, the converse proposition that the collinearity of two vectors implies their linear dependence is proved.

Combining the direct and converse propositions proved above, one can formulate them as a single theorem.

THEOREM 13.1. *A system of two vectors* $\mathbf{a}_1$ *and* $\mathbf{a}_2$ *is linearly dependent if and only if these vectors are collinear, i. e.* $\mathbf{a}_1 \parallel \mathbf{a}_2$.

§ 14. Linear dependence for $n = 3$. Coplanartity of vectors.

Let;s consider a system composed of three vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$. Assume that it is linearly dependent. Applying the item 3 of the theorem 11.1 to this system, we get that one of the three vectors is linearly expressed through the other two vectors. Taking into account the possibility of renumerating our vectors, we can assume that the vector $\mathbf{a}_1$ is expressed through the vectors $\mathbf{a}_2$ and $\mathbf{a}_3$:

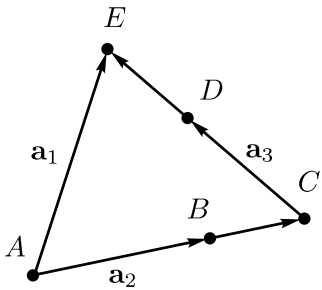

Fig. 14.1

$$\mathbf{a}_1 = \beta_2\,\mathbf{a}_2 + \beta_3\,\mathbf{a}_3. \qquad (14.1)$$

Let A be some arbitrary point of the space $\mathbb{E}$. Let's build the geometric realizations $\mathbf{a}_2 = \overrightarrow{AB}$ and $\beta_2\,\mathbf{a}_2 = \overrightarrow{AC}$. Then at the point C we build the geometric realizations of the vectors $\mathbf{a}_3 = \overrightarrow{CD}$ and $\beta_3\,\mathbf{a}_3 = \overrightarrow{CE}$. The vectors $\overrightarrow{AC}$ and $\overrightarrow{CE}$ constitute two sides of the triangle ACE (see Fig. 14.1). Then the sum of the vectors (14.1) is presented by the third side $\mathbf{a}_1 = \overrightarrow{AE}$.

The triangle ACE is a planar form. The geometric realizations of the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ lie on the sides of this triangle. Therefore they lie on the plane ACE. Instead of $\mathbf{a}_1 = \overrightarrow{AE}$, $\mathbf{a}_2 = \overrightarrow{AB}$, and $\mathbf{a}_3 = \overrightarrow{CD}$ by means of parallel translations we can build some other geometric realizations of these three vectors. These geometric realizations do not lie on the plane ACE, but they keep parallelism to this plane.

DEFINITION 14.1. Three free vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ are called *coplanar* if their geometric realizations are parallel to some plane common to all three of them.

LEMMA 14.1. *The linear dependence of three vecctors* $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ *implies their coplanarity.*

EXERCISE 14.1. *The above considerations yield a proof for the lemma 14.1 on the base of the formula* (14.1) *in the case where*

$$\mathbf{a}_2 \neq \mathbf{0}, \qquad \mathbf{a}_3 \neq \mathbf{0}, \qquad \mathbf{a}_2 \nparallel \mathbf{a}_3. \tag{14.2}$$

Consider special cases where one or several conditions (14.2) *are not fulfilled and derive the lemma 14.1 from the formula* (14.1) *in those special cases.*

LEMMA 14.2. *The coplanarity of three vectors* $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ *imply their linear dependence.*

PROOF. If $\mathbf{a}_2 = \mathbf{0}$ or $\mathbf{a}_3 = \mathbf{0}$, then the propositions of the lemma 14.2 follows from the first item of the theorem 11.1. If $\mathbf{a}_2 \neq \mathbf{0}$, $\mathbf{a}_3 \neq \mathbf{0}$, $\mathbf{a}_2 \parallel \mathbf{a}_3$, then the proposition of the lemma 14.2 follows from the theorem 13.1 and from the item 2 of the theorem 11.1. Therefore, in order to complete the proof of the lemma 14.2 we should consider the case where all of the three conditions (14.2) are fulfilled.

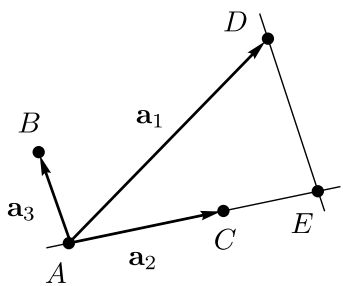

Fig. 14.2

Let A be some arbitrary point of the space $\mathbb{E}$. At this point we build the geometric realizations of the vectors $\mathbf{a}_1 = \overrightarrow{AD}$, $\mathbf{a}_2 = \overrightarrow{AC}$, and $\mathbf{a}_3 = \overrightarrow{AB}$ (see Fig. 14.2). Due to the coplanarity of the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ their geometric realizations $\overrightarrow{AB}$, $\overrightarrow{AC}$, and $\overrightarrow{AD}$ lie on a plane. Let's denote this plane through

α. Through the point D we draw a line parallel to the vector $\mathbf{a}_3 \neq \mathbf{0}$. Such a line is unique and it lies on the plane α. This line intersects the line comprising the vector $\mathbf{a}_2 = \overrightarrow{AC}$ at some unique point E since $\mathbf{a}_2 \neq \mathbf{0}$ and $\mathbf{a}_2 \nparallel \mathbf{a}_3$. Considering the points A, E, and D in Fig. 14.2, we derive the equality

$$\mathbf{a}_1 = \overrightarrow{AD} = \overrightarrow{AE} + \overrightarrow{ED}. \tag{14.3}$$

The vector $\overrightarrow{AE}$ is collinear to the vector $\overrightarrow{AC} = \mathbf{a}_2 \neq \mathbf{0}$ since these vectors lie on the same line. For this reason there is a number β_2 such that $\overrightarrow{AE} = \beta_2\,\mathbf{a}_2$. The vector $\overrightarrow{ED}$ is collinear to the vector $\overrightarrow{AB} = \mathbf{a}_3 \neq \mathbf{0}$ since these vectors lie on parallel lines. Hence $\overrightarrow{ED} = \beta_3\,\mathbf{a}_3$ for some number β_3. Upon substituting

$$\overrightarrow{AE} = \beta_2\,\mathbf{a}_2, \qquad\qquad \overrightarrow{ED} = \beta_3\,\mathbf{a}_3$$

into the equality (14.3) this equality takes the form of (14.1).

The last step in proving the lemma 14.2 consists in writing the equality (14.1) in the following form:

$$1 \cdot \mathbf{a}_1 + (-\beta_2) \cdot \mathbf{a}_2 + (-\beta_3) \cdot \mathbf{a}_3 = \mathbf{0}. \tag{14.4}$$

Since $1 \neq 0$, the left hand side of the equality (14.4) is a non-trivial linear combination of the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ which is equal to zero. The existence of such a linear combination means that the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ are linearly dependent. □

The following theorem is derived from the lemmas 14.1 and 14.2.

THEOREM 14.1. *A system of three vectors* $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ *is linearly dependent if and only if these vectors are coplanar.*

§ 15. Linear dependence for $n \geqslant 4$.

THEOREM 15.1. *Any system consisting of four vectors* $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, $\mathbf{a}_4$ *in the space* $\mathbb{E}$ *is linearly dependent.*

THEOREM 15.2. *Any system consisting of more than four vectors in the space* $\mathbb{E}$ *is linearly dependent.*

The theorem 15.2 follows from the theorem 15.1 due to the item 3 of the theorem 11.1. Therefore it is sufficient to prove the theorem 15.1. The theorem 15.1 itself expresses a property of the three-dimensional space $\mathbb{E}$.

PROOF OF THE THEOREM 15.1. Let's choose the subsystem composed by three vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ within the system of four vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, $\mathbf{a}_4$. If these three vectors are linearly

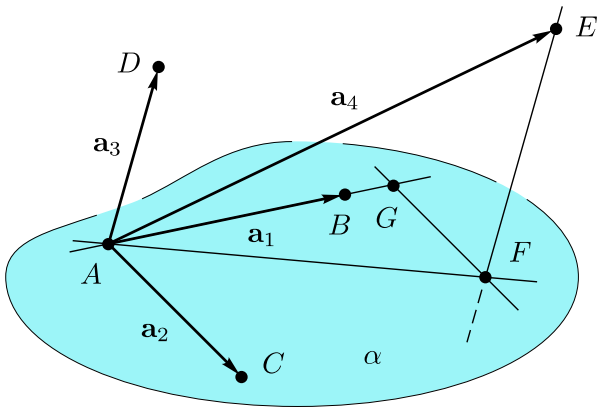

Fig. 15.1

dependent, then in order to prove the linear dependence of the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, $\mathbf{a}_4$ it is sufficient to apply the item 3 of the theorem 11.1. Therefore in what fallows we consider the case where the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ are linearly independent.

From the linear independence of the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, according to the theorem 14.1, we derive their non-coplanarity. Moreove, from the linear independence of $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ due to the item 3 of the theorem 11.1 we derive the linear independence of any smaller subsystem within the system of these three vectors. In particular, the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ are nonzero and the vectors

$\mathbf{a}_1$ and $\mathbf{a}_1$ are not collinear (see Theorems 12.1 and 13.1), i. e.

$$\mathbf{a}_1 \neq \mathbf{0}, \qquad \mathbf{a}_2 \neq \mathbf{0}, \qquad \mathbf{a}_3 \neq \mathbf{0}, \qquad \mathbf{a}_1 \nparallel \mathbf{a}_2. \tag{15.1}$$

Let A be some arbitrary point of the space $\mathbb{E}$. Let's build the geometric realizations $\mathbf{a}_1 = \overrightarrow{AB}$, $\mathbf{a}_2 = \overrightarrow{AC}$, $\mathbf{a}_3 = \overrightarrow{AD}$, $\mathbf{a}_4 = \overrightarrow{AE}$ with the initial point A. Due to the condition $\mathbf{a}_1 \nparallel \mathbf{a}_2$ in (15.1) the vectors $\overrightarrow{AB}$ and $\overrightarrow{AC}$ define a plane (see Fig. 15.1). Let's denothe this plane through α. The vector $\overrightarrow{AD}$ does not lie on the plane α and it is not parallel to this plane since the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ are not coplanar.

Let's draw a line passing through the terminal point of the vector $\mathbf{a}_4 = \overrightarrow{AE}$ and being parallel to the vector $\mathbf{a}_3$. Since $\mathbf{a}_3 \nparallel \alpha$, this line crosses the plane α at some unique point F and we have

$$\mathbf{a}_4 = \overrightarrow{AE} = \overrightarrow{AF} + \overrightarrow{FE}. \tag{15.2}$$

Now let's draw a line passing through the point F and being parallel to the vector $\mathbf{a}_2$. Such a line lies on the plane α. Due to $\mathbf{a}_1 \nparallel \mathbf{a}_2$ this line intersects the line AB at some unique point G. Hence we have the following equality:

$$\overrightarrow{AF} = \overrightarrow{AG} + \overrightarrow{GF}. \tag{15.3}$$

Note that the vector $\overrightarrow{AG}$ lies on the same line as the vector $\mathbf{a}_1 = \overrightarrow{AB}$. From (15.1) we get $\mathbf{a}_1 \neq \mathbf{0}$. Hence there is a number β_1 such that $\overrightarrow{AG} = \beta_1\,\mathbf{a}_1$. Similarly, from $\overrightarrow{GF} \parallel \mathbf{a}_2$ and $\mathbf{a}_2 \neq \mathbf{0}$ we derive $\overrightarrow{GF} = \beta_2\,\mathbf{a}_2$ for some number β_2 and from $\overrightarrow{FE} \parallel \mathbf{a}_3$ and $a_3 \neq \mathbf{0}$ we derive that $\overrightarrow{FE} = \beta_3\,\mathbf{a}_3$ for some number β_3. The rest is to substitute the obtained expressions for $\overrightarrow{AG}$, $\overrightarrow{GF}$, and $\overrightarrow{FE}$ into the formulas (15.3) and (15.2). This yields

$$\mathbf{a}_4 = \beta_1\,\mathbf{a}_1 + \beta_2\,\mathbf{a}_2 + \beta_3\,\mathbf{a}_3. \tag{15.4}$$

The equality (15.4) can be rewritten as follows:

$$1 \cdot \mathbf{a}_4 + (-\beta_1) \cdot \mathbf{a}_1 + (-\beta_2) \cdot \mathbf{a}_2 + (-\beta_3) \cdot \mathbf{a}_3 = \mathbf{0}. \tag{15.5}$$

Since $1 \neq 0$, the left hand side of the equality (15.5) is a non-trivial linear combination of the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, $\mathbf{a}_4$ which is equal to zero. The existence of such a linear combination means that the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, $\mathbf{a}_3$ are linearly dependent. The theorem 15.1 is proved. □

§ 16. Bases on a line.

Let a be some line in the space $\mathbb{E}$. Let's consider free vectors parallel to the line a. They have geometric realizations lying on the line a. Restricting the freedom of moving such vectors, i. e. forbidding geometric realizations outside the line a, we obtain partially free vectors lying on the line a.

Definition 16.1. A system consisting of one non-zero vector $\mathbf{e} \neq \mathbf{0}$ lying on a line a is called a *basis* on this line. The vector $\mathbf{e}$ is called the *basis vector* of this basis.

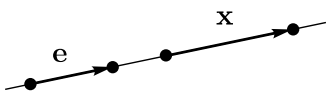

Fig. 16.1

Let $\mathbf{e}$ be the basis vector of some basis on the line a and let $\mathbf{x}$ be some other vector lying on this line (see Fig. 16.1). Then $\mathbf{x} \parallel \mathbf{e}$ and hence there is a number x such that the vector $\mathbf{x}$ is expressed through $\mathbf{a}$ by means of the formula

$$\mathbf{x} = x\,\mathbf{e}. \tag{16.1}$$

The number x in the formula (16.1) is called the *coordinate* of the vector $\mathbf{x}$ in the basis $\mathbf{e}$, while the formula (16.1) itself is called the *expansion* of the vector $\mathbf{x}$ in this basis.

When writing the coordinates of vectors extracted from their expansions (16.1) in a basis these coordinates are usually sur-

rounded with double vertical lines

$$\mathbf{x} \mapsto \|x\|. \tag{16.2}$$

Then these coordinated turn to matrices (see [7]). The mapping (16.2) implements the basic idea of analytical geometry. This idea consists in replacing geometric objects by their numeric presentations. Bases in this case are tools for such a transformation.

§ 17. Bases on a plane.

Let α be some plane in the space $\mathbb{E}$. Let's consider free vectors parallel to the plane α. They have geometric realizations lying on the plane α. Restricting the freedom of moving such vectors, i. e. forbidding geometric realizations outside the plane α, we obtain partially free vectors lying on the plane α.

Definition 17.1. An ordered pair of two non-collinear vectors $\mathbf{e}_1$, $\mathbf{e}_2$ lying on a plane α is called a *basis* on this plane. The vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are called the *basis vectors* of this basis.

In the definition 17.1 the term "ordered system of vectors" is used. This term means a system of vectors in which some ordering of vectors is fixed: $\mathbf{e}_1$ is the first vector, $\mathbf{e}_2$ is the second vector. If we exchange the vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ and take $\mathbf{e}_2$ for the first vector, while $\mathbf{e}_1$ for the second vector, that would be another basis $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$ different from the basis $\mathbf{e}_1$, $\mathbf{e}_2$:

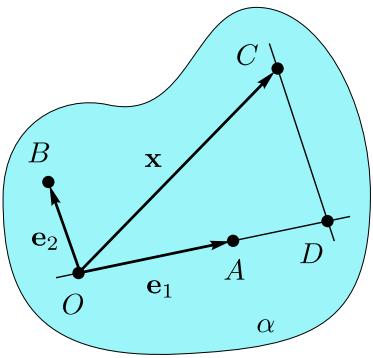

Fig. 17.1

$$\tilde{\mathbf{e}}_1 = \mathbf{e}_2, \qquad \tilde{\mathbf{e}}_2 = \mathbf{e}_1.$$

Let $\mathbf{e}_1, \mathbf{e}_2$ be a basis on a plane α and let $\mathbf{x}$ be some vector

lying on this place. Let's choose some arbitrary point $O \in \alpha$ and let's build the geometric realizations of the three vectors $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{x}$ with the initial point O:

$$\mathbf{e}_1 = \overrightarrow{OA}, \qquad \mathbf{e}_2 = \overrightarrow{OB}, \qquad \mathbf{x} = \overrightarrow{OC}. \tag{17.1}$$

Due to our choice of the vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{x}$ and due to $O \in \alpha$ the geometric realizations (17.1) lie on the plane α. Let's draw a line passing through the terminal point of the vector $\mathbf{x}$, i. e. through the point C, and being parallel to the vector $\mathbf{e}_2 = \overrightarrow{OB}$. Due to non-collinearity of the vectors $\mathbf{e}_1 \nparallel \mathbf{e}_2$ such a line intersects the line comprising the vector $\mathbf{e}_1 = \overrightarrow{OA}$ at some unique point D (see Fig. 17.1). This yields the equality

$$\overrightarrow{OC} = \overrightarrow{OD} + \overrightarrow{DC}. \tag{17.2}$$

The vector $\overrightarrow{OD}$ in (17.2) is collinear with the vector $\mathbf{e}_1 = \overrightarrow{OA}$, while the vector $\overrightarrow{DC}$ is collinear with the vector $\mathbf{e}_2 = \overrightarrow{OB}$. For this reason there are two numbers x_1 and x_2 such that

$$\overrightarrow{OD} = x_1\,\mathbf{e}_1, \qquad \overrightarrow{DC} = x_2\,\mathbf{e}_2. \tag{17.3}$$

Upon substituting (17.3) into (17.2) and taking into account the formulas (17.1) we get the equality

$$\mathbf{x} = x_1\,\mathbf{e}_1 + x_2\,\mathbf{e}_2. \tag{17.4}$$

The formula (17.4) is analogous to the formula (16.1). The numbers x_1 and x_2 are called the *coordinates* of the vector $\mathbf{x}$ in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, while the formula (17.4) itself is called the *expansion* of the vector $\mathbf{x}$ in this basis. When writing the coordinates of vectors they are usually arranged into columns and surrounded with double vertical lines

$$\mathbf{x} \mapsto \begin{Vmatrix} x_1 \\ x_2 \end{Vmatrix}. \tag{17.5}$$

The column of two numbers x_1 and x_2 in (17.5) is called the *coordinate column* of the vector $\mathbf{x}$.

§ 18. Bases in the space.

Definition 18.1. An ordered system of three non-coplanar vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ is called a *basis in the space* $\mathbb{E}$.

Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ be a basis in the space $\mathbb{E}$ and let $\mathbf{x}$ be some vector. Let's choose some arbitrary point O and build the geometric realizations of all of the four vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, and $\mathbf{x}$ with the common initial point O:

$$\begin{aligned} \mathbf{e}_1 &= \overrightarrow{OA}, & \mathbf{e}_2 &= \overrightarrow{OB}, \\ \mathbf{e}_3 &= \overrightarrow{OC}, & \mathbf{x} &= \overrightarrow{OD}. \end{aligned} \tag{18.1}$$

The vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are not collinear since otherwise the whole system of three vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ would be coplanar. For this reason the vectors $\mathbf{e}_1 = \overrightarrow{OA}$ and $\mathbf{e}_2 = \overrightarrow{OB}$ define a plane (this plane is denoted through α in Fig 18.1) and they lie on this plane. The vector $\mathbf{e}_3 = \overrightarrow{OC}$ does not lie on the plane α and it is not parallel to this plane (see Fig. 18.1).

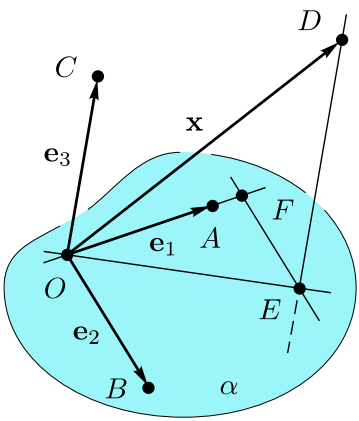

Fig. 18.1

Let's draw a line passing through the terminal point of the vector $\mathbf{x} = \overrightarrow{OD}$ and being parallel to the vector $\mathbf{e}_3$. Such a line is not parallel to the plane α since $\mathbf{e}_3 \nparallel \alpha$. It crosses the plane α at some unique point E. As a result we get the equality

$$\overrightarrow{OD} = \overrightarrow{OE} + \overrightarrow{ED}. \tag{18.2}$$

Now let's draw a line passing through the point E and being parallel to the vector $\mathbf{e}_2$. The vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ in (18.1) are not collinear. For this reason such a line crosses the line comprising the vector $\mathbf{e}_1$ at some unique point F. Considering the sequence of points O, F, E, we derive

$$\overrightarrow{OE} = \overrightarrow{OF} + \overrightarrow{FE}. \tag{18.3}$$

Combining the equalities (18.3) and (18.2), we obtain

$$\overrightarrow{OD} = \overrightarrow{OF} + \overrightarrow{FE} + \overrightarrow{ED}. \tag{18.4}$$

Note that, according to our geometric construction, the following collinearity conditions are fulfilled:

$$\overrightarrow{OF} \parallel \mathbf{e}_1, \qquad \overrightarrow{FE} \parallel \mathbf{e}_2, \qquad \overrightarrow{ED} \parallel \mathbf{e}_3. \tag{18.5}$$

From the collinearity condition $\overrightarrow{OF} \parallel \mathbf{e}_1$ in (18.5) we derive the existence of a number x_1 such that $\overrightarrow{OF} = x_1\,\mathbf{e}_1$. From the other two collinearity conditions in (18.5) we derive the existence of two other numbers x_2 and x_3 such that $\overrightarrow{FE} = x_2\,\mathbf{e}_2$ and $\overrightarrow{ED} = x_3\,\mathbf{e}_3$. As a result the equality (18.4) is written as

$$\mathbf{x} = x_1\,\mathbf{e}_1 + x_2\,\mathbf{e}_2 + x_3\,\mathbf{e}_3. \tag{18.6}$$

The formula (18.6) is analogous to the formulas (16.1) and (17.4). The numbers x_1, x_2, x_3 are called the *coordinates* of the vector $\mathbf{x}$ in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, while the formula (18.6) itself is called the *expansion* of the vector $\mathbf{x}$ in this basis. In writing the coordinates of a vector they are usually arranged into a column surrounded with two double vertical lines:

$$\mathbf{x} \mapsto \left\| \begin{matrix} x_1 \\ x_2 \\ x_3 \end{matrix} \right\|. \tag{18.7}$$

The column of the numbers x_1, x_2, x_3 in (18.7) is called the *coordinate column* of a vector $\mathbf{x}$.

§ 19. Uniqueness of the expansion of a vector in a basis.

Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ be some basis in the space $\mathbb{E}$. The geometric construction shown in Fig. 18.1 can be applied to an arbitrary vector $\mathbf{x}$. It yields an expansion of (18.6) of this vector in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. However, this construction is not a unique way foe expanding a vector in a basis. For example, instead of the plane OAB the plane OAC can be taken for α, while the line can be directed parallel to the vector $\mathbf{e}_2$. The the line EF is directed parallel to the vector $\mathbf{e}_3$ and the point F is obtained in its intersection with the line OA comprising the geometric realization of the vector $\mathbf{e}_1$. Such a construction potentially could yield some other expansion of a vector $\mathbf{x}$ in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, i. e. an expansion with the coefficients different from those of (18.6). The fact that actually this does not happen should certainly be proved.

THEOREM 19.1. *The expansion of an arbitrary vector* $\mathbf{x}$ *in a given basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ *is unique.*

PROOF. The proof is by contradiction. Assume that the expansion (18.6) is not unique and we have some other expansion of the vector $\mathbf{x}$ in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$:

$$\mathbf{x} = \tilde{x}_1\,\mathbf{e}_1 + \tilde{x}_2\,\mathbf{e}_2 + \tilde{x}_3\,\mathbf{e}_3. \tag{19.1}$$

Let's subtract th expansion (18.6) from the expansion (19.1). Upon collecting similar term we get

$$(\tilde{x}_1 - x_1)\,\mathbf{e}_1 + (\tilde{x}_2 - x_2)\,\mathbf{e}_2 + (\tilde{x}_3 - x_3)\,\mathbf{e}_3 = \mathbf{0}. \tag{19.2}$$

According to the definition 18.1, the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ is a triple of non-coplanar vectors. Due to the theorem 14.1 such a triple of vectors is linearly independent. Looking at (19.2), we see that

there we have a linear combination of the basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ which is equal to zero. Applying the theorem 10.1, we conclude that this linear combination is trivial, i. e.

$$\tilde{x}_1 - x_1 = 0, \qquad \tilde{x}_2 - x_2 = 0, \qquad \tilde{x}_3 - x_3 = 0. \tag{19.3}$$

The equalities (19.3) mean that the coefficients in the expansions (19.1) and (18.6) do coincide, which contradicts the assumption that these expansions are different. The contradiction obtained proves the theorem 19.1. □

Exercise 19.1. *By analogy to the theorem 19.1 formulate and prove uniqueness theorems for expansions of vectors in bases on a plane and in bases on a line.*

§ 20. Index setting convention.

The theorem 19.1 on the uniqueness of an expansion of vector in a basis means that the mapping (18.7), which associates vectors with their coordinates in some fixed basis, is a bijective mapping. This makes bases an important tool for quantitative description of geometric objects. This tool was improved substantially when a special index setting convention was admitted. This convention is known as Einstein's tensorial notation.

Definition 20.1. The *index setting convention*, which is also known as *Einstein's tensorial notation*, is a set of rules for placing indices in writing components of numeric arrays representing various geometric objects upon choosing some basis or some coordinate system.

Einstein's tensorial notation is not a closed set of rules. When new types of geometric objects are designed in science, new rules are added. For this reason below I formulate the index setting rules as they are needed.

Definition 20.2. Basis vectors in a basis are enumerated by lower indices, while the coordinates of vectors expanded in a basis are enumerated by upper indices.

The rule formulated in the definition 20.2 belongs to Einstein's tensorial notation. According to this rule the formula (18.6) should be rewritten as

$$\mathbf{x} = x^1\,\mathbf{e}_1 + x^2\,\mathbf{e}_2 + x^3\,\mathbf{e}_3 = \sum_{i=1}^{3} x^i\,\mathbf{e}_i, \tag{20.1}$$

whire the mapping (18.7) should be written in the following way:

$$\mathbf{x} \mapsto \begin{Vmatrix} x^1 \\ x^2 \\ x^3 \end{Vmatrix}. \tag{20.2}$$

Exercise 20.1. *The rule from the definition 20.2 is applied for bases on a line and for bases on a plane. Relying on this rule rewrite the formulas* (16.1), (16.2), (17.4), *and* (17.5).

§ 21. Usage of the coordinates of vectors.

Vectors can be used in solving various geometric problems, where the basic algebraic operations with them are performed. These are the operation of vector addition and the operation of multiplication of vectors by numbers. The usage of bases and coordinates of vectors in bases relies on the following theorem.

Theorem 21.1. *Let some basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ *in the space* $\mathbb{E}$ *be chosen and fixed. In this situation when adding vectors their coordinates are added, while when multiplying a vector by a number its coordinates are multiplied by this number, i. e. if*

$$\mathbf{x} \mapsto \begin{Vmatrix} x^1 \\ x^2 \\ x^3 \end{Vmatrix}, \qquad \mathbf{y} \mapsto \begin{Vmatrix} y^1 \\ y^2 \\ y^3 \end{Vmatrix}, \tag{21.1}$$

then for $\mathbf{x}+\mathbf{y}$ *and* $\alpha\cdot\mathbf{x}$ *we have the relationships*

$$\mathbf{x}+\mathbf{y}\mapsto\begin{Vmatrix}x^1+y^1\\x^2+y^2\\x^3+y^3\end{Vmatrix},\qquad\qquad \alpha\cdot\mathbf{x}\mapsto\begin{Vmatrix}\alpha\,x^1\\\alpha\,x^2\\\alpha\,x^3\end{Vmatrix}. \tag{21.2}$$

Exercise 21.1. *Prove the theorem 21.1, using the formulas (21.1) and (21.2) for this purpose, and prove the theorem 19.1.*

Exercise 21.2. *Formulate and prove theorems analogous to the theorem 21.1 in the case of bases on a line and on a plane.*

§ 22. Changing a basis. Transition formulas and transition matrices.

When solving geometric problems sometimes one needs to change a basis replacing it with another basis. Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ and $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$ be two bases in the space $\mathbb{E}$. If the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ is replaced by the basis $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$, then $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ is usually called the *old basis*, while $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$ is called the *new basis*. The procedure of changing an old basis for a new one can be understood as a *transition* from an old basis to a new basis or, in other words, as a *direct transition*. Conversely, changing a new basis for an old one is understood as an *inverse transition*.

Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ and $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$ be two bases in the space $\mathbb{E}$, where $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ is an old basis and $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$ is a new basis. In the direct transition procedure vectors of a new basis are expanded in an old basis, i. e. we have

$$\begin{aligned}\tilde{\mathbf{e}}_1&=S_1^1\,\mathbf{e}_1+S_1^2\,\mathbf{e}_2+S_1^3\,\mathbf{e}_3,\\ \tilde{\mathbf{e}}_2&=S_2^1\,\mathbf{e}_1+S_2^2\,\mathbf{e}_2+S_2^3\,\mathbf{e}_3,\\ \tilde{\mathbf{e}}_3&=S_3^1\,\mathbf{e}_1+S_3^2\,\mathbf{e}_2+S_3^3\,\mathbf{e}_3.\end{aligned} \tag{22.1}$$

The formulas (22.1) are called the *direct transition formulas*. The numeric coefficients S_1^1, S_1^2, S_1^3 in (22.1) are the coordinates

of the vector $\tilde{\mathbf{e}}_1$ expanded in the old basis. According to the definition 20.2, they are enumerated by an upper index. The lower index 1 of them is the number of the vector $\tilde{\mathbf{e}}_1$ of which they are the coordinates. It is used in order to distinguish the coordinates of the vector $\tilde{\mathbf{e}}_1$ from the coordinates of $\tilde{\mathbf{e}}_2$ and $\tilde{\mathbf{e}}_3$ in the second and in the third formulas (22.1).

Let's apply the first mapping (20.2) to the transition formulas and write the coordinates of the vectors $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$ as columns:

$$\tilde{\mathbf{e}}_1 \mapsto \begin{Vmatrix} S_1^1 \\ S_1^2 \\ S_1^3 \end{Vmatrix}, \qquad \tilde{\mathbf{e}}_2 \mapsto \begin{Vmatrix} S_2^1 \\ S_2^2 \\ S_2^3 \end{Vmatrix}, \qquad \tilde{\mathbf{e}}_3 \mapsto \begin{Vmatrix} S_3^1 \\ S_3^2 \\ S_3^3 \end{Vmatrix}. \tag{22.2}$$

The columns (22.2) are usually glued into a single matrix. Such a matrix is naturally denoted through S:

$$S = \begin{Vmatrix} S_1^1 & S_2^1 & S_3^1 \\ S_1^2 & S_2^2 & S_3^2 \\ S_1^3 & S_2^3 & S_3^3 \end{Vmatrix}. \tag{22.3}$$

Definition 22.1. The matrix (22.3) whose components are determined by the direct transition formulas (22.1) is called the *direct transition matrix*.

Note that the components of the direct transition matrix S_j^i are enumerated by two indices one of which is an upper index, while the other is a lower index. These indices define the position of the element S_j^i in the matrix S: the upper index i is a row number, while the lower index j is a column number. This notation is a part of a general rule.

Definition 22.2. If elements of a double index array are enumerated by indices on different levels, then in composing a matrix of these elements the upper index is used as a row number, while the lower index is used a column number.

Definition 22.3. If elements of a double index array are enumerated by indices on the same level, then in composing a matrix of these elements the first index is used as a row number, while the second index is used a column number.

The definitions 22.2 and 22.3 can be considered as a part of the index setting convention from the definition 20.1, though formally they are not since they do not define the positions of array indices, but the way to visualize the array as a matrix.

The direct transition formulas (22.1) can be written in a concise form using the summation sign for this purpose:

$$\tilde{\mathbf{e}}_j = \sum_{i=1}^{3} S_j^i \, \mathbf{e}_i, \quad \text{where} \quad j = 1,\, 2,\, 3. \tag{22.4}$$

There is another way to write the formulas (22.1) concisely. It is based on the matrix multiplication (see [7]):

$$\| \, \tilde{\mathbf{e}}_1 \quad \tilde{\mathbf{e}}_2 \quad \tilde{\mathbf{e}}_3 \, \| = \| \, \mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3 \, \| \cdot \begin{Vmatrix} S_1^1 & S_2^1 & S_3^1 \\ S_1^2 & S_2^2 & S_3^2 \\ S_1^3 & S_2^3 & S_3^3 \end{Vmatrix}. \tag{22.5}$$

Note that the basis vectors $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$ and $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ in the formula (22.5) are written in rows. This fact is an instance of a general rule.

Definition 22.4. If elements of a single index array are enumerated by lower indices, then in matrix presentation they are written in a row, i. e. they constitute a matrix whose height is equal to unity.

Definition 22.5. If elements of a single index array are enumerated by upper indices, then in matrix presentation they are written in a column, i. e. they constitute a matrix whose width is equal to unity.

Note that writing the components of a vector $\mathbf{x}$ as a column in the formula (20.2) is concordant with the rule from the definition 22.5, while the formula (18.7) violates two rules at once — the rule from the definition 20.2 and the rule from the definition 22.5.

Now let's consider the inverse transition from the new basis $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$ to the old basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. In the inverse transition procedure the vectors of an old basis are expanded in a new basis:

$$\begin{aligned}\mathbf{e}_1 &= T_1^1\,\tilde{\mathbf{e}}_1 + T_1^2\,\tilde{\mathbf{e}}_2 + T_1^3\,\tilde{\mathbf{e}}_3,\\ \mathbf{e}_2 &= T_2^1\,\tilde{\mathbf{e}}_1 + T_2^2\,\tilde{\mathbf{e}}_2 + T_2^3\,\tilde{\mathbf{e}}_3,\\ \mathbf{e}_3 &= T_3^1\,\tilde{\mathbf{e}}_1 + T_3^2\,\tilde{\mathbf{e}}_2 + T_3^3\,\tilde{\mathbf{e}}_3.\end{aligned} \tag{22.6}$$

The formulas (22.6) are called the *inverse transition formulas*. The numeric coefficients in the formulas (22.6) are coordinates of the vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ in their expansions in the new basis. These coefficients are arranged into columns:

$$\mathbf{e}_1 \mapsto \begin{Vmatrix} T_1^1\\ T_1^2\\ T_1^3 \end{Vmatrix}, \qquad \mathbf{e}_2 \mapsto \begin{Vmatrix} T_2^1\\ T_2^2\\ T_2^3 \end{Vmatrix}, \qquad \mathbf{e}_3 \mapsto \begin{Vmatrix} T_3^1\\ T_3^2\\ T_3^3 \end{Vmatrix}. \tag{22.7}$$

Then the columns (22.7) are united into a matrix:

$$T = \begin{Vmatrix} T_1^1 & T_2^1 & T_3^1\\ T_1^2 & T_2^2 & T_3^2\\ T_1^3 & T_2^3 & T_3^3 \end{Vmatrix}. \tag{22.8}$$

Definition 22.6. The matrix (22.8) whose components are determined by the inverse transition formulas (22.6) is called the *inverse transition matrix*.

The inverse transition formulas (22.6) have a concise form, analogous to the formula (22.4):

$$\mathbf{e}_j = \sum_{i=1}^{3} T_j^i\,\tilde{\mathbf{e}}_i, \quad \text{where} \quad j = 1,\,2,\,3. \tag{22.9}$$

There is also a matrix form of the formulas (22.6):

$$\| \mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3 \| = \| \tilde{\mathbf{e}}_1 \quad \tilde{\mathbf{e}}_2 \quad \tilde{\mathbf{e}}_3 \| \cdot \begin{Vmatrix} T_1^1 & T_2^1 & T_3^1 \\ T_1^2 & T_2^2 & T_3^2 \\ T_1^3 & T_2^3 & T_3^3 \end{Vmatrix}. \tag{22.10}$$

The formula (22.10) is analogous to the formula (22.5).

EXERCISE 20.1. *By analogy to* (22.1) *and* (22.6) *write the transition formulas for bases on a plane and for bases on a line (see Definitions 16.1 and 17.1). Write also the concise and matrix versions of these formulas.*

§ 23. Some information on transition matrices.

THEOREM 23.1. *The matrices S and T whose components are determined by the transition formulas* (22.1) *and* (22.6) *are inverse to each other, i. e. $T = S^{-1}$ and $S = T^{-1}$.*

I do not prove the theorem 23.1 in this book. The reader can find this theorem and its proof in [1].

The relationships $T = S^{-1}$ and $S = T^{-1}$ from the theorem 23.1 mean that the product of S by T and the product of T by S both are equal to the unit matrix (see [7]):

$$S \cdot T = 1 \qquad\qquad T \cdot S = 1. \tag{23.1}$$

Let's recall that the unit matrix is a square $n \times n$ matrix that has ones on the main diagonal and zeros in all other positions. Such a matrix is often denoted through the same symbol 1 as the numeric unity. Therefore we can write

$$1 = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}. \tag{23.2}$$

In order to denote the components of the unit matrix (23.2) the symbol δ is used. The indices enumerating rows and columns can be placed either on the same level or on different levels:

$$\delta^{ij} = \delta^i_j = \delta_{ij} = \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j. \end{cases} \tag{23.3}$$

Definition 23.1. The double index numeric array δ determined by the formula (23.3) is called the *Kronecker symbol*[1] or the *Kronecker delta*.

The positions of indices in the Kronecker symbol are determined by a context where it is used. For example, the relationships (23.1) can be written in components. In this particular case the indices of the Kronecker symbol are placed on different levels:

$$\sum_{j=1}^{3} S^i_j\, T^j_k = \delta^i_k, \qquad\qquad \sum_{j=1}^{3} T^i_j\, S^j_k = \delta^i_k. \tag{23.4}$$

Such a placement of the indices in the Kronecker symbol in (23.4) is inherited from the transition matrices S and T.

Noter that the transition matrices S and T are square matrices. For such matrices the concept of the *determinant* is introduced (see [7]). This is a number calculated through the components of a matrix according to some special formulas. In the case of the unit matrix (23.2) these formulas yield

$$\det 1 = 1. \tag{23.5}$$

The following fact is also well known. Its proof can be found in the book [7].

Theorem 23.2. *The determinant of a product of matrices is equal to the product of their determinants.*

[1] Don't mix with the Kronecker symbol used in number theory (see [8]).

Let's apply the theorem 23.2 and the formula (23.5) to the products of the matrices S and T in (23.1). This yields

$$\det S \cdot \det T = 1. \tag{23.6}$$

DEFINITION 23.2. A matrix with zero determinant is called *degenerate*. If the determinant of a matrix is nonzero, such a matrix is called *non-degenerate*.

From the formula (23.6) and the definition 23.2 we immediately derive the following theorem.

THEOREM 23.3. *For any two bases in the space* $\mathbb{E}$ *the corresponding transition matrices* S *and* T *are non-degenerate and the product of their determinants is equal to the unity.*

THEOREM 23.4. *Each non-degenerate* 3×3 *matrix* S *is a transition matrix relating some basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ *in the space* $\mathbb{E}$ *with some other basis* $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$ *in this space.*

The theorem 23.4 is a strengthened version of the theorem 23.3. Its proof can be found in the book [1].

EXERCISE 23.1. *Formulate theorems analogous to the theorems 23.1, 23.2, and 23.4 in the case of bases on a plane and in the case of bases on a line.*

§ 24. Index setting in sums.

As we have already seen, in dealing with coordinates of vectors formulas with sums arise. It is convenient to write these sums in a concise form using the summation sign. Writing sums in this way one should follow some rules, which are listed below as definitions.

DEFINITION 24.1. Each summation sign in a formula has its scope. This scope begins immediately after the summation sign to the right of it and ranges up to some delimiter:

1) the end of the formula;

2) the equality sign;
3) the plus sign «+» or the minus sign «−» not enclosed into brackets opened after a summation sign in question;
4) the closing bracket whose opening bracket precedes the summation sign in question.

Let's recall that a summation sign is present in a formula and if some variable is used as a cycling variable in this summation sign, such a variable is called a *summation index* (see Formula (8.3) and the comment to it).

Definition 24.2. Each summation index can be used only within the scope of the corresponding summation sign.

Apart from simple sums, multiple sums can be used in formulas. They obey the following rule.

Definition 24.3. A variable cannot be used as a summation index in more than one summation signs of a multiple sum.

Definition 24.4. Variables which are not summation indices are called *free variables*.

Summation indices as well as free variables can be used as indices enumerating basis vectors and array components. The following terminology goes along with this usage.

Definition 24.5. A free variable which is used as an index is called a *free index*.

In the definitions 24.1, 24.2 and 24.3 the commonly admitted rules are listed. Apart from them there are more special rules which are used within the framework of Einstein's tensorial notation (see Definition 20.1).

Definition 24.6. If an expression is a simple sum or a multiple sum and if each summand of it does not comprise other sums, then each free index should have exactly one entry in this expression, while each summation index should enter twice — once as an upper index and once as a lower index.

Definition 24.7. The expression built according to the definition 24.6 can be used for composing sums with numeric coefficients. Then all summands in such sums should have the same set of free indices and each free index should be on the same level (upper or lower) in all summands. Regardless to the number of summands, in counting the number of entries to the whole sum each free index is assumed to be entering only once. The level of a free index in the sum (upper or lower) is determined by its level in each summand.

Lets consider some expressions as examples:

$$\sum_{i=1}^{3} a^i\,b_i, \qquad \sum_{i=1}^{3} a^i\,g_{ij}, \qquad \sum_{i=1}^{3}\sum_{k=1}^{3} a^i\,b^k\,g_{ik}, \tag{24.1}$$

$$\sum_{k=1}^{3} 2\,A^i_k\,b^m\,u^k\,v_q + \sum_{j=1}^{3} 3\,C^{ijm}_{jq} - \sum_{r=1}^{3}\sum_{s=1}^{3} C^r_{qs}\,B^{ism}\,v_r. \tag{24.2}$$

Exercise 24.1. *Verify that each expression in* (24.1) *satisfies the definition 24.6, while the expression* (24.2) *satisfies the definition 24.7.*

Definition 24.8. Sums composed according to the definition 24.7 can enter as subexpressions into simple and multiple sums which will be external sums with respect to them. Then some of their free indices or all of their free indices can turn into summation indices. Those of free indices that remain free are included into the list of free indices of the whole expression.

In counting the number of entries of an index in a sum included into an external simple or multiple sum the rule from the definition 24.7 is applied. Taking into account this rue, each free index of the ultimate expression should enter it exactly once, while each summation index should enter it exactly twice — once as an upper index and once as a lower index. In counting the number of entries of an index in a sum included into an

external simple or multiple sum the rule from the definition 24.7 is applied. Taking into account this rue, each free index of the ultimate expression should enter it exactly once, while each summation index should enter it exactly twice — once as an upper index and once as a lower index. In counting the number of entries of an index in a sum included into an external simple or multiple sum the rule from the definition 24.7 is applied. Taking into account this rue, each free index of the ultimate expression should enter it exactly once, while each summation index should enter it exactly twice — once as an upper index and once as a lower index. In counting the number of entries of an index in a sum included into an external simple or multiple sum the rule from the definition 24.7 is applied. Taking into account this rue, each free index of the ultimate expression should enter it exactly once, while each summation index should enter it exactly twice — once as an upper index and once as a lower index.

As an example we consider the following expression which comprises inner and outer sums:

$$\sum_{k=1}^{3} A_k^i \left(B_q^k + \sum_{i=1}^{3} C_{iq}^k \, u^i \right). \tag{24.3}$$

Exercise 24.2. *Make sure that the expression* (24.3) *satisfies the definition 24.8.*

Exercise 24.3. *Open the brackets in* (24.3) *and verify that the resulting expression satisfies the definition 24.7.*

The expressions built according to the definitions 24.6, 24.7, and 24.8 can be used for composing equalities. In composing equalities the following rule is applied.

Definition 24.9. Both sides of an equality should have the same set of free indices and each free index should have the same position (upper or lower) in both sides of an equality.

§ 25. Transformation of the coordinates of vectors under a change of a basis.

Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ and $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$ be two bases in the space $\mathbb{E}$ and let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ be changed for the basis $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$. As we already mentioned, in this case the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ is called an old basis, while $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$ is called a new basis.

