A. S. Solodovnikov
and G. A. Toropova

Linear Algebra

with Elements of Analytic Geometry

Mir Publishers Moscow

Linear Algebra with Elements of Analytic Geometry

А. С. СОЛОДОВНИКОВ, Г. А. ТОРОПОВА

ЛИНЕЙНАЯ АЛГЕБРА С ЭЛЕМЕНТАМИ АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ

ИЗДАТЕЛЬСТВО «ВЫСШАЯ ШКОЛА» МОСКВА

A. S. Solodovnikov
and G. A. Toropova

Linear Algebra
with Elements of Analytic Geometry

Mir Publishers
Moscow

Translated from the Russian by
Tamara Baranovskaya

First published 1990
Revised from the 1987 Russian edition

На английском языке

Printed in the Union of Soviet Socialist Republics

ISBN 5-03-001789-5

CONTENTS

PREFACE

Recent decades have seen extensive application of mathematics in all spheres of man's activity. The use of mathematics as a tool for simulating the world around us has increased, and as a result a mathematical education has become ever more necessary.

"Linear algebra with elements of analytic geometry" is a basic course for mathematics students in technical schools. The topic is important because most applied problems are either "linear" in nature or admit of "linearization".

The concept of a *vector* underlies linear algebra. In a narrow sense, a vector is an arrow or a directed line segment in the plane or in space. However, mathematicians have extended the concept and this in more general interpretation covers a wider circle of objects.

The vector is primarily used in analytic geometry. The basic idea in this branch of mathematics is the coordinates, that is, numbers defining the position of a geometric object. The coordinate method was developed to solve geometrical problems and then extended to all fields of mathematics and so it became a universal tool of its application.

The union of two main ideas, those of coordinates and a vector, underlies the course on analytic algebra with the elements of analytic geometry. The mathematical methods a student masters when studying the course will be of practical use in all applications of mathematics.

This study aid is in two parts. The first part presents the elements of analytic geometry and the theory of determinants, while the second part is devoted to linear algebra, namely, systems of linear equations, matrix theory, and vector, Euclidean, and affine spaces. The concepts of linear algebra are mainly studied using arithmetic vector spaces $\mathbf{R}^n$ ($n = 1, 2, 3$), although Chapter 10 also presents abstract vector spaces. All problems of linear algebra are in any case reduced to analyzing systems of linear equations. We use *Gaussian elimination* for this purpose. We believe that this approach best serves the aim of the course, which is to teach the student methods for solving linear problems avoiding wherever possible abstract notions.

The book is oriented to high-school graduates. Nevertheless, a small part of the material demands more of the student, namely Chapter 5 (nth-order determinants), Section 10.5 (abstract vector spaces), and Chapter 13 (affine spaces). Each chapter (except for Chapter 8) ends with exercises.

When selecting the material for this study aid, we were guided by the course curriculum and a desire for simplicity of presentation.

We would like to express our appreciation to the reviewers rofessor O. V. Manturov and Z. M. Egozar'yan, who compiled the course curriculum, for their constructive and valuable suggestions.

The authors

Part One

ANALYTIC GEOMETRY

Chapter 1

VECTORS IN THE PLANE AND IN SPACE. CARTESIAN COORDINATE SYSTEM

1.1. VECTORS

1. Notion of a vector.

Many physical quantities encountered in mathematics and its applications, such as the length of a line segment, the area of a figure, the volume and mass of a body, are completely described by real numbers, their magnitudes. These quantities are called *scalars.* Other physical quantities, such as force and velocity are not completely determined by a number, they also possess direction. Such quantities are called *vectors.*

Definition. A *directed line segment* or *vector* is a line segment whose endpoints are specified. A *vector is directed* from its initial to its terminal point.

Vectors are represented geometrically as arrows (Fig. 1). The tail of the arrow is called the *initial point* of the vector, and the tip of the arrow the *terminal point.* A vector whose initial point is A and whose terminal point is B is denoted $\overrightarrow{AB}$, and its length or magnitude is denoted $|\overrightarrow{AB}|$. Another notation of a vector is lowercase boldface type such as **a** and its magnitude is denoted $|\mathbf{a}|$ or a. The initial point of the vector $\overrightarrow{AB}$ (point A) is sometimes called the *point of its application.*

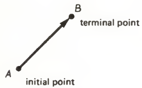


Figure 1

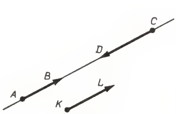


Figure 2

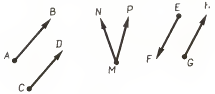


Figure 3

Definition. Vectors lying on the same line or on parallel lines are called *collinear vectors.*

Any two of the three vectors in Fig. 2 are collinear to each other.

Definition. Vectors which are collinear, have the same length, and are in the same direction are called *equivalent.*

If vectors are equivalent, we write $\vec{AB} = \vec{CD}$.

The definition of vector equality implies that the point of application of the vector can be arbitrarily chosen. In this sense, vectors are often called *free.* The vectors $\vec{AB}$ and $\vec{CD}$ shown in Fig. 3 are equivalent, while the vectors $\vec{MN}$ and $\vec{MP}$ and $\vec{EF}$ and $\vec{GH}$ are not.

If the initial point of a vector coincides with its terminal point, the vector is represented by a point and have no specific direction. Such a vector is called the *zero vector* and is denoted by $\mathbf{0}$. Thus, the equality $\mathbf{a} = \mathbf{0}$ is read as "the vector **a** equals zero", which means that the initial and terminal points of **a** coincide. The magnitude of the zero vector is zero.

From the definition of collinear vectors we infer that the zero vector is collinear with any vector.

2. Addition of vectors.

A course of high-school mathematics covers the operation of vector addition. Let us recall the geometric interpretation of vector addition, that is, the *triangle law of addition.*

If **a** and **b** are any two vectors, then the sum $\mathbf{a} + \mathbf{b}$ is the vector constructed as follows (Fig. 4). Place the initial point of a vector $\mathbf{a}'$, which is equivalent to **a**, at any point O and position a vector $\mathbf{b}'$, which is equivalent to **b**, so that its initial

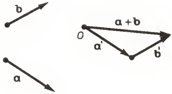


Figure 4

point coincides with the terminal point of **a**′. The vector **a** + **b** is the arrow from the initial point of **a**′ (i.e. the point O) to the terminal point of **b**′.

We emphasize that the vector **a** + **b** so constructed does not depend on the position of the point O. If we take any other point, say O^*, and perform the same construction, we obtain the same vector **a** + **b** (Fig. 5).

If vectors **a** and **b** are not collinear, we can apply the *parallelogram law of addition.* To obtain the vector **a** + **b**, we place the initial points of **a** and **b** at the same point O and construct a parallelogram (Fig. 6). The diagonal from the point O is the vector **a** + **b**.

The parallelogram law is often applied in physics problems, for instance, when composing forces, i.e. to find the resultant force.

Here are some properties of vector addition familiar from high-school mathematics.

1°. *Commutativity,*

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a} \tag{1}$$

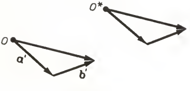


Figure 5

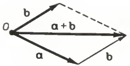


Figure 6

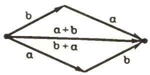


Figure 7

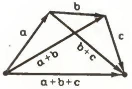


Figure 8

The property is illustrated in Fig. 7 when **a** and **b** are not collinear.

2^0. *Associativity,*

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}) \tag{2}$$

(see Fig. 8).

The two properties imply that the sum of any number of vectors is not affected by the order of addition of the methods of grouping.

To construct the sum $\mathbf{a}_1 + \mathbf{a}_2 + \ldots + \mathbf{a}_n$, it is most convenient to apply the *rule of closure of an open polygon,* which is a generalization of the triangle law and is as follows. If we go from the initial point of the vector $\mathbf{a}_1$ to the terminal point of the vector $\mathbf{a}_n$ by way of vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, then the vector from the initial point of $\mathbf{a}_1$ to the terminal point of $\mathbf{a}_n$ is their sum $\mathbf{a}_1 + \mathbf{a}_2 + \ldots + \mathbf{a}_n$ (Fig. 9).

3. Vector subtraction.

Definition. Given any two vectors **a** and **b**, their *difference* $\mathbf{b} - \mathbf{a}$ is a vector **d**, such that when added to the vector **a** yields the vector **b**.

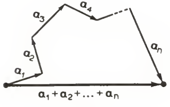


Figure 9

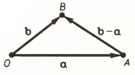


Figure 10

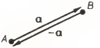


Figure 11

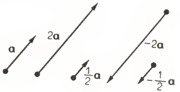


Figure 12

Hence follows the rule for vector subtraction. We position vectors **a** and **b** so that their initial points coincide (point O, see Fig. 10), the vector from the terminal point (A) of **a** to the terminal point (B) of **b** is the vector **b** − **a**. Indeed, if $\mathbf{a} = \overrightarrow{OA}$ and $\mathbf{b} = \overrightarrow{OB}$, then $\overrightarrow{OA} + \overrightarrow{AB} = \overrightarrow{OB}$, i.e. $\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA}$, or

$$\overrightarrow{AB} = \mathbf{b} - \mathbf{a}$$

The vector with the same length as the vector **a** but oppositely directed is called the *negative of* (*opposite* to) **a** and is denoted by −**a**. We can easily see that

$$\mathbf{a} + (-\mathbf{a}) = 0 \tag{3}$$

Obviously, if $\mathbf{a} = \overrightarrow{AB}$, then $-\mathbf{a} = \overrightarrow{BA}$, and we have (see Fig. 11)

$$\mathbf{a} + (-\mathbf{a}) = \overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = 0$$

Formula (3) yields the following for the difference of two vectors

$$\mathbf{b} - \mathbf{a} = \mathbf{b} + (-a) \tag{4}$$

Indeed, if we add the vector **a** and the vector **b** + (−**a**), we get

$$\mathbf{a} + [\mathbf{b} + (-\mathbf{a})] = \mathbf{b} + [\mathbf{a} + (-\mathbf{a})] = \mathbf{b} + 0 = \mathbf{b}$$

Formula (4) is most useful when we add or subtract more than two vectors. For instance, to find the difference **a** − **b** − **c** − **d**, we should use vectors **a**, −**b**, −**c**, −**d** to construct their sum according to the rule of closure of an open polygon.

4. Multiplication of a vector by a scalar.

If **a** is a vector and λ is a number (scalar), then the *product*

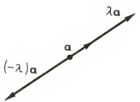


Figure 13

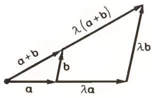


Figure 14

$\lambda\mathbf{a}$ is a vector which is collinear to $\mathbf{a}$, has the length $|\lambda|\cdot|\mathbf{a}|$, and has the same direction as $\mathbf{a}$ if $\lambda > 0$ and is opposite to $\mathbf{a}$ if $\lambda < 0$. Here $|\lambda|$ denotes the absolute value of the number λ. The product $0\mathbf{a}$ is equal to zero (zero vector).

If $\lambda > 0$, the multiplication of $\mathbf{a}$ by λ can be visualized as follows: the vector $\lambda\mathbf{a}$ is the result of "stretching" the vector $\mathbf{a}\lambda$ times. However, we mean a mathematical stretching, so that if, for instance, $\lambda = 1/2$, stretching λ times means that the length of $\mathbf{a}$ is reduced to one half. If $\lambda < 0$, stretching $|\lambda|$ times is accompanied by changing the direction of $\mathbf{a}$ to the opposite (Fig. 12).

The operation of vector multiplication by scalars possesses the following property:

$$\lambda(\mu\mathbf{a}) = (\lambda\mu)\,\mathbf{a} \tag{5}$$

The vectors $\lambda(\mu\mathbf{a})$ and $(\lambda\mu)\,\mathbf{a}$ have the same magnitude equal to $|\lambda||\mu||\mathbf{a}|$, have the same direction as $\mathbf{a}$ if λ and μ are of the same sign and have the opposite direction if λ and μ are of opposite signs. If either λ or μ is zero, both vectors $\lambda(\mu\mathbf{a})$ and $(\lambda\mu)\,\mathbf{a}$ are zero.

Here is another property of the operation of vector multiplication by a scalar:

$$-(\lambda\mathbf{a}) = (-\lambda)\,\mathbf{a} \tag{6}$$

This property is illustrated in Fig. 13.

Theorem (on collinear vectors). *If* $\mathbf{a}$ *and* $\mathbf{b}$ *are two collinear vectors and* $\mathbf{a}$ *is nonzero, then there is a unique number* λ *such that* $\mathbf{b} = \lambda\mathbf{a}$.

□ If $\mathbf{b} = 0$, then $\lambda = 0$. Now let $\mathbf{b} \neq 0$. Let $\lambda = |\mathbf{b}|/|\mathbf{a}|$ if the vectors $\mathbf{b}$ and $\mathbf{a}$ are in the same direction, and $\lambda = -|\mathbf{b}|/|\mathbf{a}|$ if they have opposite directions. Then $\mathbf{b} = \lambda\mathbf{a}$.

If $\mathbf{b} = \mu\mathbf{a}$ along with $\mathbf{b} = \lambda\mathbf{a}$, then $(\lambda - \mu)\,\mathbf{a} = 0$, and since $\mathbf{a} \neq 0$, we have $\lambda - \mu = 0$, i.e. $\lambda = \mu$. ■

5. Linear operations on vectors.

The operations of vector addition and multiplication by a number are called *linear operations*. (The operation of subtraction is defined via addition and thereby is considered to be "secondary".)

In the school course the following two properties, each involving both linear operations, were proved:

$$(\lambda + \mu)\,\mathbf{a} = \lambda\mathbf{a} + \mu\mathbf{a} \tag{7}$$

$$\lambda(\mathbf{a} + \mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b} \tag{8}$$

Here $\mathbf{a}$, $\mathbf{b}$, λ, and μ are arbitrary. Figure 14 illustrates Eq. (8) for $\lambda > 1$.

Properties (1) through (8) are important since they make it possible to perform calculations in vector algebra basically as in common algebra, for instance, we can use the rules of arithmetic for removing brackets and factoring out.

6. A Linear combination of several vectors.

Suppose we have several vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$.

Definition. *A linear combination of vectors* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ is the sum of the products of these vectors multiplied by any numbers $c_1, c_2, \ldots, c_n$.

For instance, $3\mathbf{a} - \frac{1}{2}\mathbf{b} + 7\mathbf{c}$ is a linear combination of the vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$.

Example. A quadrilateral $ABCD$ is given, and P and Q are the midpoints of the sides BC and AD respectively (Fig. 15), express the vector $\overrightarrow{PQ}$ via the vectors $\overrightarrow{AB}$, $\overrightarrow{BC}$, $\overrightarrow{CD}$.