Let's consider some arbitrary vector $\mathbf{x}$ in the space $\mathbb{E}$. Expanding this vector in the old basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ and in the new basis $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$, we get two sets of its coordinates:

$$\mathbf{x} \mapsto \begin{Vmatrix} x^1 \\ x^2 \\ x^3 \end{Vmatrix}, \qquad \mathbf{x} \mapsto \begin{Vmatrix} \tilde{x}^1 \\ \tilde{x}^2 \\ \tilde{x}^3 \end{Vmatrix}. \tag{25.1}$$

Both mappings (25.1) are bijective. For this reason there is a bijective correspondence between two sets of numbers x^1, x^2, x^3 and $\tilde{x}^1$, $\tilde{x}^2$, $\tilde{x}^3$. In order to get explicit formulas expressing the coordinates of the vector $\mathbf{x}$ in the new basis through its coordinates in the old basis we use the expansion (20.1):

$$\mathbf{x} = \sum_{j=1}^{3} x^j \, \mathbf{e}_j. \tag{25.2}$$

Let's apply the inverse transition formula (22.9) in order to express the vector $\mathbf{e}_j$ in (25.2) through the vectors of the new basis $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$. Upon substituting (22.9) into (25.2), we get

$$\begin{aligned} \mathbf{x} &= \sum_{j=1}^{3} x^j \left(\sum_{i=1}^{3} T_j^i \, \tilde{\mathbf{e}}_i \right) = \\ &= \sum_{j=1}^{3} \sum_{i=1}^{3} x^j \, T_j^i \, \tilde{\mathbf{e}}_i = \sum_{i=1}^{3} \left(\sum_{j=1}^{3} T_j^i \, x^j \right) \tilde{\mathbf{e}}_i. \end{aligned} \tag{25.3}$$

The formula (25.3) expresses the vector $\mathbf{x}$ as a linear combination of the basis vectors $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$, i. e. it is an expansion of the vector

$\mathbf{x}$ in the new basis. Due to the uniqueness of the expansion of a vector in a basis (see Theorem 19.1) the coefficients of such an expansion should coincide with the coordinates of the vector $\mathbf{x}$ in the new basis $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$:

$$\tilde{x}^i = \sum_{j=1}^{3} T_j^i \, x^j, \quad \text{there} \quad i = 1, 2, 3. \tag{25.4}$$

The formulas (25.4) expressing the coordinates of an arbitrary vector $\mathbf{x}$ in a new basis through its coordinates in an old basis are called the *direct transformation formulas*. Accordingly, the formulas expressing the coordinates of an arbitrary vector $\mathbf{x}$ in an old basis through its coordinates in a new basis are called the *inverse transformation formulas*. The latter formulas need not be derived separately. It is sufficient to move the tilde sign from the left hand side of the formulas (25.4) to their right hand side and replace T_j^i with S_j^i. This yields

$$x^i = \sum_{j=1}^{3} S_j^i \, \tilde{x}^j, \quad \text{where} \quad i = 1, 2, 3. \tag{25.5}$$

The direct transformation formulas (25.4) have the expanded form where the summation is performed explicitly:

$$\begin{aligned} \tilde{x}^1 &= T_1^1 \, x^1 + T_2^1 \, x^2 + T_3^1 \, x^3, \\ \tilde{x}^2 &= T_1^2 \, x^1 + T_2^2 \, x^2 + T_3^2 \, x^3, \\ \tilde{x}^3 &= T_1^3 \, x^1 + T_2^3 \, x^2 + T_3^3 \, x^3. \end{aligned} \tag{25.6}$$

The same is true for the inverse transformation formulas (25.5):

$$\begin{aligned} x^1 &= S_1^1 \, \tilde{x}^1 + S_2^1 \, \tilde{x}^2 + S_3^1 \, \tilde{x}^3, \\ x^2 &= S_1^2 \, \tilde{x}^1 + S_2^2 \, \tilde{x}^2 + S_3^2 \, \tilde{x}^3, \\ x^3 &= S_1^3 \, \tilde{x}^1 + S_2^3 \, \tilde{x}^2 + S_3^3 \, \tilde{x}^3. \end{aligned} \tag{25.7}$$

Along with (25.6), there is the matrix form of the formulas (25.4):

$$\begin{Vmatrix} \tilde{x}^1 \\ \tilde{x}^2 \\ \tilde{x}^3 \end{Vmatrix} = \begin{Vmatrix} T_1^1 & T_2^1 & T_3^1 \\ T_1^2 & T_2^2 & T_3^2 \\ T_1^3 & T_2^3 & T_3^3 \end{Vmatrix} \cdot \begin{Vmatrix} x^1 \\ x^2 \\ x^3 \end{Vmatrix} \tag{25.8}$$

Similarly, the inverse transformation formulas (25.5), along with the expanded form (25.7), have the matrix form either

$$\begin{Vmatrix} x^1 \\ x^2 \\ x^3 \end{Vmatrix} = \begin{Vmatrix} T_1^1 & T_2^1 & T_3^1 \\ T_1^2 & T_2^2 & T_3^2 \\ T_1^3 & T_2^3 & T_3^3 \end{Vmatrix} \cdot \begin{Vmatrix} \tilde{x}^1 \\ \tilde{x}^2 \\ \tilde{x}^3 \end{Vmatrix} \tag{25.9}$$

EXERCISE 25.1. *Write the analogs of the transformation formulas* (25.4), (25.5), (25.6), (25.7), (25.8), (25.9) *for vectors on a plane and for vectors on a line expanded in corresponding bases on a plane and on a line.*

§ 26. Scalar product.

Let $\mathbf{a}$ and $\mathbf{b}$ be two nonzero free vectors. Let's build their geometric realizations $\mathbf{a} = \overrightarrow{OA}$ and $\mathbf{b} = \overrightarrow{OB}$ at some arbitrary point O. The smaller of two angles formed by the rays $[OA)$ and $[OB)$ at the point O is called the *angle between vectors* $\overrightarrow{OA}$ and $\overrightarrow{OB}$. In Fig. 26.1 this angle is denoted through φ. The value of the angle φ ranges from 0 to π:

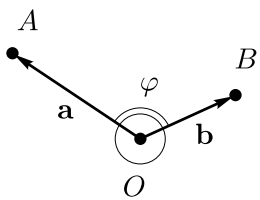

Fig. 26.1

$$0 \leqslant \varphi \leqslant \pi.$$

The lengths of the vectors $\overrightarrow{OA}$ and $\overrightarrow{OB}$ do not depend on the choice of a point O (see Definitions 3.1 and 4.2). The same is true for the angle between them. Therefore

we can deal with the lengths of the free vectors $\mathbf{a}$ and $\mathbf{b}$ and with the angle between them:

$$|\mathbf{a}| = |OA|, \qquad |\mathbf{b}| = |OB|, \qquad \widehat{\mathbf{ab}} = \widehat{AOB} = \varphi.$$

In the case where $\mathbf{a} = \mathbf{0}$ or $\mathbf{b} = \mathbf{0}$ the lengths of the vectors $\mathbf{a}$ and $\mathbf{b}$ are defined, but the angle between these vectors is not defined.

DEFINITION 26.1. The *scalar product* of two nonzero vectors $\mathbf{a}$ and $\mathbf{b}$ is a number equal to the product of their lengths and the cosine of the angle between them:

$$(\mathbf{a}, \mathbf{b}) = |\mathbf{a}|\,|\mathbf{b}|\,\cos\varphi. \tag{26.1}$$

In the case where $\mathbf{a} = \mathbf{0}$ or $\mathbf{b} = \mathbf{0}$ the scalar product $(\mathbf{a}, \mathbf{b})$ is assumed to be equal to zero by definition.

A comma is the multiplication sign in the writing the scalar product, not by itself, but together with round brackets surrounding the whole expression. These brackets are natural delimiters for multiplicands: the first multiplicand is an expression between the opening bracket and the comma, while the second multiplicand is an expression between the comma and the closing bracket. Therefore in complicated expressions no auxiliary delimiters are required. For example, in the formula

$$(\mathbf{a} + \mathbf{b}, \mathbf{c} + \mathbf{d})$$

the sums $\mathbf{a} + \mathbf{b}$ and $\mathbf{c} + \mathbf{d}$ are calculated first, then the scalar multiplication is performed.

A remark. Often the scalar product is written as $\mathbf{a} \cdot \mathbf{b}$. Even the special term «dot product» is used. However, to my mind, this notation is not good. It is misleading since the dot sign is used for denoting the product of a vector and a number and for denoting the product of two numbers.

§ 27. Orthogonal projection onto a line.

Let $\mathbf{a}$ and $\mathbf{b}$ be two free vectors such that $\mathbf{a} \neq \mathbf{0}$. Let's build their geometric realizations $\mathbf{a} = \overrightarrow{OA}$ and $\mathbf{b} = \overrightarrow{OB}$ at some arbitrary point O. The nonzero vector $\overrightarrow{OA}$ defines a line. Let's drop the perpendicular from the terminal point of the vector $\overrightarrow{OB}$, i. e. from the point B, to this line and let's denote through C the base of this perpendicular (see Fig. 27.1). In the special case where $\mathbf{b} \parallel \mathbf{a}$ and where the point B lies on the line OA we choose the point C coinciding with the point B.

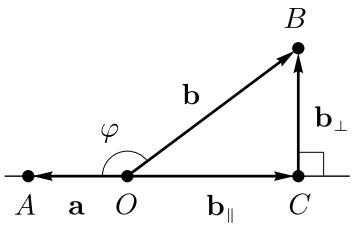

Fig. 27.1

The point C determines two vectors $\overrightarrow{OC}$ and $\overrightarrow{CB}$. The vector $\overrightarrow{OC}$ is collinear to the vector $\mathbf{a}$, while the vector $\overrightarrow{CB}$ is perpendicular to it. By means of parallel translations one can replicate the vectors $\overrightarrow{OC}$ and $\overrightarrow{CB}$ up to free vectors $\mathbf{b}_\parallel$ and $\mathbf{b}_\perp$ respectively. Note that the point C is uniquely determined by the point B and by the line OA (see Theorem 6.5 in Chapter III of the book [7]). For this reason the vectors $\mathbf{b}_\parallel$ and $\mathbf{b}_\perp$ do not depend on the choice of a point O and we can formulate the following theorem.

THEOREM 27.1. *For any nonzero vector* $\mathbf{a} \neq \mathbf{0}$ *and for any vector* $\mathbf{b}$ *there two unique vectors* $\mathbf{b}_\parallel$ *and* $\mathbf{b}_\perp$ *such that the vector* $\mathbf{b}_\parallel$ *is collinear to* $\mathbf{a}$*, the vector* $\mathbf{b}_\perp$ *is perpendicular to* $\mathbf{a}$*, and they both satisfy the equality being the expansion of the vector* $\mathbf{b}$*:*

$$\mathbf{b} = \mathbf{b}_\parallel + \mathbf{b}_\perp. \tag{27.1}$$

One should recall the special case where the point C coincides with the point B. In this case $\mathbf{b}_\perp = \mathbf{0}$ and we cannot verify visually the orthogonality of the vectors $\mathbf{b}_\perp$ and $\mathbf{a}$. In order

to extend the theorem 27.1 to this special case the following definition is introduced.

DEFINITION 27.1. All null vectors are assumed to be perpendicular to each other and each null vector is assumed to be perpendicular to any nonzero vector.

Like the definition 3.2, the definition 27.1 is formulated for geometric vectors. Upon passing to free vectors it is convenient to unite the definition 3.2 with the definition 27.1 and then formulate the following definition.

DEFINITION 27.2. A free null vector $\mathbf{0}$ of any physical nature is codirected to itself and to any other vector. A free null vector $\mathbf{0}$ of any physical nature is perpendicular to itself and to any other vector.

When taking into account the definition 27.2, the theorem 27.1 is proved by the constructions preceding it, while the expansion (27.1) follows from the evident equality

$$\overrightarrow{OB} = \overrightarrow{OC} + \overrightarrow{CB}.$$

Assume that a vector $\mathbf{a} \neq \mathbf{0}$ is fixed. In this situation the theorem 27.1 provides a mapping $\pi_{\mathbf{a}}$ that associates each vector $\mathbf{b}$ with its parallel component $\mathbf{b}_{\parallel}$.

DEFINITION 27.3. The mapping $\pi_{\mathbf{a}}$ that associates each free vector $\mathbf{b}$ with its parallel component $\mathbf{b}_{\parallel}$ in the expansion (27.1) is called the *orthogonal projection onto a line given by the vector* $\mathbf{a} \neq \mathbf{0}$ or, more exactly, the *orthogonal projection onto the direction of the vector* $\mathbf{a} \neq \mathbf{0}$.

The orthogonal projection $\pi_{\mathbf{a}}$ is closely related to the scalar product of vectors. This relation is established by the following theorem.

THEOREM 27.2. *For each nonzero vector* $\mathbf{a} \neq \mathbf{0}$ *and for any*

vector $\mathbf{b}$ *the vector* $\pi_{\mathbf{a}}(\mathbf{b})$ *is calculated by means of the formula*

$$\pi_{\mathbf{a}}(\mathbf{b}) = \frac{(\mathbf{b},\mathbf{a})}{|\mathbf{a}|^2}\,\mathbf{a}. \tag{27.2}$$

Proof. If $\mathbf{b} = \mathbf{0}$ both sides of the equality (27.2) do vanish and it is trivially fulfilled. Therefore we can assume that $\mathbf{b} \neq \mathbf{0}$.

It is easy to see that the vectors in two sides of the equality (27.2) are collinear. For the beginning let's prove that the lengths of these two vectors are equal to each other. The length of the vector $\pi_{\mathbf{a}}(\mathbf{b})$ is calculated according to Fig. 27.1:

$$|\pi_{\mathbf{a}}(\mathbf{b})| = |\mathbf{b}_{\parallel}| = |\mathbf{b}|\,|\cos\varphi|. \tag{27.3}$$

The length of the vector in the right hand side of the formula (27.2) is determined by the formula itself:

$$\left|\frac{(\mathbf{b},\mathbf{a})}{|\mathbf{a}|^2}\,\mathbf{a}\right| = \frac{|(\mathbf{b},\mathbf{a})|}{|\mathbf{a}|^2}\,|\mathbf{a}| = \frac{|\mathbf{b}|\,|\mathbf{a}|\,|\cos\varphi|}{|\mathbf{a}|} = |\mathbf{b}|\,|\cos\varphi|. \tag{27.4}$$

Comparing the results of (27.3) and (27.4), we conclude that

$$|\pi_{\mathbf{a}}(\mathbf{b})| = \left|\frac{(\mathbf{b},\mathbf{a})}{|\mathbf{a}|^2}\,\mathbf{a}\right|. \tag{27.5}$$

Due to (27.5) in order to prove the equality (27.2) it is sufficient to prove the codirectedness of vectors

$$\pi_{\mathbf{a}}(\mathbf{b}) \uparrow\uparrow \frac{(\mathbf{b},\mathbf{a})}{|\mathbf{a}|^2}\,\mathbf{a}. \tag{27.6}$$

Since $\pi_{\mathbf{a}}(\mathbf{b}) = \mathbf{b}_{\parallel}$, again applying Fig. 27.1, we consider the following three possible cases:

$$0 \leqslant \varphi < \pi/2, \qquad \varphi = \pi/2, \qquad \pi/2 < \varphi \leqslant \pi.$$

In the first case both vectors (27.6) are codirected with the vector $\mathbf{a} \neq \mathbf{0}$. Hence they are codirected with each other.

In the second case both vectors (27.6) are equal to zero. They are codirected according to the definition 3.2.

In the third case both vectors (27.6) are opposite to the vector $\mathbf{a} \neq \mathbf{0}$. Therefore they are again codirected with each other. The relationship (27.6) and the theorem 27.2 in whole are proved. □

DEFINITION 27.4. A mapping f acting from the set of all free vectors to the set of all free vectors is called a *linear mapping* if it possesses the following two properties:

1) $f(\mathbf{a}+\mathbf{b}) = f(\mathbf{a}) + f(\mathbf{b})$;
2) $f(\alpha\,\mathbf{a}) = \alpha\, f(\mathbf{a})$.

The properties 1) and 2), which should be fulfilled for any two vectors $\mathbf{a}$ and $\mathbf{b}$ and for any number α, constitute a property which is called the *linearity*.

THEOREM 27.3. *For any nonzero vector* $\mathbf{a} \neq \mathbf{0}$ *the orthogonal projection* $\pi_{\mathbf{a}}$ *onto a line given by the vector* $\mathbf{a}$ *is a linear mapping.*

In order to prove the theorem 27.3 we need the following auxiliary lemma.

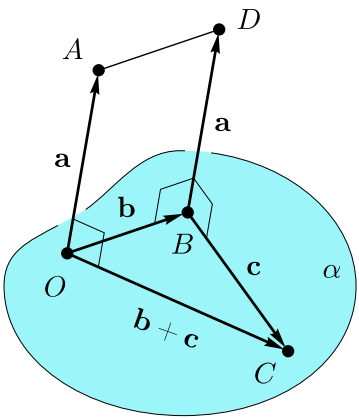

Fig. 27.2

LEMMA 27.1. *For any nonzero vector* $\mathbf{a} \neq \mathbf{0}$ *the sum of two vectors collinear to* $\mathbf{a}$ *is a vector collinear to* $\mathbf{a}$ *and the sum of two vectors perpendicular to* $\mathbf{a}$ *is a vector perpendicular to* $\mathbf{a}$.

PROOF OF THE LEMMA 27.1. The first proposition of the lemma is obvious. It follows immediately from the definition 5.1.

Let's prove the second proposition. Let $\mathbf{b}$ and $\mathbf{c}$ be two vectors, such that $\mathbf{b} \perp \mathbf{a}$ and $\mathbf{c} \perp \mathbf{a}$.

In the cases $\mathbf{b} = \mathbf{0}$, $\mathbf{c} = \mathbf{0}$, $\mathbf{b} + \mathbf{c} = \mathbf{0}$, and in the case $\mathbf{b} \parallel \mathbf{c}$ the second proposition of the lemma is also obvious. In these cases geometric realizations of all the three vectors $\mathbf{b}$, $\mathbf{c}$, and $\mathbf{b} + \mathbf{c}$ can be chosen such that they lie on the same line. Such a line is perpendicular to the vector $\mathbf{a}$.

Let's consider the case where $\mathbf{b} \nparallel \mathbf{c}$. Let's build a geometric realization of the vector $\mathbf{b} = \overrightarrow{OB}$ with initial point at some arbitrary point O. Then we lay the vector $\mathbf{c} = \overrightarrow{BC}$ at the terminal point B of the vector $\overrightarrow{OB}$. Since $\mathbf{b} \nparallel \mathbf{c}$, the points O, B, and C do not lie on a single straight line altogether. Hence they determine a plane. We denote this plane through α. The sum of vectors $\overrightarrow{OC} = \overrightarrow{OB} + \overrightarrow{BC}$ lies on this plane. It is a geometric realization for the vector $\mathbf{b} + \mathbf{c}$, i.. e. $\overrightarrow{OC} = \mathbf{b} + \mathbf{c}$.

We build two geometric realizations $\overrightarrow{OC}$ and $\overrightarrow{BD}$ for the vector $\mathbf{a}$. These are two different geometric vectors. From the equality $\overrightarrow{OA} = \overrightarrow{BD}$ we derive that the lines OA and BD are parallel. Due to $\mathbf{b} \perp \mathbf{a}$ and $\mathbf{c} \perp \mathbf{a}$ the line BD is perpendicular to the pair of crosswise intersecting lines OB and BC lying on the plane α. Hence it is perpendicular to this plane. From $BD \perp \alpha$ and $BD \parallel OA$ we derive $OA \perp \alpha$, while from $OA \perp \alpha$ we derive $\overrightarrow{OA} \perp \overrightarrow{OC}$. Hence the sum of vectors $\mathbf{b} + \mathbf{c}$ is perpendicular to the vector $\mathbf{a}$. The lemma 27.1 is proved. □

Proof of the theorem 27.3. According to the definition 27.4, in order to prove the theorem we need to verify two linearity conditions for the mapping $\pi_{\mathbf{a}}$. The first of these conditions in our particular case is written as the equality

$$\pi_{\mathbf{a}}(\mathbf{b} + \mathbf{c}) = \pi_{\mathbf{a}}(\mathbf{b}) + \pi_{\mathbf{a}}(\mathbf{c}). \tag{27.7}$$

Let's denote $\mathbf{d} = \mathbf{b} + \mathbf{c}$. According to the theorem 27.2, there are expansions of the form (27.1) for the vectors $\mathbf{b}$, $\mathbf{c}$, and $\mathbf{d}$:

$$\mathbf{b} = \mathbf{b}_{\parallel} + \mathbf{b}_{\perp}, \tag{27.8}$$

$$\mathbf{c} = \mathbf{c}_{\parallel} + \mathbf{c}_{\perp}, \tag{27.9}$$

$$\mathbf{d} = \mathbf{d}_{\parallel} + \mathbf{d}_{\perp}. \tag{27.10}$$

According the same theorem 27.2, the components of the expansions (27.8), (27.9), (27.10) are uniquely fixed by the conditions

$$\mathbf{b}_{\parallel} \parallel \mathbf{a}, \qquad \mathbf{b}_{\perp} \perp \mathbf{a}, \tag{27.11}$$

$$\mathbf{c}_{\parallel} \parallel \mathbf{a}, \qquad \mathbf{c}_{\perp} \perp \mathbf{a}, \tag{27.12}$$

$$\mathbf{d}_{\parallel} \parallel \mathbf{a}, \qquad \mathbf{d}_{\perp} \perp \mathbf{a}. \tag{27.13}$$

Adding the equalities (27.8) and (27.9), we get

$$\mathbf{d} = \mathbf{b} + \mathbf{c} = (\mathbf{b}_{\parallel} + \mathbf{c}_{\parallel}) + (\mathbf{b}_{\perp} + \mathbf{c}_{\perp}). \tag{27.14}$$

Due to (27.11) and (27.12) we can apply the lemma 27.1 to the components of the expansion (27.14). This yields

$$(\mathbf{b}_{\parallel} + \mathbf{c}_{\parallel}) \parallel \mathbf{a}, \qquad (\mathbf{b}_{\perp} + \mathbf{c}_{\perp}) \perp \mathbf{a}. \tag{27.15}$$

The rest is to compare (27.14) with (27.10) and (27.15) with (27.13). From this comparison, applying the theorem 27.2, we derive the following relationships:

$$\mathbf{d}_{\parallel} = \mathbf{b}_{\parallel} + \mathbf{c}_{\parallel}, \qquad \mathbf{d}_{\perp} = \mathbf{b}_{\perp} + \mathbf{c}_{\perp}. \tag{27.16}$$

According to the definition 27.3, the first of the above relationships (27.16) is equivalent to the equality (27.7) which was to be verified.

Let's proceed to proving the second linearity condition for the mapping $\pi_{\mathbf{a}}$. It is written as follows:

$$\pi_{\mathbf{a}}(\alpha\,\mathbf{b}) = \alpha\,\pi_{\mathbf{a}}(\mathbf{b}). \tag{27.17}$$

Let's denote $\mathbf{e} = \alpha\,\mathbf{b}$ and then, applying the theorem 27.2, write

$$\mathbf{b} = \mathbf{b}_{\parallel} + \mathbf{b}_{\perp}, \tag{27.18}$$

$$\mathbf{e} = \mathbf{e}_{\parallel} + \mathbf{e}_{\perp}. \tag{27.19}$$

According to the theorem 27.2, the components of the expansions (27.18) and (27.19) are uniquely fixed by the conditions

$$\mathbf{b}_{\parallel} \parallel \mathbf{a}, \qquad \mathbf{b}_{\perp} \perp \mathbf{a}, \tag{27.20}$$

$$\mathbf{e}_{\parallel} \parallel \mathbf{a}, \qquad \mathbf{e}_{\perp} \perp \mathbf{a}. \tag{27.21}$$

Let's multiply both sides of (27.18) by α. Then we get

$$\mathbf{e} = \alpha\mathbf{b} = \alpha\,\mathbf{b}_{\parallel} + \alpha\mathbf{b}_{\perp}. \tag{27.22}$$

Multiplying a vector by the number α, we get a vector collinear to the initial vector. For this reason from (27.20) we derive

$$(\alpha\,\mathbf{b}_{\parallel}) \parallel \mathbf{a}, \qquad (\alpha\,\mathbf{b}_{\perp}) \perp \mathbf{a}, \tag{27.23}$$

Let's compare (27.22) with (27.19) and (27.23) with (27.21). Then, applying the theorem 27.2, we obtain

$$\mathbf{e}_{\parallel} = \alpha\,\mathbf{b}_{\parallel}, \qquad \mathbf{e}_{\perp} = \alpha\,\mathbf{b}_{\perp}. \tag{27.24}$$

According to the definition 27.3 the first of the equalities (27.24) is equivalent to the required equality (27.17). The theorem 27.3 is proved. □

§ 28. Properties of the scalar product.

THEOREM 28.1. *The scalar product of vectors possesses the following four properties which are fulfilled for any three vectors* $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ *and for any number* α*:*

1) $(\mathbf{a}, \mathbf{b}) = (\mathbf{b}, \mathbf{a})$;
2) $(\mathbf{a} + \mathbf{b}, \mathbf{c}) = (\mathbf{a}, \mathbf{c}) + (\mathbf{b}, \mathbf{c})$;
3) $(\alpha\,\mathbf{a}, \mathbf{c}) = \alpha\,(\mathbf{a}, \mathbf{c})$;
4) $(\mathbf{a}, \mathbf{a}) \geqslant 0$ *and* $(\mathbf{a}, \mathbf{a}) = 0$ *implies* $\mathbf{a} = \mathbf{0}$.

DEFINITION 28.1. The property 1) in the theorem 28.1 is called the property of *symmetry*; the properties 2) and 3) are called the properties of *linearity with respect to the first multiplicand*; the property 4) is called the property of *positivity*.

PROOF OF THE THEOREM 28.1. The property of symmetry 1) is immediate from the definition 26.1 and the formula (26.1) in this definition. Indeed, if one of the vectors $\mathbf{a}$ or $\mathbf{b}$ is equal to zero, then both sides of the equality $(\mathbf{a},\mathbf{b}) = (\mathbf{b},\mathbf{a})$ equal to zero. Hence the equality is fulfilled in this case.

In the case of the nonzero vectors $\mathbf{a}$ and $\mathbf{b}$ the angle φ is determined by the pair of vectors $\mathbf{a}$ and $\mathbf{b}$ according to Fig. 26.1, it does not depend on the order of vectors in this pair. Therefore the equality $(\mathbf{a},\mathbf{b}) = (\mathbf{b},\mathbf{a})$ in this case is reduced to the equality

$$|\mathbf{a}|\,|\mathbf{b}|\,\cos\varphi = |\mathbf{b}|\,|\mathbf{a}|\,\cos\varphi.$$

It is obviously fulfilled since $|\mathbf{a}|$ and $|\mathbf{b}|$ are numbers complemented by some measure units depending on the physical nature of the vectors $\mathbf{a}$ and $\mathbf{b}$.

Let's consider the properties of linearity 2) and 3). If $\mathbf{c} = \mathbf{0}$, then both sides of the equalities $(\mathbf{a}+\mathbf{b},\mathbf{c}) = (\mathbf{a},\mathbf{c}) + (\mathbf{b},\mathbf{c})$ and $(\alpha\,\mathbf{a},\mathbf{b}) = \alpha\,(\mathbf{a},\mathbf{b})$ are equal to zero. Hence these equalities are fulfilled in this case.

If $\mathbf{c} \neq \mathbf{0}$, then we apply the theorems 27.2 and 27.3. From the theorems 27.2 we derive the following equalities:

$$\pi_{\mathbf{c}}(\mathbf{a}+\mathbf{b}) - \pi_{\mathbf{c}}(\mathbf{a}) - \pi_{\mathbf{c}}(\mathbf{b}) = \frac{(\mathbf{a}+\mathbf{b},\mathbf{c}) - (\mathbf{a},\mathbf{c}) - (\mathbf{b},\mathbf{c})}{|\mathbf{c}|^2}\,\mathbf{c},$$

$$\pi_{\mathbf{c}}(\alpha\,\mathbf{a}) - \alpha\,\pi_{\mathbf{c}}(\mathbf{a}) = \frac{(\alpha\,\mathbf{a},\mathbf{c}) - \alpha\,(\mathbf{a},\mathbf{c})}{|\mathbf{c}|^2}\,\mathbf{c}.$$

Due to the theorem 27.3 the mapping $\pi_{\mathbf{c}}$ is a linear mapping (see Definition 27.4). Therefore the left hand sides of the above equalities are zero. Now due to $\mathbf{c} \neq \mathbf{0}$ we conclude that the

numerators of the fractions in their right hand sides are also zero. This fact proves the properties 2) and 3) from the theorem 28.1 are valid in the case $\mathbf{c} \neq \mathbf{0}$.

According to the definition 26.1 the scalar product $(\mathbf{a}, \mathbf{a})$ is equal to zero for $\mathbf{a} = \mathbf{0}$. Otherwise, if $\mathbf{a} \neq \mathbf{0}$, the formula (26.1) is applied where we should set $\mathbf{b} = \mathbf{a}$. This yields $\varphi = 0$ and

$$(\mathbf{a}, \mathbf{a}) = |\mathbf{a}|^2 > 0.$$

This inequality proves the property 4) and completes the proof of the theorem 28.1 in whole. □

THEOREM 28.2. *Apart from the properties 1)–4), the scalar product of vectors possesses the following two properties fulfilled for any three vectors* $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ *and for any number* α*:*

5) $(\mathbf{c}, \mathbf{a} + \mathbf{b}) = (\mathbf{c}, \mathbf{a}) + (\mathbf{a}, \mathbf{b})$;
6) $(\mathbf{c}, \alpha\,\mathbf{a}) = \alpha\,(\mathbf{c}, \mathbf{a})$.

DEFINITION 28.2. The properties 5) and 6) in the theorem 28.2 are called the properties of *linearity with respect to the second multiplicand.*

The properties 5) and 6) are easily derived from the properties 2) and 3) by applying the property 1). Indeed, we have

$$(\mathbf{c}, \mathbf{a} + \mathbf{b}) = (\mathbf{a} + \mathbf{b}, \mathbf{c}) = (\mathbf{a}, \mathbf{c}) + (\mathbf{b}, \mathbf{c}) = (\mathbf{c}, \mathbf{a}) + (\mathbf{c}, \mathbf{b}),$$
$$(\mathbf{c}, \alpha\,\mathbf{a}) = (\alpha\,\mathbf{a}, \mathbf{c}) = \alpha\,(\mathbf{a}, \mathbf{c}) = \alpha\,(\mathbf{c}, \mathbf{a}).$$

These calculations prove the theorem 28.2.

§ 29. Calculation of the scalar product through the coordinates of vectors in a skew-angular basis.

Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ be some arbitrary basis in the space $\mathbb{E}$. According to the definition 18.1, this is an ordered triple of non-coplanar vectors. The arbitrariness of a basis means that no auxiliary restrictions are imposed onto the vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, except for

non-coplanarity. In particular, this means that the angles between the vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ in an arbitrary basis should not be right angles. For this reason such a basis is called a *skew-angular basis* and abbreviated as SAB.

DEFINITION 29.1. In this book a *skew-angular basis* (SAB) is understood as an *arbitrary basis*.

Thus, let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ be some skew-angular basis in the space $\mathbb{E}$ and let $\mathbf{a}$ and $\mathbf{b}$ be two free vectors given by its coordinates in this basis. We write this fact as follows:

$$\mathbf{a} = \begin{Vmatrix} a^1 \\ a^2 \\ a^3 \end{Vmatrix}, \qquad \mathbf{b} = \begin{Vmatrix} b^1 \\ b^2 \\ b^3 \end{Vmatrix} \tag{29.1}$$

Unlike (21.1), instead of the arrow sign in (29.1) we use the equality sign. Doing this, we emphasize the fact that once a basis is fixed, vectors are uniquely identified with their coordinates.

The conditional writing (29.1) means that the vectors $\mathbf{a}$ and $\mathbf{b}$ are presented by the following expansions:

$$\mathbf{a} = \sum_{i=1}^{3} a^i\,\mathbf{e}_i, \qquad \mathbf{b} = \sum_{j=1}^{3} b^j\,\mathbf{e}_j. \tag{29.2}$$

Substituting (29.2) into the scalar product $(\mathbf{a}, \mathbf{b})$, we get

$$(\mathbf{a}, \mathbf{b}) = \Bigl(\sum_{i=1}^{3} a^i\,\mathbf{e}_i, \sum_{j=1}^{3} b^j\,\mathbf{e}_j\Bigr). \tag{29.3}$$

In order to transform the formulas (29.3) we apply the properties 2) and 5) of the scalar product from the theorems 28.1 and 28.2. Due to these properties we can take the summation signs over i and j out of the brackets of the scalar product:

$$(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{3}\sum_{j=1}^{3} (a^i\,\mathbf{e}_i, b^j\,\mathbf{e}_j). \tag{29.4}$$

Then we apply the properties 3) and 6) from the theorems 28.1 and 28.2. Due to these properties we can take the numeric factors a^i and b^j out of the brackets of the scalar product in (29.4):

$$(\mathbf{a},\mathbf{b}) = \sum_{i=1}^{3}\sum_{j=1}^{3} a^i\, b^j\, (\mathbf{e}_i,\mathbf{e}_j). \tag{29.5}$$

The quantities $(\mathbf{e}_i,\mathbf{e}_j)$ in the formula (29.5) depend on a basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, namely on the lengths of the basis vectors and on the angles between them. They do not depend on the vectors $\mathbf{a}$ and $\mathbf{b}$. The quantities $(\mathbf{e}_i,\mathbf{e}_j)$ constitute an array of nine numbers

$$g_{ij} = (\mathbf{e}_i,\mathbf{e}_j) \tag{29.6}$$

enumerated by two lower indices. The components of the array (29.6) are usually arranged into a square matrix:

$$G = \begin{Vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{Vmatrix} \tag{29.7}$$

Definition 29.2. The matrix (29.7) with the components (29.6) is called the Gram matrix of a basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$.

Taking into account the notations (29.6), we write the formula (29.5) in the following way:

$$(\mathbf{a},\mathbf{b}) = \sum_{i=1}^{3}\sum_{j=1}^{3} a^i\, b^j\, g_{ij}. \tag{29.8}$$

Definition 29.3. The formula (29.8) is called the *formula for calculating the scalar product through the coordinates of vectors in a skew-angular basis.*

The formula (29.8) can be written in the matrix form

$$(\mathbf{a},\mathbf{b}) = \| a^1 \quad a^2 \quad a^3 \| \cdot \left\| \begin{matrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{matrix} \right\| \cdot \left\| \begin{matrix} b^1 \\ b^2 \\ b^3 \end{matrix} \right\| \tag{29.9}$$

Note that the coordinate column of the vector $\mathbf{b}$ in the formula (29.9) is used as it is, while the coordinate column of the vector $\mathbf{a}$ is transformed into a row. Such a transformation is known as *matrix transposing* (see [7]).

Definition 29.4. A transformation of a rectangular matrix under which the element in the intersection of i-th row and j-th column is taken to the intersection of j-th row and i-th column is called the *matrix transposing*. It is denoted by means of the sign $\top$. In the TEX and LaTEX computer packages this sign is coded by the operator `\top`.

The operation of matrix transposing can be understood as the mirror reflection with respect to the main diagonal of a matrix

Taking into account the notations (29.1), (29.7), and the definition 29.4, we can write the matrix formula (29.9) as follows:

$$(\mathbf{a},\mathbf{b}) = \mathbf{a}^\top \cdot G \cdot \mathbf{b}. \tag{29.10}$$

In the right hand side of the formula (29.10) the vectors $\mathbf{a}$ and $\mathbf{b}$ are presented by their coordinate columns, while the transformation of one of them into a row is written through the matrix transposing.

Exercise 29.1. *Show that for an arbitrary rectangular matrix A the equality $(A^\top)^\top = A$ is fulfilled.*

Exercise 29.2. *Show that for the product of two matrices A and B the equality $(A \cdot B)^\top = B^\top \cdot A^\top$ is fulfilled.*

EXERCISE 29.3. *Define the Gram matrices for bases on a line and for bases on a plane. Write analogs of the formulas* (29.8), (29.9), *and* (29.10) *for the scalar product of vectors lying on a line and on a plane.*

§ 30. Symmetry of the Gram matrix.

DEFINITION 30.1. A square matrix A is called *symmetric*, if it is preserved under transposing, i. e. if the following equality is fulfilled: $A^\top = A$.

Gram matrices possesses many important properties. One of these properties is their symmetry.

THEOREM 30.1. *The Gram matrix G of any basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ in the space $\mathbb{E}$ is symmetric.*

PROOF. According to the definition 30.1 the symmetry of G is expressed by the formula $G^\top = G$. According to the definition 29.4, the equality $G^\top = G$ is equivalent to the relationship

$$g_{ij} = g_{ji} \tag{30.1}$$

for the components of the matrix G. As for the relationship (30.1), upon applying (29.6), it reduces to the equality

$$(\mathbf{e}_i, \mathbf{e}_j) = (\mathbf{e}_j, \mathbf{e}_i)$$

which is fulfilled due to the symmetry of the scalar product (see Theorem 28.1 and Definition 28.1). □

Note that the coordinate columns of the vectors $\mathbf{a}$ and $\mathbf{b}$ enter the right hand side of the formula (29.10) in somewhat unequal way — one of them is transposed, the other is not transposed. The symmetry of the matrix G eliminates this difference. Redesignating the indices i and j in the double sum

(29.8) and taking into account the relationship (30.1) for the components of the Gram matrix, we get

$$(\mathbf{a},\mathbf{b}) = \sum_{j=1}^{3}\sum_{i=1}^{3} a^j\, b^i\, g_{ji} = \sum_{i=1}^{3}\sum_{j=1}^{3} b^i\, a^j\, g_{ij}. \tag{30.2}$$

In the matrix form the formula (30.2) is written as follows:

$$(\mathbf{a},\mathbf{b}) = \mathbf{b}^{\top}\cdot G\cdot \mathbf{a}. \tag{30.3}$$

The formula (30.3) is analogous to the formula (29.10), but in this formula the coordinate column of the vector $\mathbf{b}$ is transposed, while the coordinate column of the vector $\mathbf{a}$ is not transposed.

EXERCISE 30.1. *Formulate and prove a theorem analogous to the theorem 30.1 for bases on a plane. Is it necessary to formulate such a theorem for bases on a line.*

§ 31. Orthonormal basis.

DEFINITION 31.1. A basis on a straight line consisting of a nonzero vector $\mathbf{e}$ is called an *orthonormal basis*, if $\mathbf{e}$ is a unit vector, i. e. if $|\mathbf{e}| = 1$.

DEFINITION 31.2. A basis on a plane, consisting of two non-collinear vectors $\mathbf{e}_1, \mathbf{e}_2$, is called an *orthonormal basis*, if the vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are two vectors of the unit lengths perpendicular to each other.

DEFINITION 31.3. A basis in the space $\mathbb{E}$ consisting of three non-coplanar vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ is called an *orthonormal basis* if the vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ are three vectors of the unit lengths perpendicular to each other.

In order to denote an orthonormal basis in each of the there cases listed above we use the abbreviation ONB. According to

the definition 29.1 the orthonormal basis is not opposed to a skew-angular basis SAB, it is a special case of such a basis.

Note that the unit lengths of the basis vectors of an orthonormal basis in the definitions 31.1, 31.2, and 31.2 mean that their lengths are not one centimeter, not one meter, not one kilometer, but the pure numeric unity. For this reason all geometric realizations of such vectors are conditionally geometric (see § 2). Like velocity vectors, acceleration vectors, and many other physical quantities, basis vectors of an orthonormal basis can be drawn only upon choosing some scaling factor. Such a factor in this particular case is needed for to transform the numeric unity into a unit of length.

§ 32. Gram matrix of an orthonormal basis.

Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ be some orthonormal basis in the space $\mathbb{E}$. According to the definition 31.3 the vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ satisfy the following relationships:

$$\begin{array}{lll} |\mathbf{e}_1| = 1, & |\mathbf{e}_2| = 1, & |\mathbf{e}_3| = 1, \\ \mathbf{e}_1 \perp \mathbf{e}_2, & \mathbf{e}_2 \perp \mathbf{e}_3, & \mathbf{e}_3 \perp \mathbf{e}_1. \end{array} \tag{32.1}$$

Applying (32.1) and (29.6), we find the components of the Gram matrix for the orthonormal basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$:

$$g_{ij} = \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j. \end{cases} \tag{32.2}$$

From (32.2) we immediately derive the following theorem.

Theorem 32.1. *The Gram matrix* (29.7) *of any orthonormal basis is a unit matrix:*

$$G = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix} = 1. \tag{32.3}$$

Let's recall that the components of a unit matrix constitute a numeric array δ which is called the Kronecker symbol (see Definition 23.1). Therefore, taking into account (23.3), the equality (32.2) can be written as:

$$g_{ij} = \delta_{ij}. \tag{32.4}$$

The Kronecker symbol in (32.4) inherits the lower position of indices from g_{ij}. Therefore it is different from the Kronecker symbol in (23.4). Despite being absolutely identical, the components of the unit matrix in (32.3) and in (23.1) are of absolutely different nature. Having negotiated to use indices on upper and lower levels (see Definition 20.1), now we are able to reflect this difference in denoting these components.

§ 33. Calculation of the scalar product through the coordinates of vectors in an orthonormal basis.

According to the definition 29.1 the term *skew-angular basis* is used as a synonym of an arbitrary basis. For this reason an orthonormal basis is a special case of a skew-angular basis and we can use the formula (29.8), taking into account (32.4):

$$(\mathbf{a},\mathbf{b}) = \sum_{i=1}^{3}\sum_{j=1}^{3} a^i\, b^i\, \delta_{ij}. \tag{33.1}$$

In calculating the sum over j in (33.1), it is the inner sum here, the index j runs over three values and only for one of these three values, where $j = i$, the Kronecker symbol δ_{ij} is nonzero. For this reason wee can retain only one summand of the inner sum over j in (33.1), omitting other two summands:

$$(\mathbf{a},\mathbf{b}) = \sum_{i=1}^{3} a^i\, b^i\, \delta_{ii}. \tag{33.2}$$

We know that $\delta_{ii} = 1$. Therefore the formula (33.2) turns to

$$(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{3} a^i\, b^i. \tag{33.3}$$

DEFINITION 33.1. The formula (33.3) is called the *formula for calculating the scalar product through the coordinates of vectors in an orthonormal basis.*

Note that the sums in the formula (29.8) satisfy the index setting rule from the definition 24.8, while the sum in the formula (33.3) breaks this rule. In this formula the summation index has two entries and both of them are in the upper positions. This is a peculiarity of an orthonormal basis. It is more symmetric as compared to a general skew-angular basis and this symmetry hides some rules that reveal in a general non-symmetric bases.

The formula (33.3) has the following matrix form:

$$(\mathbf{a}, \mathbf{b}) = \| \, a^1 \quad a^2 \quad a^3 \, \| \cdot \begin{Vmatrix} b^1 \\ b^2 \\ b^3 \end{Vmatrix}. \tag{33.4}$$

Taking into account the notations (29.1) and taking into account the definition 29.4, the formula (33.4) can be abbreviated to

$$(\mathbf{a}, \mathbf{b}) = \mathbf{a}^{\top} \cdot \mathbf{b}. \tag{33.5}$$

The formula (33.4) can be derived from the formula (29.9), while the formula (33.5) can be derived from the formula (29.10).

§ 34. Right and left triples of vectors. The concept of orientation.

DEFINITION 34.1. An *ordered triple of vectors* is a list of three vectors for which the order of listing vectors is fixed.

Definition 34.2. An ordered triple of non-coplanar vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ is called a *right triple* if, when observing from the end of the third vector, the shortest rotation from the first vector toward the second vector is seen as a counterclockwise rotation.

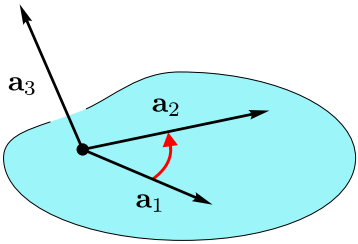

Fig. 34.1

In the definition 34.2 we implicitly assume that the geometric realizations of the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ with some common initial point are considered as it is shown in Fig. 34.1.

Definition 34.3. An ordered triple of non-coplanar vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ is called a *left triple* if, when observing from the end of the third vector, the shortest rotation from the first vector toward the second vector is seen as a clockwise rotation.

A given rotation about a given axis when observing from a given position could be either a clockwise rotation or a counterclockwise rotation. No other options are available. For this reason each ordered triple of non-coplanar vectors is either left or right. No other triples sorted by this criterion are available.

Definition 34.4. The property of ordered triples of non-coplanar vectors to be left or right is called their *orientation*.

§ 35. Vector product.

Let $\mathbf{a}$ and $\mathbf{b}$ be two non-collinear free vectors. Let's lay their geometric realizations $\mathbf{a} = \overrightarrow{OA}$ and $\mathbf{b} = \overrightarrow{OB}$ at some arbitrary point O. In this case the vectors $\mathbf{a}$ and $\mathbf{b}$ define a plane AOB and lie on this plane. The angle φ between the vectors $\mathbf{a}$ and $\mathbf{b}$ is determined according to Fig. 26.1. Due to $\mathbf{a} \nparallel \mathbf{b}$ this angle ranges in the interval $0 < \varphi < \pi$ and hence $\sin\varphi \neq 0$.

Let's draw a line through the point O perpendicular to the plane AOB and denote this line through c. The line c is

perpendicular to the vectors $\mathbf{a} = \overrightarrow{OA}$ and $\mathbf{b} = \overrightarrow{OB}$:

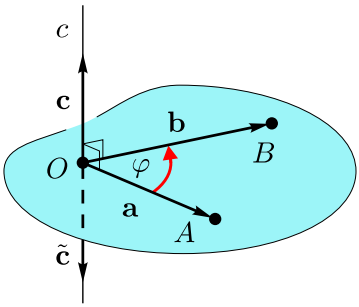

Fig. 35.1

$$\begin{aligned} c &\perp \mathbf{a}, \\ c &\perp \mathbf{b}. \end{aligned} \tag{35.1}$$

It is clear that the conditions (35.1) fix a unique line c passing through the point O (see Theorems 1.1 and 1.3 in Chapter IV of the book [6]).

There are two directions on the line c. In Fig 35.1 they are given by the vectors $\mathbf{c}$ and $\tilde{\mathbf{c}}$. The vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ constitute a right triple, while $\mathbf{a}$, $\mathbf{b}$, $\tilde{\mathbf{c}}$ is a left triple. So, specifying the orientation chooses one of two possible directions on the line c.

DEFINITION 35.1. The *vector product* of two non-collinear vectors $\mathbf{a}$ and $\mathbf{b}$ is a vector $\mathbf{c} = [\mathbf{a}, \mathbf{b}]$ which is determined by the following three conditions:

1) $\mathbf{c} \perp \mathbf{a}$ and $\mathbf{c} \perp \mathbf{b}$;
2) the vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ form a right triple;
3) $|\mathbf{c}| = |\mathbf{a}|\,|\mathbf{b}|\sin\varphi$.

In the case of collinear vectors $\mathbf{a}$ and $\mathbf{b}$ their vector product $[\mathbf{a}, \mathbf{b}]$ is taken to be zero by definition.

A comma is the multiplication sign in the writing the vector product, not by itself, but together with square brackets surrounding the whole expression. These brackets are natural delimiters for multiplicands: the first multiplicand is an expression between the opening bracket and the comma, while the second multiplicand is an expression between the comma and the closing bracket. Therefore in complicated expressions no auxiliary delimiters are required. For example, in the formula

$$[\mathbf{a} + \mathbf{b}, \mathbf{c} + \mathbf{d}]$$

the sums $\mathbf{a}+\mathbf{b}$ and $\mathbf{c}+\mathbf{d}$ are calculated first, then the vector multiplication is performed.

A remark. Often the vector product is written as $\mathbf{a}\times\mathbf{b}$. Even the special term «cross product» is used. However, to my mind, this notation is not good. It is misleading since the cross sign is sometimes used for denoting the product of numbers when a large formula is split into several lines.

A remark. The physical nature of the vector product $[\mathbf{a},\mathbf{b}]$ often differs from the nature of its multiplicands $\mathbf{a}$ and $\mathbf{b}$. Even if the lengths of the vectors $\mathbf{a}$ and $\mathbf{b}$ are measured in length units, the length of their product $[\mathbf{a},\mathbf{b}]$ is measured in units of area.

EXERCISE 35.1. *Show that the vector product* $\mathbf{c}=[\mathbf{a},\mathbf{b}]$ *of two free vectors* $\mathbf{a}$ *and* $\mathbf{b}$ *is a free vector and, being a free vector, it does not depend on where the point* O *in Fig. 35.1 is placed.*

§ 36. Orthogonal projection onto a plane.

Let $\mathbf{a}\neq\mathbf{0}$ be some nonzero free vector. According to the theorem 27.1, each free vector $\mathbf{b}$ has the expansion

$$\mathbf{b}=\mathbf{b}_{\parallel}+\mathbf{b}_{\perp} \tag{36.1}$$

relative to the vector $\mathbf{a}$, where the vector $\mathbf{b}_{\parallel}$ is collinear to the vector $\mathbf{a}$, while the vector $\mathbf{b}_{\perp}$ is perpendicular to the vector $\mathbf{a}$. Recall that through $\pi_{\mathbf{a}}$ we denoted a mapping that associates each vector $\mathbf{b}$ with its component $\mathbf{b}_{\parallel}$ in the expansion (36.1). Such a mapping was called the orthogonal projection onto the direction of the vector $\mathbf{a}\neq\mathbf{0}$ (see Definition 27.3).

DEFINITION 36.1. The mapping $\pi_{\perp\mathbf{a}}$ that associates each free vector $\mathbf{b}$ with its perpendicular component $\mathbf{b}_{\perp}$ in the expansion (36.1) is called the *orthogonal projection onto a plane perpendicular to the vector* $\mathbf{a}\neq\mathbf{0}$ or, more exactly, the *orthogonal projection onto the orthogonal complement of the vector* $\mathbf{a}\neq\mathbf{0}$.

DEFINITION 36.2. The *orthogonal complement* of a free vector

a is the collection of all free vectors **x** perpendicular to **a**:

$$\alpha = \{\mathbf{x}\colon\ \mathbf{x} \perp \mathbf{a}\}. \tag{36.2}$$

The orthogonal complement (36.2) of a nonzero vector $\mathbf{a} \neq \mathbf{0}$ can be visualized as a plane if we choose one of its geometric realizations $\mathbf{a} = \overrightarrow{OA}$. Indeed, let's lay various vectors perpendicular to **a** at the point O. The ending points of such vectors fill the plane α shown in Fig. 36.1.

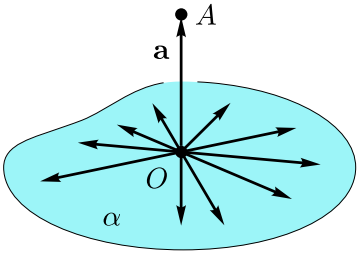

Fig. 36.1

The properties of the orthogonal projections onto a line $\pi_{\mathbf{a}}$ from the definition 27.1 and the orthogonal projections onto a plane $\pi_{\perp\mathbf{a}}$ from the definition 36.1 are very similar. Indeed, we have the following theorem.

THEOREM 36.1. *For any nonzero vector* $\mathbf{a} \neq \mathbf{0}$ *the orthogonal projection* $\pi_{\perp\mathbf{a}}$ *onto a plane perpendicular to the vector* **a** *is a linear mapping.*

PROOF. In order to prove the theorem 36.1 we write the relationship (36.1) as follows:

$$\mathbf{b} = \pi_{\mathbf{a}}(\mathbf{b}) + \pi_{\perp\mathbf{a}}(\mathbf{b}). \tag{36.3}$$

The relationship (36.3) is an identity, it is fulfilled for any vector **b**. First we replace the vector **b** by $\mathbf{b} + \mathbf{c}$ in (36.3), then we replace **b** by $\alpha\,\mathbf{b}$ in (36.3). As a result we get two relationships

$$\pi_{\perp\mathbf{a}}(\mathbf{b} + \mathbf{c}) = \mathbf{b} + \mathbf{c} - \pi_{\mathbf{a}}(\mathbf{b} + \mathbf{c}), \tag{36.4}$$

$$\pi_{\perp\mathbf{a}}(\alpha\,\mathbf{b}) = \alpha\,\mathbf{b} - \pi_{\mathbf{a}}(\alpha\,\mathbf{b}). \tag{36.5}$$

Due to the theorem 27.3 the mapping $\pi_{\mathbf{a}}$ is a linear mapping. For this reason the relationships (36.4) and (36.5) can be

transformed into the following two relationships:

$$\pi_{\perp\mathbf{a}}(\mathbf{b}+\mathbf{c}) = \mathbf{b} - \pi_{\mathbf{a}}(\mathbf{b}) + \mathbf{c} - \pi_{\mathbf{a}}(\mathbf{c}), \tag{36.6}$$

$$\pi_{\perp\mathbf{a}}(\alpha\,\mathbf{b}) = \alpha\,(\mathbf{b} - \pi_{\mathbf{a}}(\mathbf{b})). \tag{36.7}$$

The rest is to apply the identity (36.3) to the relationships (36.6) and (36.7). As a result we get

$$\pi_{\perp\mathbf{a}}(\mathbf{b}+\mathbf{c}) = \pi_{\perp\mathbf{a}}(\mathbf{b}) + \pi_{\perp\mathbf{a}}(\mathbf{c}), \tag{36.8}$$

$$\pi_{\perp\mathbf{a}}(\alpha\,\mathbf{b}) = \alpha\,\pi_{\perp\mathbf{a}}(\mathbf{b}). \tag{36.9}$$

The relationships (36.8) and (36.9) are exactly the linearity conditions from the definition 27.4 written for the mapping $\pi_{\perp\mathbf{a}}$. The theorem 36.1 is proved. □

§ 37. Rotation about an axis.

Let $\mathbf{a} \neq \mathbf{0}$ be some nonzero free vector and let $\mathbf{b}$ be some arbitrary free vector. Let's lay the vector $\mathbf{b} = \overrightarrow{BO}$ at some arbitrary point B. Then we lay the vector $\mathbf{a} = \overrightarrow{OA}$ at the terminal point of the vector $\overrightarrow{BO}$. The vector $\mathbf{a} = \overrightarrow{OA}$ is nonzero. For this reason it defines a line OA. We take this line for the rotation axis. Let's denote through $\theta^{\varphi}_{\mathbf{a}}$ the rotation of the space $\mathbb{E}$ about the axis OA by the angle φ (see Fig. 37.1). The vector $\mathbf{a} = \overrightarrow{OA}$ fixes one of two directions on the rotation axis. At the same tame this vector fixes the positive direction of rotation about the axis OA.

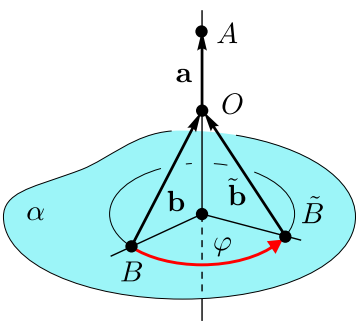

Fig. 37.1

Definition 37.1. The rotation about an axis OA with with fixed direction $\mathbf{a} = \overrightarrow{OA}$ on it is called a positive rotation if, being observed from the terminal point of the vector $\overrightarrow{OA}$, i. e.

when looking from the point A toward the point O, it occurs in the counterclockwise direction.

Taking into account the definition 37.1, we can consider the rotation angle φ as a signed quantity. If $\varphi > 0$, the rotation $\theta^{\varphi}_{\mathbf{a}}$ occurs in the positive direction with respect to the vector $\mathbf{a}$, if $\varphi < 0$, it occurs in the negative direction.

Let's apply the rotation mapping $\theta^{\varphi}_{\mathbf{a}}$ to the vectors $\mathbf{a} = \overrightarrow{OA}$ and $\mathbf{b} = \overrightarrow{BO}$ in Fig. 37.1. The points A and O are on the rotation axis. For this reason under the rotation $\theta^{\varphi}_{\mathbf{a}}$ the points A and O stay at their places and the vector $\mathbf{a} = \overrightarrow{OA}$ does not change. As for the vector $\mathbf{b} = \overrightarrow{BO}$, it is mapped onto another vector $\overrightarrow{\tilde{B}O}$. Now, applying parallel translations, we can replicate the vector $\overrightarrow{\tilde{B}O}$ up to a free vector $\tilde{\mathbf{b}} = \overrightarrow{\tilde{B}O}$ (see Definitions 4.1 and 4.2). The vector $\tilde{\mathbf{b}}$ is said to be produced from the vector $\mathbf{b}$ by applying the mapping $\theta^{\varphi}_{\mathbf{a}}$ and is written as

$$\tilde{\mathbf{b}} = \theta^{\varphi}_{\mathbf{a}}(\mathbf{b}). \tag{37.1}$$

LEMMA 37.1. *The free vector $\tilde{\mathbf{b}} = \theta^{\varphi}_{\mathbf{a}}(\mathbf{b})$ in (37.1) produced from a free vector $\mathbf{b}$ by means of the rotation mapping $\theta^{\varphi}_{\mathbf{a}}$ does not depend on the choice of a geometric realization of the vector $\mathbf{a}$ defining the rotation axis and on a geometric realization of the vector $\mathbf{b}$ itself.*

DEFINITION 37.2. The mapping $\theta^{\varphi}_{\mathbf{a}}$ acting upon free vectors of the space $\mathbb{E}$ and taking them to other free vectors in $\mathbb{E}$ is called the rotation by the angle φ about the vector a.

EXERCISE 37.1. *Rotations and parallel translations belong to the class of mappings preserving lengths of segments and measures of angles. They take each segment to a congruent segment and each angle to a congruent angle (see [6]). Let p be some parallel translation, let p^{-1} be its inverse parallel translation, and let θ be a rotation by some angle about some axis. Prove that the*

composite mapping $\tilde{\theta} = p \circ \theta \circ p^{-1}$ *is the rotation by the same angle about the axis produced from the axis of* θ *by applying the parallel translation* p *to it.*

EXERCISE 37.2. *Apply the result of the exercise 37.1 for proving the lemma 37.1.*

THEOREM 37.1. *For any nonzero free vector* $\mathbf{a} \neq \mathbf{0}$ *and for any angle* φ *the rotation* $\theta_{\mathbf{a}}^{\varphi}$ *by the angle* φ *about the vector* $\mathbf{a}$ *is a linear mapping of free vectors.*

PROOF. In order to prove the theorem we need to inspect the conditions 1) and 2) from the definition 27.4 for the mapping $\theta_{\mathbf{a}}^{\varphi}$. Let's begin with the first of these conditions. Assume that $\mathbf{b}$ and $\mathbf{c}$ are two free vectors. Let's build their geometric realizations $\mathbf{b} = \overrightarrow{BC}$ and $\mathbf{c} = \overrightarrow{CO}$. Then the vector $\overrightarrow{BO}$ is the geometric realization for the sum of vectors $\mathbf{b} + \mathbf{c}$.

Now let's choose a geometric realization for the vector $\mathbf{a}$. It determines the rotation axis. According to the lemma 37.1 the actual place of such a geometric realization does not matter for the ultimate definition of the mapping $\theta_{\mathbf{a}}^{\varphi}$ as applied to free vectors. But for the sake of certainty we choose $\mathbf{a} = \overrightarrow{OA}$.

Let's apply the rotation by the angle φ about the axis OA to the points B, C, and O. The point O is on the rotation axis. Therefore it is not moved. The points B and C are moved to the points $\tilde{B}$ and $\tilde{C}$ respectively, while the triangle BCO is moved to the triangle $\tilde{B}\tilde{C}O$. As a result we get the following relationships:

$$\begin{aligned} \theta_{\mathbf{a}}^{\varphi}\big(\overrightarrow{BC}\big) &= \overrightarrow{\tilde{B}\tilde{C}}, \\ \theta_{\mathbf{a}}^{\varphi}\big(\overrightarrow{CO}\big) &= \overrightarrow{\tilde{C}O}, \\ \theta_{\mathbf{a}}^{\varphi}\big(\overrightarrow{BO}\big) &= \overrightarrow{\tilde{B}O}. \end{aligned} \tag{37.2}$$

But the vectors $\overrightarrow{\tilde{B}\tilde{C}}$ and $\overrightarrow{\tilde{C}O}$ in (37.2) are geometric realizations for the vectors $\tilde{\mathbf{b}} = \theta_{\mathbf{a}}^{\varphi}(\mathbf{b})$ and $\tilde{\mathbf{c}} = \theta_{\mathbf{a}}^{\varphi}(\mathbf{c})$, while $\overrightarrow{\tilde{B}O}$ is a geometric

realization for the vector $\theta^{\varphi}_{\mathbf{a}}(\mathbf{b}+\mathbf{c})$. Hence we have

$$\theta^{\varphi}_{\mathbf{a}}(\mathbf{b}+\mathbf{c}) = \overrightarrow{\tilde{B}O} = \overrightarrow{\tilde{B}\tilde{C}} + \overrightarrow{\tilde{C}O} = \theta^{\varphi}_{\mathbf{a}}(\mathbf{b}) + \theta^{\varphi}_{\mathbf{a}}(\mathbf{c}). \tag{37.3}$$

The chain of equalities (37.3) proves the first linearity condition from the definition 27.4 as applied to $\theta^{\varphi}_{\mathbf{a}}$.

Let's proceed to proving the second linearity condition. It is more simple than the first one. Multiplying a vector $\mathbf{b}$ by a number α, we make the length of its geometric realizations $|\alpha|$ times as greater. If $\alpha > 0$, geometric realizations of the vector $\alpha\,\mathbf{b}$ are codirected to geometric realizations of $\mathbf{b}$. If $\alpha < 0$, they are opposite to geometric realizations of $\mathbf{b}$. And if $\alpha = 0$, geometric realizations of the vector $\alpha\,\mathbf{b}$ do vanish. Let's apply the rotation by the angle φ about some geometric realization of the vector $\mathbf{a} \neq \mathbf{0}$ some geometric realizations of the vectors $\mathbf{b}$ and $\alpha\,\mathbf{b}$. Such a mapping preserves the lengths of vectors. Hence it preserves all the relations of their lengths. Moreover it maps straight lines to straight lines and preserves the order of points on that straight lines. Hence codirected vectors are mapped to codirected ones and opposite vectors to opposite ones respectively. As a result

$$\theta^{\varphi}_{\mathbf{a}}(\alpha\,\mathbf{b}) = \alpha\,\theta^{\varphi}_{\mathbf{a}}(\mathbf{b}). \tag{37.4}$$

The relationship (37.4) completes the proof of the linearity for the mapping $\theta^{\varphi}_{\mathbf{a}}$ as applied to free vectors. □

§ 38. The relation of the vector product with projections and rotations.

Let's consider two non-collinear vectors $\mathbf{a} \nparallel \mathbf{b}$ and their vector product $\mathbf{c} = [\mathbf{a}, \mathbf{b}]$. The length of the vector $\mathbf{c}$ is determined by the lengths of $\mathbf{a}$ and $\mathbf{b}$ and by the angle φ between them:

$$|\mathbf{c}| = |\mathbf{a}|\,|\mathbf{b}|\,\sin\varphi. \tag{38.1}$$

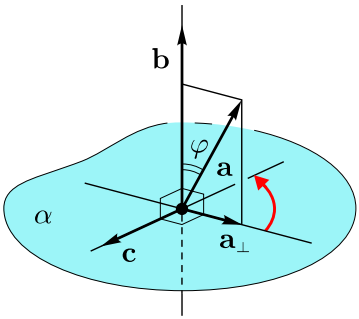

Fig. 38.1

The vector $\mathbf{c}$ lies on the plane α perpendicular to the vector $\mathbf{b}$ (see Fig. 38.1). Let's denote through $\mathbf{a}_\perp$ the orthogonal projection of the vector $\mathbf{a}$ onto the plane α, i e. we set

$$\mathbf{a}_\perp = \pi_{\perp\mathbf{b}}(\mathbf{a}). \tag{38.2}$$

The length of the above vector (38.2) is determined by the formula $|\mathbf{a}_\perp| = |\mathbf{a}| \sin\varphi$. Comparing this formula with (38.1) and taking into account Fig. 38.1, we conclude that in order to superpose the vector $\mathbf{a}_\perp$ with the vector $\mathbf{c}$ one should first rotate it counterclockwise by the right angle about the vector $\mathbf{b}$ and then multiply by the negative number $-|\mathbf{b}|$. This yields the formula

$$[\mathbf{a}, \mathbf{b}] = -|\mathbf{b}| \cdot \theta_{\mathbf{b}}^{\pi/2}\big(\pi_{\perp\mathbf{b}}(\mathbf{a})\big). \tag{38.3}$$

The formula (38.3) sets the relation of the vector product with the two mappings $\pi_{\perp\mathbf{b}}$ and $\theta_{\mathbf{b}}^{\pi/2}$. One of them is the projection onto the orthogonal complement of the vector $\mathbf{b}$, while the other is the rotation by the angle $\pi/2$ about the vector $\mathbf{b}$. The formula (38.3) is applicable provided the vector $\mathbf{b}$ is nonzero:

$$\mathbf{b} \neq \mathbf{0}, \tag{38.4}$$

while the condition $\mathbf{a} \nparallel \mathbf{b}$ can be broken. If $\mathbf{a} \parallel \mathbf{b}$ both sides of the formula (38.4) vanish, but the formula itself remains valid.

§ 39. Properties of the vector product.

THEOREM 39.1. *The vector product of vectors possesses the following four properties fulfilled for any three vectors* $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ *and*

for any number α:

1) $[\mathbf{a}, \mathbf{b}] = -[\mathbf{b}, \mathbf{a}]$;
2) $[\mathbf{a} + \mathbf{b}, \mathbf{c}] = [\mathbf{a}, \mathbf{c}] + [\mathbf{b}, \mathbf{c}]$;
3) $[\alpha\,\mathbf{a}, \mathbf{c}] = \alpha\,[\mathbf{a}, \mathbf{c}]$;
4) $[\mathbf{a}, \mathbf{b}] = 0$ *if and only if the vectors* $\mathbf{a}$ *and* $\mathbf{b}$ *are collinear, i. e. if* $\mathbf{a} \parallel \mathbf{b}$.

DEFINITION 39.1. The property 1) in the theorem 39.1 is called *anticommutativity*; the properties 2) and 3) are called the properties of *linearity with respect to the first multiplicand*; the property 4) is called the *vanishing condition*.

PROOF OF THE THEOREM 39.1. The property of anticommutativity 1) is derived immediately from the definition 35.1. Let $\mathbf{a} \nparallel \mathbf{b}$. Exchanging the vectors $\mathbf{a}$ and $\mathbf{b}$, we do not violate the first and the third conditions for the triple of vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ in the definition 35.1, provided they are initially fulfilled. As for the direction of rotation in Fig. 35.1, it changes for the opposite one. Therefore, if the triple $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ is right, the triple $\mathbf{b}$, $\mathbf{a}$, $\mathbf{c}$ is left. In order to get a right triple the vectors $\mathbf{b}$ and $\mathbf{a}$ should be complemented with the vector $-\mathbf{c}$. This yield the equality

$$[\mathbf{a}, \mathbf{b}] = -[\mathbf{b}, \mathbf{a}] \tag{39.1}$$

for the case $\mathbf{a} \nparallel \mathbf{b}$. If $\mathbf{a} \parallel \mathbf{b}$, both sides of the equality (39.1) do vanish. So the equality remains valid in this case too.

Let $\mathbf{c} \neq \mathbf{0}$. The properties of linearity 1) and 2) for this case are derived with the use of the formula (38.3) and the theorems 36.1 and 37.1. Let's write the formula (38.3) as

$$[\mathbf{a}, \mathbf{c}] = -|\mathbf{c}| \cdot \theta_{\mathbf{c}}^{\pi/2}\big(\pi_{\perp\mathbf{c}}(\mathbf{a})\big). \tag{39.2}$$

Then we change the vector $\mathbf{a}$ in (39.2) for the sum of vectors $\mathbf{a} + \mathbf{b}$ and apply the theorems 36.1 and 37.1. This yields

$$[\mathbf{a} + \mathbf{b}, \mathbf{c}] = -|\mathbf{c}| \cdot \theta_{\mathbf{c}}^{\pi/2}\big(\pi_{\perp\mathbf{c}}(\mathbf{a} + \mathbf{b})\big) = -|\mathbf{c}| \cdot$$

$$\cdot\,\theta_{\mathbf{c}}^{\pi/2}\big(\pi_{\perp\mathbf{c}}(\mathbf{a})+\pi_{\perp\mathbf{c}}(\mathbf{b})\big)=-|\mathbf{c}|\cdot\theta_{\mathbf{c}}^{\pi/2}\big(\pi_{\perp\mathbf{c}}(\mathbf{a})\big)-$$
$$-\,|\mathbf{c}|\cdot\theta_{\mathbf{c}}^{\pi/2}\big(\pi_{\perp\mathbf{c}}(\mathbf{b})\big)=[\mathbf{a},\mathbf{c}]+[\mathbf{b},\mathbf{c}].$$

Now we change the vector $\mathbf{a}$ in (39.2) for the product $\alpha\,\mathbf{a}$ and then apply the theorems 36.1 and 37.1 again:

$$[\alpha\,\mathbf{a},\mathbf{c}]=-|\mathbf{c}|\cdot\theta_{\mathbf{c}}^{\pi/2}\big(\pi_{\perp\mathbf{c}}(\alpha\,\mathbf{a})\big)=-|\mathbf{c}|\cdot\theta_{\mathbf{c}}^{\pi/2}\big(\alpha\,\pi_{\perp\mathbf{c}}(\mathbf{a})\big)=$$
$$=-\alpha\,|\mathbf{c}|\cdot\theta_{\mathbf{c}}^{\pi/2}\big(\pi_{\perp\mathbf{c}}(\mathbf{a})\big)=\alpha\,[\mathbf{a},\mathbf{c}].$$

The calculations which are performed above prove the equalities 2) and 3) in the theorem 39.1 for the case $\mathbf{c}\neq\mathbf{0}$. If $\mathbf{c}=\mathbf{0}$, both sides of these equalities do vanish and they appear to be trivially fulfilled.

Let's proceed to proving the fourth item in the theorem 39.1. For $\mathbf{a}\parallel\mathbf{b}$ the vector product $[\mathbf{a},\mathbf{b}]$ vanishes by the definition 35.1. Let $\mathbf{a}\nparallel\mathbf{b}$. In this case both vectors $\mathbf{a}$ and $\mathbf{b}$ are nonzero, while the angle φ between them differs from 0 and π. For this reason $\sin\varphi\neq 0$. Summarizing these restrictions and applying the item 3) of the definition 35.1, we get

$$|[\mathbf{a},\mathbf{b}]|=|\mathbf{a}|\,|\mathbf{b}|\,\sin\varphi\neq 0,$$

i. e. for $\mathbf{a}\nparallel\mathbf{b}$ the vector product $[\mathbf{a},\mathbf{b}]$ cannot vanish. The proof of the theorem 39.1 is over. □

THEOREM 39.2. *Apart from the properties 1)–4), the vector product possesses the following two properties which are fulfilled for any vectors* $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ *and for any number* α*:*

5) $[\mathbf{c},\mathbf{a}+\mathbf{b}]=[\mathbf{c},\mathbf{a}]+[\mathbf{a},\mathbf{b}]$;
6) $[\mathbf{c},\alpha\,\mathbf{a}]=\alpha\,[\mathbf{c},\mathbf{a}]$.

DEFINITION 39.2. The properties 5) and 6) in the theorem 39.2 are called the properties of *linearity with respect to the second multiplicand*.

The properties 5) and 6) are easily derived from the properties 2) and 3) by applying the property 1). Indeed, we have

$$[\mathbf{c}, \mathbf{a}+\mathbf{b}] = -[\mathbf{a}+\mathbf{b}, \mathbf{c}] = -[\mathbf{a}, \mathbf{c}] - [\mathbf{b}, \mathbf{c}] = [\mathbf{c}, \mathbf{a}] + [\mathbf{c}, \mathbf{b}],$$
$$[\mathbf{c}, \alpha\,\mathbf{a}] = -[\alpha\,\mathbf{a}, \mathbf{c}] = -\alpha\,[\mathbf{a}, \mathbf{c}] = \alpha\,[\mathbf{c}, \mathbf{a}].$$

These calculations prove the theorem 39.2.

§ 40. Structural constants of the vector product.

Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ be some arbitrary basis in the space $\mathbb{E}$. Let's take two vectors $\mathbf{e}_i$ and $\mathbf{e}_j$ of this basis and consider their vector product $[\mathbf{e}_i, \mathbf{e}_j]$. The vector $[\mathbf{e}_i, \mathbf{e}_j]$ can be expanded in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. Such an expansion is usually written as follows:

$$[\mathbf{e}_i, \mathbf{e}_j] = C^1_{ij}\,\mathbf{e}_1 + C^2_{ij}\,\mathbf{e}_2 + C^3_{ij}\,\mathbf{e}_3. \tag{40.1}$$

The expansion (40.1) contains three coefficients C^1_{ij}, C^2_{ij} and C^3_{ij}. However, the indices i and j in it run independently over three values 1, 2, 3. For this reason, actually, the formula (40.1) represent nine expansions, the total number of coefficients in it is equal to twenty seven.

The formula (40.1) can be abbreviated in the following way:

$$[\mathbf{e}_i, \mathbf{e}_j] = \sum_{k=1}^{3} C^k_{ij}\,\mathbf{e}_k. \tag{40.2}$$

Let's apply the theorem 19.1 on the uniqueness of the expansion of a vector in a basis to the expansions of $[\mathbf{e}_i, \mathbf{e}_j]$ in (40.1) or in (40.2). As a result we can formulate the following theorem.

THEOREM 40.1. *Each basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ *in the space* $\mathbb{E}$ *is associated with a collection of twenty seven constants* C^k_{ij} *which are determined uniquely by this basis through the expansions* (40.2).

DEFINITION 40.1. The constants C^k_{ij}, which are uniquely de-

termined by a basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ through the expansions (40.2), are called the *structural constants of the vector product* in this basis.

The structural constants of the vector product are similar to the components of the Gram matrix for a basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ (see Definition 29.2). But they are more numerous and form a three index array with two lower indices and one upper index. For this reason they cannot be placed into a matrix.

§ 41. Calculation of the vector product through the coordinates of vectors in a skew-angular basis.

Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ be some skew-angular basis. According to the definition 29.1 the term *skew-angular basis* in this book is used as a synonym of an arbitrary basis. Let's choose some arbitrary vectors $\mathbf{a}$ and $\mathbf{b}$ in the space $\mathbb{E}$ and consider their expansions

$$\mathbf{a} = \sum_{i=1}^{3} a^i\,\mathbf{e}_i, \qquad\qquad \mathbf{b} = \sum_{j=1}^{3} b^j\,\mathbf{e}_j \tag{41.1}$$

in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. Substituting (41.1) into the vector product $[\mathbf{a},\mathbf{b}]$, we get the following formula:

$$[\mathbf{a},\mathbf{b}] = \left[\sum_{i=1}^{3} a^i\,\mathbf{e}_i,\ \sum_{j=1}^{3} b^j\,\mathbf{e}_j\right]. \tag{41.2}$$

In order to transform the formula (41.2) we apply the properties 2) and 5) of the vector product (see Theorems 39.1 and 39.2). Due to these properties we can bring the summation signs over i and j outside the brackets of the vector product:

$$[\mathbf{a},\mathbf{b}] = \sum_{i=1}^{3}\sum_{j=1}^{3}[a^i\,\mathbf{e}_i, b^j\,\mathbf{e}_j]. \tag{41.3}$$

Now let's apply the properties 3) and 6) from the theorems 39.1 and 39.2. Due to these properties we can bring the numeric

factors a^i and b^j outside the brackets of the vector product (41.3):

$$[\mathbf{a},\mathbf{b}] = \sum_{i=1}^{3}\sum_{j=1}^{3} a^i\, b^j\, [\mathbf{e}_i,\mathbf{e}_j]. \tag{41.4}$$

The vector products $[\mathbf{e}_i,\mathbf{e}_j]$ in the formula (41.4) can be replaced by their expansions (40.2). Upon substituting (40.2) into (41.4) the formula (41.4) is written as follows:

$$[\mathbf{a},\mathbf{b}] = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} a^i\, b^j\, C_{ij}^{k}\, \mathbf{e}_k. \tag{41.5}$$

DEFINITION 41.1. The formula (41.5) is called the *formula for calculating the vector product through the coordinates of vectors in a skew-angular basis.*

§ 42. Structural constants of the vector product in an orthonormal basis.

Let's recall that an orthonormal basis (ONB) in the space $\mathbb{E}$ is a basis composed by three unit vectors perpendicular to each other (see Definition 31.3). By their orientation, triples of non-coplanar vectors in the space $\mathbb{E}$ are subdivided into right and left triples (see Definition 34.4). Therefore all bases in the space $\mathbb{E}$ are subdivided into right bases and left bases, which applies to orthonormal bases as well.

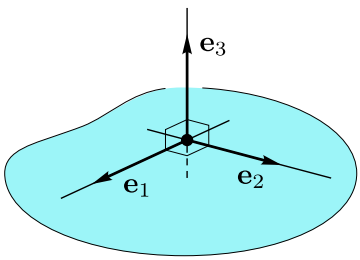

Fig. 42.1

Let's consider some right orthonormal basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. It is shown in Fig. 42.1. Using the definition 35.1, one can calculate various pairwise vector products of the vectors composing this basis. Since the geometry of a right

ONB is rather simple, we can perform these calculations up to an explicit result and compose the multiplication table for $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$:

$$\begin{array}{lll} [\mathbf{e}_1, \mathbf{e}_1] = \mathbf{0}, & [\mathbf{e}_1, \mathbf{e}_2] = \mathbf{e}_3, & [\mathbf{e}_1, \mathbf{e}_3] = -\mathbf{e}_2, \\ [\mathbf{e}_2, \mathbf{e}_1] = -\mathbf{e}_3, & [\mathbf{e}_2, \mathbf{e}_2] = \mathbf{0}, & [\mathbf{e}_2, \mathbf{e}_3] = \mathbf{e}_1, \\ [\mathbf{e}_3, \mathbf{e}_1] = \mathbf{e}_2, & [\mathbf{e}_3, \mathbf{e}_2] = -\mathbf{e}_1, & [\mathbf{e}_3, \mathbf{e}_3] = \mathbf{0}. \end{array} \tag{42.1}$$

Let's choose the first of the relationships (42.1) and write its right hand side in the form of an expansion in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$:

$$[\mathbf{e}_1, \mathbf{e}_1] = 0\,\mathbf{e}_1 + 0\,\mathbf{e}_2 + 0\,\mathbf{e}_3. \tag{42.2}$$

Let's compare the expansion (42.2) with the expansion (40.1) written for the case $i = 1$ and $j = 1$:

$$[\mathbf{e}_1, \mathbf{e}_1] = C^1_{11}\,\mathbf{e}_1 + C^2_{11}\,\mathbf{e}_2 + C^3_{11}\,\mathbf{e}_3. \tag{42.3}$$

Due to the uniqueness of the expansion of a vector in a basis (see Theorem 19.1) from (42.2) and (42.3) we derive

$$C^1_{11} = 0, \qquad C^2_{11} = 0, \qquad C^3_{11} = 0. \tag{42.4}$$

Now let's choose the second relationship (42.1) and write its right hand side in the form of an expansion in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$:

$$[\mathbf{e}_1, \mathbf{e}_2] = 0\,\mathbf{e}_1 + 0\,\mathbf{e}_2 + 1\,\mathbf{e}_3. \tag{42.5}$$

Comparing (42.5) with the expansion (40.1) written for the case $i = 1$ and $j = 2$, we get the values of the following constants:

$$C^1_{12} = 0, \qquad C^2_{12} = 0, \qquad C^3_{12} = 1. \tag{42.6}$$

Repeating this procedure for all relationships (42.1), we can get the complete set of relationships similar to (42.4) and (42.6).

Then we can organize them into a single list:

$$
\begin{array}{lll}
C^1_{11} = 0, & C^2_{11} = 0, & C^3_{11} = 0, \\
C^1_{12} = 0, & C^2_{12} = 0, & C^3_{12} = 1, \\
C^1_{13} = 0, & C^2_{13} = -1, & C^3_{13} = 0, \\
C^1_{21} = 0, & C^2_{21} = 0, & C^3_{21} = -1, \\
C^1_{22} = 0, & C^2_{22} = 0, & C^3_{22} = 0, \\
C^1_{23} = 1, & C^2_{23} = 0, & C^3_{23} = 0, \\
C^1_{31} = 0, & C^2_{31} = 1, & C^3_{31} = 0, \\
C^1_{32} = -1, & C^2_{32} = 0, & C^3_{32} = 1, \\
C^1_{33} = 0, & C^2_{33} = 0, & C^3_{33} = 0.
\end{array}
\tag{42.7}
$$

The formulas (42.7) determine all of the 27 structural constants of the vector product in a right orthonormal basis. Let's write this result as a theorem.

THEOREM 42.1. *For any right orthonormal basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ *in the space* $\mathbb{E}$ *the structural constants of the vector product are determined by the formulas* (42.7).

THEOREM 42.2. *For any left orthonormal basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ *in the space* $\mathbb{E}$ *the structural constants of the vector product are derived from* (42.7) *by changing signs «+» for «−» and vise versa.*

EXERCISE 42.1. *Draw a left orthonormal basis and, applying the definition 35.1 to the pairwise vector products of the basis vectors, derive the relationships analogous to* (42.1). *Then prove the theorem 42.2.*

§ 43. Levi-Civita symbol.

Let's examine the formulas (42.7) for the structural constants of the vector product in a right orthonormal basis. One can

easily observe the following pattern in them:

$$C_{ij}^k = 0 \text{ if there are coinciding values of the indices } i, j, k. \tag{43.1}$$

The condition (43.1) describes all of the cases where the structural constants in (42.7) do vanish. The cases where $C_{ij}^k = 1$ are described by the condition

$$C_{ij}^k = 1 \text{ if the indices } i, j, k \text{ take the values } (1,2,3),\ (2,3,1), \text{ or } (3,1,2). \tag{43.2}$$

Finally, the cases where $C_{ij}^k = -1$ are described by the condition

$$C_{ij}^k = -1 \text{ if the indices } i, j, k \text{ take the values } (1,3,2),\ (3,2,1), \text{ or } (2,1,3). \tag{43.3}$$

The triples of numbers in (43.2) and (43.3) constitute the complete set of various permutations of the numbers 1, 2, 3:

$$\begin{array}{lll} (1,2,3), & (2,3,1), & (3,1,2), \\ (1,3,2), & (3,2,1), & (2,1,3). \end{array} \tag{43.4}$$

The first three permutations in (43.4) are called *even permutations*. They are produced from the right order of the numbers 1, 2, 3 by applying an even number of pairwise transpositions to them. Indeed, we have

$$\begin{aligned} &(1,2,3); \\ &(1,2,3) \xrightarrow{1} (2,1,3) \xrightarrow{2} (2,3,1); \\ &(1,2,3) \xrightarrow{1} (1,3,2) \xrightarrow{2} (3,1,2). \end{aligned}$$

The rest three permutations in (43.4) are called *odd permutations*. In the case of these three permutations we have

$$(1,2,3) \xrightarrow{1} (1,3,2);$$

$$(1,2,3) \xrightarrow{1} (3,2,1);$$
$$(1,2,3) \xrightarrow{1} (2,1,3).$$

DEFINITION 43.1. The permutations (43.4) constitute a set, which is usually denoted through S_3. If $\sigma \in S_3$, then $(-1)^\sigma$ means the parity of the permutation σ:

$$(-1)^\sigma = \begin{cases} 1 \text{ if the permutation } \sigma \text{ is even;} \\ -1 \text{ if the permutation } \sigma \text{ is odd.} \end{cases}$$

Zeros, unities, and minus unities from (43.1), (43.2), and (43.3) are usually united into a single numeric array:

$$\varepsilon^{ijk} = \varepsilon_{ijk} = \begin{cases} 0 & \text{if there are coinciding values of the indices } i,j,k; \\ 1 & \text{if the values of the indices } i,j,k \text{ form an even permutation of the numbers } 1,2,3; \\ -1 & \text{if the values of the indices } i,j,k \text{ form an odd permutation of the numbers } 1,2,3. \end{cases} \tag{43.5}$$

DEFINITION 43.2. The numeric array ε determined by the formula (43.5) is called the Levi-Civita symbol.

When writing the components of the Levi-Civita symbol either three upper indices or three lower indices are used. Thus we emphasize the equity of all these three indices. Placing indices on different levels in the Levi-Civita symbol is nor welcome. Summarizing what was said above, the formulas (43.1), (43.2), and (43.3) are written as follows:

$$C_{ij}^k = \varepsilon_{ijk}. \tag{43.6}$$

THEOREM 43.1. *For any right orthonormal basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ *the*

structural constants of the vector product in such a basis are determined by the equality (43.6).

In the case of a left orthonormal basis we have the theorem 42.2. It yields the equality

$$C_{ij}^{k} = -\varepsilon_{ijk}. \tag{43.7}$$

THEOREM 43.2. *For any left orthonormal basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ *the structural constants of the vector product in such a basis are determined by the equality* (43.7).

Note that the equalities (43.6) and (43.7) violate the index setting rule given in the definition 24.9. The matter is that the structural constants of the vector product C_{ij}^{k} are the components of a geometric object. The places of their indices are determined by the index setting convention, which is known as Einstein's tensorial notation (see Definition 20.1). As for the Levi-Civita symbol, it is an array of purely algebraic origin.

The most important property of the Levi-Civita symbol is its *complete skew symmetry* or *complete antisymmetry*. It is expressed by the following equalities:

$$\begin{aligned} &\varepsilon_{ijk} = -\varepsilon_{jik}, \qquad \varepsilon_{ijk} = -\varepsilon_{ikj}, \qquad \varepsilon_{ijk} = -\varepsilon_{kji}, \\ &\varepsilon^{ijk} = -\varepsilon^{jik}, \qquad \varepsilon^{ijk} = -\varepsilon^{ikj}, \qquad \varepsilon^{ijk} = -\varepsilon^{kji}. \end{aligned} \tag{43.8}$$

The equalities (43.8) mean, that under the transposition of any two indices the quantity $\varepsilon_{ijk} = \varepsilon^{ijk}$ changes its sign. These equalities are easily derived from (43.5).

§ 44. Calculation of the vector product through the coordinates of vectors in an orthonormal basis.

Let's recall that the term *skew-angular basis* in this book is used as a synonym of an arbitrary basis (see Definition 29.1). Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ be a right orthonormal basis. It can be treated as a

special case of a skew-angular basis. Substituting (43.6) into the formula (41.5), we obtain the formula

$$[\mathbf{a},\mathbf{b}]=\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}a^i\,b^j\,\varepsilon_{ijk}\,\mathbf{e}_k. \tag{44.1}$$

Here a^i and b^j are the coordinates of the vectors $\mathbf{a}$ and $\mathbf{b}$ in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$.

In order to simplify the formulas (44.1) note that the majority components of the Levi-Civita symbol are equal to zero. Only six of its twenty seven components are nonzero. Applying the formula (43.5), we can bring the formula (44.1) to

$$\begin{aligned}[\mathbf{a},\mathbf{b}]=a^1\,b^2\,\mathbf{e}_3+a^2\,b^3\,\mathbf{e}_1+a^3\,b^1\,\mathbf{e}_2-\\-a^2\,b^1\,\mathbf{e}_3-a^3\,b^2\,\mathbf{e}_1-a^1\,b^3\,\mathbf{e}_2.\end{aligned} \tag{44.2}$$

Upon collecting similar terms the formula (44.2) yields

$$\begin{aligned}[\mathbf{a},\mathbf{b}]=\mathbf{e}_1\,(a^2\,b^3-a^3\,b^2)-\\-\mathbf{e}_2\,(a^1\,b^3-a^3\,b^1)+\mathbf{e}_3\,(a^1\,b^2-a^2\,b^1).\end{aligned} \tag{44.3}$$

Now from the formula (44.3) we derive

$$[\mathbf{a},\mathbf{b}]=\mathbf{e}_1\begin{vmatrix}a^2 & a^3\\ b^2 & b^3\end{vmatrix}-\mathbf{e}_2\begin{vmatrix}a^1 & a^3\\ b^1 & b^3\end{vmatrix}+\mathbf{e}_3\begin{vmatrix}a^1 & a^2\\ b^1 & b^2\end{vmatrix}. \tag{44.4}$$

In deriving the formula (44.4) we used the formula for the determinant of a 2×2 matrix (see [7]).