○ We have

$$\overrightarrow{PQ} = \overrightarrow{PB} + \overrightarrow{BA} + \overrightarrow{AQ} = -\frac{1}{2}\overrightarrow{BC} - \overrightarrow{AB} + \frac{1}{2}\overrightarrow{AD}$$

Since $\overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{CD} + \overrightarrow{DA} = \overrightarrow{AA} = 0$ and hence $\overrightarrow{DA} =$

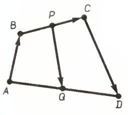


Figure 15

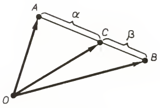


Figure 16

$-(\vec{AB} + \vec{BC} + \vec{CD})$, we get

$$\vec{PQ} = -\frac{1}{2}\vec{BC} - \vec{AB} + \frac{1}{2}(\vec{AB} + \vec{BC} + \vec{CD})$$

$$= -\frac{1}{2}\vec{AB} + \frac{1}{2}\vec{CD}$$

We see that the expression does not contain the vector $\vec{BC}$. ●

7. A line segment divided in a ratio.

Let a point C divide a line segment AB (see Fig. 16) in the ratio $\frac{\alpha}{\beta}$, i.e.

$$\frac{|AC|}{|CB|} = \frac{\alpha}{\beta}$$

Then,

$$\vec{AC} = \frac{\alpha}{\beta}\vec{CB} \tag{9}$$

We connect points A, B, and C with a point O and formulate the following problem: *express vector* $\vec{OC}$ *in terms of vectors* $\vec{OA}$ *and* $\vec{OB}$.

○ We have $\vec{AC} = \vec{OC} - \vec{OA}$ and $\vec{CB} = \vec{OB} - \vec{OC}$. Multiplying both sides of the first equation by β and both sides of the second by α, we have

$$\beta\vec{AC} = \beta(\vec{OC} - \vec{OA}) \quad \text{and} \quad \alpha\vec{CB} = \alpha(\vec{OB} - \vec{OC})$$

Since, by virtue of (9), the left-hand sides of the equalities are

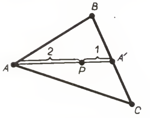


Figure 17

equal, their right-hand sides are also equal

$$\beta(\vec{OC} - \vec{OA}) = \alpha(\vec{OB} - \vec{OC})$$

We write this equation as $(\alpha + \beta)\,\vec{OC} = \beta\vec{OA} + \alpha\vec{OB}$, whence

$$\vec{OC} = \frac{\beta}{\alpha + \beta}\,\vec{OA} + \frac{\alpha}{\alpha + \beta}\,\vec{OB} \quad \bullet \tag{10}$$

Formula (10) is widely used in calculations.

Example. Prove that the medians of a triangle intersect at one point, which is the centroid of the triangle, and that this point divides each median in the ratio 2:1 reckoning from the vertex along the median.

○ We choose an arbitrary point O in the plane of the triangle. Then consider point P which divides median AA' in the ratio 2:1 (Fig. 17). By formula (10) we have

$$\vec{OP} = \frac{1}{2+1}\,\vec{OA} + \frac{2}{2+1}\,\vec{OA'}$$

$$\vec{OA'} = \frac{1}{2}\,\vec{OB} + \frac{1}{2}\,\vec{OC}$$

Whence

$$\vec{OP} = \frac{1}{3}\,\vec{OA} + \frac{2}{3}\left(\frac{1}{2}\,\vec{OB} + \frac{1}{2}\,\vec{OC}\right)$$

$$= \frac{1}{3}\,(\vec{OA} + \vec{OB} + \vec{OC})$$

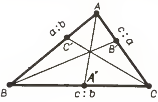


Figure 18

Vectors $\vec{OA}$, $\vec{OB}$, and $\vec{OC}$ are in equal proportion in this expression. Hence it is clear that if we take points Q and R which divide medians BB' and CC' respectively in the ratio 2:1, we obtain similar expressions for vectors $\vec{OQ}$ and $\vec{OR}$. Consequently, $\vec{OP} = \vec{OQ} = \vec{OR}$, which proves that points P, Q, and R coincide. ●

Here is another example. It generalizes the preceding example and represents Ceva's theorem from elementary geometry, viz. *if the sides of a triangle are divided in the ratios* $\alpha:\beta$, $\gamma:\alpha$, $\beta:\gamma$ (see Fig. 18), *then the straight lines joining the vertices to the points of division of the opposite sides intersect at one point.*

□ We consider point P which divides the line segment AA' in the ratio $(\beta + \gamma):\alpha$. From formula (10) we have

$$\vec{OP} = \frac{\alpha}{\alpha + \beta + \gamma}\vec{OA} + \frac{\beta + \gamma}{\alpha + \beta + \gamma}\vec{OA'}$$

$$= \frac{\alpha}{\alpha + \beta + \gamma}\vec{OA} + \frac{\beta + \gamma}{\alpha + \beta + \gamma}\left(\frac{\beta}{\beta + \gamma}\vec{OB} + \frac{\gamma}{\beta + \gamma}\vec{OC}\right)$$

$$= \frac{1}{\alpha + \beta + \gamma}(\alpha\vec{OA} + \beta\vec{OB} + \gamma\vec{OC})$$

Thus it is clear that if we take point Q, which divides line segment BB' in the ratio $(\alpha + \gamma):\beta$, and point R, which divides line segment CC' in the ratio $(\alpha + \beta):\gamma$, then we obtain similar expressions for vectors $\vec{OQ}$ and $\vec{OR}$. Consequently, $\vec{OP} = \vec{OQ} = \vec{OR}$, which proves that points P, Q, and R coincide. ■

The theorems on the intersection of bisectors and on the intersection of the altitudes of a triangle at one point are special cases of Ceva's theorem.

1.2. VECTOR BASIS IN THE PLANE AND IN SPACE

1. Vector basis in the plane. Representation of a vector in terms of the basis.

Lemma. *If vectors* **a** *and* **b** *are not collinear, then the equation*

$$\alpha \mathbf{a} + \beta \mathbf{b} = 0 \tag{1}$$

holds true if and only if both α *and* β *are zero.*

□ Let, say, $\alpha \neq 0$. Then (1) yields

$$\mathbf{a} = -\frac{\beta}{\alpha}\mathbf{b}$$

and this contradicts the hypothesis that **a** and **b** are not collinear. Therefore, $\alpha = 0$. We can prove that $\beta = 0$ in a similar way. ■

Let us now consider a plane in space.

Definition. A *vector basis* in a given plane is a set of any two noncollinear vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ in the plane.

The vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are termed the first and the second basis vectors respectively.

We prove the following important theorem.

Theorem. *Let* $\mathbf{e}_1$, $\mathbf{e}_2$ *be a vector basis in the plane. Then any vector* **a** *in the plane can be uniquely represented as a linear combination of the basis vectors* $\mathbf{e}_1$ *and* $\mathbf{e}_2$

$$\mathbf{a} = X\mathbf{e}_1 + Y\mathbf{e}_2 \tag{2}$$

□ We reduce the vectors $\mathbf{e}_1$, $\mathbf{e}_2$, and **a** to a common origin O and draw a straight line through point A (the terminus of **a**) and parallel to $\mathbf{e}_2$. Let A_1 be the point of intersection of the straight line and the axis of $\mathbf{e}_1$ (Fig. 19); point A_1 does exist since the vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are not collinear. We have $\mathbf{a} = \overrightarrow{OA_1} + \overrightarrow{A_1A}$. By the theorem on collinear vectors, $\overrightarrow{OA_1} = X\mathbf{e}_1$ and $\overrightarrow{A_1A} = Y\mathbf{e}_2$, where X and Y are numbers, whence $\mathbf{a} = X\mathbf{e}_1 + Y\mathbf{e}_2$, and this is the representation of the vector **a** as a linear combination of the vectors $\mathbf{e}_1$ and $\mathbf{e}_2$.

We now prove that this representation is unique. Let the equality

$$\mathbf{a} = X^*\mathbf{e}_1 + Y^*\mathbf{e}_2 \tag{3}$$

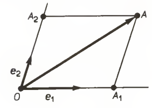


Figure 19

as well as Eq. (2) be true. Subtracting (2) from (3) we have

$$(X - X^*)\,\mathbf{e}_1 + (Y - Y^*)\,\mathbf{e}_2 = 0$$

According to the above lemma this means that $X = X^*$ and $Y = Y^*$. ∎

Definition. Equality (2) is the representation of the vector **a** in terms of the basis $\mathbf{e}_1$, $\mathbf{e}_2$ and the numbers X and Y are called the coordinates of **a** in the basis $\mathbf{e}_1$, $\mathbf{e}_2$ (or in terms of the basis $\mathbf{e}_1$, $\mathbf{e}_2$).

This definition of vector coordinates implies that *equal vectors have equal coordinates.* Indeed, if $\mathbf{a} = X\mathbf{e}_1 + Y\mathbf{e}_2$ and $\mathbf{a}' = X'\mathbf{e}_1 + Y'\mathbf{e}_2$, then the equality $\mathbf{a} = \mathbf{a}'$ implies that

$$(X' - X)\,\mathbf{e}_1 + (Y' - Y)\,\mathbf{e}_2 = 0$$

whence from the lemma we have $X' = X$ and $Y' = Y$.

For the sake of brevity we accept the following notation: if X and Y are the coordinates of the vector **a** in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, we shall write $\mathbf{a} = \langle X, Y\rangle$ in the basis $\mathbf{e}_1$, $\mathbf{e}_2$; if the basis is known, we shall simply write $\mathbf{a} = \langle X, Y\rangle$.

Let us return to the proof of the representation theorem. A direct corollary of the proof is the following proposition: *let vectors* $\mathbf{e}_1$, $\mathbf{e}_2$, *and* **a** *be reduced to a common origin O. Through the terminus of* **a** *we draw a straight line parallel to the vector* $\mathbf{e}_1$. *Let* A_2 *be the point of intersection of the straight line and the axis of the vector* $\mathbf{e}_2$. *Then* $\overrightarrow{OA_2} = Y\mathbf{e}_2$, *where Y is the second coordinate of* **a** *in the basis* $\mathbf{e}_1$, $\mathbf{e}_2$.

Indeed, we can see from Fig. 19 that $\overrightarrow{OA_2} = \overrightarrow{A_1A}$; consequently, $\overrightarrow{OA_2} = Y\mathbf{e}_2$.

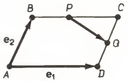


Figure 20

By changing the indices of the basis vectors we can always get $\mathbf{e}_1$ to be the second basis vector (we assume that $\mathbf{e}_1' = \mathbf{e}_2$, $\mathbf{e}_2' = \mathbf{e}_1$). Hence it is clear that we have an analogous proposition for the first coordinate of the vector $\mathbf{a}$: *through the terminus of* $\mathbf{a}$ *we draw a straight line parallel to the vector* $\mathbf{e}_2$. *Let* A_1 *be the point where the straight line intersects the axis of the vector* $\mathbf{e}_1$. *Then* $\overrightarrow{OA}_1 = X\mathbf{e}_1$, *where* X *is the first coordinate of* $\mathbf{a}$.

Example. Given a parallelogram $ABCD$ (Fig. 20). The points P and Q are the midpoints of the sides BC and CD respectively, find the coordinates of the vector $\overrightarrow{PQ}$ if we assume that $\mathbf{e}_1 = \overrightarrow{AD}$ and $\mathbf{e}_2 = \overrightarrow{AB}$ are the basis vectors.

○ We have $\overrightarrow{PQ} = \overrightarrow{PC} + \overrightarrow{CQ} = \frac{1}{2}\overrightarrow{AD} - \frac{1}{2}\overrightarrow{AB} = \frac{1}{2}\mathbf{e}_1 - \frac{1}{2}\mathbf{e}_2$. Hence $\overrightarrow{PQ} = \langle 1/2, -1/2\rangle$ in the basis $\mathbf{e}_1$, $\mathbf{e}_2$. ●

2. Coplanar vectors.

We say that the vector $\overrightarrow{AB}$ is *parallel* to a given plane α if the straight line AB containing this vector is parallel to the plane. A zero vector is assumed to be parallel to any plane.

Definition. Several vectors **a**, **b**, **c**, ... in space are called *coplanar* if they are parallel to the same plane.

If vectors $\overrightarrow{OA}$, $\overrightarrow{OB}$, $\overrightarrow{OC}$, ... with the common origin O are coplanar, then the points A, B, C, ... are obviously in the same plane. In this sense we can say that coplanar vectors can be *translated to the same plane*.

Example. Consider a pyramid with the vertices A, B, C, D (Fig. 21). The vectors $\overrightarrow{BC}$, $\overrightarrow{CD}$, $\overrightarrow{DB}$ are obviously coplanar, while the vectors $\overrightarrow{AD}$, $\overrightarrow{AB}$, $\overrightarrow{AC}$ are not, neither are the vectors $\overrightarrow{AB}$, $\overrightarrow{AD}$, $\overrightarrow{DC}$: otherwise the point C would lie in the plane ABD.

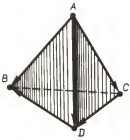


Figure 21

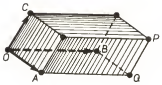


Figure 22

Two vectors are always coplanar, while three vectors can be noncoplanar. If the vectors $\overrightarrow{OA}$, $\overrightarrow{OB}$, $\overrightarrow{OC}$ are not coplanar, we can obtain three distinct planes by drawing through the terminus of each vector a plane parallel to the two remaining vectors. These three planes together with three other planes, which are defined by the pairs of vectors $\overrightarrow{OA}$ and $\overrightarrow{OB}$, $\overrightarrow{OA}$ and $\overrightarrow{OC}$, $\overrightarrow{OB}$ and $\overrightarrow{OC}$, respectively, enclose a parallelepiped (Fig. 22). If P is a vertex of the parallelepiped, then

$$\overrightarrow{OP} = \overrightarrow{OQ} + \overrightarrow{QP} = \overrightarrow{OQ} + \overrightarrow{OC} = \overrightarrow{OA} + \overrightarrow{OB} + \overrightarrow{OC}.$$

This expression gives a *way of constructing the sum of three noncoplanar vectors* $\overrightarrow{OA}$, $\overrightarrow{OB}$, $\overrightarrow{OC}$: we consider the parallelepiped with vertex O and the line segments OA, OB, OC as its three edges; then the vector $\overrightarrow{OA} + \overrightarrow{OB} + \overrightarrow{OC}$ is a principal diagonal of the parallelepiped starting at the vertex O.

3. Vector basis in space. Representation of an arbitrary vector in terms of the basis.

Lemma. *If three vectors* **a**, **b**, *and* **c** *are noncoplanar, then the equality*

$$\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} = 0 \tag{4}$$

holds true if $\alpha = 0$, $\beta = 0$, $\gamma = 0$.

□ Let, for example, $\alpha \neq 0$. Then from (4) we have

$$\mathbf{a} = -\frac{\beta}{\alpha}\mathbf{b} - \frac{\gamma}{\alpha}\mathbf{c}. \tag{5}$$

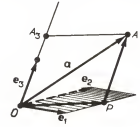


Figure 23

If vectors **b** and **c** are collinear, then (5) implies that all three vectors **a**, **b**, and **c** are collinear, and hence are coplanar. If **b** and **c** are not collinear, then after being reduced to a common origin they will define some plane; then by (5) the three vectors **a**, **b**, and **c** are parallel to that plane. Thus the assumption that $\alpha \neq 0$ led us to a contradiction with the hypothesis. This means that $\alpha = 0$. Similarly we can prove that $\beta = 0$ and $\gamma = 0$. ■

Definition. A *vector basis* in space is a set of any three non-coplanar vectors $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$.

The vectors $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ are called the first, the second, and the third basis vectors.

Theorem. *Let* $\mathbf{e}_1$, $\mathbf{e}_2$, *and* $\mathbf{e}_3$ *be a vector basis in space. Then any vector* **a** *in space can be uniquely represented as a linear combination of the basis vectors* $\mathbf{e}_1$, $\mathbf{e}_2$, *and* $\mathbf{e}_3$:

$$\mathbf{a} = X\mathbf{e}_1 + Y\mathbf{e}_2 + Z\mathbf{e}_3 \tag{6}$$

□ We reduce the vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, and **a** to a common origin O and denote the plane defined by the vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ by α. We draw through a point A (the terminus of **a**) a straight line l parallel to $\mathbf{e}_3$ (Fig. 23). We designate the point where the straight line intersects the plane α as P (the point P does exist since the straight line l is not parallel to α). We have

$$\mathbf{a} = \overrightarrow{OA} = \overrightarrow{OP} + \overrightarrow{PA}.$$

By virtue of the theorem proved in Sec. 1.2.1 the vector $\overrightarrow{OP}$ can be represented as a linear combination $\mathbf{e}_1$ and $\mathbf{e}_2$: $\overrightarrow{OP} = X\mathbf{e}_1 + Y\mathbf{e}_2$; by the theorem on collinear vectors,

$\vec{PA} = Z\mathbf{e}_3$. Hence we have $\mathbf{a} = X\mathbf{e}_1 + Y\mathbf{e}_2 + Z\mathbf{e}_3$, and this proves that the vector $\mathbf{a}$ can be represented as a linear combination of the vectors $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$.

The uniqueness of the representation can be proved in the same way as in the preceding theorem. ■

Definition. Expression (6) is the *representation of vector* $\mathbf{a}$ *in the basis* $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$, and the numbers X, Y, and Z are called the *coordinates of* $\mathbf{a}$ *in the basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$.

As in the case of the plane, here we accept the following notation: if X, Y, Z are the coordinates of the vector $\mathbf{a}$ in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, then we shall write $\mathbf{a} = \langle X, Y, Z\rangle$ in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$.

If the basis is predetermined, then we simply write $\mathbf{a} = \langle X, Y, Z\rangle$.

We now turn to the proof of the last theorem again. A direct consequence of the proof is the following proposition: *let vectors* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, *and* $\mathbf{a}$ *be reduced to a common origin O. Through the terminus of* $\mathbf{a}$ *we draw a plane parallel to the vectors* $\mathbf{e}_1$ *and* $\mathbf{e}_2$ *and designate by* A_3 *the point of intersection of this plane and the axis of* $\mathbf{e}_3$. *Then* $\vec{OA_3} = Z\mathbf{e}_3$, *where Z is the third coordinate of the vector* $\mathbf{a}$ *in the basis* $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$.

Indeed, we can see from Fig. 23 that $\vec{OA_3} = \vec{PA}$; consequently, $\vec{OA_3} = Z\mathbf{e}_3$.

By changing the indices of the basis vectors we can always get $\mathbf{e}_2$ or $\mathbf{e}_1$ to be the third basis vector (for example, assume that $\mathbf{e}_1' = \mathbf{e}_1$, $\mathbf{e}_2' = \mathbf{e}_3$, $\mathbf{e}_3' = \mathbf{e}_2$). Hence it follows that propositions similar to the one just proved are also valid for two other coordinates of the vector $\mathbf{a}$.

Let a plane be drawn through the terminus of the vector $\mathbf{a}$ *parallel to vectors* $\mathbf{e}_1$ *and* $\mathbf{e}_3$, *and* A_2 *be the point where that plane intersects the axis of* $\mathbf{e}_2$. *Then* $\vec{OA_2} = Y\mathbf{e}_2$, *where Y is the second coordinate of* $\mathbf{a}$.

Let a plane be drawn through the terminus of the vector $\mathbf{a}$ *parallel to vectors* $\mathbf{e}_2$ *and* $\mathbf{e}_3$, *and let* A_1 *be the point where that plane intersects the axis of* $\mathbf{e}_1$. *Then* $\vec{OA_1} = X\mathbf{e}_1$, *where X is the first coordinate of* $\mathbf{a}$.

Example. A pyramid with vertices A, B, C, and D is given

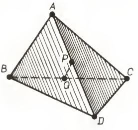


Figure 24

and P and Q are the midpoints of edges AD and BC respectively (Fig. 24), find the coordinates of $\vec{PQ}$ in the basis $\mathbf{e}_1 = \vec{AB}$, $\mathbf{e}_2 = \vec{AC}$, $\mathbf{e}_3 = \vec{AD}$.

We have

$$\vec{PQ} = \vec{PA} + \vec{AC} + \vec{CQ} = -\frac{1}{2}\vec{AD} + \vec{AC} + \frac{1}{2}\vec{CB}$$

$$= -\frac{1}{2}\vec{AD} + \vec{AC} + \frac{1}{2}(\vec{CA} + \vec{AB})\frac{1}{2}\mathbf{e}_1 + \frac{1}{2}\mathbf{e}_2 - \frac{1}{2}\mathbf{e}_3$$

Whence $\vec{PQ} = \langle 1/2,\ 1/2,\ -1/2 \rangle$ in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. ●

4. Operations on vectors defined by their coordinates. Assume that we have two vectors $\mathbf{a}$ and $\mathbf{b}$ such that $\mathbf{a} = \langle X, Y \rangle$, $\mathbf{b} = \langle X', Y' \rangle$, in the vector basis $\mathbf{e}_1$, $\mathbf{e}_2$. Then we have $\mathbf{a} = X\mathbf{e}_1 + Y\mathbf{e}_2$, $\mathbf{b} = X'\mathbf{e}_1 + Y'\mathbf{e}_2$, and therefore

$$\mathbf{a} + \mathbf{b} = (X + X')\,\mathbf{e}_1 + (Y + Y')\,\mathbf{e}_2$$

$$\mathbf{a} - \mathbf{b} = (X - X')\,\mathbf{e}_1 + (Y - Y')\,\mathbf{e}_2$$

Thus

$$\mathbf{a} + \mathbf{b} = \langle X + X',\ Y + Y' \rangle, \quad \mathbf{a} - \mathbf{b} = \langle X - X',\ Y - Y' \rangle$$

that is, *when adding or subtracting vectors we add or subtract their respective coordinates.*

Similarly, if $\mathbf{a} = \langle X, Y \rangle$, we can write $\mathbf{a} = X\mathbf{e}_1 + Y\mathbf{e}_2$, $\lambda\mathbf{a} = \lambda X\mathbf{e}_1 + \lambda Y\mathbf{e}_2$, whence it follows that

$$\lambda\mathbf{a} = \langle \lambda X, \lambda Y \rangle$$

that is, *when multiplying a vector by a number we multiply its coordinates by that number.*

The rules of operations on vectors in space are the same as those in the plane; if $\mathbf{a} = \langle X, Y, Z \rangle$ and $\mathbf{b} = \langle X', Y', Z' \rangle$, then

$$\mathbf{a} + \mathbf{b} = \langle X + X', \ Y + Y', \ Z + Z' \rangle$$
$$\mathbf{a} - \mathbf{b} = \langle X - X', \ Y - Y', \ Z - Z' \rangle$$
$$\lambda\mathbf{a} = \langle \lambda X, \ \lambda Y, \ \lambda Z \rangle$$

We can prove these equalities in the same way we did in the case of the plane.

Example. Let $\mathbf{a} = \langle -1, 2, 5 \rangle$, $\mathbf{b} = \langle 1, 3, 7 \rangle$. Find the vector $7\mathbf{a} - 5\mathbf{b}$.

○ We have $7\mathbf{a} = \langle -7, 14, 35 \rangle$, $5\mathbf{b} = \langle 5, 15, 35 \rangle$, $7\mathbf{a} - 5\mathbf{b} = \langle -12, \ -1, \ 0 \rangle$. ●

5. The condition of the collinearity of vectors in coordinates. Suppose that $\mathbf{a} = \langle X, Y \rangle$ and $\mathbf{b} = \langle X', Y' \rangle$ are two vectors in the plane, with $\mathbf{a} \neq 0$. If $\mathbf{b}$ is not collinear with $\mathbf{a}$, then $\mathbf{b} = \lambda\mathbf{a}$, where λ is a number. Since equal vectors have equal coordinates, we have

$$X' = \lambda X, \quad Y' = \lambda Y \tag{7}$$

Conversely: If equalities (7) are true, then $\mathbf{b} = \lambda\mathbf{a}$, that is $\mathbf{b}$ is collinear with $\mathbf{a}$.

Thus, *the vector* $\mathbf{b}$ *is collinear with the nonzero vector* $\mathbf{a}$ *if and only if the coordinates of* $\mathbf{b}$ *are proportional to the respective coordinates of* $\mathbf{a}$.

The same conclusion is also valid for vectors in space.

If none of the coordinates of vector $\mathbf{a}$ is zero, the condition that vector $\mathbf{b}$ is collinear with vector $\mathbf{a}$ can be written in the form

$$\frac{X'}{X} = \frac{Y'}{Y}$$

in the plane, and as

$$\frac{X'}{X} = \frac{Y'}{Y} = \frac{Z'}{Z}$$

in space.

For instance, the vector $\mathbf{b} = \langle -2, 6, 4 \rangle$ is collinear with the vector $\mathbf{a} = \langle -3, 9, 6 \rangle$ since $-2/-3 = 6/9 = 4/6$.

Example. Verify that the two vectors $\mathbf{a} = \langle -1, 3 \rangle$ and $\mathbf{b} = \langle 2, 2 \rangle$ in the plane are noncollinear and express the vector $\mathbf{c} = \langle 7, -5 \rangle$ in terms of the basis $\mathbf{a}$, $\mathbf{b}$.

○ Since $-1/2 \neq 3/2$, $\mathbf{a}$ and $\mathbf{b}$ are noncollinear. Consequently, they form a basis in the plane. Let $\mathbf{c} = X\mathbf{a} + Y\mathbf{b}$. To find X and Y, we equate the respective coordinates of the vectors $\mathbf{c}$ and $X\mathbf{a} + Y\mathbf{b}$ and obtain the following system of two equations in two unknowns X and Y:

$$7 = X\cdot(-1) + Y\cdot 2\cdot 2 = X\cdot 3 + Y\cdot 2$$

By solving it we find $X = -3$, $Y = 2$. Thus, $\mathbf{c} = -3\mathbf{a} + 2\mathbf{b}$. ●

1.3. CARTESIAN COORDINATE SYSTEM ON A STRAIGHT LINE, IN THE PLANE, AND IN SPACE

Suppose that O is a fixed point and we call it the *origin*. If M is an arbitrary point, then the vector $\overrightarrow{OM}$ is called the *radius vector* of M relative to the origin O, or in short, the radius vector of M.

1. Cartesian coordinates on a straight line.

Suppose that we have a straight line l in space. We choose the origin O on that line and also a nonzero vector $\mathbf{e}$ which we shall call a *basis* vector (Fig. 25).

Definition. The set $\{0, \mathbf{e}\}$ of the point O and the basis vector $\mathbf{e}$ is called a *Cartesian coordinate system on the straight line*.

We consider an arbitrary point M on the straight line l. Since the vectors $\overrightarrow{OM}$ and $\mathbf{e}$ are collinear, we have $\overrightarrow{OM} = x\mathbf{e}$, where x is a number. We call this number the *coordinate of M on the straight line*. The coordinate of the origin O is zero. All the points M on the line have either positive or negative coordinates depending on whether the directions of $\overrightarrow{OM}$ and $\mathbf{e}$ coin-

Figure 25

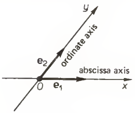


Figure 26

cide or are opposite. The straight line l on which we have introduced coordinates is called a *coordinate axis* or the *x-axis*.

Assigning coordinates on the straight line leads to each point M on the line being associated with a unique number x, the coordinate of the point. Conversely: an arbitrary number x corresponds to a unique point M for which the number is its coordinate.

2. Cartesian coordinate in the plane.

We choose an origin O in the plane and two collinear vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ which form a vector basis.

Definition. The set $\{O, \mathbf{e}_1, \mathbf{e}_2\}$ of the point O and the vector basis $\mathbf{e}_1, \mathbf{e}_2$ is called a *Cartesian coordinate system in the plane.*

Two straight lines which pass through O parallel to the respective vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are called the *coordinate axes*, the first being the *abscissa* or the *x-axis* and the second, the *ordinate* or the *y-axis.*

We shall always depict the vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ lying along the respective coordinate axes (Fig. 26).

If the Cartesian coordinate system is given in the plane, then the position of a point M with respect to that system can be defined by two numbers x and y, which are the coordinates of M in the system.

Definition. The *coordinates of the point M* in the plane with respect to the Cartesian coordinate system $\{O, \mathbf{e}_1, \mathbf{e}_2\}$ are the coordinates of its radius vector $\overrightarrow{OM}$ in the basis $\mathbf{e}_1, \mathbf{e}_2$.

In other words, to find the coordinates of a point M, we should express the vector $\overrightarrow{OM}$ in terms of the basis $\mathbf{e}_1, \mathbf{e}_2$:

$$\overrightarrow{OM} = x\mathbf{e}_1 + y\mathbf{e}_2$$

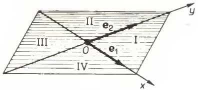


Figure 27

The numbers x and y are the coordinates of M with respect to the Cartesian coordinate system $\{O, \mathbf{e}_1, \mathbf{e}_2\}$.

The x coordinate is called the *abscissa* of the point M, and the y coordinate the *ordinate* of M. The coordinates of a point are usually given in parentheses. In compact notation, we speak of "the point $M(x, y)$".

Thus, every point M of the plane with the fixed coordinate system $\{O, \mathbf{e}_1, \mathbf{e}_2\}$ is associated with an ordered pair of numbers (x, y). Conversely, every ordered pair of numbers (x, y) is associated with the only point M in the plane; this point is the terminus of the vector $\overrightarrow{OM} = x\mathbf{e}_1 + y\mathbf{e}_2$.

The coordinates of the origin O are 0, 0. The coordinates of points lying on the abscissa axis are x, 0, and the coordinates of points of the ordinate axis are 0, y. The remaining part of the plane is separated by the coordinate axes into four compartments called *quadrants* (Fig. 27): we have $x > 0$, $y > 0$ for the points of the first quadrant, $x < 0$, $y > 0$ for those of the second quadrant, $x < 0$, $y < 0$ for those of the third quadrant, and $x > 0$, $y < 0$ for those of the fourth quadrant.

Figure 28 shows several points with their coordinates.

Coordinate systems underlie the methods of analytic geometry and reduce any geometric problem to that of arithmetic or algebra. First, we arithmetize initial data. For example, to specify a point we define its coordinates. The solution of the problem is also arithmetized (we will touch upon this subject below). The final result also has an arithmetic form; for instance, a point is found when its coordinates are found.

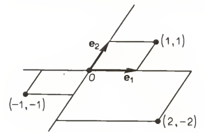


Figure 28

3. Coordinates of a vector in Cartesian coordinate systems.

Definition. The *coordinates of a vector* $\mathbf{a}$ in the plane with respect to the Cartesian coordinate system $\{O, \mathbf{e}_1, \mathbf{e}_2\}$ are the coordinates of this vector in terms of the basis $\mathbf{e}_1$, $\mathbf{e}_2$.