Note that the right hand side of the formula (44.4) is the expansion of the determinant of a 3×3 matrix by its first row (see [7]). Therefore this formula can be written as:

$$[\mathbf{a},\mathbf{b}]=\begin{vmatrix}\mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3\\ a^1 & a^2 & a^3\\ b^1 & b^2 & b^3\end{vmatrix}. \tag{44.5}$$

Here a^1, a^2, a^3 and b^1, b^2, b^3 are the coordinates of the vectors $\mathbf{a}$ and $\mathbf{b}$ in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. They fill the second and the third rows in the determinant (44.5).

DEFINITION 44.1. The formulas (44.1) and (44.5) are called the *formulas for calculating the vector product through the coordinates of vectors in a right orthonormal basis.*

Let's proceed to the case of a left orthonormal basis. In this case the structural constants of the vector product are given by the formula (43.7). Substituting (43.7) into (41.5), we get

$$[\mathbf{a},\mathbf{b}] = -\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} a^i\, b^j\, \varepsilon_{ijk}\, \mathbf{e}_k. \tag{44.6}$$

Then from (44.6) we derive the formula

$$[\mathbf{a},\mathbf{b}] = -\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a^1 & a^2 & a^3 \\ b^1 & b^2 & b^3 \end{vmatrix}. \tag{44.7}$$

DEFINITION 44.2. The formulas (44.6) and (44.7) are called the *formulas for calculating the vector product through the coordinates of vectors in a left orthonormal basis.*

§ 45. Mixed product.

DEFINITION 45.1. The *mixed product* of three free vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ is a number obtained as the scalar product of the vector $\mathbf{a}$ by the vector product of $\mathbf{b}$ and $\mathbf{c}$:

$$(\mathbf{a},\mathbf{b},\mathbf{c}) = (\mathbf{a},[\mathbf{b},\mathbf{c}]). \tag{45.1}$$

As we see in the formula (45.1), the mixed product has three multiplicands. They are separated from each other by

commas. Commas are the multiplication signs in writing the mixed product, not by themselves, but together with the round brackets surrounding the whole expression.

Commas and brackets in writing the mixed product are natural delimiters for multiplicands: the first multiplicand is an expression between the opening bracket and the first comma, the second multiplicand is an expression placed between two commas, and the third multiplicand is an expression between the second comma and the closing bracket. Therefore in complicated expressions no auxiliary delimiters are required. Foe example, in

$$(\mathbf{a}+\mathbf{b},\mathbf{c}+\mathbf{d},\mathbf{e}+\mathbf{f})$$

the sums $\mathbf{a}+\mathbf{b}$, $\mathbf{c}+\mathbf{d}$, and $\mathbf{e}+\mathbf{f}$ are calculated first, then the mixed product itself is calculated.

§ 46. Calculation of the mixed product through the coordinates of vectors in an orthonormal basis.

The formula (45.1) reduces the calculation of the mixed product to successive calculations of the vectorial and scalar products. In the case of the vectorial and scalar products we already have rather efficient formulas for calculating them through the coordinates of vectors in an orthonormal basis.

Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ be a right orthonormal basis and let $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ be free vectors given by their coordinates in this basis:

$$\mathbf{a}=\begin{Vmatrix} a^1\\ a^2\\ a^3\end{Vmatrix},\qquad \mathbf{b}=\begin{Vmatrix} b^1\\ b^2\\ b^3\end{Vmatrix},\qquad \mathbf{c}=\begin{Vmatrix} c^1\\ c^2\\ c^3\end{Vmatrix}. \tag{46.1}$$

Let's denote $\mathbf{d}=[\mathbf{b},\mathbf{c}]$. Then the formula (45.1) is written as

$$(\mathbf{a},\mathbf{b},\mathbf{c})=(\mathbf{a},\mathbf{d}). \tag{46.2}$$

In order to calculate the vector $\mathbf{d}=[\mathbf{b},\mathbf{c}]$ we apply the formula

(44.1) which now is written as follows:

$$\mathbf{d} = \sum_{k=1}^{3}\left(\sum_{i=1}^{3}\sum_{j=1}^{3} b^i\, c^j\, \varepsilon_{ijk}\right)\mathbf{e}_k. \tag{46.3}$$

The formula (46.3) is an expansion of the vector $\mathbf{d}$ in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. Hence the coefficients in this expansion should coincide with the coordinates of the vector $\mathbf{d}$:

$$d^k = \sum_{i=1}^{3}\sum_{j=1}^{3} b^i\, c^j\, \varepsilon_{ijk}. \tag{46.4}$$

The next step consists in using the coordinates of the vector $\mathbf{d}$ from (46.4) for calculating the scalar product in the right hand side of the formula (46.2). The formula (33.3) now is written as

$$(\mathbf{a}, \mathbf{d}) = \sum_{k=1}^{3} a^k\, d^k. \tag{46.5}$$

Let's substitute (46.4) into (46.5) and take into account (46.2). As a result we get the formula

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \sum_{i=1}^{3} a^k\left(\sum_{i=1}^{3}\sum_{j=1}^{3} b^i\, c^j\, \varepsilon_{ijk}\right). \tag{46.6}$$

Expanding the right hand side of the formula (46.6) and changing the order of summations in it, we bring it to

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} b^i\, c^j\, \varepsilon_{ijk}\, a^k. \tag{46.7}$$

Note that the right hand side of the formula (46.7) differs from that of the formula (44.1) by changing a^i for b^i, changing b^j for

c^j, and changing $\mathbf{e}_k$ for a^k. For this reason the formula (46.7) can be brought to the following form analogous to (44.5):

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \begin{vmatrix} a^1 & a^2 & a^3 \\ b^1 & b^2 & b^3 \\ c^1 & c^2 & c^3 \end{vmatrix}. \tag{46.8}$$

Another way for transforming the formula (46.7) is the use of the complete antisymmetry of the Levi-Civita symbol (43.8). Applying this property, we derive the identity $\varepsilon_{ijk} = \varepsilon_{kij}$. Due to this identity, upon changing the order of multiplicands and redesignating the summation indices in the right hand side of the formula (46.7), we can bring this formula to the following form:

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} a^i\, b^j\, c^k\, \varepsilon_{ijk}. \tag{46.9}$$

Definition 46.1. The formulas (46.8) and (46.9) are called the *formulas for calculating the mixed product through the coordinates of vectors in a right orthonormal basis.*

The coordinates of the vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ used in the formulas (46.8) and (46.9) are taken from (46.1).

Let's proceed to the case of a left orthonormal basis. Analogs of the formulas (46.8) and (46.9) for this case are obtained by changing the sign in the formulas (46.8) and (46.9):

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = -\begin{vmatrix} a^1 & a^2 & a^3 \\ b^1 & b^2 & b^3 \\ c^1 & c^2 & c^3 \end{vmatrix}. \tag{46.10}$$

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = -\sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} a^i\, b^j\, c^k\, \varepsilon_{ijk}. \tag{46.11}$$

DEFINITION 46.2. The formulas (46.10) and (46.11) are called the *formulas for calculating the mixed product through the coordinates of vectors in a left orthonormal basis.*

The formulas (46.10) and (46.11) can be derived using the theorem 42.2 or comparing the formulas (43.6) and (43.7).

§ 47. Properties of the mixed product.

THEOREM 47.1. *The mixed product possesses the following four properties which are fulfilled for any four vectors* $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$ *and for any number* α*:*

1) $(\mathbf{a},\mathbf{b},\mathbf{c}) = -(\mathbf{a},\mathbf{c},\mathbf{b})$,
   $(\mathbf{a},\mathbf{b},\mathbf{c}) = -(\mathbf{a},\mathbf{b},\mathbf{a})$,
   $(\mathbf{a},\mathbf{b},\mathbf{c}) = -(\mathbf{b},\mathbf{a},\mathbf{c})$;
2) $(\mathbf{a}+\mathbf{b},\mathbf{c},\mathbf{d}) = (\mathbf{a},\mathbf{c},\mathbf{d}) + (\mathbf{b},\mathbf{c},\mathbf{d})$;
3) $(\alpha\,\mathbf{a},\mathbf{c},\mathbf{d}) = \alpha\,(\mathbf{a},\mathbf{c},\mathbf{d})$;
4) $(\mathbf{a},\mathbf{b},\mathbf{c}) = 0$ *if and only if the vectors* $\mathbf{a}$, $\mathbf{b}$, *and* $\mathbf{c}$ *are coplanar.*

DEFINITION 47.1. The property 1) expressed by three equalities in the theorem 47.1 is called the property of *complete skew symmetry* or *complete antisymmetry*, the properties 2) and 3) are called the properties of *linearity with respect to the first multiplicand*, the property 4) is called the *vanishing condition.*

PROOF OF THE THEOREM 47.1. The first of the three equalities composing the property of complete antisymmetry 1) follows from the formula (45.1) and the theorems 39.1 and 28.1:

$$(\mathbf{a},\mathbf{b},\mathbf{c}) = (\mathbf{a},[\mathbf{b},\mathbf{c}]) = (\mathbf{a},-[\mathbf{c},\mathbf{b}]) = -(\mathbf{a},[\mathbf{c},\mathbf{b}]) = -(\mathbf{a},\mathbf{c},\mathbf{b}).$$

The other two equalities entering the property 1) cannot be derived in this way. Therefore we need to use the formula (46.8). Transposition of two vectors in the left hand side of this formula corresponds to the transposition of two rows in the determinant in the right hand side of this formula. It is well known that the

transposition of any two rows in a determinant changes its sign. This observation proves all of the three equalities composing the property of complete antisymmetry for the mixed product.

The properties of linearity 2) and 3) of the mixed product in the theorem 47.1 are derived from the corresponding properties of the scalar and vectorial products due to the formula (45.1):

$$(\mathbf{a}+\mathbf{b},\mathbf{c},\mathbf{d}) = (\mathbf{a}+\mathbf{b},[\mathbf{c},\mathbf{d}]) = (\mathbf{a},[\mathbf{c},\mathbf{d}]) + \\ + (\mathbf{b},[\mathbf{c},\mathbf{d}]) = (\mathbf{a},\mathbf{c},\mathbf{d}) + (\mathbf{b},\mathbf{c},\mathbf{d}),$$

$$(\alpha\,\mathbf{a},\mathbf{c},\mathbf{d}) = (\alpha\,\mathbf{a},[\mathbf{c},\mathbf{d}]) = \alpha\,(\mathbf{a},[\mathbf{c},\mathbf{d}]) = \alpha\,(\mathbf{a},\mathbf{c},\mathbf{d}).$$

Let's proceed to proving the fourth property of the mixed product in the theorem 47.1. Assume that the vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are coplanar. In this case they are parallel to some plane α in the space $\mathbb{E}$ and one can choose their geometric realizations lying on this plane α. If $\mathbf{b} \nparallel \mathbf{c}$, then the vector product $\mathbf{d} = [\mathbf{b},\mathbf{c}]$ is nonzero and perpendicular to the plane α. As for the vector $\mathbf{a}$, it is parallel to this plane. Hence $\mathbf{d} \perp \mathbf{a}$, which yields the equalities $(\mathbf{a},\mathbf{b},\mathbf{c}) = (\mathbf{a},[\mathbf{b},\mathbf{c}]) = (\mathbf{a},\mathbf{d}) = 0$.

If $\mathbf{b} \parallel \mathbf{c}$, then the vector product $[\mathbf{b},\mathbf{c}]$ is equal to zero and the equality $(\mathbf{a},\mathbf{b},\mathbf{c}) = 0$ is derived from $[\mathbf{b},\mathbf{c}] = 0$ with use of the initial formula (45.1) for the scalar product.

Now, conversely, assume that $(\mathbf{a},\mathbf{b},\mathbf{c}) = 0$. If $\mathbf{b} \parallel \mathbf{c}$, then the vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ determine not more than two directions in the space $\mathbb{E}$. For any two lines in this space always there is a plane to which these lines are parallel. In this case the vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are coplanar regardless to the equality $(\mathbf{a},\mathbf{b},\mathbf{c}) = 0$.

If $\mathbf{b} \nparallel \mathbf{c}$, then $\mathbf{d} = [\mathbf{b},\mathbf{c}] \neq \mathbf{0}$. Choosing geometric realizations of the vectors $\mathbf{b}$ and $\mathbf{c}$ with some common initial point O, we easily build a plane α comprising both of these geometric realizations. The vector $\mathbf{d} \neq \mathbf{0}$ in perpendicular to this plane α. Then from $(\mathbf{a},\mathbf{b},\mathbf{c}) = (\mathbf{a},[\mathbf{b},\mathbf{c}]) = (\mathbf{a},\mathbf{d}) = 0$ we derive $\mathbf{a} \perp \mathbf{d}$, which yields $\mathbf{a} \parallel \alpha$. The vectors $\mathbf{b}$ and $\mathbf{c}$ are also parallel to the plane α since their geometric realizations lie on this plane.

Hence all of the three vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are parallel to the plane α, which means that they are coplanar. The theorem 47.1 is completely proved. $\square$

THEOREM 47.2. *Apart from the properties 1)–4), the mixed product possesses the following four properties which are fulfilled for any four vectors* $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$ *and for any number* α*:*

5) $(\mathbf{c}, \mathbf{a}+\mathbf{b}, \mathbf{d}) = (\mathbf{c}, \mathbf{a}, \mathbf{d}) + (\mathbf{c}, \mathbf{b}, \mathbf{d})$;
6) $(\mathbf{c}, \alpha\,\mathbf{a}, \mathbf{d}) = \alpha\,(\mathbf{c}, \mathbf{a}, \mathbf{d})$;
7) $(\mathbf{c}, \mathbf{d}, \mathbf{a}+\mathbf{b}) = (\mathbf{c}, \mathbf{d}, \mathbf{a}) + (\mathbf{c}, \mathbf{d}, \mathbf{b})$;
8) $(\mathbf{c}, \mathbf{d}, \alpha\,\mathbf{a}) = \alpha\,(\mathbf{c}, \mathbf{d}, \mathbf{a})$.

DEFINITION 47.2. The properties 5) and 6) in the theorem 47.2 are called the properties of *linearity with respect to the second multiplicand*, the properties 7) and 8) are called the properties of *linearity with respect to the third multiplicand*.

The property 5) is derived from the property 2) in the theorem 47.1 in the following way:

$$\begin{aligned}(\mathbf{c}, \mathbf{a}+\mathbf{b}, \mathbf{d}) &= -(\mathbf{a}+\mathbf{b}, \mathbf{c}, \mathbf{d}) = -((\mathbf{a}, \mathbf{c}, \mathbf{d}) + (\mathbf{b}, \mathbf{c}, \mathbf{d})) = \\ &= -(\mathbf{a}, \mathbf{c}, \mathbf{d}) - (\mathbf{b}, \mathbf{c}, \mathbf{d}) = (\mathbf{c}, \mathbf{a}, \mathbf{d}) + (\mathbf{c}, \mathbf{b}, \mathbf{d}).\end{aligned}$$

The property 1) from this theorem is also used in the above calculations. As for the properties 6), 7), and 8) in the theorem 47.2, they are also easily derived from the properties 2) and 3) with the use of the property 1). Indeed, we have

$$(\mathbf{c}, \alpha\,\mathbf{a}, \mathbf{d}) = -(\alpha\,\mathbf{a}, \mathbf{c}, \mathbf{d}) = -\alpha\,(\mathbf{a}, \mathbf{c}, \mathbf{d}) = \alpha\,(\mathbf{c}, \mathbf{a}, \mathbf{d}),$$

$$\begin{aligned}(\mathbf{c}, \mathbf{d}, \mathbf{a}+\mathbf{b}) &= -(\mathbf{a}+\mathbf{b}, \mathbf{d}, \mathbf{c}) = -((\mathbf{a}, \mathbf{d}, \mathbf{c}) + (\mathbf{b}, \mathbf{d}, \mathbf{c})) = \\ &= -(\mathbf{a}, \mathbf{d}, \mathbf{c}) - (\mathbf{b}, \mathbf{d}, \mathbf{c}) = (\mathbf{c}, \mathbf{d}, \mathbf{a}) + (\mathbf{c}, \mathbf{d}, \mathbf{b}),\end{aligned}$$

$$(\mathbf{c}, \mathbf{d}, \alpha\,\mathbf{a}) = -(\alpha\,\mathbf{a}, \mathbf{d}, \mathbf{c}) = -\alpha\,(\mathbf{a}, \mathbf{d}, \mathbf{c}) = \alpha\,(\mathbf{c}, \mathbf{d}, \mathbf{a}).$$

The calculations performed prove the theorem 47.2.

§ 48. The concept of the oriented volume.

Let $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ be a right triple of non-coplanar vectors in the space $\mathbb{E}$. Let's consider their mixed product $(\mathbf{a}, \mathbf{b}, \mathbf{c})$. Due to the item 4) from the theorem 47.1 the non-coplanarity of the vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ means $(\mathbf{a}, \mathbf{b}, \mathbf{c}) \neq 0$, which in turn due to (45.1) implies $[\mathbf{b}, \mathbf{c}] \neq \mathbf{0}$.

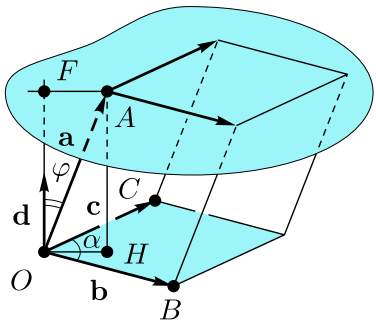

Fig. 48.1

Due to the item 4) from the theorem 39.1 the non-vanishing condition $[\mathbf{b}, \mathbf{c}] \neq \mathbf{0}$ means $\mathbf{b} \nparallel \mathbf{c}$. Let's build the geometric realizations of the non-collinear vectors $\mathbf{b}$ and $\mathbf{c}$ at some common initial point O and denote them $\mathbf{b} = \overrightarrow{OB}$ and $\mathbf{c} = \overrightarrow{OC}$. Then we build the geometric realization of the vector $\mathbf{a}$ at the same initial point O and denote it through $\mathbf{a} = \overrightarrow{OA}$. Let's complement the vectors $\overrightarrow{OA}$, $\overrightarrow{OB}$, and $\overrightarrow{OC}$ up to a skew-angular parallelepiped as shown in Fig. 48.1.

Let's denote $\mathbf{d} = [\mathbf{b}, \mathbf{c}]$. The vector $\mathbf{d}$ is perpendicular to the base plane of the parallelepiped, its length is calculated by the formula $|\mathbf{d}| = |\mathbf{b}|\,|\mathbf{c}|\,\sin\alpha$. It is easy to see that the length of $\mathbf{d}$ coincides with the base area of our parallelepiped, i. e. with the area of the parallelogram built on the vectors $\mathbf{b}$ and $\mathbf{c}$:

$$S = |\mathbf{d}| = |\mathbf{b}|\,|\mathbf{c}|\,\sin\alpha. \tag{48.1}$$

THEOREM 48.1. *For any two vectors the length of their vector product coincides with the area of the parallelogram built on these two vectors.*

The fact formulated in the theorem 48.1 is known as the *geometric interpretation of the vector product.*

Now let's return to Fig. 48.1. Applying the formula (45.1), for the mixed product $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ we derive

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = (\mathbf{a}, [\mathbf{b}, \mathbf{c}]) = (\mathbf{a}, \mathbf{d}) = |\mathbf{a}|\,|\mathbf{d}|\,\cos\varphi. \tag{48.2}$$

Note that $|\mathbf{a}|\cos\varphi$ is the length of the segment $[OF]$, which coincides with the length of $[AH]$. The segment $[AH]$ is parallel to the segment $[OF]$ and to the vector $\mathbf{d}$, whivh is perpendicular to the base plane of the skew-angular parallelepiped shown in Fig. 48.1. Hence the segment $[AH]$ represents the height of this parallelepiped ad we have the formula

$$h = |AH| = |\mathbf{a}|\,\cos\varphi. \tag{48.3}$$

Now from (48.1), (48.3), and (48.2) we derive

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = S\,h = V, \tag{48.4}$$

i. e. the mixed product $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ in our case coincides with the volume of the skew-angular parallelepiped built on the vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$.

In the general case the value of the mixed product of three non-coplanar vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ can be either positive or negative, while the volume of a parallelepiped is always positive. Therefore in the general case the formula (48.4) should be written as

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \begin{cases} V, & \text{if } \mathbf{a},\ \mathbf{b},\ \mathbf{c} \text{ is a right triple of vectors;} \\ -V, & \text{if } \mathbf{a},\ \mathbf{b},\ \mathbf{c} \text{ is a left triple of vectors.} \end{cases} \tag{48.5}$$

Definition 48.1. The oriented volume of an ordered triple of non-coplanar vectors is a quantity which is equal to the volume of the parallelepiped built on these vectors in the case where these vectors form a right triple and which is equal to the volume of

this parallelepiped taken with the minus sign in the case where these vectors form a left triple.

The formula (48.5) can be written as a theorem.

THEOREM 48.2. *The mixed product of any triple of non-coplanar vectors coincides with their oriented volume.*

DEFINITION 48.2. If $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ is a basis in the space $\mathbb{E}$, then the oriented volume of the triple of vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ is called the oriented volume of this basis.

§ 49. Structural constants of the mixed product.

Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ be some basis in the space $\mathbb{E}$. Let's consider various mixed products composed by the vectors of this basis:

$$c_{ijk} = (\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k). \tag{49.1}$$

DEFINITION 49.1. For any basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ in the space $\mathbb{E}$ the quantities c_{ijk} given by the formula (49.1) are called the *structural constants of the mixed product* in this basis.

The formula (49.1) is similar to the formula (29.6) for the components of the Gram matrix. However, the structural constants of the mixed product c_{ijk} in (49.1) constitute a three index array which cannot be laid into a matrix.

An important property of the structural constants c_{ijk} is their *complete skew symmetry* or *complete antisymmetry*. This property is expressed by the following equalities:

$$c_{ijk} = -c_{jik}, \qquad c_{ijk} = -c_{ikj}, \qquad c_{ijk} = -c_{kji}, \tag{49.2}$$

The relationships (49.2) mean that under the transposition of any two indices the quantity c_{ijk} changes its sign. These relationships are easily derived from (43.5) by applying the item 1) from the theorem 47.1 to the right hand side of (49.1).

The following relationships are an immediate consequence of the property of complete antisymmetry of the structural constants of the mixed product c_{ijk}:

$$c_{iik} = -c_{iik}, \qquad c_{ijj} = -c_{ijj}, \qquad c_{iji} = -c_{iji}, \tag{49.3}$$

They are derived by substituting $j = i$, $k = j$, and $k = i$ into (49.2). From the relationships (49.3) the vanishing condition for the structural constants c_{ijk} is derived:

$$c_{ijk} = 0, \text{ if there are coinciding values of the indices } i, j, k. \tag{49.4}$$

Now assume that the values of the indices i, j, k do not coincide. In this case, applying the relationships (49.2), we derive

$$c_{ijk} = c_{123} \text{ if the indices } i, j, k \text{ take the values } (1,2,3),\ (2,3,1), \text{ or } (3,1,2); \tag{49.5}$$

$$c_{ijk} = -c_{123} \text{ if the indices } i, j, k \text{ take the values } (1,3,2),\ (3,2,1), \text{ or } (2,1,3). \tag{49.6}$$

The next step consists in comparing the relationships (49.4), (49.5), and (49.6) with the formula (43.5) that determines the Levi-Civita symbol ε_{ijk}. Such a comparison yields

$$c_{ijk} = c_{123}\ \varepsilon_{ijk}. \tag{49.7}$$

Note that $c_{123} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. This formula follows from (49.1). Therefore the formula (49.7) can be written as

$$c_{ijk} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)\ \varepsilon_{ijk}. \tag{49.8}$$

THEOREM 49.1. *In an arbitrary basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ *the structural constants of the mixed product are expressed by the formula* (49.8) *through the only one constant — the oriented volume of the basis.*

§ 50. Calculation of the mixed product through the coordinates of vectors in a skew-angular basis.

Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ be a skew-angular basis in the space $\mathbb{E}$. Let's recall that the term *skew-angular basis* in this book is used as a synonym of an arbitrary basis (see Definition 29.1). Let $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ be free vectors given by their coordinates in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$:

$$\mathbf{a} = \begin{Vmatrix} a^1 \\ a^2 \\ a^3 \end{Vmatrix}, \qquad \mathbf{b} = \begin{Vmatrix} b^1 \\ b^2 \\ b^3 \end{Vmatrix}, \qquad \mathbf{c} = \begin{Vmatrix} c^1 \\ c^2 \\ c^3 \end{Vmatrix}. \tag{50.1}$$

The formulas (50.1) mean that we have the expansions

$$\mathbf{a} = \sum_{i=1}^{3} a^i\,\mathbf{e}_i, \qquad \mathbf{b} = \sum_{j=1}^{3} b^j\,\mathbf{e}_j, \qquad \mathbf{c} = \sum_{k=1}^{3} c^k\,\mathbf{e}_k. \tag{50.2}$$

Let's substitute (50.2) into the mixed product $(\mathbf{a}, \mathbf{b}, \mathbf{c})$:

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \Bigl(\sum_{i=1}^{3} a^i\,\mathbf{e}_i, \sum_{j=1}^{3} b^j\,\mathbf{e}_j, \sum_{k=1}^{3} c^k\,\mathbf{e}_k\Bigr). \tag{50.3}$$

In order to transform the formula (50.3) we apply the properties of the mixed product 2), 5), and 7) from the theorems 47.1 and 47.2. Due to these properties we can bring the summation signs over i, j, and k outside the brackets of the mixed product:

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} (a^i\,\mathbf{e}_i, b^j\,\mathbf{e}_j, c^k\,\mathbf{e}_k). \tag{50.4}$$

Now we apply the properties 3), 6), 8) from the theorems 47.1 and 47.2. Due to these properties we can bring the numeric factors a^i, b^j, c^k outside the brackets of the mixed product (50.4):

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} a^i\,b^j\,c^k\,(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k). \tag{50.5}$$

The quantities $(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k)$ are structural constants of the mixed product in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ (see (49.1)). Therefore the formula (50.5) can be written as follows:

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} a^i\, b^j\, c^k\, c_{ijk}. \tag{50.6}$$

Let's substitute (49.8) into (50.6) and take into account that the oriented volume $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ does not depend on the summation indices i, j, and k. Therefore the oriented volume $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ can be brought outside the sums as a common factor:

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} a^i\, b^j\, c^k\, \varepsilon_{ijk}. \tag{50.7}$$

Note that the formula (50.7) differs from the formula (46.9) only by the extra factor $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ in its right hand side. As for the formula (46.9), it is brought to the form (46.8) by applying the properties of the Levi-Civita symbol ε_{ijk} only. For this reason the formula (50.7) can be brought to the following form:

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \begin{vmatrix} a^1 & a^2 & a^3 \\ b^1 & b^2 & b^3 \\ c^1 & c^2 & c^3 \end{vmatrix}. \tag{50.8}$$

Definition 50.1. The formulas (50.6), (50.7), and (50.8) are called the *formulas for calculating the mixed product through the coordinates of vectors in a skew-angular basis.*

§ 51. The relation of structural constants of the vectorial and mixed products.

The structural constants of the mixed product are determined by the formula (49.1). Let's apply the formula (45.1) in order to

transform (49.1). As a result we get the formula

$$c_{ijk} = (\mathbf{e}_i, [\mathbf{e}_j, \mathbf{e}_k]). \tag{51.1}$$

Now we can apply the formula (40.2). Let's write it as follows:

$$[\mathbf{e}_j, \mathbf{e}_k] = \sum_{q=1}^{3} C_{jk}^q \, \mathbf{e}_q. \tag{51.2}$$

Substituting (51.2) into (51.1) and taking into account the properties 5) and 6) from the theorem 28.2, we derive

$$c_{ijk} = \sum_{q=1}^{3} C_{jk}^q \, (\mathbf{e}_i, \mathbf{e}_q). \tag{51.3}$$

Let's apply the formulas (29.6) and (30.1) to (51.3) and bring the formula (51.3) to the form

$$c_{ijk} = \sum_{q=1}^{3} C_{jk}^q \, g_{qi}. \tag{51.4}$$

The following formula is somewhat more beautiful:

$$c_{ijk} = \sum_{q=1}^{3} C_{ij}^q \, g_{qk}. \tag{51.5}$$

In order to derive the formula (51.5) we apply the identity $c_{ijk} = c_{jki}$ to the left hand side of the formula (51.4). This identity is derived from (49.2). Then we perform the cyclic redesignation of indices $i \to k \to j \to i$.

The formula (51.5) is the first formula relating the structural constants of the vectorial and mixed products. It is important from the theoretical point of view, but this formula is of little

use practically. Indeed, it expresses the structural constants of the mixed product through the structural constants of the vector product. But for the structural constants of the mixed product we already have the formula (49.8) which is rather efficient. As for the structural constants of the vector product, we have no formula yet, except the initial definition (40.2). For this reason we need to invert the formula (51.5) and express C_{ij}^q through c_{ijk}. In order to reach this goal we need some auxiliary information on the Gram matrix.

THEOREM 51.1. *The Gram matrix G of any basis in the space $\mathbb{E}$ is non-degenerate, i. e. its determinant is nonzero:* $\det G \neq 0$.

THEOREM 51.2. *For any basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ in the space $\mathbb{E}$ the determinant of the Gram matrix G is equal to the square of the oriented volume of this basis:*

$$\det G = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)^2. \tag{51.6}$$

The theorem 51.1 follows from the theorem 51.2. Indeed, a basis is a triple of non-coplanar vectors. From the non-coplanarity of the vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ due to item 4) of the theorem 47.1 we get $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \neq 0$. Then the formula (51.6) yields

$$\det G > 0, \tag{51.7}$$

while the theorem 51.1 follows from the inequality (51.7).

I will not prove the theorem 51.2 right now at this place. This theorem is proved below in § 56.

Let's proceed to deriving consequences from the theorem 51.1. It is known that each non-degenerate matrix has an inverse matrix (see [7]). Let's denote through G^{-1} the matrix inverse to the Gram matrix G. In writing the components of the matrix G^{-1} the following convention is used. Let's proceed to deriving consequences from the theorem 51.1. It is known that each non-degenerate matrix has an inverse matrix (see [7]). Let's

denote through G^{-1} the matrix inverse to the Gram matrix G. In writing the components of the matrix G^{-1} the following convention is used. Let's proceed to deriving consequences from the theorem 51.1. It is known that each non-degenerate matrix has an inverse matrix (see [7]). Let's denote through G^{-1} the matrix inverse to the Gram matrix G. In writing the components of the matrix G^{-1} the following convention is used.

Definition 51.1. For denoting the components of the matrix G^{-1} inverse to the Gram matrix G the same symbol g as for the components of the matrix G itself is used, but the components of the *inverse Gram matrix* are enumerated with two upper indices:

$$G^{-1} = \begin{Vmatrix} g^{11} & g^{12} & g^{13} \\ g^{21} & g^{22} & g^{23} \\ g^{31} & g^{32} & g^{33} \end{Vmatrix} \tag{51.8}$$

The matrices G and G^{-1} are inverse to each other. Their product in any order is equal to the unit matrix:

$$G \cdot G^{-1} = 1, \qquad G^{-1} \cdot G = 1. \tag{51.9}$$

From the regular course of algebra we know that each of the equalities (51.9) fixes the matrix G^{-1} uniquely once the matrix G is given (see. [7]). Now we apply the matrix transposition operation to both sides of the matrix equalities (51.9):

$$(G \cdot G^{-1})^{\top} = 1^{\top} = 1. \tag{51.10}$$

Then we use the identity $(A \cdot B)^{\top} = B^{\top} \cdot A^{\top}$ from the exercise 29.2 in order to transform the formula (51.10) and take into account the symmetry of the matrix G (see Theorem 30.1):

$$(G^{-1})^{\top} \cdot G^{\top} = (G^{-1})^{\top} \cdot G = 1. \tag{51.11}$$

The rest is to compare the equality (51.11) with the second matrix equality (51.9). This yields

$$(G^{-1})^{\top} = G^{-1} \tag{51.12}$$

The formula (51.12) can be written as a theorem.

THEOREM 51.3. *For any basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ *in the space* $\mathbb{E}$ *the matrix* G^{-1} *inverse to the Gram matrix of this basis is symmetric.*

In terms of the matrix components of the matrix (51.7) the equality (51.12) is written as an equality similar to (30.1):

$$g^{ij} = g^{ji}. \tag{51.13}$$

The second equality (51.9) is written as the following relationships for the components of the matrices (51.8) and (29.7):

$$\sum_{k=1}^{3} g^{sk}\, g_{kq} = \delta^s_q. \tag{51.14}$$

Here δ^s_q is the Kronecker symbol defined by the formula (23.3). Using the symmetry identities (51.13) and (30.1), we write the relationship (51.14) in the following form:

$$\sum_{k=1}^{3} g_{qk}\, g^{ks} = \delta^s_q. \tag{51.15}$$

Now let's multiply both ides of the equality (51.5) by g^{ks} and then perform summation over k in both sides of this equality:

$$\sum_{k=1}^{3} c_{ijk}\, g^{ks} = \sum_{k=1}^{3}\sum_{q=1}^{3} C^q_{ij}\, g_{qk}\, g^{ks}. \tag{51.16}$$

If we take into account the identity (51.15), then we can bring

the formula (51.16) to the following form:

$$\sum_{k=1}^{3} c_{ijk}\, g^{ks} = \sum_{q=1}^{3} C_{ij}^{q}\, \delta_{q}^{s}. \tag{51.17}$$

When summing over q in the right hand side of the equality (51.17) the index q runs over three values 1, 2, 3. Then the Kronecker symbol δ_q^s takes the values 0 and 1, the value 1 is taken only once when $q = s$. This means that only one of three summands in the right hand side of (51.17) is nonzero. This nonzero summand is equal to C_{ij}^{s}. Hence the formula (51.17) can be written in the following form:

$$\sum_{k=1}^{3} c_{ijk}\, g^{ks} = C_{ij}^{s}. \tag{51.18}$$

Let's change the symbol k for q and the symbol s for k in (51.18). Then we transpose left and right hand sides of this formula:

$$C_{ij}^{k} = \sum_{q=1}^{3} c_{ijq}\, g^{qk}. \tag{51.19}$$

The formula (51.19) is the second formula relating the structural constants of the vectorial and mixed products. On the base of the relationships (51.5) and (51.19) we formulate a theorem.

THEOREM 51.4. *For any basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ *in the space* $\mathbb{E}$ *the structural constants of the vectorial and mixed products in this basis are related to each other in a one-to-one manner by means of the formulas* (51.5) *and* (51.19).

§ 52. Effectivization of the formulas for calculating vectorial and mixed products.

Let's consider the formula (29.8) for calculating the scalar product in a skew-angular basis. Apart from the coordinates of

vectors, this formula uses the components of the Gram matrix (29.7). In order to get the components of this matrix one should calculate the mutual scalar products of the basis vectors (see formula (29.6)), for this purpose one should measure their lengths and the angles between them (see Definition 26.1). No other geometric constructions are required. For this reason the formula (29.8) is recognized to be effective.

Now let's consider the formula (41.5) for calculating the vector product in a skew-angular basis. This formula uses the structural constants of the vector product which are defined by means of the formula (40.2). According to this formula, in order to calculate the structural constants one should calculate the vector products $[\mathbf{e}_i, \mathbf{e}_j]$ in its left hand side. For this purpose one should construct the normal vectors (perpendiculars) to the planes given by various pairs of basis vectors $\mathbf{e}_i$, $\mathbf{e}_j$. (see Definition 35.1). Upon calculating the vector products $[\mathbf{e}_i, \mathbf{e}_j]$ one should expand them in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, which require some auxiliary geometric constructions (see formula (18.4) and Fig. 18.1). For this reason the efficiency of the formula (41.5) is much less than the efficiency of the formula (29.8).

And finally we consider the formula (50.7) for calculating the mixed product in a skew-angular basis. In order to apply this formula one should know the value of the mixed product of the three basis vectors $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. It is called the oriented volume of a basis (see Definition 48.2). Due to the theorem 48.2 and the definition 48.1 for this purpose one should calculate the volume of the skew-angular parallelepiped built on the basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. In order to calculate the volume of this parallelepiped one should know its base area and its height. The area of its base is effectively calculated by the lengths of two basis vectors and the angle between them (see formula (48.1)). As for the height of the parallelepiped, in order to find it one should drop a perpendicular from one of its vertices to its base plane. Since we need such an auxiliary geometric construction, the formula (50.7) is less effective as compared to the formula (29.8) in the case of

the scalar product.

In order to make the formulas (41.5) and (50.7) effective we use the formula (51.6). It leads to the following relationship:

$$(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = \pm\sqrt{\det G}. \tag{52.1}$$

The sign in (52.1) is determined by the orientation of a basis:

$$(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = \begin{cases} \sqrt{\det G} & \text{if the basis } \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \\ & \text{is right;} \\ -\sqrt{\det G} & \text{if the basis } \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \\ & \text{is left.} \end{cases} \tag{52.2}$$

Let's substitute the expression (52.1) into the formula for the mixed product (50.7). As a result we get

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \pm\sqrt{\det G} \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} a^i\, b^j\, c^k\, \varepsilon_{ijk}. \tag{52.3}$$

Similarly, substituting (52.1) into the formula (50.8), we get

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \pm\sqrt{\det G} \begin{vmatrix} a^1 & a^2 & a^3 \\ b^1 & b^2 & b^3 \\ c^1 & c^2 & c^3 \end{vmatrix}. \tag{52.4}$$

Definition 52.1. The formulas (52.3) and (52.4) are called the effectivized formulas for calculating the mixed product through the coordinates of vectors in a skew-angular basis.

In the case of the formula (41.5), in order to make it effective we need the formulas (49.8) and (51.19). Substituting the expression (52.1) into these formulas, we obtain

$$c_{ijk} = \pm\sqrt{\det G}\; \varepsilon_{ijk}, \tag{52.5}$$

$$C_{ij}^{k} = \pm\sqrt{\det G} \sum_{q=1}^{3} \varepsilon_{ijq}\, g^{qk}. \tag{52.6}$$

DEFINITION 52.2. The formulas (52.5) and (52.6) are called the effectivized formulas for calculating the structural constants of the mixed and vectorial products.

Now let's substitute the formula (52.6) into (41.5). This leads to the following formula for the vector product:

$$[\mathbf{a},\mathbf{b}] = \pm\sqrt{\det G}\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{q=1}^{3} a^i\, b^j\, \varepsilon_{ijq}\, g^{qk}\, \mathbf{e}_k. \tag{52.7}$$

DEFINITION 52.3. The formula (52.7) is called the effectivized formula for calculating the vector product through the coordinates of vectors in a skew-angular basis.

§ 53. Orientation of the space.

Let's consider the effectivized formulas (52.1), (52.3), (52.4), (52.5), (52.6), and (52.7) from § 52. Almost all information on a basis in these formulas is given by the Gram matrix G. The Gram matrix arising along with the choice of a basis reflects an important property of the space $\mathbb{E}$ — its metric.

DEFINITION 53.1. The *metric* of the space $\mathbb{E}$ is its structure (its feature) that consists in possibility to measure the lengths of segments and the numeric values of angles in it.

The only non-efficiency remaining in the effectivized formulas (52.1), (52.3), (52.4), (52.5), (52.6), and (52.7) is the choice of sign in them. As was said above, the sign in these formulas is determined by the orientation of a basis (see formula (52.2)). There is absolutely no possibility to determine the orientation of a basis through the information comprised in its Gram matrix. The matter is that the mathematical space $\mathbb{E}$ described by Euclid's axioms (see [6]), comprises the possibility to distinguish a pair of bases with different orientations from a pair of bases with coinciding orientations. However, it does not contain any reasons for to prefer bases with one of two possible orientations.

The concept of right triple of vectors (see Definition 34.2) and the possibility to distinguish right triples of vectors from left ones is due to the presence of people, it is due to their ability to observe vectors and compare their rotation with the rotation of clock hands. Is there a *fundamental asymmetry between left and right* not depending on the presence of people and on other non-fundamental circumstances? Is the space *fundamentally oriented*? This is a question on the nature of the physical space $\mathbb{E}$. Some research in the field of elementary particle physics says that such an asymmetry does exist. As for me, as the author of this book I cannot definitely assert that this problem is finally resolved.

§ 54. Contraction formulas.

Contraction formulas is a collection of four purely algebraic identities relating the Levi-Civita symbol and the Kronecker symbol with each other.