In other words, in order to find the coordinates of the vector $\mathbf{a}$, we should express it in terms of the basis $\mathbf{e}_1$, $\mathbf{e}_2$:

$$\mathbf{a} = X\mathbf{e}_1 + Y\mathbf{e}_2$$

the coefficients X and Y are the coordinates of $\mathbf{a}$ relative to the Cartesian coordinate system $\{O, \mathbf{e}_1, \mathbf{e}_2\}$.

Let us consider the following important problem.

Problem. Given two points $A(x, y)$, $B(x', y')$, find the coordinates X and Y of the vector $\overrightarrow{AB}$.

○ We have $\overrightarrow{OA} = x\mathbf{e}_1 + y\mathbf{e}_2$, $\overrightarrow{OB} = x'\mathbf{e}_1 + y'\mathbf{e}_2$. By subtracting the first equality from the second we get $\overrightarrow{AB} = (x' - x)\mathbf{e}_1 + (y' - y)\mathbf{e}_2$. Consequently,

$$X = x' - x, \quad Y = y' - y \; \bullet \tag{1}$$

Thus, *the coordinates of a vector are equal to the difference of the respective coordinates of the terminus and origin of the vector.*

Using (1) and the condition for two vectors to be collinear we can derive a condition when three points of the plane, $A(x, y)$, $B(x', y')$, and $C(x'', y'')$, lie on the same straight line. For the points A, B, and C to lie on one straight line, it is necessary and sufficient that the vector $\overrightarrow{AC}$ be collinear to the vector $\overrightarrow{AB}$,

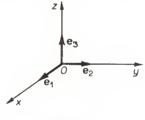


Figure 29

that is, the coordinates $x'' - x$, $y'' - y$ of $\overrightarrow{AC}$ be proportional to the respective coordinates $x' - x$, $y' - y$ of $\overrightarrow{AB}$. Thus, the *condition for three points A, B, C to lie on one straight line is that the numbers* $x'' - x$ *and* $y'' - y$ *are proportional to* $x' - x$, $y' - y$.

Example. Points $A(1, 1)$, $B(0, -3)$, $C(3, 9)$ lie on one straight line since the numbers $3 - 1$, $9 - 1$ are proportional to $0 - 1$, $-3 - 1$.

4. Cartesian coordinate system in space.

Suppose we have a fixed point O (the origin) and a vector basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ in space.

Definition. A *Cartesian coordinate system in space* is the set $\{O, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$.

Three straight lines passing through O parallel to the corresponding vectors $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ are called the *axes of coordinates* and designated the x-, y-, and z-axes respectively. We shall always depict vectors $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ along the corresponding coordinate axes (Fig. 29).

Definition. The *coordinates of a point M in space* relative to the Cartesian coordinate system $\{O, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ are the coordinates of its radius vector $\overrightarrow{OM}$ in this system.

In other words, the coordinates of the point M are three numbers x, y, z such that

$$\overrightarrow{OM} = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3$$

Just as in the plane, x and y are the abscissa and ordinate of the point M respectively, the third coordinate z is the appli-

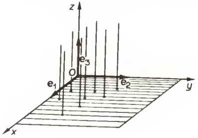


Figure 30

cate of M. In compact notation, we speak of "the point $M(x, y, z)$".

A Cartesian coordinate system in space yields a one-to-one correspondence between point M and ordered triples (x, y, z) of real numbers.

The coordinates of the origin O are 0, 0, 0. The coordinates of the points lying on the x-, y-, and z-axes are $(x, 0, 0)$, $(0, y, 0)$, and $(0, 0, z)$ respectively, and the coordinates of the points belonging to the xOy, xOz, and yOz planes are $(x, y, 0)$, $(x, 0, z)$, and $(0, y, z)$ respectively. The coordinate planes partition space into eight compartments called *octants.* In the first octant, $x > 0$, $y > 0$, $z > 0$, and it is shown in Fig. 30 by the straight lines which are parallel to the z-axis.

Definition. The *coordinates of a vector* $\mathbf{a}$ *in space* relative to the Cartesian coordinate system $\{O, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ are the coordinates of this vector in the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$.

As in the plane, we can find the coordinates of a vector from the coordinates of its origin and terminus: if $A(x, y, z)$ and $B(x', y', z')$, then the coordinates X, Y, Z of the vector $\overrightarrow{AB}$ are $X = x' - x$, $Y = y' - y$, $Z = z' - z$ respectively.

Finally, the condition that the three points $A(x, y, z)$, $B(x', y', z')$, $C(x'', y'', z'')$ lie on one straight line in space is analogous to condition (2) derived for the plane: the *numbers* $x'' - x$, $y'' - y$, $z'' - z$ *are proportional to* $x' - x$, $y' - y$, $z' - z$.

5. The coordinates of the point which divides a line segment in a given ratio.

Given two points in space, find the point C which divides the line segment AB in the ratio $\alpha:\beta$. In analytic geometry, to find a point always means to find its coordinates.

We use the following formula

$$\vec{OC} = \frac{\beta}{\alpha + \beta}\vec{OA} + \frac{\alpha}{\alpha + \beta}\vec{OB}$$

which was derived in Sec. 1.1., Item 7. We label the coordinates of point C with an asterisk and get

$$x^* = \frac{\beta}{\alpha + \beta}x + \frac{\alpha}{\alpha + \beta}x'$$

$$y^* = \frac{\beta}{\alpha + \beta}y + \frac{\alpha}{\alpha + \beta}y' \qquad (3)$$

$$z^* = \frac{\beta}{\alpha + \beta}z + \frac{\alpha}{\alpha + \beta}z'$$

which is the solution of the problem.

The same problem can be solved in the plane; then we have $A(x, y)$ and $B(x', y')$, and the solution is the first two formulas in (3).

In a special case when $\alpha = \beta$, formulas (3) assume the following form

$$x^* = \frac{1}{2}(x + x'), \quad y^* = \frac{1}{2}(y + y'), \quad z^* = \frac{1}{2}(z + z')$$

that is, the *coordinates of the midpoint of a line segment are equal to half-sums of the coordinates of the endpoints.*

Example 1. The line segment AB with endpoints $A(7, 1)$ and $B(4, -5)$ is divided into three equal parts. Find the coordinates of the points of division.

○ Suppose P is the division point closest to A. Then $\alpha:\beta = 1:2$ and the coordinates of P are

$$x^* = \frac{2}{3}\cdot 7 + \frac{1}{3}\cdot 4 = 6, \quad y^* = \frac{2}{3}\cdot 1 + \frac{1}{3}\cdot(-5) = -1$$

For the second point of division Q we have $\alpha:\beta = 2:1$, consequently,

$$x^* = \frac{1}{3}\cdot 7 + \frac{2}{3}\cdot 4 = 5, \quad y^* = \frac{1}{3}\cdot 1 + \frac{2}{3}\cdot(-5) = -3$$

Thus, the division points are $P(6, -1)$ and $Q(5, -3)$. ●

Example 2. Find the point which is symmetric to the point $A(-2, 0, 7)$ relative to the point $B(5, -1, 2)$.

○ If C is the required point, then B is the midpoint of the line segment AC. Therefore, denoting the coordinates of C by x, y, z, we get

$$5 = (-2 + x)/2, \quad -1 = (0 + y)/2, \quad 2 = (7 + z)/2$$

whence $x = 12$, $y = -2$, $z = -3$. ●

Exercises to Chapter 1

1.1.

1. Find three arbitrary vectors a, b, and c in the plane and construct (a) $3\mathbf{a}$, (b) $-\frac{1}{2}\mathbf{b}$, (c) $2\mathbf{a} + 3\mathbf{b}$, (d) $\frac{1}{2}\mathbf{a} - 3\mathbf{b}$, (e) $\mathbf{a} + 2\mathbf{b} - \frac{1}{3}\mathbf{c}$, (f) $-\frac{1}{2}\mathbf{a} - \mathbf{b} + 2\mathbf{c}$.

2. We have a parallelogram $ABCD$ with $\vec{AB} = \mathbf{a}$ and $\vec{AD} = \mathbf{b}$. Express the vectors $\vec{PA}$, $\vec{PB}$, $\vec{PC}$, and $\vec{PD}$ in terms of $\mathbf{a}$ and $\mathbf{b}$, where P is the point where the diagonals intersect.

3. Two vectors $\mathbf{a}$ and $\mathbf{b}$ having the same length are laid off from a common origin. Prove that the vector $\mathbf{a} + \mathbf{b}$ laid off from the same origin is directed along the bisector of the angle $\widehat{\mathbf{a}, \mathbf{b}}$.

4. Two vectors $\mathbf{a}$ and $\mathbf{b}$ are laid off from a common origin. Prove that the vector $\frac{1}{|\mathbf{a}|}\mathbf{a} + \frac{1}{|\mathbf{b}|}\mathbf{b}$, laid off from the same origin, is directed along the bisector of the angle $\widehat{\mathbf{a}, \mathbf{b}}$.

5. Prove that for any two vectors $\mathbf{a}$ and $\mathbf{b}$ the inequality $|\mathbf{a} + \mathbf{b}| \leqslant |\mathbf{a}| + |\mathbf{b}|$ is true. Find the condition when the equality sign is valid.

6. Find the condition two vectors $\mathbf{a}$ and $\mathbf{b}$ should satisfy so that (a) $|\mathbf{a} + \mathbf{b}| = |\mathbf{a} - \mathbf{b}|$, (b) $|\mathbf{a} + \mathbf{b}| > |\mathbf{a} - \mathbf{b}|$, (c) $|\mathbf{a} + \mathbf{b}| < |\mathbf{a} - \mathbf{b}|$.

7. Given a triangle ABC and a point O in the plane, a point P is constructed such that $\vec{OP} = \frac{1}{3}(\vec{OA} + \vec{OB} + \vec{OC})$, prove that the sum $(\vec{PA} + \vec{PB} + \vec{PC})$ is zero. Try to generalize the problem to an n-gon.

8. Prove that there are no two distinct points P and Q such that $\vec{PA} + \vec{PB} + \vec{PC} = 0$ and $\vec{QA} + \vec{QB} + \vec{QC} = 0$.

9. The side BC of a triangle ABC is divided by points P and Q into three equal parts. Denoting $\overrightarrow{AB} = \mathbf{c}$ and $\overrightarrow{AC} = \mathbf{b}$, express the vectors $\overrightarrow{AP}$ and $\overrightarrow{AQ}$ in terms of $\mathbf{b}$ and $\mathbf{c}$.

10. We mark a point D on the side AB of a triangle ABC so that CD is the bisector of $\angle C$. Express the vectors $\overrightarrow{AD}$, $\overrightarrow{BD}$, and $\overrightarrow{CD}$ in terms of $\overrightarrow{CB} = \mathbf{a}$ and $\overrightarrow{CA} = \mathbf{b}$.

11. Prove that the line segments joining the midpoints of opposite edges of a tetrahedron intersect, and the point of intersection is the midpoint of these line segments.

12. Prove that a line segment joining the midpoints of the diagonals of a trapezoid is parallel to its bases.

13. In a parallelogram $ABCD$ the vertex B is connected to the midpoint Q of the side AD. Find the ratio in which the line segment BQ divides the diagonal AC.

1.2.

1. In a regular hexagon $A_1A_2A_3A_4A_5A_6$ we have $\overrightarrow{A_1A_2} = \mathbf{a}$ and $\overrightarrow{A_1A_6} = \mathbf{b}$. Express the vectors $\overrightarrow{A_1A_3}$, $\overrightarrow{A_1A_4}$, $\overrightarrow{A_1A_5}$ in the basis $\mathbf{a}$, $\mathbf{b}$.

2. Find the values α and β such that the following vectors are collinear: (a) $\langle -2.3, \beta\rangle$, $\langle \alpha, -6.2\rangle$, (b) $\langle \alpha, -2\rangle$, $\langle \beta, 2\beta\rangle$.

3. Given two vectors, $\mathbf{a} = \langle 2, 3\rangle$ and $\mathbf{b} = \langle 1, 2\rangle$ in the plane, check that they are noncollinear and represent the vector $\mathbf{c} = \langle 4, 9\rangle$ in the basis $\mathbf{a}$, $\mathbf{b}$.

1.3.

1. Find the coordinates of the point which is symmetric to a point $A(a, b)$ with respect to the origin of a Cartesian coordinate system.

2. Given four points, $A(1, 3)$, $B(4, 7)$, $C(2, 8)$, $D(-1, 4)$ in the plane, verify that the quadrilateral $ABCD$ is a parallelogram.

3. Given three points, $A(-1, 2)$, $B(0, 3)$, $C(4, -5)$, find a fourth point D such that the quadrilateral is a parallelogram.

4. Show that the four points $A(3, -1, 2)$, $B(1, 2, -1)$, $C(-1, 1, -3)$, $D(3, -5, 3)$ are the vertices of a trapezoid.

5. Check whether the following points lie on one straight line: (a) $A(3, 1)$, $B(-2, -9)$, $C(8, 11)$, (b) $A(0, 2)$, $B(-1, 5)$, $C(3, 4)$, (c) $A(1, -5, 3)$, $B(5, -1, 7)$, $C(6, 0, 8)$.

6. Find the coordinates of the midpoint of the line segment AB with endpoints $A(2, 3)$ and $B(-4, 7)$.

7. Find the point which is symmetric to the point $A(-3, 0)$ relative to the point $B(2, 9)$.

8. The line segment AB with endpoints $A(1, -3, -5)$ and $B(7, 3, 4)$ is divided into three equal parts. Find the division points.

9. A line segment AB is divided by the points $P(0, 5, 2)$ and $Q(1, 7, 2)$ into three equal parts. Find its endpoints.

10. Find the point at which the medians of a triangle are concurrent (its centroid). The vertices of the triangle are at the points $A(1, 4)$, $B(-5, 0)$, $C(-2, 1)$.

Chapter 2

RECTANGULAR CARTESIAN COORDINATES. SIMPLE PROBLEMS IN ANALYTIC GEOMETRY

2.1. PROJECTION OF A VECTOR ON AN AXIS

Suppose that a straight line is specified in space. We choose the positive direction on it.

Definition. An *axis* is a straight line having the positive direction and a unit of measurement.

The positive direction and the unit of length are usually specified simultaneously by a vector **e** having length 1 which is parallel to that straight line. This vector is called a *unit vector* or *basis vector* (Fig. 31).

We denote the straight line by l and drop the perpendicular to it from an arbitrary point A. The point A', which is the foot of the perpendicular, is the *orthogonal projection* (or simply projection) of A on the axis l.

Definition. Let l be an arbitrary axis and $\overrightarrow{AB}$ an arbitrary vector in space. The vector $\overrightarrow{A'B'}$ whose initial and terminal points are the projections of the points A and B on the l-axis is called the *vector projection* of $\overrightarrow{AB}$ on the l axis (Fig. 32).

Along with the projection of a vector there is also the *scalar projection* of a vector on an axis.

Definition. The *scalar projection* of a vector $\overrightarrow{AB}$ on the l

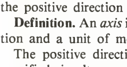


Figure 31

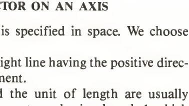


Figure 32

axis is a number λ such that

$$\overrightarrow{A'B'} = \lambda \mathbf{e}$$

We denote a scalar projection by $\text{proj}_l \overrightarrow{AB}$, or sometimes $\text{proj}\ \mathbf{a}\overrightarrow{AB}$, where $\mathbf{a}$ is a nonzero vector directed along the positive l axis. The absolute value of the number $\text{proj}_l \overrightarrow{AB}$ is equal to the length of $\overrightarrow{A'B'}$, where the sign indicates whether the vector has the same or opposite direction to the unit vector $\mathbf{e}$ of the l axis.

Below we shall mostly use scalar projections, and so when we say "projection" we shall always mean a scalar projection, unless otherwise specified.

1. Angle between vectors. Angle between a vector and an axis.

Suppose that a and b are two arbitrary vectors. We place their initial points at the same point and draw a plane through them.

The *angle* φ *between two vectors* $\mathbf{a}$ and $\mathbf{b}$ is the angle of the shortest rotation from $\mathbf{a}$ to $\mathbf{b}$ in the plane of the vectors (Fig. 33). Obviously, $0 \leqslant \varphi \leqslant \pi$. If at least one of the vectors $\mathbf{a}$, $\mathbf{b}$ is zero, the angle between them is undefined.

The angle between a vector $\mathbf{a}$ and the l axis is the angle φ between $\mathbf{a}$ and the unit vector $\mathbf{e}$ of the l axis (Fig. 34).

We denote the angle φ by $\widehat{(\mathbf{a}, \mathbf{b})}$ or $\widehat{(\mathbf{a}, l)}$.

2. Theorem on the projection of a vector on an axis.

Theorem. *The projection of a vector* $\mathbf{a}$ *on the* l *axis is equal to the product of the length of* $\mathbf{a}$ *and the cosine of the angle between* $\mathbf{a}$ *and the* l *axis*:

$$\text{proj}_l \mathbf{a} = |\mathbf{a}| \cos \widehat{(\mathbf{a}, l)} \qquad (1)$$

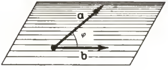


Figure 33

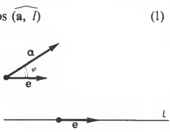


Figure 34

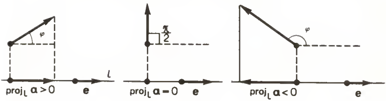


Figure 35

Formula (1) follows directly from the definition of cosine. Figure 35 shows some possible cases.

3. Properties of projections.

We shall derive the three most important properties of projections of vectors.

1^0. *Equal vectors have equal projections on the same axis.*

Indeed, if $\mathbf{a} = \mathbf{b}$, then $|\mathbf{a}| = |\mathbf{b}|$ and $\widehat{(\mathbf{a}, l)} = \widehat{(\mathbf{b}, l)}$, whence by formula (1) it follows that $\text{proj}_l \mathbf{a} = \text{proj}_l \mathbf{b}$.

2^0. *The projection of the sum of several vectors on an axis is equal to the sum of their projections on that axis*:

$$\text{proj}_l(\mathbf{a} + \mathbf{b} + \ldots) = \text{proj}_l \mathbf{a} + \text{proj}_l \mathbf{b} + \ldots$$

To ascertain the validity of this property, it is sufficient to verify it for two summands.

We construct a broken line OAB containing the directed line segments $\overrightarrow{OA} = \mathbf{a}$ and $\overrightarrow{AB} = \mathbf{b}$ (Fig. 36). We project the points O, A, B onto the l axis and get points O', A', B'.

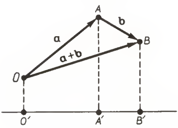


Figure 36

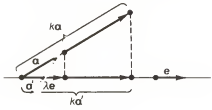


Figure 37

Suppose that λ is the projection of the vector **a** and μ is the projection of **b** on the l axis. Then $\overrightarrow{O'A'} = \lambda\mathbf{e}$, $\overrightarrow{A'B'} = \mu\mathbf{e}$, where **e** is the unit vector of the l axis. Whence $\overrightarrow{O'B'} = (\lambda + \mu)\mathbf{e}$, and this means that $\lambda + \mu$ is the projection of $\overrightarrow{OB}$ (i.e. of the vector **a** + **b**) on the l axis. Thus,

$$\text{proj}_l(\mathbf{a} + \mathbf{b}) = \text{proj}_l\,\mathbf{a} + \text{proj}_l\,\mathbf{b}$$

3^0. *When multiplying a vector into a number, we multiply its projection by that number*:

$$\text{proj}_l(k\mathbf{a}) = k\ \text{proj}_l\,\mathbf{a}$$

□ We translate the initial point of **a** to some point on the l axis. Let $\mathbf{a}'$ be the vector projection of **a** on the l axis (Fig. 37). We have $\mathbf{a}' = \lambda\mathbf{e}$, where λ is the projection of **a** on the l axis.

First we consider the case where $k > 0$. "Extension" of **a** k times obviously causes "extension" of $\mathbf{a}'$ k times, that is, it becomes a vector $k\mathbf{a}'$. But $k\mathbf{a}' = k(\lambda\mathbf{e}) = (k\lambda)\,\mathbf{e}$. Whence it follows that $\text{proj}_l(k\mathbf{a}) = k\lambda = k\cdot\text{proj}_l\mathbf{a}$.

We suggest the reader consider the case where $k < 0$ independently. When $k = 0$ property 3^0 is obvious. ■

2.2. RECTANGULAR CARTESIAN COORDINATE SYSTEM

A *rectangular* Cartesian coordinate system is the simplest of Cartesian coordinate systems.

Definition. The Cartesian coordinate system $\{O, \mathbf{e}_1, \mathbf{e}_2\}$ in the plane is called *rectangular* if $\mathbf{e}_1$ and $\mathbf{e}_2$ are perpendicular unit vectors.

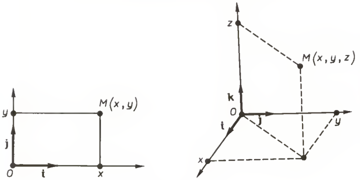


Figure 38

Similarly, we can define the rectangular Cartesian coordinate system $\{O, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ in space; here $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ should also be mutually perpendicular unit vectors.

The basis unit vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ of a rectangular Cartesian coordinate system in the plane are usually denoted by **i**, **j**, and the basis unit vectors $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ in a rectangular Cartesian coordinate system in space are usually denoted by **i**, **j**, **k**. Accordingly, an arbitrary radius vector $\overrightarrow{OM}$ can be represented in the basis as

$$\overrightarrow{OM} = x\mathbf{i} + y\mathbf{j} \quad \text{(in the plane)}$$
$$\overrightarrow{OM} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \quad \text{(in space)}$$

The coordinates of the point M are x, y, in the first case, and x, y, z in the second (Fig. 38).

In what follows we shall say a *rectangular coordinate system* rather than a rectangular Cartesian coordinate system, and coordinates x, y in the plane and x, y, z in space we shall call *rectangular coordinates.*

When considering rectangular coordinate systems we should note that there are *right-handed* and *left-handed* rectangular coordinate systems.

In a left-handed rectangular coordinate system, the 90° rotation from the vector **i** to the vector **j** appears clockwise (Fig. 39),

Figure 39

while in a right-handed coordinate system this rotation is counterclockwise. These two coordinate systems are different and no displacements in the plane can make the left-handed pair **i**, **j** coincide with the right-handed one.

In a left-handed rectangular coordinate system in space, when three vectors **i**, **j**, **k** originate from a common point, the 90° rotation from **i** to **j** when viewed from the terminus of **k** appears clockwise, while in a right-handed coordinate system such a rotation appears counterclockwise (Fig. 40). The two systems differ in the same way as left-hand and right-hand screws do.

In this book we shall *only* use *right-handed* coordinate systems.

1. Coordinates of a vector in a rectangular system as its projections on coordinate axes.

Here are some remarks. Suppose that in a rectangular system in the plane we have a point A with coordinates x and y. Then $A_1(x, 0)$ and $A_2(0, y)$ are its projections on the coordinate axes, which follows directly from our reasoning (see Sec. 1.2., Item 1). Similarly, if in a rectangular system in space the coor-

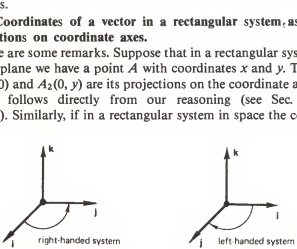


Figure 40

dinates of a point A are x, y, z, then $A_1(x, 0, 0)$, $A_2(0, y, 0)$, $A_3(0, 0, z)$ are its projections on the coordinate axes (see Sec. 1.2., Item 3).

Theorem. *The coordinates of a vector $\overrightarrow{AB}$ in a rectangular coordinate system coincide with its projections on the coordinate axes.*

□ Suppose we are given $A(x, y, z)$ and $B(x', y', z')$. We have proved that the coordinates of $\overrightarrow{AB}$ in the given coordinate system are $x' - x$, $y' - y$, $z' - z$.

Projecting the points A and B onto the coordinate axes yields the points A_1, A_2, A_3 and B_1, B_2, B_3. We have

$$\overrightarrow{A_1B_1} = \overrightarrow{OB_1} - \overrightarrow{OA_1} = x'\mathbf{i} - x\mathbf{i} = (x' - x)\,\mathbf{i}$$

and, similarly, $\overrightarrow{A_2B_2} = (y' - y)\,\mathbf{j}$, $\overrightarrow{A_3B_3} = (z' - z)\,\mathbf{k}$. These equalities imply that

$$\text{proj}_x\overrightarrow{AB} = x' - x,\ \text{proj}_y\overrightarrow{AB} = y' - y,\ \text{proj}_z\overrightarrow{AB} = z' - z. \ \blacksquare$$

2. Length of a line segment in coordinates.

Suppose that we have a coordinate axis with origin O and basis unit vector $\mathbf{e}$. We mark two points, $A(x_1)$ and $B(x_2)$ (Fig. 41). Then $\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = (x_2 - x_1)\,\mathbf{e}$. Hence it follows that the distance between A and B is

$$|AB| = |x_2 - x_1|$$

Let us now choose two points in a rectangular coordinate system, $A(x_1, y_1)$ and $B(x_2, y_2)$. We first assume that the line segment AB is not parallel to either coordinate axis. Through A and B we draw straight lines parallel to the coordinate axes (Fig. 42). By the Pythagorean theorem we have $|AB|^2 =$

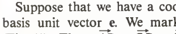


Figure 41

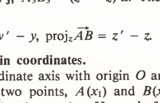


Figure 42

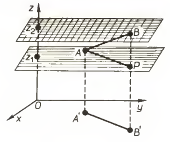


Figure 43

$|AP|^2 + |BP|^2$ from the right triangle ABP. But $|AP| = |\text{proj}_x\overrightarrow{AB}| = |x_2 - x_1|$, $|BP| = |\text{proj}_y\overrightarrow{AB}| = |y_2 - y_1|$, whence it follows that

$$|AB|^2 = |x_2 - x_1|^2 + |y_2 - y_1|^2$$

Thus,

$$|AB| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \qquad (1)$$

Formula (1) is also valid when the line segment AB is parallel to either the x-axis or y-axis.

Formula (1) *expresses the length of a line segment in the plane via the coordinates of its endpoints.*

Finally, we consider points A and B in space. We have $A(x_1, y_1, z_1)$, and $B(x_2, y_2, z_2)$. Assume first that the line segment AB is not parallel to the xy coordinate plane. We project A and B onto that plane and get the points $A'(x_1, y_1, 0)$ and $B'(x_2, y_2, 0)$. According to what we have proved, $|A'B'|^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2$.

Through A and B we draw planes parallel to the z-axis (Fig. 43). By the Pythagorean theorem, from the triangle ABP we have

$$|AB|^2 = |AP|^2 + |PB|^2 = |A'B'|^2 + |PB|^2$$

But $|PB| = |z_2 - z_1|$, since $|PB|$ is the absolute value of the projection of $\overrightarrow{AB}$ on the z-axis. Whence

$$|AB|^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$$

Thus,

$$|AB| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \qquad (2)$$

The last formula is also valid when the line segment AB is parallel to the xy plane (then $|AB| = |A'B'|$, and $z_2 - z_1 = 0$).

Formula (2) *expresses the length of a line segment in space via the coordinates of its endpoints.*

Example 1. Find the distance between the points $A(-1, 1, 5)$ and $B(1, 3, 4)$.

○ From formula (2) we have

$$|AB| = \sqrt{(1 - (-1))^2 + (1 - 3)^2 + (5 - 4)^2} = 3. \; ●$$

Example 2. Find the point in the plane equidistant from the three points, $A(11, 3)$, $B(10, 6)$, $C(-1, 9)$. In other words, find the circumcentre of the triangle ABC.

○ Suppose that $P(x, y)$ is the required point. We write the conditions $|PA| = |PB|$ and $|PA| = |PC|$ as follows:

$$\sqrt{(x - 11)^2 + (y - 3)^2} = \sqrt{(x - 10)^2 + (y - 6)^2}$$
$$\sqrt{(x - 11)^2 + (y - 3)^2} = \sqrt{(x + 1)^2 + (y - 9)^2}$$

or, after squaring and simplifying, as $x - 3y = -3$, $2x - y = 4$. Solving this system of two equations in two unknowns we find that $x = 3$, $y = 2$. The required point is $P(3, 2)$. ●

3. Length of a vector in coordinates.

In Sec. 1.3 we derived formulas for the coordinates of a vector $\overrightarrow{AB}$ in terms of the coordinates of its initial and terminal points:

$$X = x_2 - x_1, \quad Y = y_2 - y_1$$

when $A(x_1, y_1)$ and $B(x_2, y_2)$ are two points in the plane, and

$$X = x_2 - x_1, \quad Y = y_2 - y_1, \quad Z = z_2 - z_1$$

when $A(x_1, y_1, z_1)$ and $B(x_2, y_2, z_2)$ are two points in space.

Using formulas (1) and (2) we can now express the length of an arbitrary vector $\overrightarrow{AB} = \mathbf{a}$ in terms of its coordinates in a rec-

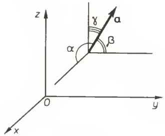


Figure 44

tangular coordinate system:

$$|\mathbf{a}| = \sqrt{X^2 + Y^2}$$

if $\mathbf{a} = \langle X, Y \rangle$ is a vector in the plane, and

$$|\mathbf{a}| = \sqrt{X^2 + Y^2 + Z^2}$$

if $\mathbf{a} = \langle X, Y, Z \rangle$ is the vector in space.

Example. The length of the vector $\mathbf{a} = \langle -3, 4 \rangle$ in the plane is $\sqrt{(-3)^2 + 4^2} = 5$. The length of the vector $\mathbf{a} = \langle 1, -2, 2 \rangle$ in space is $\sqrt{1^2 + (-2)^2 + 2^2} = 3$.