THEOREM 54.1. *The Levi-Civita symbol and the Kronecker symbol are related by the first contraction formula*

$$\varepsilon^{mnp}\,\varepsilon_{ijk} = \begin{vmatrix} \delta_i^m & \delta_j^m & \delta_k^m \\ \delta_i^n & \delta_j^n & \delta_k^n \\ \delta_i^p & \delta_j^p & \delta_k^p \end{vmatrix}. \tag{54.1}$$

PROOF. Let's denote through f_{ijk}^{mnp} the right hand side of the first contraction formula (54.1). The transposition of any two lower indices in f_{ijk}^{mnp} is equivalent to the corresponding transposition of two columns in the matrix (54.1). It is known that the transposition of two columns of a matrix results in changing the sign of its determinant. This yield the following relationships for the quantities f_{ijk}^{mnp}:

$$f_{ijk}^{mnp} = -f_{jik}^{mnp}, \quad f_{ijk}^{mnp} = -f_{ikj}^{mnp}, \quad f_{ijk}^{mnp} = -f_{kji}^{mnp}. \tag{54.2}$$

The relationships (54.2) are analogous to the relationships (49.2). Repeating the considerations used in § 49 when deriving the formulas (49.4), (49.5), (49.6), from the formula (54.2) we derive

$$f^{mnp}_{ijk} = \begin{cases} 0 & \text{if there are coinciding values of the indices } i,j,k; \\ f^{mnp}_{123} & \text{if the values of the indices } i,j,k \text{ form an even permutation of the numbers } 1,2,3; \\ -f^{mnp}_{123}, & \text{if the values of the indices } i,j,k \text{ form an odd permutation of the numbers } 1,2,3. \end{cases} \tag{54.3}$$

Let's compare the formula (54.3) with the formula (43.5) defining the Levi-Civita symbol. Such a comparison yields

$$f^{mnp}_{ijk} = f^{mnp}_{123}\,\varepsilon_{ijk}. \tag{54.4}$$

Like the initial quantities f^{mnp}_{ijk}, the factor f^{mnp}_{123} in (54.4) is defined as the determinant of a matrix:

$$f^{mnp}_{123} = \begin{vmatrix} \delta^m_1 & \delta^m_2 & \delta^m_3 \\ \delta^n_1 & \delta^n_2 & \delta^n_3 \\ \delta^p_1 & \delta^p_2 & \delta^p_3 \end{vmatrix}. \tag{54.5}$$

Transposition of any pair of the upper indices in f^{mnp}_{123} is equivalent to the corresponding transposition of rows in the matrix (54.5). Again we know that the transposition of any two rows of a matrix changes the sign of its determinant. Hence we derive the following relationships for the quantities (54.5):

$$f^{mnp}_{123} = -f^{nmp}_{123}, \quad f^{mnp}_{123} = -f^{mpn}_{123}, \quad f^{mnp}_{123} = -f^{pnm}_{123} \tag{54.6}$$

The relationships (54.6) are analogous to the relationships (54.2), which are analogous to (49.2). Repeating the considerations used

in § 49, from the relationships (54.6) we immediately derive the formula analogous to the formula (54.3):

$$f_{123}^{mnp} = \begin{cases} 0 & \text{if there are coinciding values of the indices } m, n, p; \\ f_{123}^{123} & \text{if the values of the indices } m, n, p \text{ form an even permutation of the numbers } 1, 2, 3; \\ -f_{123}^{123} & \text{if the values of the indices } m, n, p \text{ form an odd permutation of the numbers } 1, 2, 3. \end{cases} \tag{54.7}$$

Let's compare the formula (54.7) with the formula (43.5) defining the Levi-Civita symbol. This comparison yields

$$f_{123}^{mnp} = f_{123}^{123}\, \varepsilon^{mnp}. \tag{54.8}$$

Let's combine the formulas (54.4) and (54.8), i. e. we substitute (54.8) into (54.4). This substitution leads to the formula

$$f_{ijk}^{mnp} = f_{123}^{123}\, \varepsilon^{mnp}\, \varepsilon_{ijk}. \tag{54.9}$$

Now we need to calculate the coefficient f_{123}^{123} in the formula (54.9). Let's recall that the quantity f_{ijk}^{mnp} was defined as the right hand side of the formula (54.1). Therefore we can write

$$f_{ijk}^{mnp} = \begin{vmatrix} \delta_i^m & \delta_j^m & \delta_k^m \\ \delta_i^n & \delta_j^n & \delta_k^n \\ \delta_i^p & \delta_j^p & \delta_k^p \end{vmatrix} \tag{54.10}$$

and substitute $i = m = 1$, $j = n = 2$, $k = p = 3$ into (54.10):

$$f_{123}^{123} = \begin{vmatrix} \delta_1^1 & \delta_2^1 & \delta_3^1 \\ \delta_1^2 & \delta_2^2 & \delta_3^2 \\ \delta_1^3 & \delta_2^3 & \delta_3^3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1. \tag{54.11}$$

Taking into account the value of the coefficient f^{123}_{123} given by the formula (54.11), we can transform the formula (54.9) to

$$f^{mnp}_{ijk} = \varepsilon^{mnp}\,\varepsilon_{ijk}. \tag{54.12}$$

Now the required contraction formula (54.1) is obtained as a consequence of (54.10) and (54.12). The theorem 54.1 is proved. □

THEOREM 54.2. *The Levi-Civita symbol and the Kronecker symbol are related by the second contraction formula*

$$\sum_{k=1}^{3} \varepsilon^{mnk}\,\varepsilon_{ijk} = \begin{vmatrix} \delta^m_i & \delta^m_j \\ \delta^n_i & \delta^n_j \end{vmatrix}. \tag{54.13}$$

PROOF. The second contraction formula (54.13) is derived from the first contraction formula (54.1). For this purpose we substitute $p = k$ into the formula (54.1) and take into account that $\delta^k_k = 1$. This yields the formula

$$\varepsilon^{mnk}\,\varepsilon_{ijk} = \begin{vmatrix} \delta^m_i & \delta^m_j & \delta^m_k \\ \delta^n_i & \delta^n_j & \delta^n_k \\ \delta^k_i & \delta^k_j & 1 \end{vmatrix}. \tag{54.14}$$

Let's insert the summation over k to both sides of the formula (54.14). As a result we get the formula

$$\sum_{k=1}^{3} \varepsilon^{mnk}\,\varepsilon_{ijk} = \sum_{k=1}^{3} \begin{vmatrix} \delta^m_i & \delta^m_j & \delta^m_k \\ \delta^n_i & \delta^n_j & \delta^n_k \\ \delta^k_i & \delta^k_j & 1 \end{vmatrix}. \tag{54.15}$$

The left hand side of the obtained formula (54.15) coincides with the left hand side of the second contraction formula (54.13) to be proved. For this reason below we transform the right hand side of the formula (54.15) only. Let's expand the determinant in

the formula (54.15) by its last row:

$$\sum_{k=1}^{3}\varepsilon^{mnk}\,\varepsilon_{ijk}=\sum_{k=1}^{3}\left(\delta_i^k\begin{vmatrix}\delta_j^m & \delta_k^m\\ \delta_j^n & \delta_k^n\end{vmatrix}-\right.$$
$$\left.-\,\delta_j^k\begin{vmatrix}\delta_i^m & \delta_k^m\\ \delta_i^n & \delta_k^n\end{vmatrix}+\begin{vmatrix}\delta_i^m & \delta_j^m\\ \delta_i^n & \delta_j^n\end{vmatrix}\right). \tag{54.16}$$

The summation over k in the right hand side of the formula (54.16) applies to all of the three terms enclosed in the round brackets. Expanding these brackets, we get

$$\sum_{k=1}^{3}\varepsilon^{mnk}\,\varepsilon_{ijk}=\sum_{k=1}^{3}\delta_i^k\begin{vmatrix}\delta_j^m & \delta_k^m\\ \delta_j^n & \delta_k^n\end{vmatrix}-$$
$$-\sum_{k=1}^{3}\delta_j^k\begin{vmatrix}\delta_i^m & \delta_k^m\\ \delta_i^n & \delta_k^n\end{vmatrix}+\sum_{k=1}^{3}\begin{vmatrix}\delta_i^m & \delta_j^m\\ \delta_i^n & \delta_j^n\end{vmatrix}. \tag{54.17}$$

The first sum in the right hand side of (54.17) contains the factor δ_i^k. In performing the summation cycle over k this factor appears to be nonzero only once when $k=i$. For this reason only one summand of the first sum does actually survive. In this term $k=i$. Similarly, in the second sum also only one its term does actually survive, in this term $k=j$. As for the last sum in the right hand side of (54.17), the expression being summed does not depend on k. Therefore it triples upon calculating this sum. Taking into account that $\delta_i^i=\delta_j^j=1$, we get

$$\sum_{k=1}^{3}\varepsilon^{mnk}\,\varepsilon_{ijk}=\begin{vmatrix}\delta_j^m & \delta_i^m\\ \delta_j^n & \delta_i^n\end{vmatrix}-\begin{vmatrix}\delta_i^m & \delta_j^m\\ \delta_i^n & \delta_j^n\end{vmatrix}+3\begin{vmatrix}\delta_i^m & \delta_j^m\\ \delta_i^n & \delta_j^n\end{vmatrix}.$$

The first determinant in the right hand side of the above formula differs from two others by the transposition of its columns. If we perform this transposition once more, it changes its sign and

we get a formula with three coinciding determinants. Then, collecting the similar terms, we derive

$$\sum_{k=1}^{3} \varepsilon^{mnk}\, \varepsilon_{ijk} = (-1-1+3) \begin{vmatrix} \delta_i^m & \delta_j^m \\ \delta_i^n & \delta_j^n \end{vmatrix}. \tag{54.18}$$

Now it is easy to see that the formula (54.18) leads to the required formula (54.13). The theorem 54.2 is proved. □

THEOREM 54.3. *The Levi-Civita symbol and the Kronecker symbol are related by the third contraction formula*

$$\sum_{j=1}^{3}\sum_{k=1}^{3} \varepsilon^{mjk}\, \varepsilon_{ijk} = 2\,\delta_i^m. \tag{54.19}$$

PROOF. The third contraction formula (54.19) is derived from the second contraction formula (54.13). For this purpose let's substitute $n = j$ into the formula (54.13), take into account that $\delta_j^j = 1$, and insert the summation over j into both sides of the obtained equality. This yields the formula

$$\sum_{j=1}^{3}\sum_{k=1}^{3} \varepsilon^{mjk}\, \varepsilon_{ijk} = \sum_{j=1}^{3} \begin{vmatrix} \delta_i^m & \delta_j^m \\ \delta_i^j & 1 \end{vmatrix} = \sum_{j=1}^{3} (\delta_i^m - \delta_j^m\, \delta_i^j).$$

Upon expanding the brackets the sum over j in the right hand side of the above formula can be calculated explicitly:

$$\sum_{j=1}^{3}\sum_{k=1}^{3} \varepsilon^{mjk}\, \varepsilon_{ijk} = \sum_{j=1}^{3} \delta_i^m - \sum_{j=1}^{3} \delta_j^m\, \delta_i^j = 3\,\delta_i^m - \delta_i^m = 2\,\delta_i^m.$$

It is easy to see that the calculations performed prove the formula (54.19) and the theorem 54.3 in whole. □

THEOREM 54.4. *The Levi-Civita symbol and the Kronecker symbol are related by the fourth contraction formula*

$$\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\varepsilon^{ijk}\,\varepsilon_{ijk}=6. \tag{54.20}$$

PROOF. The fourth contraction formula (54.20) is derived from the third contraction formula (54.19). For this purpose we substitute $m=i$ into (54.19) and insert the summation over i into both sides of the obtained equality. This yields

$$\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\varepsilon^{ijk}\,\varepsilon_{ijk}=2\sum_{i=1}^{3}\delta_i^i=2\sum_{i=1}^{3}1=2\cdot 3=6.$$

The above calculations prove the formula (54.20), which completes the proof of the theorem 54.4. □

§ 55. The triple product expansion formula and the Jacobi identity.

THEOREM 55.1. *For any triple of free vectors* $\mathbf{a}$, $\mathbf{b}$, *and* $\mathbf{c}$ *in the space* $\mathbb{E}$ *the following identity is fulfilled:*

$$[\mathbf{a},[\mathbf{b},\mathbf{c}]]=\mathbf{b}\,(\mathbf{a},\mathbf{c})-\mathbf{c}\,(\mathbf{a},\mathbf{b}), \tag{55.1}$$

which is known as the triple product expansion formula[1].

PROOF. In order to prove the identity (55.1) we choose some right orthonormal basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ in the space $\mathbb{E}$ and let's expand the vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ in this basis:

$$\mathbf{a}=\sum_{i=1}^{3}a^i\,\mathbf{e}_i,\qquad \mathbf{b}=\sum_{j=1}^{3}b^j\,\mathbf{e}_j,\qquad \mathbf{c}=\sum_{k=1}^{3}c^k\,\mathbf{e}_k. \tag{55.2}$$

[1] In Russian literature the triple product expansion formula is known as the double vectorial product formula or the «BAC minus CAB» formula.

Let's denote $\mathbf{d} = [\mathbf{b}, \mathbf{c}]$ and use the formula (44.1) for to calculate the vector $\mathbf{d}$. We write this formula as

$$\mathbf{d} = \sum_{k=1}^{3}\left(\sum_{i=1}^{3}\sum_{j=1}^{3} b^i\, c^j\, \varepsilon_{ijk}\right)\mathbf{e}_k. \tag{55.3}$$

The formula (55.3) is the expansion of the vector $\mathbf{d}$ in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. Therefore we can get its coordinates:

$$d^k = \sum_{i=1}^{3}\sum_{j=1}^{3} b^i\, c^j\, \varepsilon_{ijk}. \tag{55.4}$$

Now we again apply the formula (44.1) in order to calculate the vector $[\mathbf{a}, [\mathbf{b}, \mathbf{c}]] = [\mathbf{a}, \mathbf{d}]$. In this case we write it as follows:

$$[\mathbf{a}, \mathbf{d}] = \sum_{n=1}^{3}\sum_{m=1}^{3}\sum_{k=1}^{3} a^m\, d^k\, \varepsilon_{mkn}\, \mathbf{e}_n. \tag{55.5}$$

Substituting (55.4) into the formula (55.5), we get

$$\begin{aligned}[\mathbf{a}, \mathbf{d}] &= \sum_{n=1}^{3}\sum_{m=1}^{3}\sum_{k=1}^{3}\sum_{i=1}^{3}\sum_{j=1}^{3} a^m\, b^i\, c^j\, \varepsilon_{ijk}\, \varepsilon_{mkn}\, \mathbf{e}_n = \\ &= \sum_{n=1}^{3}\sum_{m=1}^{3}\sum_{i=1}^{3}\sum_{j=1}^{3} a^m\, b^i\, c^j \left(\sum_{k=1}^{3} \varepsilon_{ijk}\, \varepsilon_{mkn}\right)\mathbf{e}_n.\end{aligned} \tag{55.6}$$

The upper or lower position of indices in the Levi-Civita symbol does not matter (see (43.5)). Therefore, taking into account (43.8), we can write $\varepsilon_{mkn} = \varepsilon^{mkn} = -\varepsilon^{mnk}$ and bring the formula (55.6) to the following form:

$$[\mathbf{a}, \mathbf{d}] = -\sum_{n=1}^{3}\sum_{m=1}^{3}\sum_{i=1}^{3}\sum_{j=1}^{3} a^m\, b^i\, c^j \left(\sum_{k=1}^{3} \varepsilon^{mnk}\, \varepsilon_{ijk}\right)\mathbf{e}_n. \tag{55.7}$$

The sum enclosed in the round brackets in (55.7) coincides with the left hand side of the second contraction formula (54.13). Applying (54.13), we continue transforming the formula (55.7):

$$
\begin{aligned}
[\mathbf{a},\mathbf{d}] &= -\sum_{n=1}^{3}\sum_{m=1}^{3}\sum_{i=1}^{3}\sum_{j=1}^{3} a^m\, b^i\, c^j \begin{vmatrix} \delta_i^m & \delta_j^m \\ \delta_i^n & \delta_j^n \end{vmatrix} \mathbf{e}_n = \\
&= \sum_{n=1}^{3}\sum_{m=1}^{3}\sum_{i=1}^{3}\sum_{j=1}^{3} a^m\, b^i\, c^j \left(\delta_i^n\,\delta_j^m - \delta_i^m\,\delta_j^n\right)\mathbf{e}_n = \\
&= \sum_{n=1}^{3}\sum_{m=1}^{3}\sum_{i=1}^{3}\sum_{j=1}^{3} a^m\, b^i\, c^j\, \delta_i^n\,\delta_j^m\,\mathbf{e}_n - \\
&\qquad - \sum_{n=1}^{3}\sum_{m=1}^{3}\sum_{i=1}^{3}\sum_{j=1}^{3} a^m\, b^i\, c^j\, \delta_i^m\,\delta_j^n\,\mathbf{e}_n = \\
&= \sum_{i=1}^{3}\sum_{j=1}^{3} a^j\, b^i\, c^j\,\mathbf{e}_i - \sum_{i=1}^{3}\sum_{j=1}^{3} a^i\, b^i\, c^j\,\mathbf{e}_j.
\end{aligned}
$$

The result obtained can be written as follows:

$$
[\mathbf{a},[\mathbf{b},\mathbf{c}]] = \left(\sum_{j=1}^{3} a^j\, c^j\right)\left(\sum_{i=1}^{3} b^i\,\mathbf{e}_i\right) - \left(\sum_{i=1}^{3} a^i\, b^i\right)\left(\sum_{j=1}^{3} c^j\,\mathbf{e}_j\right).
$$

Mow let's recall that the scalar product in an orthonormal basis is calculated by the formula (33.3). The formula (33.3) is easily recognized within the above relationship. Taking into account this formula, we get the equality

$$
[\mathbf{a},[\mathbf{b},\mathbf{c}]] = (\mathbf{a},\mathbf{c})\sum_{i=1}^{3} b^i\,\mathbf{e}_i - (\mathbf{a},\mathbf{b})\sum_{j=1}^{3} c^j\,\mathbf{e}_j. \tag{55.8}
$$

In order to bring (55.8) to the ultimate form (55.1) it is sufficient to find the expansions of the form (55.2) for $\mathbf{b}$ and $\mathbf{c}$ within the

above formula (55.8). As a result the formula (55.8) takes the form $[\mathbf{a},[\mathbf{b},\mathbf{c}]] = (\mathbf{a},\mathbf{c})\,\mathbf{b} - (\mathbf{a},\mathbf{b})\,\mathbf{c}$, which in essential coincides with (55.1). The theorem 55.1 is proved. □

THEOREM 55.2. *For any triple of free vectors* $\mathbf{a}$, $\mathbf{b}$, *and* $\mathbf{c}$ *in the space* $\mathbb{E}$ *the Jacobi identity is fulfilled:*

$$[\mathbf{a},[\mathbf{b},\mathbf{c}]] + [\mathbf{b},[\mathbf{c},\mathbf{a}]] + [\mathbf{c},[\mathbf{a},\mathbf{b}]] = \mathbf{0}. \tag{55.9}$$

PROOF. The Jacoby identity (55.9) is easily derived with the use of the triple product expansion formula (55.1):

$$[\mathbf{a},[\mathbf{b},\mathbf{c}]] = \mathbf{b}\,(\mathbf{a},\mathbf{c}) - \mathbf{c}\,(\mathbf{a},\mathbf{b}). \tag{55.10}$$

Let's twice perform the cyclic redesignation of vectors $\mathbf{a} \to \mathbf{b} \to \mathbf{c} \to \mathbf{a}$ in the above formula (55.10). This yields

$$[\mathbf{b},[\mathbf{c},\mathbf{a}]] = \mathbf{c}\,(\mathbf{b},\mathbf{a}) - \mathbf{a}\,(\mathbf{b},\mathbf{c}), \tag{55.11}$$

$$[\mathbf{c},[\mathbf{a},\mathbf{b}]] = \mathbf{a}\,(\mathbf{c},\mathbf{b}) - \mathbf{b}\,(\mathbf{c},\mathbf{a}). \tag{55.12}$$

By adding the above three equalities (55.10), (55.11), and (55.12) in the left hand side we get the required expression $[\mathbf{a},[\mathbf{b},\mathbf{c}]] + [\mathbf{b},[\mathbf{c},\mathbf{a}]] + [\mathbf{c},[\mathbf{a},\mathbf{b}]]$, while th right hand side of the resulting equality vanishes. The theorem 55.2 is proved. □

§ 56. The product of two mixed products.

THEOREM 56.1. *For any six free vectors* $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{x}$, $\mathbf{y}$, *and* $\mathbf{z}$ *in the space* $\mathbb{E}$ *the following formula for the product of two mixed products* $(\mathbf{a},\mathbf{b},\mathbf{c})$ *and* $(\mathbf{x},\mathbf{y},\mathbf{z})$ *is fulfilled:*

$$(\mathbf{a},\mathbf{b},\mathbf{c})\,(\mathbf{x},\mathbf{y},\mathbf{z}) = \begin{vmatrix} (\mathbf{a},\mathbf{x}) & (\mathbf{b},\mathbf{x}) & (\mathbf{c},\mathbf{x}) \\ (\mathbf{a},\mathbf{y}) & (\mathbf{b},\mathbf{y}) & (\mathbf{c},\mathbf{y}) \\ (\mathbf{a},\mathbf{z}) & (\mathbf{b},\mathbf{z}) & (\mathbf{c},\mathbf{z}) \end{vmatrix}. \tag{56.1}$$

In order to prove the formula (56.1) we need two properties of the matrix determinants which are known as the *linearity with respect to a row* and the *linearity with respect to a column* (see [7]). The first of them can be expressed by the formula

$$\sum_{i=1}^{r}\alpha_i\begin{vmatrix} x_1^1 & \dots & x_n^1 \\ \vdots & \vdots & \vdots \\ x_1^k(i) & \dots & x_n^k(i) \\ \vdots & \vdots & \vdots \\ x_1^n & \dots & x_n^n \end{vmatrix} = \begin{vmatrix} x_1^1 & \dots & x_n^1 \\ \vdots & \vdots & \vdots \\ \sum\limits_{i=1}^{r}\alpha_i\,x_1^k(i) & \dots & \sum\limits_{i=1}^{r}\alpha_i\,x_n^k(i) \\ \vdots & \vdots & \vdots \\ x_1^n & \dots & x_n^n \end{vmatrix}.$$

LEMMA 56.1. *If the k-th row of a square matrix is a linear combination of some r rows (non necessarily coinciding with the other rows of this matrix), then the determinant of such a matrix is equal to a linear combination of r separate determinants.*

The linearity with respect to a column is formulated similarly.

LEMMA 56.2. *If the k-th column of a square matrix is a linear combination of some r columns (non necessarily coinciding with the other columns of this matrix), then the determinant of such a matrix is equal to a linear combination of r separate determinants.*

I do not write the formula illustrating the lemma 56.2 since it does not fit the width of a page in this book. This formula can be obtained from the above formula by transposing the matrices in both its sides.

PROOF OF THE THEOREM 56.1. Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ be some right orthonormal basis in the space $\mathbb{E}$. Assume that the vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ are given by their coordinates in this basis. Let's denote through L the left hand side of the formula (56.1):

$$L = (\mathbf{a},\mathbf{b},\mathbf{c})\,(\mathbf{x},\mathbf{y},\mathbf{z}). \tag{56.2}$$

In order to calculate L we apply the formula (46.9). In the case of the vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ the formula (46.9) yields

$$(\mathbf{a},\mathbf{b},\mathbf{c}) = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} a^i\, b^j\, c^k\, \varepsilon_{ijk}. \tag{56.3}$$

In the case of the vectors $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ the formula (46.9) is written as

$$(\mathbf{x},\mathbf{y},\mathbf{z}) = \sum_{m=1}^{3}\sum_{n=1}^{3}\sum_{p=1}^{3} x^m\, y^n\, z^p\, \varepsilon^{mnp}. \tag{56.4}$$

Note that raising indices of the Levi-Civita symbol in (56.4) does not change its values (see (43.5)). Now, multiplying the formulas (56.3) and (56.4), we obtain the formula for L:

$$L = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{m=1}^{3}\sum_{n=1}^{3}\sum_{p=1}^{3} a^i\, b^j\, c^k\, x^m\, y^n\, z^p\, \varepsilon^{mnp}\, \varepsilon_{ijk}. \tag{56.5}$$

The product $\varepsilon^{mnp}\,\varepsilon_{ijk}$ in (56.5) can be replaced by the matrix determinant taken from the first contraction formula (54.1):

$$L = \sum_{\substack{i=1\\m=1}}^{3}\ldots\sum_{\substack{k=1\\p=1}}^{3} a^i\, b^j\, c^k\, x^m\, y^n\, z^p \begin{vmatrix} \delta_i^m & \delta_j^m & \delta_k^m \\ \delta_i^n & \delta_j^n & \delta_k^n \\ \delta_i^p & \delta_j^p & \delta_k^p \end{vmatrix}. \tag{56.6}$$

The next step consists in applying the lemmas 56.1 and 56.2 in order to transform the formula (56.6). Applying the lemma 56.2, we bring the sum over i and the associated factor a^i into the first column of the determinant. Similarly, we bring the sum over j and the factor b^j into the second column, and finally, we bring the sum over k and its associated factor c^k into the third column of the determinant. Then we apply the lemma 56.1 in order to distribute the sums over m, n, and p to the rows of the

determinant. Simultaneously, we distribute the associated factors x^m, y^n, and z^p to the rows of the determinant. As a result of our efforts we get the following formula:

$$L = \begin{vmatrix} \sum\limits_{i=1}^{3}\sum\limits_{m=1}^{3} a^i\, x^m\, \delta_i^m & \sum\limits_{j=1}^{3}\sum\limits_{m=1}^{3} b^j\, x^m\, \delta_j^m & \sum\limits_{k=1}^{3}\sum\limits_{m=1}^{3} c^k\, x^m\, \delta_k^m \\ \sum\limits_{i=1}^{3}\sum\limits_{n=1}^{3} a^i\, y^n\, \delta_i^n & \sum\limits_{j=1}^{3}\sum\limits_{n=1}^{3} b^j\, y^n\, \delta_j^n & \sum\limits_{k=1}^{3}\sum\limits_{n=1}^{3} c^k\, y^n\, \delta_k^n \\ \sum\limits_{i=1}^{3}\sum\limits_{p=1}^{3} a^i\, z^p\, \delta_i^p & \sum\limits_{j=1}^{3}\sum\limits_{p=1}^{3} b^j\, z^p\, \delta_j^p & \sum\limits_{k=1}^{3}\sum\limits_{p=1}^{3} c^k\, z^p\, \delta_k^p \end{vmatrix}.$$

Due to the Kronecker symbols the double sums in the above formula are reduced to single sums:

$$L = \begin{vmatrix} \sum\limits_{i=1}^{3} a^i\, x^i & \sum\limits_{j=1}^{3} b^j\, x^j & \sum\limits_{k=1}^{3} c^k\, x^k \\ \sum\limits_{i=1}^{3} a^i\, y^i & \sum\limits_{j=1}^{3} b^j\, y^j & \sum\limits_{k=1}^{3} c^k\, y^k \\ \sum\limits_{i=1}^{3} a^i\, z^i & \sum\limits_{j=1}^{3} b^j\, z^j & \sum\limits_{k=1}^{3} c^k\, z^k \end{vmatrix}. \tag{56.7}$$

Let's recall that our basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ is orthonormal. The scalar product in such a basis is calculated according to the formula (33.3). Comparing (33.3) with (56.7), we see that all of the sums within the determinant (56.7) are scalar products of vectors:

$$L = \begin{vmatrix} (\mathbf{a},\mathbf{x}) & (\mathbf{b},\mathbf{x}) & (\mathbf{c},\mathbf{x}) \\ (\mathbf{a},\mathbf{y}) & (\mathbf{b},\mathbf{y}) & (\mathbf{c},\mathbf{y}) \\ (\mathbf{a},\mathbf{z}) & (\mathbf{b},\mathbf{z}) & (\mathbf{c},\mathbf{z}) \end{vmatrix}. \tag{56.8}$$

Now the formula (56.1) follows from the formulas (56.2) and (56.8). The theorem 56.1 is proved. □

The formula (56.1) is valuable for us not by itself, but due to its consequences. Assume that $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ is some arbitrary basis. Let's substitute $\mathbf{a} = \mathbf{e}_1$, $\mathbf{b} = \mathbf{e}_2$, $\mathbf{c} = \mathbf{e}_3$, $\mathbf{x} = \mathbf{e}_1$, $\mathbf{y} = \mathbf{e}_2$, $\mathbf{z} = \mathbf{e}_3$ into the formula (56.1). As a result we obtain

$$(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)^2 = \begin{vmatrix} (\mathbf{e}_1, \mathbf{e}_1) & (\mathbf{e}_2, \mathbf{e}_1) & (\mathbf{e}_3, \mathbf{e}_1) \\ (\mathbf{e}_1, \mathbf{e}_2) & (\mathbf{e}_2, \mathbf{e}_2) & (\mathbf{e}_3, \mathbf{e}_2) \\ (\mathbf{e}_1, \mathbf{e}_3) & (\mathbf{e}_2, \mathbf{e}_3) & (\mathbf{e}_3, \mathbf{e}_3) \end{vmatrix}. \tag{56.9}$$

Comparing (56.9) with (29.7) and (29.6), we see that the matrix (56.9) differs from the Gram matrix G by transposing. If we take into account the symmetry of the Gram matrix (see theorem 30.1), than we find that they do coincide. Therefore the formula (56.9) is written as follows:

$$(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)^2 = \det G. \tag{56.10}$$

The formula (56.10) coincides with the formula (51.6). This fact proves the theorem 51.2, which was unproved in § 51.

CHAPTER II

GEOMETRY OF LINES AND SURFACES.

In this Chapter the tools of the vector algebra developed in Chapter I are applied for describing separate points of the space $\mathbb{E}$ and for describing some geometric forms which are composed by these points.

§ 1. Cartesian coordinate systems.

DEFINITION 1.1. A *Cartesian coordinate system* in the space $\mathbb{E}$ is a basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ complemented by some fixed point O of this space. The point O being a part of the Cartesian coordinate system O, $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ is called the *origin* of this coordinate system.

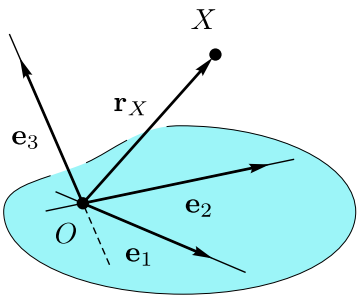

Fig. 1.1

DEFINITION 1.2. The vector $\overrightarrow{OX}$ binding the origin O of a Cartesian coordinate system O, $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ with a point X of the space $\mathbb{E}$ is called the *radius vector* of the point X in this coordinate system.

The free vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ being constituent parts of a Cartesian coordinate system O, $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ remain free. However they are often represented by geometric realizations attached to the origin O (see Fig. 1.1). These geometric realizations are extended

up to the whole lines which are called the *coordinate axes* of the Cartesian coordinate system O, $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$.

DEFINITION 1.3. The *coordinates* of a point X in a Cartesian coordinate system O, $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ are the coordinates of its radius vector $\mathbf{r}_X = \overrightarrow{OX}$ in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$.

Like other vectors, radius vectors of points are covered by the index setting convention (see Definition 20.1 in Chapter I). The coordinates of radius vectors are enumerated by upper indices, they are usually arranged into columns:

$$\mathbf{r}_X = \begin{Vmatrix} x^1 \\ x^2 \\ x^3 \end{Vmatrix}. \tag{1.1}$$

However, in those cases where a point X is represented by its coordinates these coordinates are arranged into a comma-separated row and placed within round brackets just after the symbol denoting the point X itself:

$$X = X(x^1, x^2, x^3). \tag{1.2}$$

The upper position of indices in the formula (1.2) is inherited from the formula (1.1).

Cartesian coordinate systems can be either *rectangular* or *skew-angular* depending on the type of basis used for defining them. In this book I follow a convention similar to that of the definition 29.1 in Chapter I.

DEFINITION 1.4. In this book a *skew-angular coordinate system* is understood as an *arbitrary coordinate system* where no restrictions for the basis are imposed.

DEFINITION 1.5. A rectangular coordinate system O, $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ whose basis is orthonormal is called a *rectangular coordinate system with unit scales along the axes*.

A remark. The radius vector of a point X written as $\overrightarrow{OX}$ is a geometric vector whose position in the space is fixed. The radius vector of a point X written as $\mathbf{r}_X$ is a free vector We can perform various operations of vector algebra with this vector: we can add it with other free vectors, multiply it by numbers, compose scalar products etc. But the vector $\mathbf{r}_X$ has a special mission — to be a pointer to the point X. It can do this mission only if its initial point is placed to the origin.

EXERCISE 1.1. *Relying on the definition 1.1, formulate analogous definitions for Cartesian coordinate systems on a line and on a plane.*

§ 2. Equations of lines and surfaces.

Coordinates of a single fixed point X in the space $\mathbb{E}$ are three fixed numbers (three constants). If these coordinates are changing, we get a moving point that runs over some set within the space $\mathbb{E}$. In this book we consider the cases where this set is some line or some surface. The case of a line differs from the case of a surface by its *dimension* or, in other words, by the *number of degrees of freedom*. Each surface is *two-dimensional* — a point on a surface has two degrees of freedom. Each line is *one-dimensional* — a point on a line has one degree of freedom.

Lines and surfaces contain infinitely many points. Therefore they cannot be described by enumeration where each point is described separately. Lines and surfaces are described by means of equations. The equations of lines and surfaces are subdivided into several types. If the radius vector of a point enters an equation as a whole without subdividing it into separate coordinates, such an equation is called a *vectorial equation*. If the radius vector of a point enters an equation through its coordinates, such an equation is called a *coordinate equation*.

Another attribute of the equations is the method of implementing the degrees of freedom. One or two degrees of freedom can be implemented in an equation explicitly when the radius

vector of a point is given as a function of one or two variables, which are called *parameters*. In this case the equation is called *parametric*. *Non-parametric equations* behave as obstructions decreasing the number of degrees of freedom from the initial three to one or two.

§ 3. A straight line on a plane.

Assume that some plane α in the space $\mathbb{E}$ is chosen and fixed. Then the number of degrees of freedom of a point immediately decreases from three to two. In order to study various forms of equations defining a straight line on the plane α we choose some coordinate system O, $\mathbf{e}_1$, $\mathbf{e}_2$ on this plane. Then we can define the points of the plane α and the points of a line on it by means of their radius-vectors.

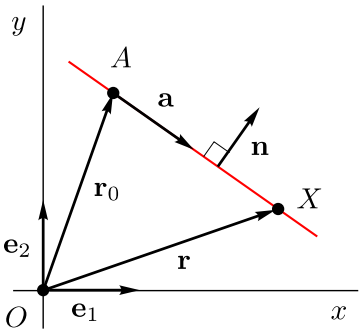

Fig. 3.1

1. Vectorial parametric equation of a line on a plane. Let;s consider a line on a plane with some coordinate system O, $\mathbf{e}_1$, $\mathbf{e}_2$. Let X be some arbitrary point on this line (see Fig. 3.1) and let A be some fixed point of this line. The position of the point X relative to the point A is marked by the vector $\overrightarrow{AX}$, while the position of the point A itself is determined by its radius vector $\mathbf{r}_0 = \overrightarrow{OA}$. Therefore we have

$$\mathbf{r} = \overrightarrow{OX} = \mathbf{r}_0 + \overrightarrow{AX}. \tag{3.1}$$

Let's choose and fix some nonzero vector $\mathbf{a} \neq \mathbf{0}$ directed along the line in question. The vector $\overrightarrow{AX}$ is expressed through $\mathbf{a}$ by means of the following formula:

$$\overrightarrow{AX} = \mathbf{a} \cdot t. \tag{3.2}$$

From the formulas (3.1) and (3.2) we immediately derive:

$$\mathbf{r} = \mathbf{r}_0 + \mathbf{a} \cdot t. \tag{3.3}$$

DEFINITION 3.1. The equality (3.3) is called the *vectorial parametric* equation of a line on a plane. The constant vector $\mathbf{a} \neq \mathbf{0}$ in it is a *directional vector* of the line, while the variable t is a *parameter*. The constant vector $\mathbf{r}_0$ in (3.3) is the *radius vector of an initial point*.

Each particular value of the parameter t corresponds to some definite point on the line. The initial point A with the radius vector $\mathbf{r}_0$ is associated with the value $t = 0$.

2. Coordinate parametric equations of a line on a plane. Let's determine the vectors $\mathbf{r}$, $\mathbf{r}_0$, and $\mathbf{a}$ in the vectorial parametric equation (3.3) through their coordinates:

$$\mathbf{r} = \begin{Vmatrix} x \\ y \end{Vmatrix}, \qquad \mathbf{r}_0 = \begin{Vmatrix} x_0 \\ y_0 \end{Vmatrix}, \qquad \mathbf{a} = \begin{Vmatrix} a_x \\ a_y \end{Vmatrix}. \tag{3.4}$$

Due to (3.4) the equation (3.3) is written as two equations:

$$\begin{cases} x = x_0 + a_x\, t, \\ y = y_0 + a_y\, t. \end{cases} \tag{3.5}$$

DEFINITION 3.2. The equalities (3.5) are called the *coordinate parametric* equations of a straight line on a plane. The constants a_x and a_y in them cannot vanish simultaneously.

3. Normal vectorial equation of a line on a plane. Let $\mathbf{n} \neq \mathbf{0}$ be a vector lying on the plane α in question and being perpendicular to the line in question (see Fig. 3.1). Let's apply the scalar multiplication by the vector $\mathbf{n}$ to both sides of the equation (3.3). As a result we get

$$(\mathbf{r}, \mathbf{n}) = (\mathbf{r}_0, \mathbf{n}) + (\mathbf{a}, \mathbf{n})\, t. \tag{3.6}$$

But $\mathbf{a} \perp \mathbf{n}$. Fort this reason the second term in the right hand side of (3.6) vanishes and the equation actually does not have the parameter t. The resulting equation is usually written as

$$(\mathbf{r} - \mathbf{r}_0, \mathbf{n}) = 0. \tag{3.7}$$

The scalar product of two constant vectors $\mathbf{r}_0$ and $\mathbf{n}$ is a numeric constant. If we denote $D = (\mathbf{r}_0, \mathbf{n})$, then the equation (3.7) can be written in the following form:

$$(\mathbf{r}, \mathbf{n}) = D. \tag{3.8}$$

DEFINITION 3.3. Any one of the two equations (3.7) and (3.8) is called the *normal vectorial* equation of a line on a plane. The constant vector $\mathbf{n} \neq \mathbf{0}$ in these equations is called a *normal vector* of this line.

4. Canonical equation of a line on a plane. Let's consider the case where $a_x \neq 0$ and $a_y \neq 0$ in the equations (3.5). In this case the parameter t can be expressed through x and y:

$$t = \frac{x - x_0}{a_x}, \qquad\qquad t = \frac{y - y_0}{a_y}. \tag{3.9}$$

From the equations (3.9) we derive the following equality:

$$\frac{x - x_0}{a_x} = \frac{y - y_0}{a_y}. \tag{3.10}$$

If $a_x = 0$, then the the first of the equations (3.5) turns to

$$x = x_0. \tag{3.11}$$

If $a_y = 0$, then the second equation (3.5) turns to

$$y = y_0. \tag{3.12}$$

Like the equation (3.10), the equations (3.11) and (3.12) do not comprise the parameter t.

DEFINITION 3.4. Any one of the three equalities (3.10), (3.11), and (3.12) is called the *canonical* equation of a line on a plane. The constants a_x and a_y in the equation (3.10) should be nonzero.

5. The equation of a line passing through two given points on a plane. Assume that two distinct points $A \neq B$ on a plane are given. We write their coordinates

$$A = A(x_0, y_0), \qquad B = B(x_1, y_1). \tag{3.13}$$

The vector $\mathbf{a} = \overrightarrow{AB}$ can be used for the directional vector of the line passing through the points A and B in (3.13). Then from (3.13) we derive the coordinates of $\mathbf{a}$:

$$\mathbf{a} = \begin{Vmatrix} a_x \\ a_y \end{Vmatrix} = \begin{Vmatrix} x_1 - x_0 \\ y_1 - y_0 \end{Vmatrix}. \tag{3.14}$$

Due to (3.14) the equation (3.10) can be written as

$$\frac{x - x_0}{x_1 - x_0} = \frac{y - y_0}{y_1 - y_0}. \tag{3.15}$$

The equation (3.15) corresponds to the case where $x_1 \neq x_0$ and $y_1 \neq y_0$. If $x_1 = x_0$, then we write the equation (3.11):

$$x = x_0 = x_1. \tag{3.16}$$

If $y_1 = y_0$, we write the equation (3.12):

$$y = y_0 = y_1. \tag{3.17}$$

The conditions $x_1 = x_0$ and $y_1 = y_0$ cannot be fulfilled simultaneously since the points A and B are distinct, i. e. $A \neq B$.

DEFINITION 3.5. Any one of the three equalities (3.15), (3.16), and (3.17) is called the *equation of a line passing through two given points* (3.13) on a plane.

6. General equation of a line on a plane. Let's apply the formula (29.8) from Chapter I in order to calculate the scalar product in (3.8). In this particular case it is written as

$$(\mathbf{r},\mathbf{n}) = \sum_{i=1}^{2}\sum_{j=1}^{2} r^i\, n^j\, g_{ij}, \tag{3.18}$$

where g_{ij} are the components of the Gram matrix for the basis $\mathbf{e}_1$, $\mathbf{e}_2$ on our plane (see Fig. 3.1). In Fig. 3.1 the basis $\mathbf{e}_1$, $\mathbf{e}_2$ is drawn to be rectangular. However, in general case it could be skew-angular as well. Let's introduce the notations

$$n_i = \sum_{j=1}^{2} n^j\, g_{ij}. \tag{3.19}$$

The quantities n^1 and n^2 in (3.18) and in (3.19) are the coordinates of the normal vector $\mathbf{n}$ (see Fig. 3.1).

DEFINITION 3.6. The quantities n_1 and n_2 produced from the coordinates of the normal vector $\mathbf{n}$ by mens of the formula (3.19) are called the *covariant coordinates* of the vector $\mathbf{n}$.

Taking into account the notations (3.19), the formula (3.18) is written in the following simplified form:

$$(\mathbf{r},\mathbf{n}) = \sum_{i=1}^{2} r^i\, n_i. \tag{3.20}$$

Let's recall the previous notations (3.4) and introduce new ones:

$$A = n_1, \qquad\qquad B = n_2. \tag{3.21}$$

Due to (3.4), (3.20), and (3.21) the equation (3.8) is written as

$$A\,x + B\,y - D = 0. \tag{3.22}$$

DEFINITION 3.7. The equation (3.22) is called the *general equation* of a line on a plane.

7. Double intersect equation of a line on a plane. Let's consider a line on a plane that does not pass through the origin and intersects with both of the coordinate axes. These conditions mean that $D \neq 0$, $A \neq 0$, and $B \neq 0$ in the equation (3.22) of this line. Through X and Y in Fig. 3.2 two intercept points are denoted:

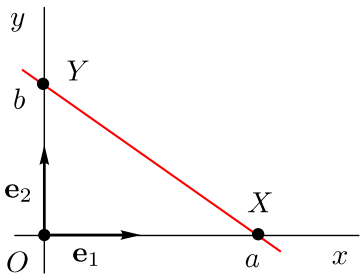

Fig. 3.2

$$\begin{aligned} X &= X(a,0),\\ Y &= Y(0,b). \end{aligned} \tag{3.23}$$

The quantities a and b in (3.23) are expressed through the constant parameters A, B, and D of the equation (3.22) by means of the following formulas:

$$a = D/A, \qquad\qquad b = D/B. \tag{3.24}$$

The equation (3.22) itself in our present case can be written as

$$\frac{x}{D/A} + \frac{y}{D/B} = 1. \tag{3.25}$$

If we take into account (3.24), the equality (3.25) turns to

$$\frac{x}{a} + \frac{y}{b} = 1. \tag{3.26}$$

DEFINITION 3.8. The equality (3.26) is called the *double intercept equation* of a line on a plane.

The name of the equation (3.26) is due to the parameters a and b being the lengths of the segments $[OX]$ and $[OY]$ which the line intercepts on the coordinate axes.

§ 4. A plane in the space.

Assume that some plane α in the space $\mathbb{E}$ is chosen and fixed. In order to study various equations determining this plane we choose some coordinate system O, $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ in the space $\mathbb{E}$. Then we can describe the points of this plane α by their radius vectors.

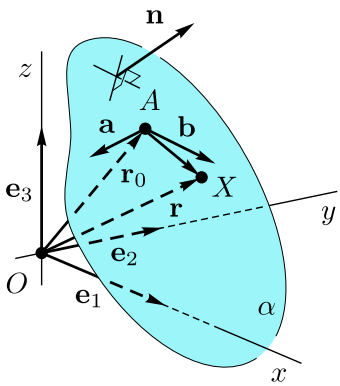

Fig. 4.1

1. Vectorial parametric equation of a plane in the space. Let's denote through A some fixed point of the plane α (see Fig. 4.1) and denote through X some arbitrary point of this plane. The position of the point X relative to the point A is marked by the vector $\overrightarrow{AX}$, while the position of the point A is determined by its radius vector $\mathbf{r}_0 = \overrightarrow{OA}$. For this reason

$$\mathbf{r} = \overrightarrow{OX} = \mathbf{r}_0 + \overrightarrow{AX}. \tag{4.1}$$

Let's choose and fix some pair of non-collinear vectors $\mathbf{a} \nparallel \mathbf{b}$ lying on the plane α. Such vectors constitute a basis on this plane (see Definition 17.1 from Chapter I). The vector $\overrightarrow{AX}$ is expanded in the basis of the vectors $\mathbf{a}$ and $\mathbf{b}$:

$$\overrightarrow{AX} = \mathbf{a} \cdot t + \mathbf{b} \cdot \tau. \tag{4.2}$$

Since X is an arbitrary point of the plane α, the numbers t and τ in (4.2) are two variable parameters. Upon substituting (4.2) into the formula (4.1), we get the following equality:

$$\mathbf{r} = \mathbf{r}_0 + \mathbf{a} \cdot t + \mathbf{b} \cdot \tau. \tag{4.3}$$

DEFINITION 4.1. The equality (4.3) is called the *vectorial parametric equation* of a plane in the space. The non-collinear vectors $\mathbf{a}$ and $\mathbf{b}$ in it are called *directional vectors* of a plane, while t and τ are called *parameters*. The fixed vector $\mathbf{r}_0$ is the *radius vector of an initial point*.

2. Coordinate parametric equation of a plane in the space. Let's determine the vectors $\mathbf{r}$, $\mathbf{r}_0$, $\mathbf{a}$, $\mathbf{b}$, in the vectorial parametric equation (4.3) through their coordinates:

$$\mathbf{r} = \begin{Vmatrix} x \\ y \\ z \end{Vmatrix}, \quad \mathbf{r}_0 = \begin{Vmatrix} x_0 \\ y_0 \\ z_0 \end{Vmatrix}, \quad \mathbf{a} = \begin{Vmatrix} a_x \\ a_y \\ a_z \end{Vmatrix}, \quad \mathbf{b} = \begin{Vmatrix} b_x \\ b_y \\ b_z \end{Vmatrix}. \tag{4.4}$$

Due to (4.4) the equation (4.3) is written as three equations

$$\begin{cases} x = x_0 + a_x\, t + b_x\, \tau, \\ y = y_0 + a_y\, t + b_y\, \tau, \\ z = z_0 + a_z\, t + b_z\, \tau. \end{cases} \tag{4.5}$$

DEFINITION 4.2. The equalities (4.5) are called the *coordinate parametric equations* of a plane in the space. The triples of constants a_x, a_y, a_z and b_x, b_y, b_z in these equations cannot be proportional to each other.

3. Normal vectorial equation of a plane in the space. Let $\mathbf{n} \neq \mathbf{0}$ be a vector perpendicular to the plane α (see Fig. 4.1). Let's apply the scalar multiplication by the vector $\mathbf{n}$ to both sides of the equation (4.3). As a result we get

$$(\mathbf{r}, \mathbf{n}) = (\mathbf{r}_0, \mathbf{n}) + (\mathbf{a}, \mathbf{n})\, t + (\mathbf{b}, \mathbf{n})\, \tau. \tag{4.6}$$

But $\mathbf{a} \perp \mathbf{n}$ and $\mathbf{b} \perp \mathbf{n}$. For this reason the second and the third terms in the right hand side of (4.6) vanish and the equation actually does not have the parameters t and τ. The resulting equation is usually written as

$$(\mathbf{r} - \mathbf{r}_0, \mathbf{n}) = 0. \tag{4.7}$$

The scalar product of two constant vectors $\mathbf{r}_0$ and $\mathbf{n}$ is a numeric constant. If we denote $D = (\mathbf{r}_0, \mathbf{n})$, then the equation (4.7) can be written in the following form:

$$(\mathbf{r}, \mathbf{n}) = D. \tag{4.8}$$

Definition 4.3. Any one of the two equalities (4.7) and (4.8) is called the *normal vectorial equation* of a plane in the space. The constant vector $\mathbf{n} \neq \mathbf{0}$ in these equations is called a *normal vector* of this plane.