4. Direction cosines.

Suppose that **a** is a nonzero vector in space. We denote by α, β, γ the angles between the vector and the x-, y-, and z-axes of the rectangular coordinate system (Fig. 44). Since the projections of a vector on the coordinate axes coincide (according to what we have proved) with respective coordinates X, Y, Z of this vector, we can write

$$X = |\mathbf{a}| \cos \alpha, \quad Y = |\mathbf{a}| \cos \beta, \quad Z = |\mathbf{a}| \cos \gamma \tag{3}$$

Thus, each coordinate of the vector is equal to the product of its length and the cosine of the angle between the vector and the corresponding coordinate axis.

Definition. The numbers $\cos \alpha$, $\cos \beta$, $\cos \gamma$ are called the *direction cosines* of the vector **a**.

The direction cosines are related as follows:

$$\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1 \tag{4}$$

Indeed, $|\mathbf{a}|^2 = X^2 + Y^2 + Z^2$, whence according to (3)

$$|\mathbf{a}|^2 = |\mathbf{a}|^2(\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma)$$

canceling out $|\mathbf{a}|^2$ we arrive at equality (4).

Example 1. Find the direction cosines of the vector $\overrightarrow{AB}$ if $A(1, -1, 3)$ and $B(2, 1, 1)$.

○ The coordinates of $\overrightarrow{AB}$ are $X = 2 - 1 = 1$, $Y = 1 - (-1) = 2$, $Z = 1 - 3 = -2$. Whence, $|\overrightarrow{AB}|^2 = \sqrt{1^2 + 2^2 + (-2)^2} = 3$. We now find the cosine vectors from formulas (3):

$$\cos \alpha = \frac{X}{|\mathbf{a}|} = \frac{1}{3}, \quad \cos \beta = \frac{Y}{|\mathbf{a}|} = \frac{2}{3}, \quad \cos \gamma = \frac{Z}{|\mathbf{a}|} = -\frac{2}{3}. \ ●$$

Example 2. The vector **a** makes angles of 60° with the *x*- and *y*-axes. Find the angle between **a** and the *z*-axis.

○ Denoting the required angle by γ we can write $\cos^2 60° + \cos^2 60° + \cos^2 \gamma = 1$, whence it follows that $\cos \gamma = \pm 1/\sqrt{2}$. We have two solutions: $\gamma = 45°$ and $\gamma = 135°$. Obviously, they correspond to two symmetric vectors relative to the *xy* coordinate plane. ●

Example 3. A vector **a** makes equal acute angles with the coordinate axes. Find the angles.

○ In this case $\alpha = \beta = \gamma$, and therefore formula (4) gives $3 \cos^2 \alpha = 1$. Whence we have $\cos \alpha = \pm 1/\sqrt{3}$. Since, by hypothesis, the angle α is acute, we take the positive value: $\cos \alpha = 1/\sqrt{3}$. From the table of cosines we find that $\alpha \approx 55°$. ●

We note in conclusion that if l is an axis in space and **e** is a basis unit vector on the axis, then the direction cosines of **e** are called the *direction cosines of the axis itself.*

2.3. SCALAR PRODUCT OF VECTORS

1. Definition of a scalar product.

Consider the following problem. A constant force **F** exerted along the line of motion moves a particle of mass 1 from point P to point Q. Find the work done. We know from the physics course that work is the product of the force and the distance moved and the cosine of the angle between the directions of the force and motion. In vector notation we can write

$$A = |\mathbf{F}||\overrightarrow{PQ}| \cos (\widehat{\mathbf{F}, \overrightarrow{PQ}})$$

where A is the required work.

There is a single operation which takes two vectors (or three numbers, viz. the lengths of the vectors and the cosine of the angle between them) to produce a scalar A. This operation is called the *scalar multiplication of two vectors.* We shall see below that this operation is widely used.

Definition. The *scalar* (or *dot*) *product* of two nonzero vectors $\mathbf{a}$ and $\mathbf{b}$ is the number

$$|\mathbf{a}||\mathbf{b}| \cos (\widehat{\mathbf{a}, \mathbf{b}})$$

If one of the vectors $\mathbf{a}$ or $\mathbf{b}$ is zero, their scalar product is zero.

The scalar product of vectors $\mathbf{a}$ and $\mathbf{b}$ is denoted as $\mathbf{a}\cdot\mathbf{b}$. Thus,

$$\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \varphi \tag{1}$$

where φ is the angle between $\mathbf{a}$ and $\mathbf{b}$. The cases where $\mathbf{a} = 0$ or $\mathbf{b} = 0$ can also be described by formula (1) since then $|\mathbf{a}| = 0$ or $|\mathbf{b}| = 0$ (the angle φ is undefined).

The scalar product is closely related to the projections of vectors. If $\mathrm{proj}_{\mathbf{a}}\mathbf{b}$ is the projection of vector $\mathbf{b}$ on the axis whose direction is the same as that of $\mathbf{a}$, then, according to the theorem on projections, we have $\mathrm{proj}_{\mathbf{a}}\mathbf{b} = |\mathbf{b}| \cos \varphi$, $\mathrm{proj}_{\mathbf{b}}\mathbf{a} = |\mathbf{a}| \cos \varphi$, which together with formula (1) imply that

$$\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}|\ \mathrm{proj}_{\mathbf{a}}\mathbf{b}, \quad \mathbf{a}\cdot\mathbf{b} = |\mathbf{b}|\ \mathrm{proj}_{\mathbf{b}}\mathbf{a} \tag{2}$$

that is, *the scalar product of two vectors is equal to the product of the length of one of them multiplied by the projection of the other on the direction of the first vector.*

Note too that *the scalar product of a vector with itself is the square of its length*

$$\mathbf{a}\cdot\mathbf{a} = |\mathbf{a}||\mathbf{a}| \cos 0 = |\mathbf{a}|^2$$

since the vector $\mathbf{a}$ forms the angle $\varphi = 0$ with itself.

2. Properties of the scalar product.

1^0. *Commutativity*:

$$\mathbf{a}\cdot\mathbf{b} = \mathbf{b}\cdot\mathbf{a}$$

This property follows directly from the definition of the scalar product.

2^0. *Associativity relative to the multiplication of a vector by a scalar*:

$$(k\mathbf{a})\cdot\mathbf{b} = k(\mathbf{a}\cdot\mathbf{b})$$

The property follows from the properties of projections since

$$(k\mathbf{a})\cdot\mathbf{b} = |\mathbf{b}|\,\mathrm{proj}_{\mathbf{b}}(k\mathbf{a}) = |\mathbf{b}|\,k\,\mathrm{proj}_{\mathbf{b}}\mathbf{a} = k|\mathbf{b}|\,\mathrm{proj}_{\mathbf{b}}\mathbf{a} = k(\mathbf{a}\cdot\mathbf{b})$$

3^0. *Distributivity over vector addition*:

$$(\mathbf{a} + \mathbf{b})\cdot\mathbf{c} = \mathbf{a}\cdot\mathbf{c} + \mathbf{b}\cdot\mathbf{c}$$

We can prove this property using the properties of projections

$$\begin{aligned}(\mathbf{a} + \mathbf{b})\cdot\mathbf{c} &= |\mathbf{c}|\,\mathrm{proj}_{\mathbf{c}}(\mathbf{a} + \mathbf{b}) = |\mathbf{c}|(\mathrm{proj}_{\mathbf{c}}\mathbf{a} + \mathrm{proj}_{\mathbf{c}}\mathbf{b})\\ &= |\mathbf{c}|\,\mathrm{proj}_{\mathbf{c}}\mathbf{a} + |\mathbf{c}|\,\mathrm{proj}_{\mathbf{c}}\mathbf{b} = \mathbf{a}\cdot\mathbf{c} + \mathbf{b}\cdot\mathbf{c}\end{aligned}$$

4^0. *The property of sign.* If $\mathbf{a}$ and $\mathbf{b}$ are two nonzero vectors and φ is the angle between them, then

$\mathbf{a}\cdot\mathbf{b} > 0$ if the angle φ is acute,
$\mathbf{a}\cdot\mathbf{b} < 0$ if the angle φ is obtuse,
$\mathbf{a}\cdot\mathbf{b} = 0$ if the angle φ is right.

Indeed, since $|\mathbf{a}| > 0$ and $|\mathbf{b}| > 0$, we have

the inequality $\mathbf{a}\cdot\mathbf{b} > 0$ is equivalent to $\cos\varphi > 0$,
the inequality $\mathbf{a}\cdot\mathbf{b} < 0$ is equivalent to $\cos\varphi < 0$,
the equality $\mathbf{a}\cdot\mathbf{b} = 0$ is equivalent to $\cos\varphi = 0$.

5^0. *The vectors* $\mathbf{a}$ *and* $\mathbf{b}$ *are perpendicular if and only if their scalar product is zero.*

Indeed, if $\mathbf{a} \perp \mathbf{b}$, then $\varphi = \pi/2$, whence it follows that $\cos\varphi = 0$ and therefore $\mathbf{a}\cdot\mathbf{b} = 0$. Conversely, if $\mathbf{a}\cdot\mathbf{b} = 0$, then either $|\mathbf{a}| = 0$ or $|\mathbf{b}| = 0$ or $\cos\varphi = 0$. In the first and second cases $\mathbf{a} \perp \mathbf{b}$ since the zero vector is perpendicular to any vector, in the third case $\varphi = \pi/2$ and $\mathbf{a} \perp \mathbf{b}$.

Here is an example of how to apply the properties of the scalar product.

Example. *PQ* is a principal diagonal of a parallelepiped (Fig. 45), *a*, *b*, and *c* are the edges emanating from the vertex *P*, and the angles between these edges are α (between *b* and *c*), β (between *a* and *c*), γ (between *a* and *b*). Find the length of the diagonal *PQ*.

○ We designate the vectors (edges) emanating from the vertex *P* as **a**, **b**, and **c** ($|\mathbf{a}| = a$, $|\mathbf{b}| = b$, $|\mathbf{c}| = c$). Obviously, $PQ = a + b + c$. Whence it follows

$$\vec{PQ}^2 = (\mathbf{a} + \mathbf{b} + \mathbf{c})\cdot(\mathbf{a} + \mathbf{b} + \mathbf{c}) = \mathbf{a}\cdot\mathbf{a} + \mathbf{a}\cdot\mathbf{b} + \mathbf{a}\cdot\mathbf{c} + \mathbf{b}\cdot\mathbf{a} + \mathbf{b}\cdot\mathbf{b} + \mathbf{b}\cdot\mathbf{c} + \mathbf{c}\cdot\mathbf{a} + \mathbf{c}\cdot\mathbf{b} + \mathbf{c}\cdot\mathbf{c}$$

Since $\mathbf{a}\cdot\mathbf{a} = a^2$, $\mathbf{b}\cdot\mathbf{b} = b^2$, $\mathbf{c}\cdot\mathbf{c} = c^2$, $\mathbf{a}\cdot\mathbf{b} = \mathbf{b}\cdot\mathbf{a} = ab\cos\gamma$, $\mathbf{a}\cdot\mathbf{c} = \mathbf{c}\cdot\mathbf{a} = ac\cos\beta$, $\mathbf{b}\cdot\mathbf{c} = \mathbf{c}\cdot\mathbf{b} = bc\cos\alpha$, finally we get

$$|\vec{PQ}|^2 = a^2 + b^2 + c^2 + 2ab\cos\gamma + 2ac\cos\beta + 2bc\cos\alpha. \quad ●$$

3. Scalar product and the test for a straight line to be perpendicular to a plane.

We can use the properties of the scalar product for moving some geometric theorems.

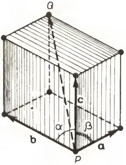


Figure 45

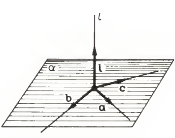


Figure 46

Theorem (the test for a straight line to be perpendicular to a plane). *If a straight line l in space is perpendicular to any two intersecting straight lines of a plane α, then l is perpendicular to any straight line in that plane.*

□ We denote by a and b two intersecting straight lines of the plane α which are perpendicular to a straight line l. Let c be a straight line belonging to the plane and passing through the point of intersection of a and b (Fig. 46). We choose nonzero vectors $\mathbf{l}$, $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ on the straight lines l, a, b, and c respectively. By the hypothesis we have $\mathbf{a}\cdot\mathbf{l} = 0$, $\mathbf{b}\cdot\mathbf{l} = 0$.

The vectors $\mathbf{a}$ and $\mathbf{b}$ are noncollinear (the straight lines a and b intersect), and therefore we can represent the vector $\mathbf{c}$ as their linear combination

$$\mathbf{c} = \lambda\mathbf{a} + \mu\mathbf{b}$$

where λ and μ are some numbers. Whence

$$\mathbf{c}\cdot\mathbf{l} = (\lambda\mathbf{a} + \mu\mathbf{b})\cdot\mathbf{l} = \lambda(\mathbf{a}\cdot\mathbf{l}) + \mu(\mathbf{b}\cdot\mathbf{l}) = 0$$

Consequently, the vector $\mathbf{c}$ is perpendicular to the straight line l. Noting that $\mathbf{c}$ is any one of vectors of the plane α, we conclude that any straight line of the plane α is perpendicular to l. ■

4. Representation of the scalar product in terms of the coordinates of vectors.

Suppose that the coordinates of vectors $\mathbf{a}$ and $\mathbf{b}$ are specified in a rectangular coordinate system in the plane: $\mathbf{a} = \langle X_1, Y_1\rangle$, $\mathbf{b} = \langle X_2, Y_2\rangle$. Then the following important formula is valid:

$$\mathbf{a}\cdot\mathbf{b} = X_1\cdot X_2 + Y_1\cdot Y_2$$

that is, *the scalar product is equal to the sum of the products of the corresponding coordinates.*

To derive this formula, we express these vectors in terms of the basis unit vectors:

$$\mathbf{a} = X_1\mathbf{i} + Y_1\mathbf{j}, \quad \mathbf{b} = X_2\mathbf{i} + Y_2\mathbf{j}$$

and then take the scalar product. We have

$$\mathbf{a}\cdot\mathbf{b} = (X_1X_2)\,\mathbf{i}\cdot\mathbf{i} + (X_1Y_2)\,\mathbf{i}\cdot\mathbf{j} + (Y_1X_2)\,\mathbf{j}\cdot\mathbf{i} + (Y_1Y_2)\,\mathbf{j}\cdot\mathbf{j}$$

Since $\mathbf{i}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{j} = 1$ (the length of $\mathbf{i}$ and $\mathbf{j}$ is unity) and $\mathbf{i}\cdot\mathbf{j} = \mathbf{j}\cdot\mathbf{i} = 0$ (**i** is perpendicular to **j**), the last expression reduces to the required form

$$\mathbf{a}\cdot\mathbf{b} = X_1X_2 + Y_1Y_2$$

If **a** and **b** are two vectors in space and $\mathbf{a} = \langle X_1, Y_1, Z_1\rangle$, $\mathbf{b} = \langle X_2, Y_2, Z_2\rangle$, then similar reasoning leads to the following equality:

$$\mathbf{a}\cdot\mathbf{b} = X_1X_2 + Y_1Y_2 + Z_1Z_2$$

which can be interpreted in the same way as in the plane.