4. Canonical equation of a plane in the space. The vector product of two non-coplanar vectors $\mathbf{a} \nparallel \mathbf{b}$ lying on the plane α can be chosen for the normal vector $\mathbf{n}$ in (4.7). Substituting $\mathbf{n} = [\mathbf{a}, \mathbf{b}]$ into the equation (4.7), we get the relationship

$$(\mathbf{r} - \mathbf{r}_0, [\mathbf{a}, \mathbf{b}]) = 0. \tag{4.9}$$

Applying the definition of the mixed product (see formula (45.1) in Chapter I), the scalar product in the formula (4.9) can be transformed to a mixed product:

$$(\mathbf{r} - \mathbf{r}_0, \mathbf{a}, \mathbf{b}) = 0. \tag{4.10}$$

Let's transform the equation (4.10) into a coordinate form. For this purpose we use the coordinate presentations of the vectors $\mathbf{r}$, $\mathbf{r}_0$, $\mathbf{a}$, and $\mathbf{b}$ taken from (4.4). Applying the formula (50.8) from Chapter I) and taking into account the fact that the oriented

volume of a basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ is nonzero, we derive

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} = 0. \tag{4.11}$$

If we use not the equation (4.7), but the equation (4.8), then instead of the equation (4.10) we get the following relationship:

$$(\mathbf{r}, \mathbf{a}, \mathbf{b}) = D. \tag{4.12}$$

In the coordinate form the relationship (4.12) is written as

$$\begin{vmatrix} x & y & z \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} = D. \tag{4.13}$$

DEFINITION 4.4. Any one of the two equalities (4.11) and (4.13) is called the *canonical equation* of a plane in the space. The triples of constants a_x, a_y, a_z and b_x, b_y, b_z in these equations should not be proportional to each other.

DEFINITION 4.5. The equation (4.10) and the equation (4.12), where $\mathbf{a} \nparallel \mathbf{b}$, are called the *vectorial forms of the canonical equation* of a plane in the space.

5. The equation of a plane passing through three given points. Assume that three points A, B, and C in the space $\mathbb{E}$ nit lying on a straight line are given. We write their coordinates

$$\begin{aligned} A &= A(x_0, y_0, z_0), \\ B &= B(x_1, y_1, z_1), \\ C &= C(x_2, y_2, z_2). \end{aligned} \tag{4.14}$$

The vectors $\mathbf{a} = \overrightarrow{AB}$ and $\mathbf{b} = \overrightarrow{AC}$ can be chosen for the directional vectors of a plane passing through the points A, B,

and C. Then from the formulas (4.14) we derive

$$\mathbf{a} = \left\| \begin{matrix} a_x \\ a_y \\ a_z \end{matrix} \right\| = \left\| \begin{matrix} x_1 - x_0 \\ y_1 - y_0 \\ z_1 - z_0 \end{matrix} \right\|, \quad \mathbf{b} = \left\| \begin{matrix} b_x \\ b_y \\ b_z \end{matrix} \right\| = \left\| \begin{matrix} x_2 - x_0 \\ y_2 - y_0 \\ z_2 - z_0 \end{matrix} \right\|. \tag{4.15}$$

Due to (4.15) the equation (4.11) can be written as

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ x_1 - x_0 & y_1 - y_0 & z_1 - z_0 \\ x_2 - x_0 & y_2 - y_0 & z_2 - z_0 \end{vmatrix} = 0. \tag{4.16}$$

If we denote through $\mathbf{r}_1$ and $\mathbf{r}_2$ the radius vectors of the points B and C from (4.14), then (4.15) can be written as

$$\mathbf{a} = \mathbf{r}_1 - \mathbf{r}_0, \qquad \mathbf{b} = \mathbf{r}_2 - \mathbf{r}_0. \tag{4.17}$$

Substituting (4.17) into the equation (4.10), we get

$$(\mathbf{r} - \mathbf{r}_0, \mathbf{r}_1 - \mathbf{r}_0, \mathbf{r}_2 - \mathbf{r}_0) = 0. \tag{4.18}$$

DEFINITION 4.6. The equality (4.16) is called the *equation of a plane passing through three given points.*

DEFINITION 4.7. The equality (4.18) is called the *vectorial form of the equation of a plane passing through three given points.*

6. General equation of a plane in the space. Let's apply the formula (29.8) from Chapter I in order to calculate the scalar product $(\mathbf{r}, \mathbf{n})$ in the equation (4.8). This yields

$$(\mathbf{r}, \mathbf{n}) = \sum_{i=1}^{3} \sum_{j=1}^{3} r^i \, n^j \, g_{ij}, \tag{4.19}$$

where g_{ij} are the components of the Gram matrix for the basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ (see Fig. 4.1). In Fig. 4.1 the basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ is

drawn to be rectangular. However, in general case it could be skew-angular as well. Let's introduce the notations

$$n_i = \sum_{j=1}^{3} n^j \, g_{ij}. \tag{4.20}$$

The quantities n^1, n^2, and n^3 in (4.19) and (4.20) are the coordinates of the normal vector $\mathbf{n}$ (see Fig. 4.1).

DEFINITION 4.8. The quantities n_1, n_2, and n_3 produced from the coordinates of the normal vector $\mathbf{n}$ by mens of the formula (4.20) are called the *covariant coordinates* of the vector $\mathbf{n}$.

Taking into account the notations (4.20), the formula (4.19) is written in the following simplified form:

$$(\mathbf{r}, \mathbf{n}) = \sum_{i=1}^{2} r^i \, n_i. \tag{4.21}$$

Let's recall the previous notations (4.4) and introduce new ones:

$$A = n_1, \qquad B = n_2, \qquad C = n_3. \tag{4.22}$$

Due to (4.4), (4.21), and (4.22) the equation (4.8) is written as

$$A\,x + B\,y + C\,y - D = 0. \tag{4.23}$$

DEFINITION 4.9. The equation (4.23) is called the *general equation* of a plane in the space.

7. Triple intercept equation of a plane in the space. Let's consider a plane in the space that does not pass through the origin and intersects with each of the three coordinate axes. These conditions mean that $D \neq 0$, $A \neq 0$, $B \neq 0$, and $C \neq 0$ in (4.23). Through X, Y, and Z in Fig. 4.2 below we denote three

intercept points of the plane. Here are the coordinates of these three intercept points X, Y, and Z produced by our plane:

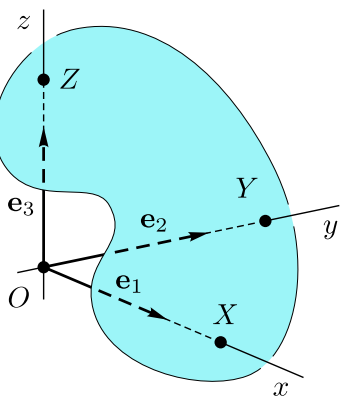

Fig. 4.2

$$\begin{aligned} X &= X(a,0,0),\\ Y &= Y(0,b,0),\\ Z &= Y(0,0,c). \end{aligned} \tag{4.24}$$

The quantities a, b, and c in (4.24) are expressed through the constant parameters A, B, C, and D of the equation (4.23) by means of the formulas

$$\begin{aligned} a &= D/A,\\ b &= D/B,\\ c &= D/C. \end{aligned} \tag{4.25}$$

The equation of the plane (4.23) itself can be written as

$$\frac{x}{D/A} + \frac{y}{D/B} + \frac{z}{D/C} = 1. \tag{4.26}$$

If we take into account (4.25), the equality (4.26) turns to

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1. \tag{4.27}$$

Definition 4.10. The equality (4.27) is called the *triple intercept equation* of a plane in the space.

§ 5. A straight line in the space.

Assume that some straight line a in the space $\mathbb{E}$ is chosen and fixed. In order to study various equations determining this line we choose some coordinate system O, $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ in the space $\mathbb{E}$. Then we can describe the points of the line by their radius

vectors relative to the origin O.

1. Vectorial parametric equation of a line in the space. Let's denote through A some fixed point on the line (see Fig. 5.1) and denote through X an arbitrary point of this line. The position of the point X relative to the point A is marked by the vector $\overrightarrow{AX}$, while the position of the point A itself is determined by its radius vector $\mathbf{r}_0 = \overrightarrow{OA}$. Therefore we have

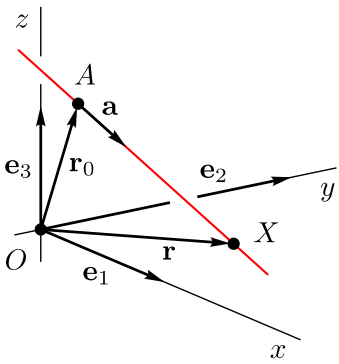

Fig. 5.1

$$\mathbf{r} = \mathbf{r}_0 + \overrightarrow{AX}. \tag{5.1}$$

Let's choose and fix some non-zero vector $\mathbf{a} \neq \mathbf{0}$ directed along the line in question. The vector $\overrightarrow{AX}$ is expressed through the vector $\mathbf{a}$ by means of the formula

$$\overrightarrow{AX} = \mathbf{a} \cdot t. \tag{5.2}$$

From the formulas (5.1) and (5.2) we immediately derive

$$\mathbf{r} = \mathbf{r}_0 + \mathbf{a} \cdot t. \tag{5.3}$$

Definition 5.1. The equality (5.3) is called the *vectorial parametric equation* of the line in the space. The constant vector $\mathbf{a} \neq \mathbf{0}$ in this equation is called a *directional vector* of the line, while t is a *parameter*. The constant vector $\mathbf{r}_0$ is the *radius vector of an initial point*.

Each particular value of the parameter t corresponds to some definite point on the line. The initial point A with the radius vector $\mathbf{r}_0$ is associated with the value $t = 0$.

2. Coordinate parametric equations of a line in the space. Let's determine the vectors $\mathbf{r}$, $\mathbf{r}_0$, and $\mathbf{a}$ in the vectorial parametric equation (5.3) through their coordinates:

$$\mathbf{r} = \begin{Vmatrix} x \\ y \\ z \end{Vmatrix}, \qquad \mathbf{r}_0 = \begin{Vmatrix} x_0 \\ y_0 \\ z_0 \end{Vmatrix}, \qquad \mathbf{a} = \begin{Vmatrix} a_x \\ a_y \\ a_z \end{Vmatrix}. \tag{5.4}$$

Due to (5.4) the equation (5.3) is written as three equations:

$$\begin{cases} x = x_0 + a_x\, t, \\ y = y_0 + a_y\, t, \\ z = z_0 + a_z\, t. \end{cases} \tag{5.5}$$

Definition 5.2. The equalities (5.5) are called the *coordinate parametric equations* of a line in the space. The constants a_x, a_y, a_z in them cannot vanish simultaneously

3. Vectorial equation of a line in the space. Let's apply the vectorial multiplication by the vector $\mathbf{a}$ to both sides of the equation (5.3). As a result we get

$$[\mathbf{r}, \mathbf{a}] = [\mathbf{r}_0, \mathbf{a}] + [\mathbf{a}, \mathbf{a}]\, t. \tag{5.6}$$

Due to the item 4 of the theorem 39.1 in Chapter I the vector product $[\mathbf{a}, \mathbf{a}]$ in (5.6) is equal to zero. For this reason the equation (5.6) actually does not contain the parameter t. This equation is usually written as follows:

$$[\mathbf{r} - \mathbf{r}_0, \mathbf{a}] = 0. \tag{5.7}$$

The vector product of two constant vectors $\mathbf{r}_0$ and $\mathbf{a}$ is a constant vector. If we denote this vector $\mathbf{b} = [\mathbf{r}_0, \mathbf{a}]$, then the equation of the line (5.7) can be written as

$$[\mathbf{r}, \mathbf{a}] = \mathbf{b}, \quad \text{where} \quad \mathbf{b} \perp \mathbf{a}. \tag{5.8}$$

DEFINITION 5.3. Any one of the two equalities (5.7) and (5.8) is called the *vectorial*[1] *equation* of a line in the space. The constant vector **b** in the equation (5.8) should be perpendicular to the directional vector **a**.

4. Canonical equation of a line in the space. Let's consider the case where $a_x \neq 0$, $a_y \neq 0$, and $a_z \neq 0$ in the equations (5.5). Then the parameter t can be expressed through x, y, and z:

$$t = \frac{x - x_0}{a_x}, \qquad t = \frac{y - y_0}{a_y}, \qquad t = \frac{z - z_0}{a_z}. \tag{5.9}$$

From the equations (5.9) one can derive the equalities

$$\frac{x - x_0}{a_x} = \frac{y - y_0}{a_y} = \frac{z - z_0}{a_z}. \tag{5.10}$$

If $a_x = 0$, then instead of the first equation (5.9) from (5.5) we derive $x = x_0$. Therefore instead of (5.10) we write

$$x = x_0, \qquad \frac{y - y_0}{a_y} = \frac{z - z_0}{a_z}. \tag{5.11}$$

If $a_y = 0$, then instead of the second equation (5.9) from (5.5) we derive $y = y_0$. Therefore instead of (5.10) we write

$$y = y_0, \qquad \frac{x - x_0}{a_x} = \frac{z - z_0}{a_z}. \tag{5.12}$$

If $a_z = 0$, then instead of the third equation (5.9) from (5.5) we derive$z = z_0$. Therefore instead of (5.10) we write

$$z = z_0, \qquad \frac{x - x_0}{a_x} = \frac{y - y_0}{a_y}. \tag{5.13}$$

[1] The terms «vectorial equation» and «vectorial parametric equation» are often confused and the term «vector equation» is misapplied to (5.3).

If $a_x = 0$ and $a_y = 0$, then instead of (5.10) we write

$$x = x_0, \qquad y = y_0. \tag{5.14}$$

If $a_x = 0$ and $a_z = 0$, then instead of (5.10) we write

$$x = x_0, \qquad z = z_0. \tag{5.15}$$

If $a_y = 0$ and $a_z = 0$, then instead of (5.10) we write

$$y = y_0, \qquad z = z_0. \tag{5.16}$$

DEFINITION 5.4. Any one of the seven pairs of equations (5.10), (5.11), (5.12), (5.13), (5.14), (5.15), and (5.16) is called the *canonical equation* of a line in the space.

5. The equation of a line passing through two given points in the space. Assume that two distinct points $A \neq B$ in the space are given. We write their coordinates

$$A = A(x_0, y_0, z_0), \qquad B = B(x_1, y_1, z_1). \tag{5.17}$$

The vector $\mathbf{a} = \overrightarrow{AB}$ can be used for the directional vector of the line passing through the points A and B in (5.17). Then from (5.17) we derive the coordinates of $\mathbf{a}$:

$$\mathbf{a} = \begin{Vmatrix} a_x \\ a_y \\ a_z \end{Vmatrix} = \begin{Vmatrix} x_1 - x_0 \\ y_1 - y_0 \\ z_1 - z_0 \end{Vmatrix}. \tag{5.18}$$

Due to (5.18) the equations (5.10) can be written as

$$\frac{x - x_0}{x_1 - x_0} = \frac{y - y_0}{y_1 - y_0} = \frac{z - z_0}{z_1 - z_0}. \tag{5.19}$$

The equations (5.19) correspond to the case where the inequalities $x_1 \neq x_0$, $y_1 \neq y_0$, and $z_1 \neq z_0$ are fulfilled.

If $x_1 = x_0$, then instead of (5.19) we write the equations

$$x = x_0 = x_1, \qquad \frac{y - y_0}{y_1 - y_0} = \frac{z - z_0}{z_1 - z_0}. \tag{5.20}$$

If $y_1 = y_0$, then instead of (5.19) we write the equations

$$y = y_0 = y_1, \qquad \frac{x - x_0}{x_1 - x_0} = \frac{z - z_0}{z_1 - z_0}. \tag{5.21}$$

If $z_1 = z_0$, then instead of (5.19) we write the equations

$$z = z_0 = z_1, \qquad \frac{x - x_0}{x_1 - x_0} = \frac{y - y_0}{y_1 - y_0}. \tag{5.22}$$

If $x_1 = x_0$ and $y_1 = y_0$, then instead of (5.19) we write

$$x = x_0 = x_1, \qquad y = y_0 = y_1. \tag{5.23}$$

If $x_1 = x_0$ and $z_1 = z_0$, then instead of (5.19) we write

$$x = x_0 = x_1, \qquad z = z_0 = z_1. \tag{5.24}$$

If $y_1 = y_0$ and $z_1 = z_0$, then instead of (5.19) we write

$$y = y_0 = y_1, \qquad z = z_0 = z_1. \tag{5.25}$$

The conditions $x_1 = x_0$, $y_1 = y_0$, and $z_1 = z_0$ cannot be fulfilled simultaneously since $A \neq B$.

Definition 5.5. Any one of the seven pairs of equalities (5.19), (5.20), (5.21), (5.22), (5.23), (5.24),and (5.25) is called the equation of a line *passing through two given points* A and B with the coordinates (5.17).

6. The equation of a line in the space **as the intersection of two planes**. In the vectorial form the equations of two intersecting planes can be written as (4.8):

$$(\mathbf{r}, \mathbf{n}_1) = D_1, \qquad (\mathbf{r}, \mathbf{n}_2) = D_2. \tag{5.26}$$

For the planes given by the equations (5.26) do actually intersect their normal vectors should be non-parallel: $\mathbf{n}_1 \nparallel \mathbf{n}_2$.

In the coordinate form the equations of two intersecting planes can be written as (4.23):

$$\begin{aligned} A_1\,x + B_1\,y + C_1\,y - D_1 = 0,\\ A_2\,x + B_2\,y + C_2\,y - D_2 = 0. \end{aligned} \tag{5.27}$$

Definition 5.6. Any one of the two pairs of equalities (5.26) and (5.27) is called the equation of a line in the space obtained *as the intersection of two planes*.

§ 6. Ellipse. Canonical equation of an ellipse.

Definition 6.1. An *ellipse* is a set of points on some plane the sum of distances from each of which to some fixed points F_1 and F_2 of this plane is a constant which is the same for all points of the set. The points F_1 and F_2 are called the *foci* of the ellipse.

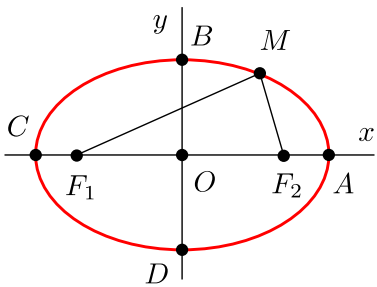

Fig. 6.1

Assume that an ellipse with the foci F_1 and F_2 is given. Let's draw the line connecting the points F_1 and F_2 and choose this line for the x-axis of a coordinate system. Let's denote through O the midpoint of the segment $[F_1F_2]$ and choose it for the origin. We choose the second coordinate axis (the y-axis) to

be perpendicular to the x-axis on the ellipse plane (see Fig. 6.1). We choose the unity scales along the axes. This means that the basis of the coordinate system constructed is orthonormal.

Let $M = M(x, y)$ be some arbitrary point of the ellipse. According to the definition 6.1 the sum $|MF_1| + |MF_2|$ is equal to a constant which does not depend on the position of M on the ellipse. Let's denoter this constant through $2\,a$ and write

$$|MF_1| + |MF_2| = 2\,a. \tag{6.1}$$

The length of the segment $[F_1F_2]$ is also a constant. Let's denote this constant through $2\,c$. This yields the relationships

$$|F_1O| = |OF_2| = c, \qquad |F_1F_2| = 2\,c. \tag{6.2}$$

From the triangle inequality $|F_1F_2| \leqslant |MF_1| + |MF_2|$ we derive the following inequality for the constants c and a:

$$c \leqslant a. \tag{6.3}$$

The case $c = a$ in (6.3) corresponds to a degenerate ellipse. In this case the triangle inequality $|F_1F_2| \leqslant |MF_1| + |MF_2|$ turns to the equality $|F_1F_2| = |MF_1| + |MF_2|$, while the MF_1F_2 itself collapses to the segment $[F_1F_2]$. Since $M \in [F_1F_2]$, a degenerate ellipse with the foci F_1 and F_2 is the segment $[F_1F_2]$. The case of a degenerate ellipse is usually excluded by setting

$$0 \leqslant c < a. \tag{6.4}$$

The formulas (6.2) determine the foci F_1 and F_2 of the ellipse in the coordinate system we have chosen:

$$F_1 = F_1(-c, 0), \qquad F_2 = F_2(c, 0). \tag{6.5}$$

Having determined the coordinates of the points F_1 and F_2 and knowing the coordinates of the point $M = M(x, y)$, we derive the

following relationships for the segments $[MF_1]$ and $[MF_2]$:

$$\begin{aligned} |MF_1| &= \sqrt{y^2 + (x+c)^2}, \\ |MF_2| &= \sqrt{y^2 + (x-c)^2}. \end{aligned} \tag{6.6}$$

Derivation of the canonical equation of an ellipse. Let's substitute (6.6) into the equality (6.1) defining the ellipse:

$$\sqrt{y^2 + (x+c)^2} + \sqrt{y^2 + (x-c)^2} = 2\,a. \tag{6.7}$$

Then we move one of the square roots to the right hand side of the formula (6.7). As a result we derive

$$\sqrt{y^2 + (x+c)^2} = 2\,a - \sqrt{y^2 + (x-c)^2}. \tag{6.8}$$

Let's square both sides of the equality (6.8):

$$\begin{gathered} y^2 + (x+c)^2 = 4\,a^2 - \\ - 4\,a\,\sqrt{y^2 + (x-c)^2} + y^2 + (x-c)^2. \end{gathered} \tag{6.9}$$

Upon expanding brackets and collecting similar terms the equality (6.9) can be written in the following form:

$$4\,a\,\sqrt{y^2 + (x-c)^2} = 4\,a^2 - 4\,x\,c. \tag{6.10}$$

Let's cancel the factor four in (6.10) and square both sides of this equality. As a result we get the formula

$$a^2\,(y^2 + (x-c)^2) = a^4 - 2\,a^2\,x\,c + x^2\,c^2. \tag{6.11}$$

Upon expanding brackets and recollecting similar terms the equality (6.11) can be written in the following form:

$$x^2\,(a^2 - c^2) + y^2\,a^2 = a^2\,(a^2 - c^2). \tag{6.12}$$

Both sides of the equality (6.12) contain the quantity $a^2 - c^2$. Due to the inequality (6.4) this quantity is positive. For this reason it can be written as the square of some number $b > 0$:

$$a^2 - c^2 = b^2. \tag{6.13}$$

Due to (6.13) the equality (6.12) can be written as

$$x^2\, b^2 + y^2\, a^2 = a^2\, b^2. \tag{6.14}$$

Since $b > 0$ and $a > 0$ (see the inequalities (6.4)), the equality (6.14) transforms to the following one:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \tag{6.15}$$

Definition 6.2. The equality (6.15) is called the *canonical* equation of an ellipse.

Theorem 6.1. *For each point $M = M(x, y)$ of the ellipse determined by the initial equation* (6.7) *its coordinates obey the canonical equation* (6.15).

The proof of the theorem 6.1 consists in the above calculations that lead to the canonical equation of an ellipse (6.15). This canonical equation yields the following important inequalities:

$$|x| \leqslant a, \qquad\qquad |y| \leqslant b. \tag{6.16}$$

Theorem 6.2. *The canonical equation of an ellipse* (6.15) *is equivalent to the initial equation* (6.7).

Proof. In order to prove the equation 6.2 we transform the expressions (6.6) relying on (6.15). From (6.15) we derive

$$y^2 = b^2 - \frac{b^2}{a^2}\, x^2. \tag{6.17}$$

Substituting (6.17) into the first formula (6.6), we get

$$\begin{aligned} |MF_1| &= \sqrt{b^2 - \frac{b^2}{a^2}\,x^2 + x^2 + 2\,x\,c + c^2} = \\ &= \sqrt{\frac{a^2 - b^2}{a^2}\,x^2 + 2\,x\,c + (b^2 + c^2)}. \end{aligned} \tag{6.18}$$

Now we take into account the relationship (6.13) and write the equality (6.18) in the following form:

$$|MF_1| = \sqrt{a^2 + 2\,x\,c + \frac{c^2}{a^2}\,x^2} = \sqrt{\Big(\frac{a^2 + c\,x}{a}\Big)^2}. \tag{6.19}$$

Upon calculating the square root the formula (6.19) yields

$$|MF_1| = \frac{|a^2 + c\,x|}{a}. \tag{6.20}$$

From the inequalities (6.4) and (6.16) we derive $a^2 + c\,x > 0$. Therefore the modulus signs in (6.20) can be omitted:

$$|MF_1| = \frac{a^2 + c\,x}{a} = a + \frac{c\,x}{a}. \tag{6.21}$$

In the case of the second formula (6.6) the considerations similar to the above ones yield the following result:

$$|MF_2| = \frac{a^2 - c\,x}{a} = a - \frac{c\,x}{a}. \tag{6.22}$$

Now it is easy to see that the equation (6.7) written as $|MF_1| + |MF_2| = 2\,a$ due to (6.6) is immediate from (6.21) and (6.22). The theorem 6.2 is proved. □

Let's consider again the inequalities (6.16). The coordinates of any point M of the ellipse obey these inequalities. The first of the inequalities (6.16) turns to an equality if M coincides with A or if M coincides with C (see Fig. 6.1). The second inequality (6.16) turns to an equality if M coincides with B or if M coincides D.

Definition 6.3. The points A, B, C, and D in Fig. 6.1 are called the *vertices* of an ellipse. The segments $[AC]$ and $[BD]$ are called the *axes* of an ellipse, while the segments $[OA]$, $[OB]$, $[OC]$, and $[OD]$ are called its *semiaxes*.

The constants a, b, and c obey the relationship (6.13). From this relationship and from the inequalities (6.4) we derive

$$0 < b \leqslant a. \tag{6.23}$$

Definition 6.4. Due to the inequalities (6.23) the semiaxis $[OA]$ in Fig. 6.1 is called the *major semiaxis* of the ellipse, while the semiaxis $[OB]$ is called its *minor semiaxis*.

Definition 6.5. A coordinate system O, $\mathbf{e}_1$, $\mathbf{e}_2$ with an orthonormal basis $\mathbf{e}_1$, $\mathbf{e}_2$ where an ellipse is given by its canonical equation (6.15) and where the inequalities (6.23) are fulfilled is called a *canonical coordinate system* of this ellipse.

§ 7. The eccentricity and directrices of an ellipse. The property of directrices.

The shape and sizes of an ellipse are determined by two constants a and b in its canonical equation (6.15). Due to the relationship (6.13) the constant b can be expressed through the constant c. Multiplying both constants a and c by the same number, we change the sizes of an ellipse, but do not change its shape. The ratio of these two constants

$$\varepsilon = \frac{c}{a}. \tag{7.1}$$

is responsible for the shape of an ellipse.

Definition 7.1. The quantity ε defined by the relationship (7.1), where a is the major semiaxis and c is the half of the interfocal distance, is called the *eccentricity* of an ellipse.

The eccentricity (7.1) is used in order to define one more numeric parameter of an ellipse. It is usually denoted through d:

$$d = \frac{a}{\varepsilon} = \frac{a^2}{c}. \tag{7.2}$$

DEFINITION 7.2. On the plane of an ellipse there are two lines parallel to its minor axis and placed at the distance d given by the formula (7.2) from its center. These lines are called *directrices* of an ellipse.

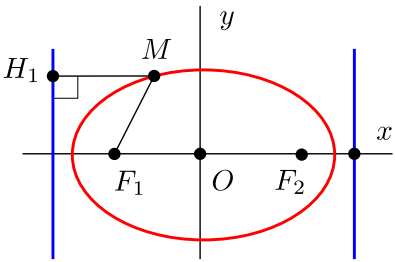

Fig. 7.1

Each ellipse has two foci and two directrices. Each directrix has the corresponding focus of it. This is that of two foci which is more close to the directrix in question. Let $M = M(x, y)$ be some arbitrary point of an ellipse. Let's connect this point with the left focus of the ellipse F_1 and drop the perpendicular from it to the left directrix of the ellipse. Let's denote through H_1 the base of such a perpendicular and calculate its length:

$$|MH_1| = |x - (-d)| = |d + x|. \tag{7.3}$$

Taking into account (7.2), the formula (7.3) can be brought to

$$|MH_1| = \left|\frac{a^2}{c} + x\right| = \frac{|a^2 + c\,x|}{c}. \tag{7.4}$$

The length of the segment MF_1 was already calculated above. Initially it was given by one of the formulas (6.6), but later the more simple expression (6.20) was derived for it:

$$|MF_1| = \frac{|a^2 + c\,x|}{a}. \tag{7.5}$$

If we divide (7.5) by (7.4), we obtain the following relationship:

$$\frac{|MF_1|}{|MH_1|} = \frac{c}{a} = \varepsilon. \tag{7.6}$$

The point M can change its position on the ellipse. Then the numerator and the denominator of the fraction (7.6) are changed, but its value remains unchanged. This fact is known as the property of directrices.

THEOREM 7.1. *The ratio of the distance from some arbitrary point M of an ellipse to its focus and the distance from this point to the corresponding directrix is a constant which is equal to the eccentricity of the ellipse.*

§ 8. The equation of a tangent line to an ellipse.

Let's consider an ellipse given by its canonical equation (6.15) in its canonical coordinate system (see Definition 6.5). Let's draw a tangent line to this ellipse and denote through $M = M(x_0, y_0)$ its tangency point (see Fig. 8.1). Our goal is to write the equation of a tangent line to an ellipse.

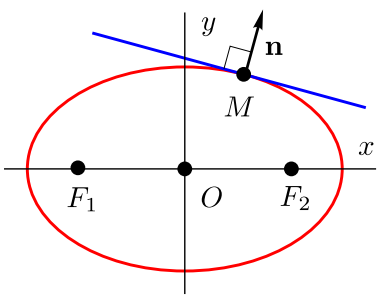

Fig. 8.1

An ellipse is a curve composed by two halves — the upper half and the lower half. Any one of these two halves can be treated as a graph of a function

$$y = f(x) \tag{8.1}$$

with the domain $(-a, a)$. The equation of a tangent line to the graph of a function is given by the following well-known formula (see [9]):

$$y = y_0 + f'(x_0)\,(x - x_0). \tag{8.2}$$

In order to apply the formula (8.2) to an ellipse we need to calculate the derivative of the function (8.1). Let's substitute the expression (8.1) for y into the formula (6.15):

$$\frac{x^2}{a^2}+\frac{(f(x))^2}{b^2}=1. \tag{8.3}$$

The equality (8.3) is fulfilled identically in x. Let's differentiate the equality (8.3) with respect to x. This yields

$$\frac{2\,x}{a^2}+\frac{2\,f(x)\,f'(x)}{b^2}=0. \tag{8.4}$$

Let's apply the formula (8.4) for to calculate the derivative $f'(x)$:

$$f'(x)=-\frac{b^2\,x}{a^2\,f(x)}. \tag{8.5}$$

In order to substitute (8.5) into the equation (8.2) we change x for x_0 and $f(x)$ for $f(x_0)=y_0$. As a result we get

$$f'(x_0)=-\frac{b^2\,x_0}{a^2\,y_0}. \tag{8.6}$$

Let's substitute (8.6) into the equation of the tangent line (8.2). This yields the following relationship

$$y=y_0-\frac{b^2\,x_0}{a^2\,y_0}\,(x-x_0). \tag{8.7}$$

Eliminating the denominator, we write the equality (8.7) as

$$a^2\,y\,y_0+b^2\,x\,x_0=a^2\,y_0^2+b^2\,x_0^2. \tag{8.8}$$

Now let's divide both sides of the equality (8.8) by $a^2\,b^2$:

$$\frac{x\,x_0}{a^2}+\frac{y\,y_0}{b^2}=\frac{x_0^2}{a^2}+\frac{y_0^2}{b^2}. \tag{8.9}$$

Note that the point $M = M(x_0, y_9)$ is on the ellipse. Therefore its coordinates satisfy the equation (6.15):

$$\frac{x_0^2}{a^2} + \frac{y_0^2}{b^2} = 1. \tag{8.10}$$

Taking into account (8.10), we can transform (8.9) to

$$\frac{x\,x_0}{a^2} + \frac{y\,y_0}{b^2} = 1. \tag{8.11}$$

This is the required equation of a tangent line to an ellipse.

THEOREM 8.1. *For an ellipse determined by its canonical equation* (6.15) *its tangent line that touches this ellipse at the point* $M = M(x_0, y_0)$ *is given by the equation* (8.11).

The equation (8.11) is a particular case of the equation (3.22) where the constants A, B, and D are given by the formulas

$$A = \frac{x_0}{a^2}, \qquad B = \frac{y_0}{b^2}, \qquad D = 1. \tag{8.12}$$

According to the definition 3.6 and the formulas (3.21) the constants A and B in (8.12) are the covariant components of the normal vector for the tangent line to an ellipse. The tangent line equation (8.11) is written in a canonical coordinate system of an ellipse. The basis of such a coordinate system is orthonormal. Therefore the formula (3.19) and the formula (32.4) from Chapter I yield the following relationships:

$$A = n_1 = n^1, \qquad B = n_2 = n^2. \tag{8.13}$$

THEOREM 8.2. *The quantities* A *and* B *in* (8.12) *are the coordinates of the normal vector* $\mathbf{n}$ *for the tangent line to an ellipse which is given by the equation* (8.11).

The relationships (8.13) prove the theorem 8.2.

§ 9. Focal property of an ellipse.

The term *focus* is well known in optics. It means a point where light rays converge upon refracting in lenses or upon reflecting in curved mirrors. In the case of an ellipse let's assume that it is manufactured of a thin strip of some flexible material and assume that its inner surface is covered with a light reflecting layer. For such a materialized ellipse one can formulate the following focal property.

THEOREM 9.1. *A light ray emitted from one of the foci of an ellipse upon reflecting on its inner surface passes through the other focus of this ellipse.*

THEOREM 9.2. *The perpendicular to a tangent line of an ellipse drawn at the tangency point is a bisector in the triangle composed by the tangency point and two foci of the ellipse.*

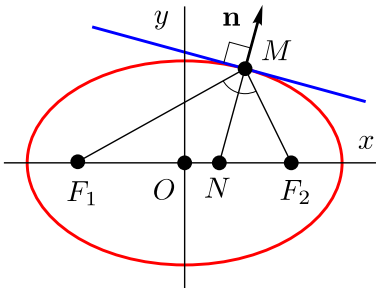

Fig. 9.1

The theorem 9.2 is a geometric version of the theorem 9.1. These theorems are equivalent due to the reflection law saying that the angle of reflection is equal to the angle of incidence.

PROOF OF THE THEOREM 9.2. Let's choose an arbitrary point $M = M(x_0, y_0)$ on the ellipse and draw the tangent line to the ellipse through this point as shown in Fig. 9.1. Then we draw the perpendicular $[MN]$ to the tangent line through the point M. The segment $[MN]$ is directed along the normal vector of the tangent line. In order to prove that this segment is a bisector of the triangle F_1MF_2 it is sufficient to prove the equality

$$\cos(\widehat{F_1MN}) = \cos(\widehat{F_2MN}). \tag{9.1}$$

The cosine equality (9.1) is equivalent to the following equality

for the scalar products of vectors:

$$\frac{(\overrightarrow{MF_1}, \mathbf{n})}{|MF_1|} = \frac{(\overrightarrow{MF_2}, \mathbf{n})}{|MF_2|}. \tag{9.2}$$

The coordinates of the points F_1 and F_2 in a canonical coordinate system of the ellipse are known (see formulas (6.5)). The coordinates of the point $M = M(x_0, y_0)$ are also known. Therefore we can find the coordinates of the vectors $\overrightarrow{MF_1}$ and $\overrightarrow{MF_2}$ used in the above formula (9.2):

$$\overrightarrow{MF_1} = \begin{Vmatrix} -c - x_0 \\ -y_0 \end{Vmatrix}, \qquad \overrightarrow{MF_2} = \begin{Vmatrix} c - x_0 \\ -y_0 \end{Vmatrix}. \tag{9.3}$$

The coordinates of the normal vector $\mathbf{n}$ of the tangent line in Fig. 9.1 are given by the formulas (8.12) and (8.13):

$$\mathbf{n} = \begin{Vmatrix} \dfrac{x_0}{a^2} \\[2ex] \dfrac{y_0}{b^2} \end{Vmatrix}. \tag{9.4}$$

Using (9.3) and (9.4), we apply the formula (33.3) from Chapter I for calculating the scalar products in (9.2):

$$\begin{aligned} (\overrightarrow{MF_1}, \mathbf{n}) &= \frac{-c\,x_0 - x_0^2}{a^2} - \frac{y_0^2}{b^2} = \frac{-c\,x_0}{a^2} - \frac{x_0^2}{a^2} - \frac{y_0^2}{b^2}, \\ (\overrightarrow{MF_2}, \mathbf{n}) &= \frac{c\,x_0 - x_0^2}{a^2} - \frac{y_0^2}{b^2} = \frac{c\,x_0}{a^2} - \frac{x_0^2}{a^2} - \frac{y_0^2}{b^2}. \end{aligned} \tag{9.5}$$

The point $M = M(x_0, y_0)$ lies on the ellipse. Therefore its coordinates should satisfy the equation of the ellipse (6.15):

$$\frac{x_0^2}{a^2} + \frac{y_0^2}{b^2} = 1. \tag{9.6}$$

Due to (9.6) the formulas (9.5) simplify and take the form

$$\begin{aligned}(\overrightarrow{MF_1},\mathbf{n}) &= \frac{-c\,x_0}{a^2} - 1 = -\frac{a^2+c\,x_0}{a^2},\\ (\overrightarrow{MF_2},\mathbf{n}) &= \frac{c\,x_0}{a^2} - 1 = -\frac{a^2-c\,x_0}{a^2}.\end{aligned} \tag{9.7}$$

In order to calculate the denominators in the formula (9.2) we use the formulas (6.21) and (6.22). In this case, when applied to the point $M = M(x_0, y_0)$, they yield

$$|MF_1| = \frac{a^2+c\,x_0}{a}, \qquad |MF_2| = \frac{a^2-c\,x_0}{a}. \tag{9.8}$$

From the formulas (9.7) and (9.8) we easily derive the equalities

$$\frac{(\overrightarrow{MF_1},\mathbf{n})}{|MF_1|} = -\frac{1}{a}, \qquad \frac{(\overrightarrow{MF_2},\mathbf{n})}{|MF_2|} = -\frac{1}{a}$$

which proves the equality (9.2). As a result the theorem 9.2 and the theorem 9.1 equivalent to it both are proved. □

§ 10. Hyperbola. Canonical equation of a hyperbola.

Definition 10.1. A *hyperbola* is a set of points on some plane the modulus of the difference of distances from each of which to some fixed points F_1 and F_2 of this plane is a constant which is the same for all points of the set. The points F_1 and F_2 are called the *foci* of the hyperbola.

Assume that a hyperbola with the foci F_1 and F_2 is given. Let's draw the line connecting the points F_1 and F_2 and choose this line for the x-axis of a coordinate system. Let's denote through O the midpoint of the segment $[F_1F_2]$ and choose it for the origin. We choose the second coordinate axis (the y-axis) to be perpendicular to the x-axis on the hyperbola plane

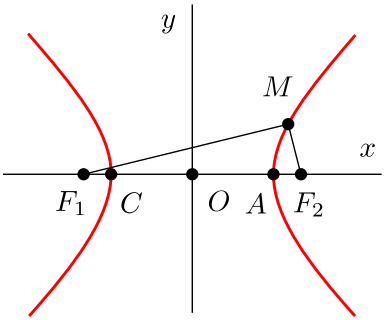

Fig. 10.1

(see Fig. 10.1). We choose the unity scales along the axes. This means that the basis of the coordinate system constructed is orthonormal.

Let $M = M(x, y)$ be some arbitrary point of the hyperbola in Fig. 10.1. According to the definition of a hyperbola 10.1 the modulus of the difference $\big||MF_1| - |MF_2|\big|$ is a constant which does not depend on a location of the point M on the hyperbola. We denote this constant through $2\,a$ and write the formula

$$\big||MF_1| - |MF_2|\big| = 2\,a. \tag{10.1}$$

According to (10.1) the points of the hyperbola are subdivided into two subsets. For the points of one of these two subsets the condition (10.1) is written as

$$|MF_1| - |MF_2| = 2\,a. \tag{10.2}$$

These points constitute the right branch of the hyperbola. For the points of the other subset the condition (10.1) is written as

$$|MF_2| - |MF_1| = 2\,a. \tag{10.3}$$

These points constitute the left branch of the hyperbola.

The length of the segment $[F_1F_2]$ is also a constant. Let's denote this constant through $2\,c$. This yields

$$|F_1O| = |OF_2| = c, \qquad |F_1F_2| = 2\,c. \tag{10.4}$$

From the triangle inequality $|F_1F_2| \geqslant \big||MF_1| - |MF_2|\big|$, from

(10.2), from (10.3), and from (10.4) we derive

$$c \geqslant a. \tag{10.5}$$

The case $c = a$ in (10.5) corresponds to a degenerate hyperbola. In this case the inequality $|F_1F_2| \geqslant \big||MF_1| - |MF_2|\big|$ turns to the equality $|F_1F_2| = \big||MF_1| - |MF_2|\big|$ which distinguishes two subcases. In the first subcase the equality $|F_1F_2| = \big||MF_1| - |MF_2|\big|$ is written as

$$|F_1F_2| = |MF_1| - |MF_2|. \tag{10.6}$$

The equality (10.6) means that the triangle F_1MF_2 collapses into the segment $[F_1M]$, i. e. the point M lies on the ray going along the x-axis to the right from the point F_2 (see Fig. 10.1).

In the second subcase of the case $c = a$ the equality $|F_1F_2| = \big||MF_1| - |MF_2|\big|$ is written as follows

$$|F_1F_2| = |MF_2| - |MF_1|. \tag{10.7}$$

The equality (10.7) means that the triangle F_1MF_2 collapses into the segment $[F_1M]$, i. e. the point M lies on the ray going along the x-axis to the left from the point F_1 (see Fig. 10.1).

As we see considering the above two subcases, if $c = a$ the degenerate hyperbola is the union of two non-intersecting rays lying on the x-axis and being opposite to each other.

Another form of a degenerate hyperbola arises if $a = 0$. In this case the branches of the hyperbola straighten and, gluing with each other, lie on the y-axis.

Both cases of degenerate hyperbolas are usually excluded from the consideration by means of the following inequalities:

$$c > a > 0. \tag{10.8}$$

The formulas (10.4) determine the coordinates of the foci F_1

and F_2 of our hyperbola in the chosen coordinate system:

$$F_1 = F_1(-c, 0), \qquad F_2 = F_2(c, 0). \tag{10.9}$$

Knowing the coordinates of the points F_1 and F_2 and knowing the coordinates of the point $M = M(x, y)$, we write the formulas

$$\begin{aligned} |MF_1| &= \sqrt{y^2 + (x+c)^2}, \\ |MF_2| &= \sqrt{y^2 + (x-c)^2}. \end{aligned} \tag{10.10}$$

Derivation of the canonical equation of a hyperbola. As we have seen above, the equality (10.1) defining a hyperbola breaks into two equalities (10.2) and (10.3) corresponding to the right and left branches of the hyperbola. Let's unite them back into a single equality of the form

$$|MF_1| - |MF_2| = \pm\, 2\, a. \tag{10.11}$$

Let's substitute the formulas (10.10) into the equality (10.11):

$$\sqrt{y^2 + (x+c)^2} - \sqrt{y^2 + (x-c)^2} = \pm\, 2\, a. \tag{10.12}$$

Then we move one of the square roots to the right hand side of the formula (10.12). As a result we derive

$$\sqrt{y^2 + (x+c)^2} = \pm\, 2\, a + \sqrt{y^2 + (x-c)^2}. \tag{10.13}$$

Squaring both sides of the equality (10.13), we get

$$\begin{aligned} y^2 + (x+c)^2 &= 4\, a^2 \pm \\ \pm\, 4\, a\, \sqrt{y^2 + (x-c)^2} &+ y^2 + (x-c)^2. \end{aligned} \tag{10.14}$$

Upon expanding brackets and collecting similar terms the equality (10.14) can be written in the following form:

$$\mp 4\, a\, \sqrt{y^2 + (x-c)^2} = 4\, a^2 - 4\, x\, c. \tag{10.15}$$

Let's cancel the factor four in (10.15) and square both sides of this equality. As a result we get the formula

$$a^2\,(y^2+(x-c)^2) = a^4 - 2\,a^2\,x\,c + x^2\,c^2. \tag{10.16}$$

Upon expanding brackets and recollecting similar terms the equality (10.16) can be written in the following form:

$$x^2\,(a^2-c^2) + y^2\,a^2 = a^2\,(a^2-c^2). \tag{10.17}$$

An attentive reader can note that the above calculations ale almost literally the same as the corresponding calculations for the case of an ellipse. As for the resulting formula (10.17), it coincides exactly with the formula (6.12). But, nevertheless, there is a difference. It consists in the inequalities (10.8), which are different from the inequalities (6.4) for an ellipse.