In particular, for one vector $\mathbf{a} = \langle X, Y, Z\rangle$ we have

$$\mathbf{a}\cdot\mathbf{a} = |\mathbf{a}|^2 = X^2 + Y^2 + Z^2$$

whence follows the familiar formula

$$|\mathbf{a}| = \sqrt{X^2 + Y^2 + Z^2}$$

5. Application of scalar products.

The following formula is widely used in practice:

$$\mathbf{a}\cdot\mathbf{b} = X_1X_2 + Y_1Y_2 + Z_1Z_2 \qquad (3)$$

(1) *The perpendicularity test for two vectors.* Two vectors **a** and **b** are perpendicular if and only if

$$X_1X_2 + Y_1Y_2 + Z_1Z_2 = 0$$

The proposition follows immediately from formula (3) and property 5^0 of the scalar product.

(2) *Angle between two vectors.* Suppose that **a** and **b** are two nonzero vectors and φ is the angle between them. From the definition of the scalar product it follows that

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

or

$$\cos\varphi = \frac{X_1X_2 + Y_1Y_2 + Z_1Z_2}{\sqrt{X_1^2 + Y_1^2 + Z_1^2}\,\sqrt{X_2^2 + Y_2^2 + Z_2^2}} \qquad (4)$$

(3) *Projection of a vector on an axis.* Suppose that we have an axis l in space with a unit vector **e** which makes angles α, β, γ with the coordinate axes. Then the projection of an arbitrary vectors $\mathbf{a} = \langle X, Y, Z\rangle$ on the l axis is

$$\text{proj}_{\mathbf{e}}\mathbf{a} = X\cos\alpha + Y\cos\beta + Z\cos\gamma \tag{5}$$

Indeed, since the length of **e** is unity, its projections on the coordinate axes are $1\cdot\cos\alpha$, $1\cdot\cos\beta$, $1\cdot\cos\gamma$. Consequently, the coordinates of **e** are also equal to $\cos\alpha$, $\cos\beta$, $\cos\gamma$. Since

$$\text{proj}_{\mathbf{e}}\mathbf{a} = |\mathbf{e}|\,\text{proj}_{\mathbf{e}}\mathbf{a} = \mathbf{a}\cdot\mathbf{e}$$

we have formula (5).

Example 1. Given the points $A(2, 0, 1)$, $B(2, 1, 0)$, $C(1, 0, 0)$. Find angle ABC.

○ We consider the vectors $\overrightarrow{BA} = \langle 2-2, 0-1, 1-0\rangle = \langle 0, -1, 1\rangle$ and $\overrightarrow{BC} = \langle 1-2, 0-1, 0-0\rangle = \langle -1, -1, 0\rangle$. The required angle ABC is the angle φ between these vectors. From the formula for the cosine of the angle between two vectors we have

$$\cos\varphi = \frac{\overrightarrow{BA}\cdot\overrightarrow{BC}}{|\overrightarrow{BA}||\overrightarrow{BC}|} = \frac{0\cdot(-1) + (-1)\cdot(-1) + 1\cdot 0}{\sqrt{0^2 + (-1)^2 + 1^2}\sqrt{(-1)^2 + (-1)^2 + 0^2}}$$

$$= \frac{1}{2}$$

Hence, $\varphi = 60°$. ●

Example 2. Find the projection of the vector $\mathbf{a} = \langle 1, 2, 3\rangle$ on the l axis which forms equal acute angles with the coordinate axes.

○ Example 3 in Sec. 2.2., Item 4 shows that the direction cosines of the l axis are $\cos\alpha = \cos\beta = \cos\gamma = 1/\sqrt{3}$. Consequently,

$$\text{proj}_l\mathbf{a} = 1\cdot\frac{1}{\sqrt{3}} + 2\cdot\frac{1}{\sqrt{3}} + 3\cdot\frac{1}{\sqrt{3}} = 2\sqrt{3} \quad ●$$

Example 3. Given two vectors $\mathbf{a} = \langle 1, 2, -1\rangle$ and $\mathbf{b} = \langle 2, -1, 3\rangle$. Find $\text{proj}_{\mathbf{a}}\mathbf{b}$.

○ Since $\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}|\,\mathrm{proj}_{\mathbf{a}}\mathbf{b}$, we have

$$\mathrm{proj}_{\mathbf{a}}\mathbf{b} = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}|} = \frac{1\cdot 2 + 2\cdot(-1) + (-1)\cdot 3}{\sqrt{1^2 + 2^2 + (-1)^2}}$$

$$= -\frac{3}{\sqrt{6}} = -\sqrt{\frac{3}{2}} \; ●$$

2.4. POLAR COORDINATES

Coordinate systems are used to locate points (in the plane or in space) by specifying their coordinates. Although the rectangular Cartesian coordinate system is most popular, other coordinates, e.g. the polar system, are also widely used. The polar coordinate system in the plane is defined by fixed point O, called the *pole* (Fig. 47), a fixed ray p emanating from O and called the *polar axis*, and a scale for length measurement.

The position of a point M different from the pole is defined in polar coordinates by two numbers, r and φ. The number $r > 0$ is the first coordinate and called the *polar radius*; it is equal to the distance from the pole O to the point M. The number φ is the second coordinate and called the *polar angle*. We choose the positive direction of reckoning the angles; as is usual in trigonometry, the positive direction is counterclockwise. Then φ is the angle through which the ray p is rotated in the positive direction for p to coincide with the ray OM. If we rotate p clockwise, then φ is a negative number. Obviously, the angle φ lies within 0 and 2π; in other words, the polar angle for point M is φ of any of the angles $\varphi + 2\pi n$, where n is an integer. The pair of numbers (r, φ) are called the *polar coordinates* of the point M.

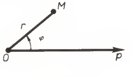


Figure 47

If M is not the pole, then $r > 0$ and φ lies within $2\pi n$. Now, if M coincides with the pole, then $r = 0$ and φ does not matter. Conversely, given the pair of numbers (r, φ) with $r > 0$, we can construct the point M for which r and φ are polar coordinates.

Exercises to Chapter 2

2.1

1. Find the projection of the unit vector $\mathbf{e}$ on the l axis with which $\mathbf{e}$ makes the following angles: (a) 30°, (b) 45°, (c) 120°, (d) 90°.

2. Given $\text{proj}_l \mathbf{a} = \frac{1}{2}|\mathbf{a}|$, find the angle between $\mathbf{a}$ and the l axis.

3. Vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ form a triangle, i.e. $\mathbf{a} + \mathbf{b} + \mathbf{c} = 0$. Find $\text{proj}_{\mathbf{a}}(\mathbf{b} + \mathbf{c})$.

4. Prove that the sum of the projections of vectors $\mathbf{a}$ and $\mathbf{b}$ on an axis having the same direction as the vector $\mathbf{a} + \mathbf{b}$ is equal to $|\mathbf{a} + \mathbf{b}|$.

2.2

1. Find the coordinates of points symmetric to the point $A(-3, 1)$ relative to the coordinate axes.

2. Find the coordinates of points which are symmetric to the point $A(-3, 1, 2)$ relative to (a) the coordinate axes, (b) the xy, xz, yz coordinate planes.

3. Find the lengths of the sides of triangle ABC if $A(3, 2)$, $B(-1, -1)$, $C(11, -6)$.

4. Find the point on the y-axis 5 units from the point $A(4, -6)$.

5. Given three points in the plane, $A(3, 7)$, $B(1, 3)$, and $C(7, 5)$, find the point which is symmetric to A relative to the straight line BC.

6. Find the center and radius of a circle passing through the point $A(-2, 4)$ and touching the coordinate axes.

7. Given two vectors $\mathbf{a} = \langle 2, -5, 3 \rangle$ and $\mathbf{b} = \langle 4, -3, -3 \rangle$. Find the lengths of the vectors $\mathbf{a} + \mathbf{b}$ and $2\mathbf{a} - 3\mathbf{b}$.

8. Given a triangle with the vertices $A(3, -5)$, $B(-3, 3)$, $C(-1, -2)$, find the length of the median AD, where D is the midpoint of the side BC.

9. Find the length and direction cosines of the vector AB with $A(-2, 1, 3)$ and $B(0, -1, 2)$.

10. Find out whether a vector in space can make the following angles with the coordinate axes: (a) 45°, 60°, 120°, (b) 45°, 135°, 60°. If the answer is affirmative, find the coordinates of the vector assuming its length to be unity.

11. Find the coordinates of the point M in space if the length of its radius vector equals 8 units and the angle of inclination to the x-axis is 45° and to the z-axis is 60°.

2.3

1. Check the validity of the following equalities: (a) $(\mathbf{a} + \mathbf{b})^2 = \mathbf{a}^2 + 2\mathbf{a}\cdot\mathbf{b} + \mathbf{b}^2$, (b) $(\mathbf{a} + \mathbf{b})\cdot(\mathbf{a} - \mathbf{b}) = \mathbf{a}^2 - \mathbf{b}^2$, (c) $(\mathbf{a} + \mathbf{b})^2 + (\mathbf{a} - \mathbf{b})^2 = 2|\mathbf{a}|^2 + 2|\mathbf{b}|^2$.

2. Three vectors **a**, **b**, and **c** form a triangle, i.e. $\mathbf{a} + \mathbf{b} + \mathbf{c} = 0$. Express the length of **c** in terms of **a** and **b**.

3. Two forces **P** and **Q** are applied to the same point. The forces act at an angle of 120° to each other and $|\mathbf{P}| = 4$ and $|\mathbf{Q}| = 7$. Find the value of the resulting force **R**.

4. Prove that a triangle with vertices (a) $A(0, 0)$, $B(3, 1)$, $C(1, 7)$ is right, (b) $A(2, -1)$, $B(4, 8)$, $C(10, 6)$ is obtuse.

5. Check whether the quadrilateral with vertices $A(5, 2, 6)$, $B(6, 4, 4)$, $C(4, 3, 2)$, $D(3, 1, -4)$ is a square.

6. Find the value of α such that the vectors $\langle -2, \alpha, 3\rangle$ and $\langle 2, \alpha, \alpha\rangle$ are perpendicular.

7. Find the angle between two vectors $\mathbf{a} = 3\mathbf{p} + 2\mathbf{q}$ and $\mathbf{b} = \mathbf{p} + 5\mathbf{q}$, where **p** and **q** are two perpendicular vectors.

8. Find the angle between the bisectors of the coordinate angles xOy and yOz.

9. Find the cosine of the angle between a diagonal and edge of the cube.

10. A parallelogram is constructed on the given vectors **a** and **b**. Express the vector of its height which is perpendicular to the side **a** in terms of the vectors **a** and **b**.

11. Find $\text{proj}_{\mathbf{a}}\mathbf{b}$ if $\mathbf{a} = \langle 1, 2, -2\rangle$ and $\mathbf{b} = \langle 0, 5, 2\rangle$.

12. Given the points $A(-3, 1, 5)$, $B(3, 9, 5)$, $C(0, 0, 1)$, $D(4, 1, -1)$, find the projection of the vector $\vec{CD}$ onto the axis having the same direction as the vector $\vec{AB}$.

13. Prove that the vector $(\mathbf{a}\cdot\mathbf{c})\,\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\,\mathbf{c}$ is perpendicular to the vector **a**.

2.4

1. Construct the points whose polar coordinates are (a) $(3, \pi/3)$, (b) $\left(\frac{1}{2}, \frac{2\pi}{3}\right)$, (c) $(\sqrt{3}, -\pi/6)$.

2. Find the polar coordinates of the point symmetric to the point $(1, \pi/4)$ relative to (a) the pole, (b) the polar axis.

3. Locate the points in the plane whose polar coordinates satisfy the following conditions: (a) $r = 1$, (b) $r = 3$, (c) $1 < r < 3$, (d) $\varphi = \pi/3$, (e) $\pi/2 < \varphi < 3\pi/2$.

Chapter 3

DETERMINANTS

3.1. SECOND-ORDER DETERMINANTS. CRAMER'S RULE

Suppose we have a square array consisting of four numbers:

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \tag{1}$$

Such an array is called a *matrix* or, to be more precise, a *matrix of order* (or *dimension*) 2×2 or a *square of order two.*

Definition. The *determinant of matrix* (1) is the number

$$a_1b_2 - a_2b_1$$

and is denoted by

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \tag{2}$$

Thus

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1$$

Example. We have

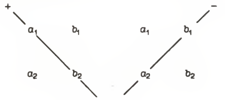
$$\begin{vmatrix} 1 & -2 \\ 3 & -5 \end{vmatrix} = 1(-5) - (-2)3 = 1$$

We should distinguish between matrix (1) and its determinant (2), namely, a matrix is an array of four numbers, while a determinant is the only number obtained from the array as indicated above.

The numbers a_1, b_1, a_2, b_2 are called the *elements* of matrix (1). The matrix has two rows and two columns. The elements a_1, b_1 and a_2, b_2 comprise the first and the second row respectively, and the elements a_1, a_2 and b_1, b_2 comprise the first and the second column respectively.

The same notions of elements, rows, and columns are also valid for determinant (2). We say that determinant (2) *is composed of the elements* a_1, b_1, a_2, b_2.

Although the expression $a_1b_2 - a_2b_1$ for the determinant is not complicated, it is advisable to memorize the following scheme for evaluating it:



The line segment marked with the + sign connects two elements whose product should be taken with the plus sign, and the line segment marked with the − sign connects two elements whose product should be taken with the minus sign.

Determinants of second-order square matrices are called *second-order determinants.*

1. Cramer's rule for solving a system of two second-order equations in two unknowns.

Suppose we have a system of two equations

$$\begin{cases} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \end{cases} \tag{3}$$

We use the coefficients of the unknowns to form the determinant

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$$

We shall call it the *determinant of system* (3) and denote it by Δ

We shall also need two other determinants, namely

$$\Delta_1 = \begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}, \quad \Delta_2 = \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}$$

The determinant Δ_1 is obtained from Δ by replacing the elements a_1, a_2 of the first column (they are the coefficients of

x in the equations of the system) by the constant terms c_1, c_2. The determinant Δ_2 is formed in a similar way by replacing the elements of the second column in Δ by the constant terms.