Both sides of the equality (10.17) comprises the quantity a^2-c^2. Due to the inequalities (10.8) this quantity is negative. For this reason the quantity a^2-c^2 can be written as the square of some positive quantity $b>0$ taken with the minus sign:

$$a^2 - c^2 = -b^2. \tag{10.18}$$

Due to (10.18) the equality (10.17) can be written as

$$-x^2\,b^2 + y^2\,a^2 = -a^2\,b^2. \tag{10.19}$$

Since $b>0$ and $a>0$ (see inequalities (10.8)), the above equality (10.19) transforms to the following one:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{10.20}$$

Definition 10.2. The equality (10.20) is called the *canonical equation* of a hyperbola.

THEOREM 10.1. *For each point $M(x, y)$ of the hyperbola determined by the initial equation* (10.12) *its coordinates satisfy the canonical equation* (10.20).

The above derivation of the canonical equation (10.20) of a hyperbola proves the theorem 10.1. The canonical equation leads (10.20) to the following important inequality:

$$|x| \geqslant a. \tag{10.21}$$

THEOREM 10.2. *The canonical equation of a hyperbola* (10.20) *is equivalent to the initial equation* (10.12).

PROOF. The proof of the theorem 10.2 is analogous to the proof of the theorem 6.2. In order to prove the theorem 10.2 we calculate the expressions (10.10) relying on the equation (10.20). The equation (10.20) itself can be written as follows:

$$y^2 = \frac{b^2}{a^2}\,x^2 - b^2. \tag{10.22}$$

Substituting (10.22) into the first formula (10.10), we get

$$\begin{aligned} |MF_1| &= \sqrt{\frac{b^2}{a^2}\,x^2 - b^2 + x^2 + 2\,x\,c + c^2} = \\ &= \sqrt{\frac{a^2 + b^2}{a^2}\,x^2 + 2\,x\,c + (c^2 - b^2)}. \end{aligned} \tag{10.23}$$

Now we take into account the relationship (10.18) and write the equality (10.23) in the following form:

$$|MF_1| = \sqrt{a^2 + 2\,x\,c + \frac{c^2}{a^2}\,x^2} = \sqrt{\Big(\frac{a^2 + c\,x}{a}\Big)^2}. \tag{10.24}$$

Upon calculating the square root the formula (10.24) yields

$$|MF_1| = \frac{|a^2 + c\,x|}{a}. \tag{10.25}$$

From the inequalities (10.8) and (10.21) we derive $|c\,x| > a^2$. Therefore the formula (10.25) can be written as follows:

$$|MF_1| = \frac{c\,|x| + \operatorname{sign}(x)\,a^2}{a} = \frac{c\,|x|}{a} + \operatorname{sign}(x)\,a. \tag{10.26}$$

In the case of the second formula (10.10) the considerations, analogous to the above ones, yield the following result:

$$|MF_2| = \frac{c\,|x| - \operatorname{sign}(x)\,a^2}{a} = \frac{c\,|x|}{a} - \operatorname{sign}(x)\,a. \tag{10.27}$$

Let's subtract the equality (10.27) from the equality (10.26). Then we get the following relationships:

$$|MF_1| - |MF_2| = 2\,\operatorname{sign}(x)\,a = \pm 2\,a. \tag{10.28}$$

The plus sign in (10.28) corresponds to the case $x > 0$, which corresponds to the right branch of the hyperbola in Fig. 10.1. The minus sign corresponds to the left branch of the hyperbola. Due to what was said the equality (10.28) is equivalent to the equality (10.11), which in turn is equivalent to the equality (10.12). The theorem 10.2 is proved. □

Let's consider again the inequality (10.21) which should be fulfilled for the x-coordinate of any point M on the hyperbola. The inequality (10.21) turns to an equality if M coincides with A or if M coincides with C (see Fig. 10.1).

Definition 10.3. The points A and C in Fig. 10.1 are called the *vertices* of the hyperbola. The segment $[AC]$ is called the *transverse axis* or the *real axis* of the hyperbola, while the segments $[OA]$ and $[OC]$ are called its *real semiaxes*.

The constant a in the equation of the hyperbola (10.20) is the length of the segment $[OA]$ in Fig. 10.1 (the length of the real semiaxis of a hyperbola). As for the constant b, there is no

segment of the length b in Fig. 10.1. For this reason the constant b is called the length of the *imaginary semiaxis* of a hyperbola, i. e. the semiaxis which does not actually exist.

Definition 10.4. A coordinate system O, $\mathbf{e}_1$, $\mathbf{e}_2$ with an orthonormal basis $\mathbf{e}_1$, $\mathbf{e}_2$ where a hyperbola is given by its canonical equation (10.20) is called a *canonical coordinate system* of this hyperbola.

§ 11. The eccentricity and directrices of a hyperbola. The property of directrices.

The shape and sizes of a hyperbola are determined by two constants a and b in its canonical equation (10.20).. Due to the relationship (10.18) the constant b can be expressed through the constant c. Multiplying both constants a and c by the same number, we change the sizes of a hyperbola, but do not change its shape. The ratio of these two constants

$$\varepsilon = \frac{c}{a}. \tag{11.1}$$

is responsible for the shape of a hyperbola.

Definition 11.1. The quantity ε defined by the relationship (11.1), where a is the real semiaxis and c is the half of the interfocal distance, is called the *eccentricity* of a hyperbola.

The eccentricity (11.1) is used in order to define one more parameter of a hyperbola. It is usually denoted through d:

$$d = \frac{a}{\varepsilon} = \frac{a^2}{c}. \tag{11.2}$$

Definition 11.2. On the plane of a hyperbola there are two lines perpendicular to its real axis and placed at the distance d given by the formula (7.2) from its center. These lines are called *directrices* of a hyperbola.

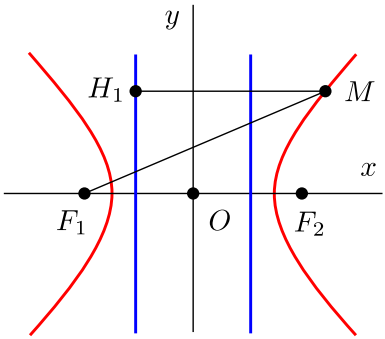

Fig. 11.1

Each hyperbola has two foci and two directrices. Each directrix has the corresponding focus of it. This is that of two foci which is more close to the directrix in question. Let $M(x, y)$ be some arbitrary point of a hyperbola. Let's connect this point with the left focus of the hyperbola F_1 and drop the perpendicular from it to the left directrix of the hyperbola. Let's denote through H_1 the base of such a perpendicular and calculate its length $|MH_1|$:

$$|MH_1| = |x - (-d)| = |d + x|. \tag{11.3}$$

Taking into account (11.2), the formula (11.3) can be brought to

$$|MH_1| = \left|\frac{a^2}{c} + x\right| = \frac{|a^2 + c\,x|}{c}. \tag{11.4}$$

The length of the segment MF_1 was already calculated above. Initially it was given by one of the formulas (10.10), but later the more simple expression (10.25) was derived for it:

$$|MF_1| = \frac{|a^2 + c\,x|}{a}. \tag{11.5}$$

If we divide (11.5) by (11.4), we obtain the following relationship:

$$\frac{|MF_1|}{|MH_1|} = \frac{c}{a} = \varepsilon. \tag{11.6}$$

The point M can change its position on the hyperbola. Then the numerator and the denominator of the fraction (11.6) are

changed, but its value remains unchanged. This fact is known as the property of directrices.

THEOREM 11.1. *The ratio of the distance from some arbitrary point M of a hyperbola to its focus and the distance from this point to the corresponding directrix is a constant which is equal to the eccentricity of the hyperbola.*

§ 12. The equation of a tangent line to a hyperbola.

Let's consider a hyperbola given by its canonical equation (10.20) in its canonical coordinate system (see Definition 10.4). Let's draw a tangent line to this hyperbola and denote through $M = M(x_0, y_0)$ its tangency point (see Fig. 12.1). Our goal is to write the equation of a tangent line to a hyperbola.

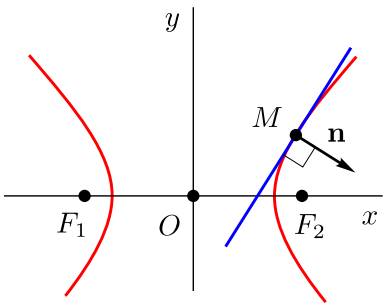

Fig. 12.1

A hyperbola consists of two branches, each branch being a curve composed by two halves — the upper half and the lower half. The upper halves of the hyperbola branches can be treated as a graph of some function of the form

$$y = f(x) \tag{12.1}$$

defined in the union of two intervals $(-\infty, -a) \cup (a, +\infty)$. Lower halves of a hyperbola can also be treated as a graph of some function of the form (12.1) with the same domain. The equation of a tangent line to the graph of the function (12.1) is given by the following well-known formula (see [9]):

$$y = y_0 + f'(x_0)\,(x - x_0). \tag{12.2}$$

In order to apply the formula (12.2) to a hyperbola we need

to calculate the derivative of the function (12.1). Let's substitute the function (12.1) into the equation (10.20):

$$\frac{x^2}{a^2} - \frac{(f(x))^2}{b^2} = 1. \tag{12.3}$$

The equality (12.3) is fulfilled identically in x. Let's differentiate the equality (12.3) with respect to x. This yields

$$\frac{2\,x}{a^2} - \frac{2\,f(x)\,f'(x)}{b^2} = 0. \tag{12.4}$$

Let's apply (12.4) for to calculate the derivative $f'(x)$:

$$f'(x) = \frac{b^2\,x}{a^2\,f(x)}. \tag{12.5}$$

In order to substitute (12.5) into the equation (12.2) we change x for x_0 and $f(x)$ for $f(x_0) = y_0$. As a result we get

$$f'(x_0) = \frac{b^2\,x_0}{a^2\,y_0}. \tag{12.6}$$

Let's substitute (12.6) into the equation of the tangent line (12.2). This yields the following relationship

$$y = y_0 + \frac{b^2\,x_0}{a^2\,y_0}\,(x - x_0). \tag{12.7}$$

Eliminating the denominator, we write the equality (12.7) as

$$a^2\,y\,y_0 - b^2\,x\,x_0 = a^2\,y_0^2 - b^2\,x_0^2. \tag{12.8}$$

Now let's divide both sides of the equality (12.8) by $a^2\,b^2$:

$$\frac{x\,x_0}{a^2} - \frac{y\,y_0}{b^2} = \frac{x_0^2}{a^2} - \frac{y_0^2}{b^2}. \tag{12.9}$$

Note that the point $M = M(x_0, y_9)$ is on the hyperbola. Therefore its coordinates satisfy the equation (10.20):

$$\frac{x_0^2}{a^2} - \frac{y_0^2}{b^2} = 1. \tag{12.10}$$

Taking into account (12.10), we can transform (12.9) to

$$\frac{x\,x_0}{a^2} - \frac{y\,y_0}{b^2} = 1. \tag{12.11}$$

THEOREM 12.1. *For a hyperbola determined by its canonical equation* (10.20) *its tangent line that touches this hyperbola at the point* $M = M(x_0, y_0)$ *is given by the equation* (12.11).

The equation (12.11) is a particular case of the equation (3.22) where the constants A, B, and D are given by the formulas

$$A = \frac{x_0}{a^2}, \qquad B = -\frac{y_0}{b^2}, \qquad D = 1. \tag{12.12}$$

According to the definition 3.6 and the formulas (3.21) the constants A and B in (12.12) are the covariant components of the normal vector for the tangent line to a hyperbola. The tangent line equation (12.11) is written in a canonical coordinate system of a hyperbola. The basis of such a coordinate system is orthonormal. Therefore the formula (3.19) and the formula (32.4) from Chapter I yield the following relationships:

$$A = n_1 = n^1, \qquad B = n_2 = n^2. \tag{12.13}$$

THEOREM 12.2. *The quantities* A *and* B *in* (12.12) *are the coordinates of the normal vector* $\mathbf{n}$ *for the tangent line to a hyperbola which is given by the equation* (12.11).

The relationships (12.13) prove the theorem 12.2.

§ 13. Focal property of a hyperbola.

Like in the case of an ellipse, assume that a hyperbola is manufactured of a thin strip of some flexible material and assume that its surface is covered with a light reflecting layer. For such a hyperbola we can formulate the following focal property.

THEOREM 13.1. *A light ray emitted in one focus of a hyperbola upon reflecting on its surface goes to infinity so that its backward extension passes through the other focus of this hyperbola.*

THEOREM 13.2. *The tangent line of a hyperbola is a bisector in the triangle composed by the tangency point and two foci of the hyperbola.*

The theorem 13.2 is a purely geometric version of the theorem 13.1. These theorems are equivalent due to the reflection law saying that the angle of reflection is equal to the angle of incidence.

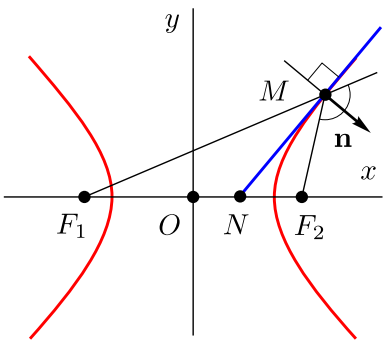

Fig. 13.1

PROOF. Let's consider some point $M = M(x_0, y_0)$ on the hyperbola and draw the tangent line to the hyperbola through this point as shown in Fig. 13.1. Let N be the x-intercept of this tangent line. Due to M and N, we have the segment $[MN]$. This segment is perpendicular to the normal vector $\mathbf{n}$ of the tangent line. In order to prove that $[MN]$ is a bisector in the triangle F_1MF_2 it is sufficient to prove that $\mathbf{n}$ is directed along the bisector for the external angle of this triangle at its vertex M. This condition can be written as

$$\frac{(\overrightarrow{F_1M}, \mathbf{n})}{|F_1M|} = \frac{(\overrightarrow{MF_2}, \mathbf{n})}{|MF_2|}. \tag{13.1}$$

The coordinates of the points F_1 and F_2 in a canonical coordinate system of the hyperbola are known (see formulas (10.9)). The coordinates of the point $M = M(x_0, y_0)$ are also known. Therefore we can find the coordinates of the vectors $\overrightarrow{F_1M}$ and $\overrightarrow{MF_2}$ used in the above formula (13.1):

$$\overrightarrow{F_1M} = \begin{Vmatrix} x_0 + c \\ y_0 \end{Vmatrix}, \qquad \overrightarrow{MF_2} = \begin{Vmatrix} c - x_0 \\ -y_0 \end{Vmatrix}. \tag{13.2}$$

The tangent line that touches the hyperbola at the point $M = M(x_0, y_0)$ is given by the equation (12.11). The coordinates of the normal vector $\mathbf{n}$ of this tangent line in Fig. 13.1 are given by the formulas (12.12) and (12.13):

$$\mathbf{n} = \begin{Vmatrix} \dfrac{x_0}{a^2} \\ -\dfrac{y_0}{b^2} \end{Vmatrix} \tag{13.3}$$

Relying upon (13.2) and (13.3), we apply the formula (33.3) from Chapter I in order to calculate the scalar products in (13.1):

$$\begin{aligned} (\overrightarrow{F_1M}, \mathbf{n}) &= \frac{c\,x_0 + x_0^2}{a^2} - \frac{y_0^2}{b^2} = \frac{c\,x_0}{a^2} + \frac{x_0^2}{a^2} - \frac{y_0^2}{b^2}, \\ (\overrightarrow{MF_2}, \mathbf{n}) &= \frac{c\,x_0 - x_0^2}{a^2} + \frac{y_0^2}{b^2} = \frac{c\,x_0}{a^2} - \frac{x_0^2}{a^2} + \frac{y_0^2}{b^2}. \end{aligned} \tag{13.4}$$

The coordinates of the point M satisfy the equation (10.20):

$$\frac{x_0^2}{a^2} - \frac{y_0^2}{b^2} = 1. \tag{13.5}$$

Due to (13.5) the formulas (13.4) simplify to

$$\begin{aligned} (\overrightarrow{F_1M}, \mathbf{n}) &= \frac{c\,x_0}{a^2} + 1 = \frac{c\,x_0 + a^2}{a^2}, \\ (\overrightarrow{MF_2}, \mathbf{n}) &= \frac{c\,x_0}{a^2} - 1 = \frac{c\,x_0 - a^2}{a^2}. \end{aligned} \tag{13.6}$$

In order to calculate the denominators in the formula (13.1) we use the formulas (10.26) and (10.27). In this case, when applied to the point $M = M(x_0, y_0)$, they yield

$$\begin{aligned} |MF_1| &= \frac{c\,|x_0| + \operatorname{sign}(x_0)\,a^2}{a}, \\ |MF_2| &= \frac{c\,|x_0| - \operatorname{sign}(x_0)\,a^2}{a}. \end{aligned} \tag{13.7}$$

Due to the purely numeric identity $|x_0| = \operatorname{sign}(x_0)\,x_0$ we can write (13.7) in the following form:

$$\begin{aligned} |MF_1| &= \frac{c\,x_0 + a^2}{a}\ \operatorname{sign}(x_0), \\ |MF_2| &= \frac{c\,x_0 - a^2}{a}\ \operatorname{sign}(x_0). \end{aligned} \tag{13.8}$$

From the formulas (13.6) and (13.8) we easily derive the equalities

$$\frac{(\overrightarrow{F_1M}, \mathbf{n})}{|MF_1|} = \frac{\operatorname{sign}(x_0)}{a}, \qquad \frac{(\overrightarrow{MF_2}, \mathbf{n})}{|MF_2|} = \frac{\operatorname{sign}(x_0)}{a}$$

that prove the equality (13.1). The theorem 13.2 is proved. □

As we noted above the theorem 13.1 is equivalent to the theorem 13.2 due to the light reflection law. Therefore the theorem 13.1 is also proved.

§ 14. Asymptotes of a hyperbola.

Asymptotes are usually understood as some straight lines to which some points of a given curve come unlimitedly close along some unlimitedly long fragments of this curve. Each hyperbola has two asymptotes (see Fig. 14.1). In a canonical coordinate system the asymptotes of the hyperbola given by the equation

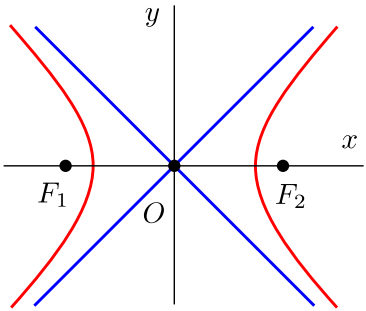

Fig. 14.1

(10.20) are determined by the following equations:

$$y = \pm\frac{b}{a}\,x. \tag{14.1}$$

One of the asymptotes is associated with the plus sign in the formula (14.1), the other asymptote is associated with the opposite minus sign.

The theory of asymptotes is closely related to the theory of limits. This theory is usually studied within the course of mathematical analysis (see [9]). For this reason I do not derive the equations (14.1) in this book.

§ 15. Parabola. Canonical equation of a parabola.

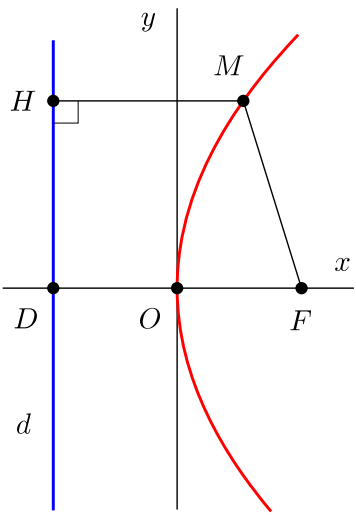

Fig. 15.1

DEFINITION 15.1. *A parabola* is a set of points on some plane each of which is equally distant from some fixed point F of this plane and from some straight line d lying on this plane. The point F is called the *focus* of this parabola, while the line d is called its *directrix*.

Assume that a parabola with the focus F and with the directrix d is given. Let's drop the perpendicular from the point F onto the line d and denote through D the base of such a perpendicular. Let's choose the

line DF for the x-axis of a coordinate system. Then we denote through O the midpoint of the segment $[DF]$ and choose this point O for the origin. And finally, we draw the y-axis of a coordinate system through the point O perpendicular to the line DF (see Fig. 15.1). Choosing the unity scales along the axes, we ultimately fix a coordinate system with an orthonormal basis on the parabola plane.

Let's denote through p the distance from the focus F of the parabola to its directrix d, i. e. we set

$$|DF| = p. \tag{15.1}$$

The point O is the midpoint of the segment $[DF]$. Therefore the equality (15.1) leads to the following equalities:

$$|DO| = |OF| = \frac{p}{2}. \tag{15.2}$$

The relationships (15.2) determine the coordinates of D and F:

$$D = D(-p/2, 0), \qquad F = F(p/2, 0). \tag{15.3}$$

Let $M = M(x, y)$ be some arbitrary point of the parabola. According to the definition 15.1, the following equality is fulfilled:

$$|MF| = |MH| \tag{15.4}$$

(see Fig. 15.1). Due to (15.3) the length of the segment $[MF]$ in the chosen coordinate system is given by the formula

$$|MF| = \sqrt{y^2 + (x - p/2)^2)}. \tag{15.5}$$

The formula for the length of $[MH]$ is even more simple:

$$|MH| = x + p/2. \tag{15.6}$$

Substituting (15.5) and (15.6) into (15.4), we get the equation

$$\sqrt{y^2 + (x - p/2)^2)} = x + p/2. \tag{15.7}$$

Let's square both sides of the equation (15.7):

$$y^2 + (x - p/2)^2) = (x + p/2)^2. \tag{15.8}$$

Upon expanding brackets and collecting similar terms in (15.8), we bring this equation to the following form:

$$y^2 = 2\,p\,x. \tag{15.9}$$

DEFINITION 15.2. The equality (15.9) is called the *canonical equation* of a parabola.

THEOREM 15.1. *For each point $M(x, y)$ of the parabola determined by the initial equation* (15.7) *its coordinates satisfy the canonical equation* (15.9).

Due to (15.1) the constant p in the equation (15.9) is a non-negative quantity. The case $p = 0$ corresponds to the degenerate parabola. From the definition 15.1 it is easy to derive that in this case the parabola turns to the straight line coinciding with the x-axis in Fig. 15.1. The case of the degenerate parabola is excluded by means of the inequality

$$p > 0. \tag{15.10}$$

Due to the inequality (15.10) from the equation (15.9) we derive

$$x \geqslant 0. \tag{15.11}$$

THEOREM 15.2. *The canonical equation of the parabola* (15.9) *is equivalent to the initial equation* (15.7).

PROOF. In order to prove the theorem 15.2 it is sufficient

to invert the calculations performed in deriving the equality (15.9) from (15.7). Note that the passage from (15.8) to (15.9) is invertible. The passage from (15.7) to (15.8) is also invertible due to the inequality (15.11), which follows from the equation (15.9). This observation completes the proof of the theorem 15.2. □

DEFINITION 15.3. The point O in Fig. 15.1 is called the *vertex* of the parabola, the line DF coinciding with the x-axis is called the *axis* of the parabola.

DEFINITION 15.4. A coordinate system O, $\mathbf{e}_1$, $\mathbf{e}_2$ with an orthonormal basis $\mathbf{e}_1$, $\mathbf{e}_2$ where a parabola is given by its canonical equation (15.9) and where the inequality (15.10) is fulfilled is called a *canonical coordinate system* of this parabola.

§ 16. The eccentricity of a parabola.

The definition of a parabola 15.1 is substantially different from the definition of an ellipse 6.1 and from the definition of a hyperbola 10.1. But it is similar to the property of directrices of an ellipse in the theorem 7.1 and to the property of directrices of a hyperbola in the theorem 11.1. Comparing the definition of a parabola 15.1 with these theorems, we can formulate the following definition.

DEFINITION 16.1. The eccentricity of a parabola is postulated to be equal to the unity: $\varepsilon = 1$.

§ 17. The equation of a tangent line to a parabola.

Let's consider a parabola given by its canonical equation (15.9) in its canonical coordinate system (see Definition 15.4). Let's draw a tangent line to this parabola and denote through $M = M(x_0, y_0)$ the tangency point (see Fig. 17.1). Our goal is to write the equation of the tangent line to the parabola through the point $M = M(x_0, y_0)$.

An parabola is a curve composed by two halves — the upper half and the lower half. Any one of these two halves of

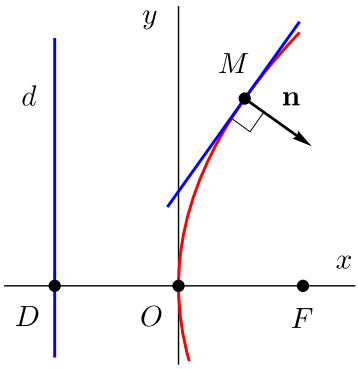

Fig. 17.1

a parabola can be treated as a graph of some function

$$y = f(x) \tag{17.1}$$

with the domain $(0, +\infty)$. The equation of a tangent line to the graph of a function (17.1) is given by the well-known formula

$$\begin{aligned} y - y_0 = \\ = f'(x_0)\,(x - x_0). \end{aligned} \tag{17.2}$$

(see [9]). In order to apply the formula (17.2) to a parabola one should calculate the derivative of the function (17.1). Let's substitute (17.1) into the equation (15.9):

$$(f(x))^2 = 2\,p\,x. \tag{17.3}$$

The equality (17.3) is fulfilled identically in x. Let's differentiate the equality (17.3) with respect to x. This yields

$$2\,f(x)\,f'(x) = 2\,p. \tag{17.4}$$

Let's apply the formula (17.4) for to calculate the derivative

$$f'(x) = \frac{p}{f(x)}. \tag{17.5}$$

In order to substitute (17.5) into the equation (17.2) we change x for x_0 and $f(x)$ for $f(x_0) = y_0$. As a result we get

$$f'(x_0) = \frac{p}{y_0}. \tag{17.6}$$

Let's substitute (17.6) into the equation of the tangent line

(17.2). This yields the following relationship:

$$y - y_0 = \frac{p}{y_0}\,(x - x_0). \tag{17.7}$$

Eliminating the denominator, we write the equality (17.7) as

$$y\,y_0 - y_0^2 = p\,x - p\,x_0. \tag{17.8}$$

Note that the point $M = M(x_0, y_9)$ is on the parabola. Therefore its coordinates satisfy the equality (15.9):

$$y_0^2 = 2\,p\,x_0. \tag{17.9}$$

Taking into account (17.9), we can transform (17.8) to

$$y\,y_0 = p\,x + p\,x_0. \tag{17.10}$$

This is the required equation of a tangent line to a parabola.

THEOREM 17.1. *For a parabola determined by its canonical equation* (15.9) *the tangent line that touches this parabola at the point* $M = M(x_0, y_0)$ *is given by the equation* (17.10).

Let's write the equation of a tangent line to a parabola in the following slightly transformed form:

$$p\,x - y\,y_0 + p\,x_0 = 0. \tag{17.11}$$

The equation (17.11) is a special instance of the equation (3.22) where the constants A, B, and D are given by the formulas

$$A = p, \qquad B = -y_0, \qquad D = p\,x_0. \tag{17.12}$$

According to the definition 3.6 and the formulas (3.21), the constants A and B in (17.12) are the covariant components of the normal vector of a tangent line to a parabola. The

equation (17.11) is written in a canonical coordinate system of the parabola. The basis of a canonical system is orthonormal (see Definition 15.4). In the case of an orthonormal basis the formula (3.19) and the formula (32.4) from Chapter I yield

$$A = n_1 = n^1, \qquad B = n_2 = n^2. \tag{17.13}$$

THEOREM 17.2. *The quantities A and B in (17.12) are the coordinates of the normal vector $\mathbf{n}$ of a tangent line to a parabola in the case where this tangent line is given by the equation (17.10).*

The relationships (17.13) prove the theorem 17.2.

§18. Focal property of a parabola.

Assume that we have a parabola manufactured of a thin strip of some flexible material covered with a light reflecting layer. For such a parabola the following focal property is formulated.

THEOREM 18.1. *A light ray emitted from the focus of a parabola upon reflecting on its surface goes to infinity parallel to the axis of this parabola.*

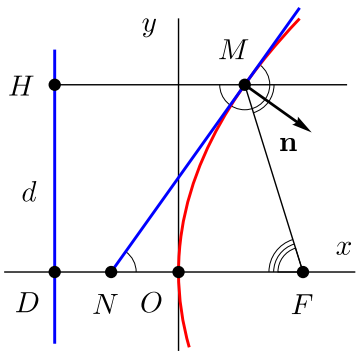

Fig. 18.1

THEOREM 18.2. *For any tangent line of a parabola the triangle formed by the tangency point M, its focus F, and by the point N at which this tangent line intersects the axis of the parabola is an isosceles triangle, i. e. the equality $|MF| = |NF|$ holds.*

As we see in Fig. 18.1, the theorem 18.2 follows from the theorem 18.1 due to the reflection law saying that the angle of reflection is equal to the angle of

incidence and due to the equality of inner crosswise lying angles in the intersection of two parallel lines by a third line (see [6]).

PROOF OF THE THEOREM 18.2. For to prove this theorem we choose some tangency point $M(x_0, y_0)$ and write the equation of the tangent line to the parabola in the form of (17.10):

$$y\,y_0 = p\,x + p\,x_0. \tag{18.1}$$

Let's find the intersection point of the tangent line (18.1) with the axis of the parabola. In a canonical coordinate system the axis of the parabola coincides with the x-axis (see Definition 15.3). Substituting $y = 0$ into the equation (18.1), we get $x = -x_0$, which determines the coordinates of the point N:

$$N = N(-x_0, 0). \tag{18.2}$$

From (18.2) and (15.3) we derive the length of the segment $[NF]$:

$$|NF| = p/2 - (-x_0) = p/2 + x_0. \tag{18.3}$$

In the case of a parabola the length of the segment $[MF]$ coincides with the length of the segment $[MH]$ (see Definition 15.1). Therefore from (15.3) we derive

$$|MF| = |MH| = x_0 - (-p/2) = x_0 + p/2. \tag{18.4}$$

Comparing (18.3) and (18.4) we get the required equality $|MF| = |NF|$. As a result the theorem 18.2 is proved. As for the theorem 18.1, it equivalent to the theorem 18.2. □

§ 19. The scale of eccentricities.

The eccentricity of an ellipse is determine by the formula (7.1), where the parameters c and a are related by the inequalities (6.4). Hence the eccentricity of an ellipse obeys the inequalities

$$0 \leqslant \varepsilon < 1. \tag{19.1}$$

The eccentricity of a parabola is equal to unity by definition. Indeed, the definition 16.1 yields

$$\varepsilon = 1. \tag{19.2}$$

The eccentricity of a hyperbola is defined by the formula (11.1), where the parameters c and a obey the inequalities (10.8). Hence the eccentricity of a hyperbola obeys the inequalities

$$1 < \varepsilon \leqslant +\infty. \tag{19.3}$$

The formulas (19.1), (19.2), and (19.3) show that the eccentricities of ellipses, parabolas, and hyperbolas fill the interval from 0 to $+\infty$ without omissions, i. e. we have the *continuous scale of eccentricities*.

§ 20. Changing a coordinate system.

Let O, $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ and $\tilde{O}$, $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$ be two Cartesian coordinate systems in the space $\mathbb{E}$. They consist of the bases $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ and $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$ complemented with two points O and $\tilde{O}$, which are called origins (see Definition 1.1). The transition from the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ to the basis $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$ and vice versa is described by two transition matrices S and T whose components are in the following transition formulas:

$$\tilde{\mathbf{e}}_j = \sum_{i=1}^{3} S_j^i \, \mathbf{e}_i, \qquad\qquad \mathbf{e}_j = \sum_{i=1}^{3} T_j^i \, \tilde{\mathbf{e}}_i \tag{20.1}$$

(see formulas (22.4) and (22.9) in Chapter I).

In order to describe the transition from the coordinate system O, $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ to the coordinate system $\tilde{O}$, $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$ and vice versa two auxiliary parameters are employed. These are the vectors $\mathbf{a} = \overrightarrow{O\tilde{O}}$ and $\tilde{\mathbf{a}} = \overrightarrow{\tilde{O}O}$.

DEFINITION 20.1. The vectors $\mathbf{a} = \overrightarrow{O\tilde{O}}$ and $\tilde{\mathbf{a}} = \overrightarrow{\tilde{O}O}$ are called the *origin displacement vectors*.

The origin displacement vector $\mathbf{a} = \overrightarrow{O\tilde{O}}$ is usually expanded in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, while the other origin displacement vector $\tilde{\mathbf{a}} = \overrightarrow{\tilde{O}O}$ is expanded in the basis $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$:

$$\mathbf{a} = \sum_{i=1}^{3} a^i\,\mathbf{e}_i, \qquad\qquad \tilde{\mathbf{a}} = \sum_{i=1}^{3} \tilde{a}^i\,\tilde{\mathbf{e}}_i. \tag{20.2}$$

The vectors $\mathbf{a} = \overrightarrow{O\tilde{O}}$ are $\tilde{\mathbf{a}} = \overrightarrow{\tilde{O}O}$ opposite to each other, i. e. the following relationships are fulfilled:

$$\mathbf{a} = -\tilde{\mathbf{a}}, \qquad\qquad \tilde{\mathbf{a}} = -\mathbf{a}. \tag{20.3}$$

Their coordinates in the expansions (20.2) are related with each other by means of the formulas

$$\tilde{a}^i = -\sum_{j=1}^{3} T^i_j\,a^j, \qquad\qquad a^i = -\sum_{j=1}^{3} S^i_j\,\tilde{a}^j. \tag{20.4}$$

The formulas (20.4) are derived from the formulas (20.3) with the use of the formulas (25.4) and (25.5) from Chapter I.

§ 21. Transformation of the coordinates of a point under a change of a coordinate system.

Let O, $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ and $\tilde{O}$, $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$, $\tilde{\mathbf{e}}_3$ be two Cartesian coordinate systems in the space $\mathbb{E}$ and let X be some arbitrary point in the space $\mathbb{E}$. Let's denote through

$$X = X(x^1, x^2, x^3), \qquad\qquad X = X(\tilde{x}^1, \tilde{x}^2, \tilde{x}^3) \tag{21.1}$$

the presentations of the point X in these two coordinate systems. The radius vectors of the point X in these systems are related with each other by means of the relationships

$$\mathbf{r}_X = \mathbf{a} + \tilde{\mathbf{r}}_X, \qquad \tilde{\mathbf{r}}_X = \tilde{\mathbf{a}} + \mathbf{r}_X, \tag{21.2}$$

The coordinates of X in (21.1) are the coordinates of its radius vectors in the bases of the corresponding coordinate systems:

$$\mathbf{r}_X = \sum_{j=1}^{3} x^j\, \mathbf{e}_j, \qquad \tilde{\mathbf{r}}_X = \sum_{j=1}^{3} \tilde{x}^j\, \tilde{\mathbf{e}}_j. \tag{21.3}$$

From (21.2), (21.3), and (20.2), applying the formulas (25.4) and (25.5) from Chapter I, we easily derive the following relationships:

$$\tilde{x}^i = \sum_{j=1}^{3} T^i_j\, x^j + \tilde{a}^i, \qquad x^i = \sum_{j=1}^{3} S^i_j\, \tilde{x}^j + a^i. \tag{21.4}$$

THEOREM 21.1. *Under a change of coordinate systems in the space* $\mathbb{E}$ *determined by the formulas* (20.1) *and* (20.2) *the coordinates of points are transformed according to the formulas* (21.4).

The formulas (21.4) are called the *direct* and *inverse transformation formulas* for the coordinates of a point *under a change of a Cartesian coordinate system.*

§ 22. Rotation of a rectangular coordinate system on a plane. The rotation matrix.

Let O, $\mathbf{e}_1$, $\mathbf{e}_2$ and $\tilde{O}$, $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$ be two Cartesian coordinate system on a plane. The formulas (20.1), (20.2), and (20.4) in this case are written as follows:

$$\tilde{\mathbf{e}}_j = \sum_{i=1}^{2} S^i_j\, \mathbf{e}_i, \qquad \mathbf{e}_j = \sum_{i=1}^{2} T^i_j\, \tilde{\mathbf{e}}_i, \tag{22.1}$$

$$\mathbf{a} = \sum_{i=1}^{2} a^i\,\mathbf{e}_i, \qquad \tilde{\mathbf{a}} = \sum_{i=1}^{2} \tilde{a}^i\,\tilde{\mathbf{e}}_i, \tag{22.2}$$

$$\tilde{a}^i = -\sum_{j=1}^{2} T^i_j\,a^j, \qquad a^i = -\sum_{j=1}^{2} S^i_j\,\tilde{a}^j. \tag{22.3}$$

Let X be some arbitrary point of the plane. Its coordinates in the coordinate systems O, $\mathbf{e}_1$, $\mathbf{e}_2$ and $\tilde{O}$, $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$ are transformed according to the formulas similar to (21.4):

$$\tilde{x}^i = \sum_{j=1}^{2} T^i_j\,x^j + \tilde{a}^i, \qquad x^i = \sum_{j=1}^{2} S^i_j\,\tilde{x}^j + a^i. \tag{22.4}$$

Assume that the origins O and $\tilde{O}$ do coincide. In this case the parameters a^1, a^2 and $\tilde{a}^1$, $\tilde{a}^2$ in the formulas (22.2), (22.3), (22.3), and (22.4) do vanish. Under the assumption $O = \tilde{O}$ we consider the special case where the bases $\mathbf{e}_1$, $\mathbf{e}_2$ and $\tilde{\mathbf{e}}_1$, $\tilde{\mathbf{e}}_2$ both are orthonormal and where one of them is produced from the other by means of the rotation by some angle φ (see Fig. 22.1). For two bases on a plane the transition matrices S and T are square matrices 2×2. The components of the direct transition matrix S are taken from the following formulas:

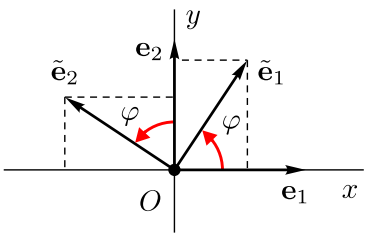

Fig. 22.1

$$\begin{aligned} \tilde{\mathbf{e}}_1 &= \cos\varphi\cdot\mathbf{e}_1 + \sin\varphi\cdot\mathbf{e}_2, \\ \tilde{\mathbf{e}}_2 &= -\sin\varphi\cdot\mathbf{e}_1 + \cos\varphi\cdot\mathbf{e}_2. \end{aligned} \tag{22.5}$$

The formulas (22.5) are derived on the base of Fig. 22.1.

Comparing the formulas (22.5) with the first relationship in (22.1), we get $S^1_1 = \cos\varphi$, $S^2_1 = \sin\varphi$, $S^2_1 = -\sin\varphi$, $S^2_2 = \cos\varphi$.

Hence we have the following formula for the matrix S:

$$S = \begin{Vmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{Vmatrix}. \tag{22.6}$$

DEFINITION 22.1. The square matrix of the form (22.6) is called the *rotation matrix by the angle* φ.

The inverse transition matrix T is the inverse matrix for S (see theorem 23.1 in Chapter I). Due to this fact and due to the formula (22.6) we can calculate the matrix T:

$$T = \begin{Vmatrix} \cos(-\varphi) & -\sin(-\varphi) \\ \sin(-\varphi) & \cos(-\varphi) \end{Vmatrix}. \tag{22.7}$$

The matrix (22.7) is also a rotation matrix by the angle φ. But the angle φ in it is taken with the minus sign, which means that the rotation is performed in the opposite direction.

Let's write the relationships (22.4) taking into account that $\mathbf{a} = 0$ and $\tilde{\mathbf{a}} = 0$, which follows from $O = \tilde{O}$, and taking into account the formulas (22.6) and (22.7):

$$\begin{aligned} \tilde{x}^1 &= \cos(\varphi)\, x^1 + \sin(\varphi)\, x^2, \\ \tilde{x}^2 &= -\sin(\varphi)\, x^1 + \cos(\varphi)\, x^2, \end{aligned} \tag{22.8}$$

$$\begin{aligned} x^1 &= \cos(\varphi)\, \tilde{x}^1 - \sin(\varphi)\, \tilde{x}^2, \\ x^2 &= \sin(\varphi)\, \tilde{x}^1 + \cos(\varphi)\, \tilde{x}^2. \end{aligned} \tag{22.9}$$

The formulas (22.8) and (22.9) are the transformation formulas for the coordinates of a point under the rotation of the rectangular coordinate system shown in Fig. 22.1.

§ 23. Curves of the second order.

DEFINITION 23.1. A *curve of the second order* or a *quadric*

on a plane is a curve which is given by a polynomial equation of the second order in some Cartesian coordinate system

$$A\,x^2 + 2\,B\,x\,y + C\,y^2 + 2\,D\,x + 2\,E\,y + F = 0. \tag{23.1}$$

Here $x = x^1$ and $y = x^2$ are the coordinates of a point on a plane. Note that the transformation of these coordinates under a change of one coordinate system for another is given by the functions of the first order in x^1 and x^2 (see formulas (22.4)). For this reason the general form of the equation of a quadric (23.1) does not change under a change of a coordinate system though the values of the parameters A, B, C, D, E, and F in it can change. From what was said we derive the following theorem.

THEOREM 23.1. *For any curve of the second order on a plane, i. e. for any quadric, there is some rectangular coordinate system with an orthonormal basis such that this curve is given by an equation of the form* (23.1) *in this coordinate system.*

§ 24. Classification of curves of the second order.

Let Γ be a curve of the second order on a plane given by an equation of the form (23.1) in some rectangular coordinate system with the orthonormal basis. Passing from one of such coordinate systems to another, one can change the constant parameters A, B, C, D, E, and F in (23.1), and one can always choose a coordinate system in which the equation (23.1) takes its most simple form.

DEFINITION 24.1. The problem of finding a rectangular coordinate system with the orthonormal basis in which the equation of a curve of the second order Γ takes its most simple form is called the problem of *bringing* the equation of a curve Γ to its *canonical form*.

An ellipse, a hyperbola, and parabola are examples of curves of the second order on a plane. The canonical forms of the

equation (23.1) for these curves are already known to us (see formulas (6.15), (10.20), and (15.9)).

DEFINITION 24.2. The problem of grouping curves by the form of their canonical equations is called the problem of *classification* of curves of the second order.

THEOREM 24.1. *For any curve of the second order* Γ *there is a rectangular coordinate system with an orthonormal basis where the constant parameter* B *of the equation* (23.1) *for* Γ *is equal to zero:* $B = 0$.

PROOF. Let O, $\mathbf{e}_1$, $\mathbf{e}_2$ be rectangular coordinate system with an orthonormal basis $\mathbf{e}_1$, $\mathbf{e}_2$ where the equation of the curve Γ has the form (23.1) (see Theorem 23.1). If $B = 0$ in (23.1), then O, $\mathbf{e}_1$, $\mathbf{e}_2$ is a required coordinate system.

If $B \neq 0$, then we perform the rotation of the coordinate system O, $\mathbf{e}_1$, $\mathbf{e}_2$ about the point O by some angle φ. Such a rotation is equivalent to the change of variables

$$\begin{aligned} x &= \cos(\varphi)\,\tilde{x} - \sin(\varphi)\,\tilde{y}, \\ y &= \sin(\varphi)\,\tilde{x} + \cos(\varphi)\,\tilde{y} \end{aligned} \tag{24.1}$$

in the equation (23.1) (see formulas (22.9)). Upon substituting (24.1) into the equation (23.1) we get the analogous equation

$$\tilde{A}\,\tilde{x}^2 + 2\,\tilde{B}\,\tilde{x}\,\tilde{y} + \tilde{C}\,\tilde{y}^2 + 2\,\tilde{D}\,\tilde{x} + 2\,\tilde{E}\,\tilde{y} + \tilde{F} = 0 \tag{24.2}$$

whose parameters $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, $\tilde{D}$, $\tilde{E}$, and $\tilde{F}$ are expressed through the parameters A, B, C, D, E, and F of the initial equation (23.1). For the parameter $\tilde{B}$ in (24.2) we derive the formula

$$\begin{aligned} \tilde{B} &= (C - A)\,\cos(\varphi)\,\sin(\varphi) + B\,\cos^2(\varphi) - \\ &- B\,\sin^2(\varphi) = \frac{C - A}{2}\,\sin(2\,\varphi) + B\,\cos(2\,\varphi). \end{aligned} \tag{24.3}$$

Since $B \neq 0$, we determine φ as a solution of the equation

$$\operatorname{ctg}(2\,\varphi) = \frac{A - C}{2\,B}. \tag{24.4}$$

The equation (24.4) is always solvable in the form of

$$\varphi = \frac{\pi}{4} - \frac{1}{2}\operatorname{arctg}\Big(\frac{A - C}{2\,B}\Big). \tag{24.5}$$

Comparing (24.4) and (24.3), we see that having determined the angle φ by means of the formula (24.5), we provide the vanishing of the parameter $\tilde{B} = 0$ in the rotated coordinate system. The theorem 24.1 is proved. □

Let's apply the theorem 24.1 and write the equation of the curve Γ in the form of the equation (23.1) with $B = 0$:

$$A\,x^2 + C\,y^2 + 2\,D\,x + 2\,E\,y + F = 0. \tag{24.6}$$

The equation (24.6) provides the subdivision of all curves of the second order on a plane into three types:

- the **elliptic type** where $A \neq 0$, $C \neq 0$ and the parameters A and C are of the same sign, i. e. $\operatorname{sign}(A) = \operatorname{sign}(C)$;
- the **hyperbolic type** where $A \neq 0$, $C \neq 0$ and the parameters A and C are of different signs, i. e. $\operatorname{sign}(A) \neq \operatorname{sign}(C)$;
- the **parabolic type** where $A = 0$ or $C = 0$.

The parameters A and C in (24.6) cannot vanish simultaneously since in this case the degree of the polynomial in (24.6) would be lower than two, which would contradict the definition 23.1.

Curves of the elliptic type. If the conditions $A \neq 0$, $C \neq 0$, and $\operatorname{sign}(A) = \operatorname{sign}(C)$ are fulfilled, without loss of generality we can assume that $A > 0$ and $C > 0$. In this case the equation (24.6) can be written as follows:

$$A\left(x + \frac{D}{A}\right)^2 + C\left(y + \frac{E}{C}\right)^2 + \left(F - \frac{D^2}{A} - \frac{E^2}{C}\right) = 0. \tag{24.7}$$

Let's perform the following change of variables in (24.7):

$$x = \tilde{x} - \frac{D}{A}, \qquad y = \tilde{y} - \frac{E}{C}. \tag{24.8}$$

The change of variables (24.8) corresponds to the displacement of the origin without rotation (the case of the unit matrices $S = 1$ and $T = 1$ in (22.4)). In addition to (24.8) we denote

$$\tilde{F} = F - \frac{D^2}{A} - \frac{E^2}{C}. \tag{24.9}$$

Taking into account (24.8) and (24.9), we write (24.7) as

$$A\,\tilde{x}^2 + C\,\tilde{y}^2 + \tilde{F} = 0, \tag{24.10}$$

where the coefficients A and C are positive: $A > 0$ and $C > 0$.

The equation (24.10) provides the subdivision of curves of the elliptic type into three subtypes:

– the case of an **ellipse** where $\tilde{F} < 0$;
– the case of an **imaginary ellipse** where $\tilde{F} > 0$;
– the case of a **point** where $\tilde{F} = 0$.

In the case of an ellipse the equation (24.10) is brought to the form (6.15) in the variables $\tilde{x}$ and $\tilde{y}$. As we know, this equation describes an ellipse.

In the case of an imaginary ellipse the equation (24.10) reduces to the equation which is similar to the equation of an ellipse:

$$\frac{\tilde{x}^2}{a^2} + \frac{\tilde{y}^2}{b^2} = -1. \tag{24.11}$$

The equation (24.11) has no solutions. Such an equation describes the empty set of points.

The case of a point is sometimes called the case of a *pair of imaginary intersecting lines*, which is somewhat not exact. In

this case the equation (24.10) describes a single point with the coordinates $\tilde{x} = 0$ and $\tilde{y} = 0$.

Curves of the hyperbolic type. If the conditions $A \neq 0$, $C \neq 0$, and $\operatorname{sign}(A) \neq \operatorname{sign}(C)$ are fulfilled, without loss of generality we can assume that $A > 0$ and $C < 0$. In this case the equation (24.6) can be written as

$$A\left(x + \frac{D}{A}\right)^2 - \tilde{C}\left(y - \frac{E}{\tilde{C}}\right)^2 + \left(F - \frac{D^2}{A} + \frac{E^2}{\tilde{C}}\right) = 0, \qquad (24.12)$$

where $-C = \tilde{C} > 0$. Let's denote

$$\tilde{F} = F - \frac{D^2}{A} + \frac{E^2}{\tilde{C}} \qquad (24.13)$$

and then perform the following change of variables, which corresponds to a displacement of the origin:

$$x = \tilde{x} - \frac{D}{A}, \qquad y = \tilde{y} + \frac{E}{\tilde{C}}. \qquad (24.14)$$

Due to (24.13) and (24.14) the equation (24.12) is written as

$$A\,\tilde{x}^2 - \tilde{C}\,\tilde{y}^2 + \tilde{F} = 0, \qquad (24.15)$$

where the coefficients A and $\tilde{C}$ are positive: $A > 0$ and $\tilde{C} > 0$.

The equation (24.15) provides the subdivision of curves of the hyperbolic type into two subtypes:

– the case of a **hyperbola** where $\tilde{F} \neq 0$;
– the case of a **pair of intersecting lines** where $\tilde{F} = 0$.

In the case of a hyperbola the equation (24.15) is brought to the form (10.20) in the variables $\tilde{x}$ and $\tilde{y}$. As we know, it describes a hyperbola.

In the case of a pair of intersecting lines the left hand side of the equation (24.15) is written as a product of two multiplicands

and the equation (24.15) is brought to

$$(\sqrt{A}\,\tilde{x} + \sqrt{\tilde{C}}\,\tilde{y})(\sqrt{A}\,\tilde{x} - \sqrt{\tilde{C}}\,\tilde{y}) = 0. \tag{24.16}$$

The equation (24.16) describes two line on a plane that intersect at the point with the coordinates $\tilde{x} = 0$ and$\tilde{y} = 0$.

Curves of the parabolic type. For curves of this type there are two options in the equation (24.6): $A = 0$, $C \neq 0$ or $C = 0$, $A \neq 0$. But the second option reduces to the first one upon changing variables $x = -\tilde{y}$, $y = \tilde{x}$, which corresponds to the rotation by the angle $\varphi = \pi/2$. Therefore without loss of generality we can assume that $A = 0$ and $C \neq 0$. Then the equation (24.6) is brought to

$$y^2 + 2\,D\,x + 2\,E\,y + F = 0. \tag{24.17}$$

In order to transform the equation (24.17) we apply the change of variables $y = \tilde{y} - E$, which corresponds to the displacement of the origin along the y-axis. Then we denote $\tilde{F} = F + E^2$. As a result the equation (24.17) is written as

$$\tilde{y}^2 + 2\,D\,x + \tilde{F} = 0. \tag{24.18}$$

The equation (24.18) provides the subdivision of curves of the parabolic type into four subtypes:

- the case of a **parabola** where $D \neq 0$;
- the case of a **pair of parallel lines** where $D = 0$ and $\tilde{F} < 0$;
- the case of a **pair of coinciding lines** where $D = 0$ and $\tilde{F} = 0$;
- the case of a **pair of imaginary parallel lines** where $D = 0$ and $\tilde{F} > 0$.

In the case of a parabola the equation (24.18) reduces to the equation (15.9) and describes a parabola.

In the case of a pair of parallel lines we introduce the notation $\tilde{F} = -y_0^2$ into the equation (24.18). As a result the equation (24.18) is written in the following form:

$$(\tilde{y} + y_0)(\tilde{y} - y_0) = 0. \tag{24.19}$$

The equation (24.19) describes a pair of lines parallel to the y-axis and being at the distance $2\,y_0$ from each other.

In the case of a pair of coinciding lines the equation (24.18) reduces to the form $\tilde{y}^2 = 0$, which describes a single line coinciding with the y-axis.

In the case of a pair of imaginary parallel lines the equation (24.18) has no solutions. It describes the empty set.

§ 25. Surfaces of the second order.

Definition 25.1. A *surface of the second order* or a *quadric* in the space $\mathbb{E}$ is a surface which in some Cartesian coordinate system is given by a polynomial equation of the second order:

$$\begin{aligned} A\,x^2 + 2\,B\,x\,y + C\,y^2 + 2\,D\,x\,z + 2\,E\,y\,z + \\ + F\,z^2 + 2\,G\,x + 2\,H\,y + 2\,I\,z + J = 0. \end{aligned} \tag{25.1}$$

Here $x = x^1$, $y = x^2$, $z = x^3$ are the coordinates of points of the space $\mathbb{E}$. Note that the transformation of the coordinates of points under a change of a coordinate system is given by functions of the first order in x^1, x^2, x^3 (see formulas (21.4)). For this reason the general form of a quadric equation (25.1) remains unchanged under a change of a coordinate system, though the values of the parameters A, B, C, D, E, F, G, H, I, and J can change. The following theorem is immediate from what was said.

Theorem 25.1. *For any surface of the second order in the space* $\mathbb{E}$, *i. e. for any quadric, there is some rectangular coordinate system with an orthonormal basis such that this surface is given*

by an equation of the form (25.1) *in this coordinate system.*

§ 26. Classification of surfaces of the second order.

The problem of classification of surfaces of the second order in $\mathbb{E}$ is solved below following the scheme explained in § 2 of Chapter VI in the book [1]. Let S be surface of the second order given by the equation (25.1) on some rectangular coordinate system with an orthonormal basis (see Theorem 25.1). Let's arrange the parameters A, B, C, D, E, F, G, H, I of the equation (25.1) into two matrices

$$\mathcal{F} = \begin{Vmatrix} A & B & D \\ B & C & E \\ D & E & F \end{Vmatrix}, \qquad \mathcal{D} = \begin{Vmatrix} G \\ H \\ I \end{Vmatrix}. \tag{26.1}$$

The matrices (26.1) are used in the following theorem.

THEOREM 26.1. *For any surface of the second order S there is a rectangular coordinate system with an orthonormal basis such that the matrix $\mathcal{F}$ in* (26.1) *is diagonal, while the matrix $\mathcal{D}$ is related to the matrix $\mathcal{F}$ by means of the formula $\mathcal{F} \cdot \mathcal{D} = 0$.*

The proof of the theorem 26.1 can be found in [1]. Applying the theorem 26.1, we can write the equation (25.1) as

$$A\,x^2 + C\,y^2 + F\,z^2 + 2\,G\,x + 2\,H\,y + 2\,I\,z + J = 0. \tag{26.2}$$

The equation (26.2) and the theorem 26.1 provide the subdivision of all surfaces of the second order in $\mathbb{E}$ into four types:

– the **elliptic type** where $A \neq 0$, $C \neq 0$, $F \neq 0$ and the quantities A, C, and F are of the same sign;
– the **hyperbolic type** where $A \neq 0$, $C \neq 0$, $F \neq 0$ and the quantities A, C and F are of different signs;
– the **parabolic type** where exactly one of the quantities A, C. and F is equal to zero and exactly one of the quantities

G, H, and I is nonzero.
– the **cylindrical type** in all other cases.

Surfaces of the elliptic type. From the conditions $A \neq 0$, $C \neq 0$, $F \neq 0$ in (26.2) and from the condition $\mathcal{F} \cdot \mathcal{D} = 0$ in the theorem 26.1 we derive $G = 0$, $H = 0$, and $I = 0$. Since the quantities A, C, and F are of the same sign, without loss of generality we can assume that all of them are positive. Hence for all surfaces of the elliptic type we can write (26.2) as

$$A\,x^2 + C\,y^2 + F\,z^2 + J = 0, \tag{26.3}$$

where $A > 0$, $C > 0$, and $F > 0$. The equation (26.3) provides the subdivision of surfaces of the elliptic type into three subtypes:

– the case of an **ellipsoid** where $J < 0$;
– the case of an **imaginary ellipsoid** where $J > 0$;
– the case of a **point** where $J = 0$.

The case of an ellipsoid is the most non-trivial of the three. In this case the equation (26.3) is brought to

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1. \tag{26.4}$$

The equation (26.4) describes the surface which is called an

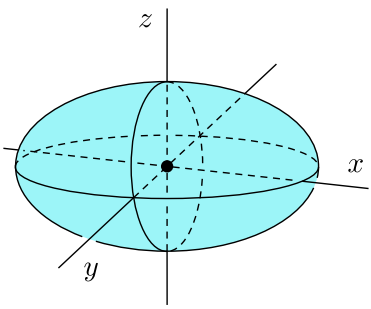

Fig. 26.1

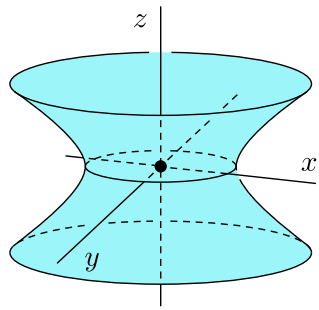

Fig. 26.2

ellipsoid. This surface is shown in Fig. 26.1.

In the case of an imaginary ellipsoid the equation (26.3) has no solutions. It describes the empty set.

In the case of a point the equation (26.3) can be written in the form very similar to the equation of an ellipsoid (26.4):

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 0. \tag{26.5}$$

The equation (26.5) describes a single point in the space $\mathbb{E}$ with the coordinates $x = 0$, $y = 0$, $z = 0$.

Surfaces of the hyperbolic type. From the three conditions $A \neq 0$, $C \neq 0$, $F \neq 0$ in (26.2) and from the condition $\mathcal{F} \cdot \mathcal{D} = 0$ in the theorem 26.1 we derive $G = 0$, $H = 0$, and $I = 0$. The quantities A, C, and F are of different signs. Without loss of generality we can assume that two of them are positive and one of them is negative. By exchanging axes, which preserves the orthogonality of coordinate systems and orthonormality of their bases, we can transform the equation (26.2) so that we would have $A > 0$, $C > 0$, and $F < 0$. As a result we conclude that for all surfaces of the hyperbolic type the initial equation (26.2) can be brought to the form

$$A\,x^2 + C\,y^2 + F\,z^2 + J = 0, \tag{26.6}$$

where $A > 0$, $C > 0$, $F < 0$. The equation (26.6) provides the subdivision of surfaces of the hyperbolic type into three subtypes:

– the case of a **hyperboloid of one sheet** where $J < 0$;
– the case of a **hyperboloid of two sheets** where $J > 0$;
– the case of a **cone** where $J = 0$.

In the case of a hyperboloid of one sheet the equation (26.6) can be written in the following form:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1. \tag{26.7}$$

The equation (26.7) describes a surface which is called the *hyperboloid of one sheet*. It is shown in Fig. 26.2.

In the case of a hyperboloid of two sheets the equation (26.6) can be written in the following form:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=-1. \tag{26.8}$$

The equation (26.8) describes a surface which is called the *hyperboloid of two sheets*. This surface is shown in Fig. 26.3.

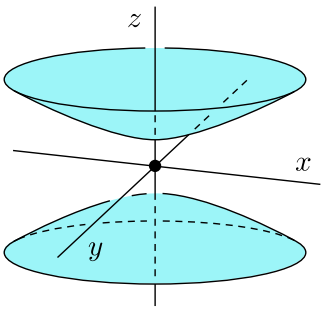

Fig. 26.3

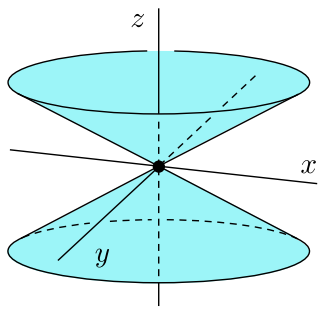

Fig. 26.4

In the case of a cone the equation (26.6) is transformed to the equation which is very similar to the equations (26.7) and (26.8), but with zero in the right hand side:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0. \tag{26.9}$$

The equation (26.9) describes a surface which is called the *cone*. This surface is shown in Fig. 26.4.

Surfaces of the parabolic type. For this type of surfaces exactly one of the three quantities A, C, and F is equal to zero. By exchanging axes, which preserves the orthogonality of coordinate systems and orthonormality of their bases, we can

transform the equation (26.2) so that we would have $A \neq 0$, $C \neq 0$, and $F = 0$. From $A \neq 0$, $C \neq 0$, and from the condition $\mathcal{F} \cdot \mathcal{D} = 0$ in the theorem 26.1 we derive $G = 0$ and $H = 0$. The value of I is not determined by the condition $\mathcal{F} \cdot \mathcal{D} = 0$. However, according to the definition of surfaces of the parabolic type exactly one of the three quantities G, H, I should be nonzero. Due to $G = 0$ and $H = 0$ we conclude that $I \neq 0$. As a result the equation (26.2) is written as

$$A\,x^2 + C\,y^2 + 2\,I\,z + J = 0, \tag{26.10}$$

where $A \neq 0$, $C \neq 0$, and $I \neq 0$. The condition $I \neq 0$ means that we can perform the displacement of the origin along the z-axis equivalent to the change of variables

$$z \to z - \frac{J}{2\,I}. \tag{26.11}$$

Upon applying the change of variables (26.11) to the equation (26.10) this equation is written as

$$A\,x^2 + C\,y^2 + 2\,I\,z = 0, \tag{26.12}$$

where $A \neq 0$, $C \neq 0$, $I \neq 0$. The equation (26.12) provides the subdivision of surfaces of the parabolic type into two subtypes:

- the case of an **elliptic paraboloid** where the quantities $A \neq 0$ and $C \neq 0$ are of the same sign;
- the case of a **hyperbolic paraboloid**, where the quantities $A \neq 0$ and $C \neq 0$ are of different signs.

In the case of an elliptic paraboloid the equation (26.12) can be written in the following form:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 2\,z. \tag{26.13}$$

The equation (26.13) describes a surface which is called the

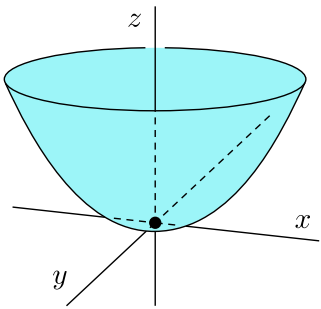

Fig. 26.5

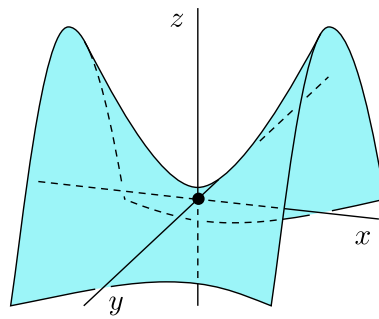

Fig. 26.6

elliptic paraboloid. This surface is shown in Fig. 26.5.

In the case of a hyperbolic paraboloid the equation (26.12) can be transformed to the following form:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 2\,z. \tag{26.14}$$

The equation (26.14) describes a saddle surface which is called the *hyperbolic paraboloid*. This surface is shown in Fig. 26.6.

Surfaces of the cylindrical type. According to the results from § 2 in Chapter VI of the book [1], in the cylindrical case the dimension reduction occurs. This means that there is a rectangular coordinate system with an orthonormal basis where the variable z drops from the equation (26.2):

$$A\,x^2 + C\,y^2 + 2\,G\,x + 2\,H\,y + J = 0. \tag{26.15}$$

The classification of surfaces of the second order described by the equation (26.15) is equivalent to the classification of curves of the second order on a plane described by the equation (24.6). The complete type list of such surfaces contains nine cases:

– the case of an **elliptic cylinder**;

– the case of an **imaginary elliptic cylinder**;
– the case of a **straight line**;

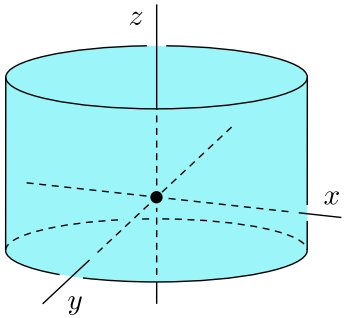

Fig. 26.7

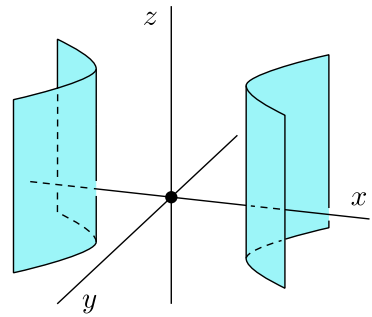

Fig. 26.8

– the case of a **hyperbolic cylinder**;
– the case of a **pair of intersecting planes**;
– the case of a **parabolic cylinder**;
– the case of a **pair of parallel planes**;
– the case of a **pair of coinciding planes**;
– the case of a **pair of imaginary parallel planes**.

In the case of an elliptic cylinder the equation (26.15) can be transformed to the following form:

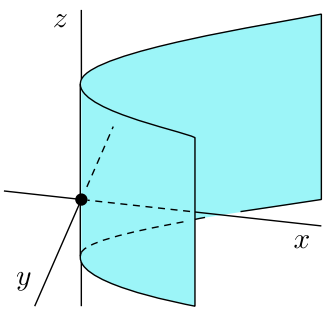

Fig. 26.9

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \qquad (26.16)$$

The equation (26.16) coincides with the equation of an ellipse on a plane (6.15). In the space $\mathbb{E}$ it describes a surface which is called the *elliptic cylinder*. This surface is shown in Fig. 26.7.

In the case of an imaginary elliptic cylinder the equation

(26.15) can be transformed to the following form:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = -1. \tag{26.17}$$

The equation (26.17) has no solutions. Such an equation describes the empty set.

In the case of a straight line the equation (26.15) can be brought to the form similar to (26.16) and (26.17):

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0. \tag{26.18}$$

The equation (26.18) describes a straight line in the space coinciding with the z-axis. In the canonical form this line is given by the equations $x = 0$ and $y = 0$ (see (5.14)).

In the case of a hyperbolic cylinder the equation (26.15) can be transformed to the following form:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{26.19}$$

The equation (26.19) coincides with the equation of a hyperbola on a plane (6.15). In the spatial case it describes a surface which is called the *hyperbolic cylinder*. It is shown in Fig. 26.8.

The next case in the list is the case of a pair of intersecting planes. In this case the equation (26.15) can be brought to

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0. \tag{26.20}$$

The equation (26.20) describes the union of two intersecting planes in the space given by the equations

$$\frac{x}{a} - \frac{y}{b} = 0, \qquad \frac{x}{a} + \frac{y}{b} = 0. \tag{26.21}$$

The planes (26.21) intersect along a line which coincides with the z-axis. This line is given by the equations $x = 0$ and $y = 0$.

In the case of a parabolic cylinder the equation (26.15) reduces to the equation coinciding with the equation of a parabola

$$y^2 = 2\,p\,x. \tag{26.22}$$

In the space the equation (26.22) describes a surface which is called the *parabolic cylinder*. This surface is shown in Fig. 26.9.

In the case of a pair of parallel planes the equation (26.15) is brought to $y^2 - y_0^2 = 0$, where $y_0 \neq 0$. It describes two parallel planes given by the equations

$$y = y_0, \qquad\qquad y = -y_0. \tag{26.23}$$

In the case of a pair of coinciding planes the equation (26.15) is also brought to $y^2 - y_0^2 = 0$, but the parameter y_0 in it is equal to zero. Due to $y_0 = 0$ two planes (26.23) are glued into a single plane which is perpendicular to the y-axis and is given by the equation $y = 0$.

In the case of a pair of imaginary parallel planes the equation (26.15) is brought to $y^2 + y_0^2 = 0$, where $y_0 \neq 0$. Such an equation has no solutions. For this reason it describes the empty set.

REFERENCES.

1. Sharipov R. A., *Course of linear algebra and multidimensional geometry*, Bashkir State University, Ufa, 1996; see also math.HO/0405323 in Electronic archive http://arXiv.org.
2. Sharipov R. A., *Course of differential geometry*, Bashkir State University, Ufa, 1996; see also e-print math.HO/0412421 at http://arXiv.org.
3. Sharipov R. A., *Theory of representations of finite groups*, Bash-NII-Stroy, Ufa, 1995; see also e-print math.HO/0612104 at http://arXiv.org.
4. Sharipov R. A., *Classical electrodynamics and theory of relativity*, Bashkir State University, Ufa, 1996; see also e-print physics/0311011 in Electronic archive http://arXiv.org.
5. Sharipov R. A., *Quick introduction to tensor analysis*, free on-line textbook, 2004; see also e-print math.HO/0403252 at http://arXiv.org.
6. Sharipov R. A., *Foundations of geometry for university students and high-school students*, Bashkir State University, 1998; see also e-print math.HO/0702029 in Electronic archive http://arXiv.org.
7. Kurosh A. G., *Course of higher algebra*, Nauka publishers, Moscow, 1968.
8. *Kronecker symbol*, Wikipedia, the Free Encyclopedia, Wikimedia Foundation Inc., San Francisco, USA.
9. Kudryavtsev L. D., *Course of mathematical analysis,* Vol. I, II, Visshaya Shkola publishers, Moscow, 1985.

CONTACTS

Address:

Ruslan A. Sharipov,
Dep. of Mathematics
and Information Techn.,
Bashkir State University,
32 Zaki Validi street,
Ufa 450074, Russia

Phone:

+7-(347)-273-67-18 (Office)
+7-(917)-476-93-48 (Cell)

URL's:

http://ruslan-sharipov.ucoz.com
http://freetextbooks.narod.ru
http://sovlit2.narod.ru

Home address:

Ruslan A. Sharipov,
5 Rabochaya street,
Ufa 450003, Russia

E-mails:

r-sharipov@mail.ru
R_Sharipov@ic.bashedu.ru

APPENDIX

List of publications by the author for the period 1986–2011.

Part 1. Soliton theory.

1. Sharipov R. A., *Finite-gap analogs of N-multiplet solutions of the KdV equation*, Uspehi Mat. Nauk **41** (1986), no. 5, 203–204.
2. Sharipov R. A., *Soliton multiplets of the Korteweg-de Vries equation*, Dokladi AN SSSR **292** (1987), no. 6, 1356–1359.
3. Sharipov R. A., *Multiplet solutions of the Kadomtsev-Petviashvili equation on a finite-gap background*, Uspehi Mat. Nauk **42** (1987), no. 5, 221–222.
4. Bikbaev R. F., Sharipov R. A., *Magnetization waves in Landau-Lifshits model*, Physics Letters A **134** (1988), no. 2, 105-108; see solv-int/9905008.
5. Bikbaev R. F. & Sharipov R. A., *Assymptotics as $t \to \infty$ for a solution of the Cauchy problem for the Korteweg-de Vries equation in the class of potentials with finite-gap behaviour as $x \to \pm\infty$*, Theor. and Math. Phys. **78** (1989), no. 3, 345–356.
6. Sharipov R. A., *On integration of the Bogoyavlensky chains*, Mat. zametki **47** (1990), no. 1, 157–160.
7. Cherdantsev I. Yu. & Sharipov R. A., *Finite-gap solutions of the Bullough-Dodd-Jiber-Shabat equation*, Theor. and Math. Phys. **82** (1990), no. 1, 155–160.
8. Cherdantsev I. Yu. & Sharipov R. A., *Solitons on a finite-gap background in Bullough-Dodd-Jiber-Shabat model*, International. Journ. of Modern Physics A **5** (1990), no. 5, 3021–3027; see math-ph/0112045.
9. Sharipov R. A. & Yamilov R. I., *Backlund transformations and the construction of the integrable boundary value problem for the equation*

$u_{xt} = e^u - e^{-2u}$, «Some problems of mathematical physics and asymptotics of its solutions», Institute of mathematics BNC UrO AN SSSR, Ufa, 1991, pp. 66–77; see solv-int/9412001.

10. Sharipov R. A., *Minimal tori in five-dimensional sphere in* $\mathbb{C}^3$, Theor. and Math. Phys. **87** (1991), no. 1, 48–56; see math.DG/0204253.
11. Safin S. S. & Sharipov R. A., *Backlund autotransformation for the equation* $u_{xt} = e^u - e^{-2u}$, Theor. and Math. Phys. **95** (1993), no. 1, 146–159.
12. Boldin A. Yu. & Safin S. S. & Sharipov R. A., *On an old paper of Tzitzeika and the inverse scattering method*, Journal of Mathematical Physics **34** (1993), no. 12, 5801–5809.
13. Pavlov M. V. & Svinolupov S. I. & Sharipov R. A., *Invariant criterion of integrability for a system of equations of hydrodynamical type*, «Integrability in dynamical systems», Inst. of Math. UrO RAN, Ufa, 1994, pp. 27–48; Funk. Anal. i Pril. **30** (1996), no. 1, 18–29; see solv-int/9407003.
14. Ferapontov E. V. & Sharipov R. A., *On conservation laws of the first order for a system of equations of hydrodynamical type*, Theor. and Math. Phys. **108** (1996), no. 1, 109–128.

Part 2. Geometry of the normal shift.

1. Boldin A. Yu. & Sharipov R. A., *Dynamical systems accepting the normal shift*, Theor. and Math. Phys. **97** (1993), no. 3, 386–395; see chao-dyn/9403003.
2. Boldin A. Yu. & Sharipov R. A., *Dynamical systems accepting the normal shift*, Dokladi RAN **334** (1994), no. 2, 165–167.
3. Boldin A. Yu. & Sharipov R. A., *Multidimensional dynamical systems accepting the normal shift*, Theor. and Math. Phys. **100** (1994), no. 2, 264–269; see patt-sol/9404001.
4. Sharipov R. A., *Problem of metrizability for the dynamical systems accepting the normal shift*, Theor. and Math. Phys. **101** (1994), no. 1, 85–93; see solv-int/9404003.
5. Sharipov R. A., *Dynamical systems accepting the normal shift*, Uspehi Mat. Nauk **49** (1994), no. 4, 105; see solv-int/9404002.
6. Boldin A. Yu. & Dmitrieva V. V. & Safin S. S. & Sharipov R. A., *Dynamical systems accepting the normal shift on an arbitrary Riemannian manifold*, «Dynamical systems accepting the normal shift», Bashkir State University, Ufa, 1994, pp. 4–19; see also Theor. and Math. Phys. **103** (1995), no. 2, 256–266 and hep-th/9405021.
7. Boldin A. Yu. & Bronnikov A. A. & Dmitrieva V. V. & Sharipov R. A., *Complete normality conditions for the dynamical systems on Riemannian manifolds*, «Dynamical systems accepting the normal shift», Bashkir

State University, 1994, pp. 20–30; see also Theor. and Math. Phys. **103** (1995), no. 2, 267–275 and astro-ph/9405049.

8. Sharipov R. A., *Higher dynamical systems accepting the normal shift*, «Dynamical systems accepting the normal shift», Bashkir State University, 1994, pp. 41–65.
9. Bronnikov A. A. & Sharipov R. A., *Axially symmetric dynamical systems accepting the normal shift in* $\mathbb{R}^n$, «Integrability in dynamical systems», Inst. of Math. UrO RAN, Ufa, 1994, pp. 62–69.
10. Sharipov R. A., *Metrizability by means of a conformally equivalent metric for the dynamical systems*, «Integrability in dynamical systems», Inst. of Math. UrO RAN, Ufa, 1994, pp. 80–90; see also Theor. and Math. Phys. **103** (1995), no. 2, 276–282.
11. Boldin A. Yu. & Sharipov R. A., *On the solution of the normality equations for the dimension* $n \geqslant 3$, Algebra i Analiz **10** (1998), no. 4, 31–61; see also solve-int/9610006.
12. Sharipov R. A., *Dynamical systems admitting the normal shift,* Thesis for the degree of Doctor of Sciences in Russia, math.DG/0002202, Electronic archive http://arXiv.org, 2000, pp. 1–219.
13. Sharipov R. A., *Newtonian normal shift in multidimensional Riemannian geometry*, Mat. Sbornik **192** (2001), no. 6, 105–144; see also math.DG/0006125.
14. Sharipov R. A., *Newtonian dynamical systems admitting the normal blow-up of points*, Zap. semin. POMI **280** (2001), 278–298; see also math.DG/0008081.
15. Sharipov R. A., *On the solutions of the weak normality equations in multidimensional case*, math.DG/0012110 in Electronic archive http://arxiv.org (2000), 1–16.
16. Sharipov R. A., *First problem of globalization in the theory of dynamical systems admitting the normal shift of hypersurfaces*, International Journal of Mathematics and Mathematical Sciences **30** (2002), no. 9, 541–557; see also math.DG/0101150.
17. Sharipov R. A., *Second problem of globalization in the theory of dynamical systems admitting the normal shift of hypersurfaces*, math.DG/0102141 in Electronic archive http://arXiv.org (2001), 1–21.
18. Sharipov R. A., *A note on Newtonian, Lagrangian, and Hamiltonian dynamical systems in Riemannian manifolds*, math.DG/0107212 in Electronic archive http://arXiv.org (2001), 1–21.
19. Sharipov R. A., *Dynamical systems admitting the normal shift and wave equations*, Theor. and Math. Phys. **131** (2002), no. 2, 244–260; see also math.DG/0108158.
20. Sharipov R. A., *Normal shift in general Lagrangian dynamics*, math.DG

/0112089 in Electronic archive http://arXiv.org (2001), 1–27.
21. Sharipov R. A, *Comparative analysis for a pair of dynamical systems one of which is Lagrangian*, math.DG/0204161 in Electronic archive http://arxiv.org (2002), 1–40.
22. Sharipov R. A., *On the concept of a normal shift in non-metric geometry*, math.DG/0208029 in Electronic archive http://arXiv.org (2002), 1–47.
23. Sharipov R. A., *V-representation for the normality equations in geometry of generalized Legendre transformation*, math.DG/0210216 in Electronic archive http://arXiv.org (2002), 1–32.
24. Sharipov R. A., *On a subset of the normality equations describing a generalized Legendre transformation*, math.DG/0212059 in Electronic archive (2002), 1–19.

Part 3. Mathematical analysis and theory of functions.

1. Sharipov R. A. & Sukhov A. B. On *CR*-mappings between algebraic Cauchy-Riemann manifolds and the separate algebraicity for holomorphic functions, Trans. of American Math. Society **348** (1996), no. 2, 767–780; see also Dokladi RAN **350** (1996), no. 4, 453–454.
2. Sharipov R. A. & Tsyganov E. N. On the separate algebraicity along families of algebraic curves, *Preprint of Baskir State University*, Ufa, 1996, pp. 1-7; see also Mat. Zametki **68** (2000), no. 2, 294–302.
3. Sharipov R. A., *Algorithms for laying points optimally on a plane and a circle*, e-print 0705.0350 in the archive http://arXiv.org (2010), 1–6.
4. Sharipov R. A., *A note on Khabibullin's conjecture for integral inequalities*, e-print 1008.0376 in Electronic archive http://arXiv.org (2010), 1–17.
5. Sharipov R. A., *Direct and inverse conversion formulas associated with Khabibullin's conjecture for integral inequalities*, e-print 1008.1572 in Electronic archive http://arXiv.org (2010), 1–7.
6. Sharipov R. A., *A counterexample to Khabibullin's conjecture for integral inequalities*, Ufa Math. Journ. **2** (2010), no. 4, 99–107; see also e-print 1008.2738 in Electronic archive http://arXiv.org (2010), 1–10.

Part 4. Symmetries and invariants.

1. Dmitrieva V. V. & Sharipov R. A., *On the point transformations for the second order differential equations*, solv-int/9703003 in Electronic archive http://arXiv.org (1997), 1–14.
2. Sharipov R. A., *On the point transformations for the equation* $y'' = P + 3\,Q\,y' + 3\,R\,y'^2 + S\,y'^3$, solv-int/9706003 in Electronic archive http://arxiv.org (1997), 1–35; see also Вестник Башкирского университета **1(I)** (1998), 5–8.

3. Mikhailov O. N. & Sharipov R. A., *On the point expansion for a certain class of differential equations of the second order*, Diff. Uravneniya **36** (2000), no. 10, 1331–1335; see also solv-int/9712001.
4. Sharipov R. A., *Effective procedure of point-classification for the equation* $y'' = P + 3\,Q\,y' + 3\,R\,y'^2 + S\,y'^3$, math.DG/9802027 in Electronic archive http://arXiv.org (1998), 1–35.
5. Dmitrieva V. V. & Gladkov A. V. & Sharipov R. A., *On some equations that can be brought to the equations of diffusion type*, Theor. and Math. Phys. **123** (2000), no. 1, 26–37; see also math.AP/9904080.
6. Dmitrieva V. V. & Neufeld E. G. & Sharipov R. A. & Tsaregorodtsev A. A., *On a point symmetry analysis for generalized diffusion type equations*, math.AP/9907130 in Electronic archive http://arXiv.org (1999), 1–52.

Part 5. General algebra.

1. Sharipov R. A, *Orthogonal matrices with rational components in composing tests for High School students*, math.GM/0006230 in Electronic archive http://arXiv.org (2000), 1–10.
2. Sharipov R. A., *On the rational extension of Heisenberg algebra*, math.RA/0009194 in Electronic archive http://arXiv.org (2000), 1–12.
3. Sharipov R. A, *An algorithm for generating orthogonal matrices with rational elements*, cs.MS/0201007 in Electronic archive http://arXiv.org (2002), 1–7.
4. Sharipov R. A, *A note on pairs of metrics in a two-dimensional linear vector space*, e-print 0710.3949 in Electronic archive http://arXiv.org (2007), 1–9.
5. Sharipov R. A, *A note on pairs of metrics in a three-dimensional linear vector space*, e-print 0711.0555 in Electronic archive http://arXiv.org (2007), 1–17.
6. Sharipov R. A., *Transfinite normal and composition series of groups*, e-print 0908.2257 in Electronic archive http://arXiv.org (2010), 1–12; Вестник Башкирского университета **15** (2010), no. 4, 1098.
7. Sharipov R. A., *Transfinite normal and composition series of modules*, e-print 0909.2068 in Electronic archive http://arXiv.org (2010), 1–12.

Part 6. Condensed matter physics.

1. Lyuksyutov S. F. & Sharipov R. A., *Note on kinematics, dynamics, and thermodynamics of plastic glassy media*, cond-mat/0304190 in Electronic archive http://arXiv.org (2003), 1–19.

2. Lyuksyutov S. F. & Paramonov P. B. & Sharipov R. A. & Sigalov G., *Exact analytical solution for electrostatic field produced by biased atomic force microscope tip dwelling above dielectric-conductor bilayer*, cond-mat/0408247 in Electronic archive http://arXiv.org (2004), 1–6.
3. Lyuksyutov S. F. & Sharipov R. A., *Separation of plastic deformations in polymers based on elements of general nonlinear theory*, cond-mat /0408433 in Electronic archive http://arXiv.org (2004), 1–4.
4. Comer J. & Sharipov R. A., *A note on the kinematics of dislocations in crystals*, math-ph/0410006 in Electronic archive http://arXiv.org (2004), 1–15.
5. Sharipov R. A., *Gauge or not gauge?* cond-mat/0410552 in Electronic archive http://arXiv.org (2004), 1–12.
6. Sharipov R. A., *Burgers space versus real space in the nonlinear theory of dislocations*, cond-mat/0411148 in Electronic archive http://arXiv.org (2004), 1–10.
7. Comer J. & Sharipov R. A., *On the geometry of a dislocated medium*, math-ph/0502007 in Electronic archive http://arXiv.org (2005), 1–17.
8. Sharipov R. A., *A note on the dynamics and thermodynamics of dislocated crystals*, cond-mat/0504180 in Electronic archive http://arXiv.org (2005), 1–18.
9. Lyuksyutov S. F., Paramonov P. B., Sharipov R. A., Sigalov G., *Induced nanoscale deformations in polymers using atomic force microscopy*, Phys. Rev. B **70** (2004), no. 174110.

Part 7. Tensor analysis.

1. Sharipov R. A., *Tensor functions of tensors and the concept of extended tensor fields*, e-print math/0503332 in the archive http://arXiv.org (2005), 1–43.
2. Sharipov R. A., *Spinor functions of spinors and the concept of extended spinor fields*, e-print math.DG/0511350 in the archive http://arXiv.org (2005), 1–56.
3. Sharipov R. A., *Commutation relationships and curvature spin-tensors for extended spinor connections*, e-print math.DG/0512396 in Electronic archive http://arXiv.org (2005), 1-22.

Part 8. Particles and fields.

1. Sharipov R. A., *A note on Dirac spinors in a non-flat space-time of general relativity*, e-print math.DG/0601262 in Electronic archive http: //arxiv.org (2006), 1–22.

2. Sharipov R. A., *A note on metric connections for chiral and Dirac spinors*, e-print math.DG/0602359 in Electronic archive http://arXiv.org (2006), 1–40.
3. Sharipov R. A., *On the Dirac equation in a gravitation field and the secondary quantization*, e-print math.DG/0603367 in Electronic archive http://arXiv.org (2006), 1–10.
4. Sharipov R. A., *The electro-weak and color bundles for the Standard Model in a gravitation field*, e-print math.DG/0603611 in Electronic archive http://arXiv.org (2006), 1–8.
5. Sharipov R. A., *A note on connections of the Standard Model in a gravitation field*, e-print math.DG/0604145 in Electronic archive http://arxiv.org (2006), 1–11.
6. Sharipov R. A., *A note on the Standard Model in a gravitation field*, e-print math.DG/0605709 in the archive http://arXiv.org (2006), 1–36.
7. Sharipov R. A., *The Higgs field can be expressed through the lepton and quark fields*, e-print hep-ph/0703001 in the archive http://arXiv.org (2007), 1–4.
8. Sharipov R. A., *Comparison of two formulas for metric connections in the bundle of Dirac spinors*, e-print 0707.0482 in Electronic archive http://arXiv.org (2007), 1–16.
9. Sharipov R. A., *On the spinor structure of the homogeneous and isotropic universe in closed model*, e-print 0708.1171 in the archive http://arXiv.org (2007), 1–25.
10. Sharipov R. A., *On Killing vector fields of a homogeneous and isotropic universe in closed model*, e-print 0708.2508 in the archive http://arXiv.org (2007), 1–19.
11. Sharipov R. A., *On deformations of metrics and their associated spinor structures*, e-print 0709.1460 in the archive http://arXiv.org (2007), 1–22.
12. Sharipov R. A., *A cubic identity for the Infeld-van der Waerden field and its application*, e-print 0801.0008 in Electronic archive http://arXiv.org (2008), 1–18.
13. Sharipov R. A., *A note on Kosmann-Lie derivatives of Weyl spinors*, e-print 0801.0622 in Electronic archive http://arXiv.org (2008), 1–22.
14. Sharipov R. A., *On operator fields in the bundle of Dirac spinors*, e-print 0802.1491 in Electronic archive http://arXiv.org (2008), 1–14.

Part 9. Number theory.

1. Sharipov R. A., *A note on a perfect Euler cuboid*, e-print 1104.1716 in Electronic archive http://arXiv.org (2011), 1–8.
2. Sharipov R. A., *A note on the* Sopfr(n) *function*, e-print 1104.5235 in Electronic archive http://arXiv.org (2011), 1–7.

3. Sharipov R. A., *Perfect cuboids and irreducible polynomials*, e-print 1108.5348 in Electronic archive http://arXiv.org (2011), 1–8.
4. Sharipov R. A., *A note on the first cuboid conjecture*, e-print 1109.2534 in Electronic archive http://arXiv.org (2011), 1–6.

Part 10. Technologies and innovations.

1. Sharipov R. A., *A polymer dental implant of the stocking type and its application*, Russian patent RU 2401082 C2 (2009).

Учебное издание

ШАРИПОВ Руслан Абдулович

КУРС АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ

Учебное пособие

Редактор Г. Г. Синайская
Корректор А. И. Николаева

Лицензия на издательскую деятелность
ЛР № 021319 от 05.01.1999 г.

Подписано в печать 02.11.2010 г.
Формат 60×84/16. Усл. печ. л. 13,11. Уч.-изд. л. 11,62.
Тираж 100. Изд. № 243. Заказ 69 а.

Редакционно-издательский центр
Башкирского государственного университета
450074, РБ, г. Уфа, ул. Заки Валиди, 32.

Отпечатано на множительном участке
Башкирского государственного университета
450074, РБ, г. Уфа, ул. Заки Валиди, 32.