Theorem (Cramer's rule for a 2 × 2 system). *If the determinant Δ of system* (3) *is nonzero, the system has a unique solution. The solution can be found from the formulas*:

$$x = \frac{\Delta_1}{\Delta}, \quad y = \frac{\Delta_2}{\Delta} \tag{4}$$

□ By multiplying both sides of the first equation in (4) by b_2 and of the second equation by $-b_1$ and adding the results, we obtain

$$(a_1b_2 - a_2b_1)\,x = c_1b_2 - c_2b_1$$

or, which is the same, $\Delta\cdot x = \Delta_1$. Whence we have

$$x = \frac{\Delta_1}{\Delta}$$

Similarly, by multiplying both sides of the first equation by $-a_2$ and of the second equation by a_1 and adding the results we arrive at the equation

$$(a_1b_2 - a_2b_1)\,y = a_1c_2 - a_2c_1$$

or, briefly, $\Delta\cdot y = \Delta_2$. Whence we have

$$y = \frac{\Delta_2}{\Delta}$$

Thus we see that if a system of equations is solvable, its solution is given by (4). Now it remains for us to check whether the numbers $x = \frac{\Delta_1}{\Delta}$, $y = \frac{\Delta_2}{\Delta}$ constitute the solution of the system, that is, whether the following equalities

$$a_1\cdot\frac{\Delta_1}{\Delta} + b_1\cdot\frac{\Delta_2}{\Delta} = c_1, \quad a_2\cdot\frac{\Delta_1}{\Delta} + b_2\cdot\frac{\Delta_2}{\Delta} = c_2$$

hold true. We have

$$\begin{aligned} a_1\Delta_1 + b_1\Delta_2 &= a_1(c_1b_2 - c_2b_1) + b_1(a_1c_2 - a_2c_1) \\ &= c_1(a_1b_2 - a_2b_1) = c_1\Delta \end{aligned}$$

and this proves the validity of the first equality. The second equality can be checked in a similar way. ■

Example. Solve the system

$$\begin{cases} 3x - 5y = 0 \\ x - 2y = 1 \end{cases}$$

○ We have

$$\Delta = \begin{vmatrix} 3 & -5 \\ 1 & -2 \end{vmatrix} = -1 \neq 0$$

consequently

$$x = \frac{\begin{vmatrix} 0 & -5 \\ 1 & -2 \end{vmatrix}}{\Delta} = -5, \quad y = \frac{\begin{vmatrix} 3 & 0 \\ 1 & 1 \end{vmatrix}}{\Delta} = -3 \; \bullet$$

The theorem we have just proved is called *Cramer's rule* (to be more accurate, Cramer's rule for a system of two equations in two unknowns). We shall discuss Cramer's rule for the general case (i.e. for a system of n equations in n unknowns) in Sec. 3.15.

3.2. THIRD-ORDER DETERMINANTS

We shall now consider a square matrix of third order, that is, an array of 3×3 numbers,

$$\begin{matrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{matrix} \tag{1}$$

The notions of element, row, and column are equivalent to those of matrices of second order.

Definition. The *determinant of matrix* (1) is the number

$$a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - c_1b_2a_3 - b_1a_2c_3 - a_1c_2b_3 \tag{2}$$

Determinants of square matrices of third order are called *third-order determinants.*

The determinant of matrix (1) is written as

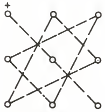
$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \tag{3}$$

Thus by definition we have

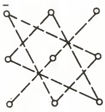
$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - c_1b_2a_3 - b_1a_2c_3 - a_1c_2b_3 \tag{4}$$

The numbers a_1, b_1, c_1, a_2, b_2, c_2, a_3, b_3, c_3 are called the elements of determinant (3). We say that determinant (3) *is formed from elements* a_1, b_1, c_1, a_2, b_2, c_2, a_3, b_3, c_3.

Although (4) seems to be complicated, it can be easily remembered. If we connect with a dashed line each three elements of the determinant whose product is taken with the plus sign in (4), we obtain scheme 1, which can be easily memorized. Similarly, for the products taken with the minus sign in (4) we have scheme 2.



Scheme 1



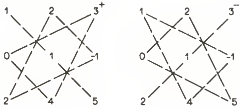
Scheme 2

The schemes illustrate the *triangle rule* for evaluating third-order determinants.

Example. Compute the third-order determinant

$$\begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \\ 2 & 4 & 5 \end{vmatrix}$$

○ We draw the schemes



and find that the determinant is equal to $1\cdot 1\cdot 5 + 2\cdot(-1)\cdot 2 + 0\cdot 4\cdot 3 - 3\cdot 1\cdot 2 - 2\cdot 0\cdot 5 - (-1)\cdot 4\cdot 1 = -1$. ●

With practice the reader will be able to apply the schemes mentally without drawing them.

Note that each of the six products in (2), say $a_1 b_2 c_3$, is called a *term* of the determinant. Any term of the determinant contains either of the letters a, b, and c as a factor; in other words, each term contains one element of the first, second, and third columns. The indices of these letters in each term are permutations of the numbers 1, 2, 3, that is, each term contains one element of the first, second, and third rows. This remark will serve as a basis for defining determinants of an arbitrary order n.

3.3. *n*th-ORDER DETERMINANTS

1. Notation.

We shall now use one letter, say a, to denote the elements in a determinant with indices showing the position of the element in the determinant rather than use different letters (a, b, c, and so on) as we did above: namely, we denote by a_{ij} an element in the ith row and jth column. For instance, a_{12} is the element of the first row and the second column (read "a one two" and not "a twelve").

In this notation, formulas for second- and third-order determinants are:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

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

2. A permutation of the numbers 1, 2, ... *n*. Even and odd permutations.

A *permutation* of the numbers 1, 2, . . ., n is their arrangement in some order (not necessarily increasing). For instance, 3, 2, 1, 4 is a permutation of the numbers 1, 2, 3, 4.

The number of permutations of 1, 2, . . ., n *is the product* $1\cdot 2\cdot 3\ldots n$.

□ We use P_n to denote the number of permutations of 1, 2, . . ., n. Obviously, $P_2 = 2 = 1\cdot 2$. We shall consider now a permutation of the numbers 1, 2, . . ., $n - 1$:

$$i_1,\ i_2,\ \ldots,\ i_{n-1} \tag{1}$$

We can use it to obtain the permutations of 1, 2, . . ., n if we write the number n either on the left of i_1, or between i_1 and i_2, or between i_2 and i_3, and so on, or on the right of i_{n-1}. In this way we can obtain n permutations of the numbers 1, 2, . . ., n from permutation (1). Since the number of permutations (1) is P_{n-1}, we have the following formula:

$$P_n = nP_{n-1}$$

Specifically,

$$P_3 = 3\cdot P_2 = 3\cdot 2\cdot 1, \quad P_4 = 4\cdot P_3 = 4\cdot 3\cdot 2\cdot 1$$
$$P_5 = 5\cdot P_4 = 5\cdot 4\cdot 3\cdot 2\cdot 1$$

In general, $P_n = n\cdot(n-1)\ldots 3\cdot 2\cdot 1$. ■

This product is written $n!$ (read "n factorial"). Thus, the number of permutations of the numbers 1, 2, . . ., n is $n!$

Suppose we have a permutation $j_1, j_2, \ldots, j_n$ of the numbers 1, 2, . . ., n. We denote it by J and write

$$J = (j_1,\ j_2,\ \ldots,\ j_n)$$

An *inversion* in a permutation J exists when a larger number precedes a smaller number.

Example. The permutation (3, 2, 1, 4) has three inversions: (3, 2), (3, 1), (2, 1).

We use $\sigma(J)$ to denote the total number of inversions in the permutation J. The permutation is said to be *even* if the number $\sigma(J)$ is even, and *odd* if $\sigma(J)$ is odd.

In the example just discussed the permutation has three inversions and consequently it is odd. Note that the permutation $(1, 2, 3, \ldots, n)$ has no inversions or, in other words, has zero inversions. Consequently, this permutation is even.

3. Definition of *n*th-order determinants.

We shall consider a square array formed from $n \times n$ numbers. Such an array is called a *square matrix of order n*. The number from the ith row and jth column of the array is denoted by a_{ij}.

We have the matrix

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \tag{2}$$

The numbers $a_{11}, \ldots, a_{nn}$ forming the matrix are called its *elements*.

Matrix (2) has n^2 elements. We choose n elements such that (1) they are in distinct rows (that is, we take one element from each row), and (2) they are in distinct columns (this means that we take one element from each column). We agree to call such a sequence of n elements of the matrix *admissible*.

An example of an admissible sequence is $a_{11}, a_{22}, \ldots, a_{nn}$. In order to obtain an admissible sequence we choose an arbitrary element of matrix (2) and then delete the row and column of the element. As a result we get a matrix of order $n - 1$. We choose an arbitrary element in this matrix and delete the row and column of this element. In the resulting matrix we again choose an element, and so on. All in all we choose n elements of matrix (2). These elements are from different rows and different columns of the matrix.

Let us consider an admissible sequence of n elements of matrix (2). We arrange the elements of the sequence in a definite order: beginning with an element a_{1j_1} from the first row, then an element a_{2j_2} from the second row, and so on. Thus we have

$$a_{1j_1},\ a_{2j_2},\ \ldots,\ a_{nj_n} \tag{3}$$

where $j_1, j_2, \ldots, j_n$ are the numbers of the rows of the elements being chosen, and by definition they are distinct. Consequently, $j_1, j_2, \ldots, j_n$ is an ordered sequence of numbers 1, 2, ..., n; we denote it by J. Thus

$$J = (j_1,\ j_2,\ \ldots,\ j_n) \tag{4}$$

is a permutation of the numbers 1, 2, ..., n.

Now if for each admissible sequence (2) we form the products $a_{1j_1}a_{2j_2}\ldots a_{nj_n}$ of all its elements, multiply the products by either $+1$ or -1 depending on whether the permutation J is even or odd, and then add the results, we obtain the expression:

$$\sum_J (-1)^{\sigma(J)} a_{1j_1}a_{2j_2}\ldots a_{nj_n} \tag{5}$$

which is the *determinant of matrix* (2) or simply an *nth-order determinant.* The letter J below the summation sign indicates that summation is carried out over all possible permutations (4) of the numbers 1, 2, ..., n. The products $a_{1j_1}a_{2j_2}\ldots a_{nj_n}$ in (5) are the *terms* of the determinant. There are $n!$ terms in an nth-order determinant.

An nth-order determinant is written as

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

The vertical lines mean that we are considering the determinant and not an array of n^2 numbers, that is, we are concerned with a number composed in a particular way from the array (expression (5)).

To summarize we formulate the following definition.

Definition. The *determinant of matrix* (2) is the expression

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum_J (-1)^{\sigma(J)} a_{1j_1} a_{2j_2} \dots a_{nj_n} \qquad (6)$$

where the summation is carried out over all permutations $J = (j_1, j_2, \dots, j_n)$ of the numbers $1, 2, \dots, n$.

Matrix (2) is often denoted by A and its determinant by $|A|$ or det A*. Sometimes we shall use the notation Δ.

4. Specific cases when $n = 2$ and $n = 3$.

Before studying nth-order determinants, it is necessary to verify that for $n = 2$ and $n = 3$ formula (6) leads to the familiar second- and third-order determinants. Products of form $a_{1j_1}a_{2j_2}$ are the terms of the second-order determinant

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

where (j_1, j_2) is any permutation of the numbers 1, 2. There are two such permutations: (1, 2) and (2, 1), the first being an even permutation (it has zero inversions). Consequently,

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

which coincides with the known expression for a second-order determinant.

The product $a_{1j_1}a_{2j_2}a_{3j_3}$ are the terms of the third-order determinant

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

where (j_1, j_2, j_3) is a permutation of the numbers 1, 2, 3. There are six such permutations, three of which are even, namely, (1, 2, 3), (2, 3, 1), (3, 1, 2), and three are odd, namely, (3, 2, 1), (2, 1, 3), (1, 3, 2). Whence it follows that

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} \quad (7)$$

This is the rule we know for evaluating third-order determinants.

3.4. TRANSPOSITION OF A DETERMINANT

We shall now consider an interesting property of a determinant.

Theorem. *The value of a determinant is left unchanged if its rows are written as columns, in the same order.*

For instance,

$$\begin{vmatrix} -1 & -2 & -3 \\ 5 & 7 & 9 \\ 0 & 4 & 6 \end{vmatrix} = \begin{vmatrix} -1 & 5 & 0 \\ -2 & 7 & 4 \\ -3 & 9 & 6 \end{vmatrix}$$

□ Suppose that

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

is the initial determinant, and

$$\Delta^* = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

is the result of interchanging of its rows and columns. Let us prove that $\Delta = \Delta^*$.

We shall now introduce an additional concept, the *inverse* permutation. Let J be a permutation of the numbers 1, 2, ..., n. Let us form a new permutation, which is *inverse* to J, and denote it by J^{-1} using the following rule. If a number j is in the ith position in the permutation J, then in the permutation J^{-1} the number i is in the jth position.

For example, the permutation (3, 6, 4, 5, 1, 2) is the inverse of the permutation (5, 6, 1, 3, 4, 2).

Let us show that $\sigma(J) = \sigma(J^{-1})$, that is, the *inverse permutation has the same number of inversions as the initial permutation.*

Indeed, suppose that numbers k and l, which are in the αth and βth positions respectively in a permutation J (with $\alpha < \beta$) form an inversion, i.e. $k > l$. The numbers α and β are in the kth and lth positions respectively in the permutation J^{-1}. Since $l < k$ but $\beta > \alpha$, the numbers α and β form an inversion in J^{-1}. Thus, each pair of numbers forming inversion in the initial permutation is associated with a similar pair in the inverse permutation. We can prove by analogy that pairs not forming an inversion in the initial permutation correspond to similar pairs in the inverse. Whence it follows that the number of inversions in the permutation J coincides with the number of inversions in the permutation J^{-1}, i.e. $\sigma(J) = \sigma(J^{-1})$.

Let us prove that $\Delta = \Delta^*$. The determinant Δ is the sum of $n!$ products of the form

$$(-1)^{\sigma(J)} a_{1j_1} a_{2j_2} \ldots a_{nj_n} \tag{1}$$

where $\sigma(J)$ is the number of inversions in the permutation $J = (j_1, j_2, \ldots, j_n)$.

Let us consider expression (1) in terms of the determinant Δ^*. We interchange the factors in the product $a_{1j_1} a_{2j_2} \ldots a_{nj_n}$ so that the first element is an element from the first row of the determinant Δ^* (i.e. from the first column of Δ), the second element is an element from the second row of Δ^*, and so on. The result will be $a_{1j_1} a_{2j_2} \ldots a_{nj_n} = a_{i_1 1} a_{i_2 2} \ldots a_{i_n n}$. The permutation $I = (i_1, i_2, \ldots, i_n)$ is the inverse of $J = (j_1, j_2, \ldots, j_n)$; indeed if in the permutation J a number q is in the pth position, then expression (1) contains the factor a_{pq}, and consequently p is in the qth position in the permutation I. Whence it follows that $\sigma(I) = \sigma(J)$, and thus

$$(-1)^{\sigma(J)} a_{1j_1} a_{2j_2} \ldots a_{nj_n} = (-1)^{\sigma(J)} a_{i_1 1} a_{i_2 2} \ldots a_{i_n n}$$

The left-hand side of this equality is a term of the determinant Δ with the sign it has in the determinant, and the right-hand side is the term of the determinant Δ^* having the corresponding sign. Thus the determinants Δ and Δ^* have the same terms taken with the same signs. Whence it follows that $\Delta = \Delta^*$. ■

The operation of interchanging the rows and columns of a determinant is called a *transposition* (Δ^* is the transpose of Δ). Therefore, the theorem just proved can be formulated thus: *the value of a determinant is left unchanged under transposition.*

The property that a determinant is unchanged upon interchanging rows and columns shows that in a certain sense rows are equivalent to columns, namely, *any statement concerning a determinant formulated in terms of rows remains valid if the word "row" is replaced by the word "column".* We shall use this statement in Sec. 3.6 when deriving the principal properties of determinants of order n.

3.5. EXPANSION OF A DETERMINANT BY ROWS AND COLUMNS

The initial formula

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum_J (-1)^{\sigma(J)} a_{1j_1} a_{2j_2} \dots a_{nj_n} \qquad (1)$$

is not suitable for evaluating nth order determinants since it has $n!$ terms and as n increases this number grows rapidly (for instance, whereas a fourth-order determinant has 24 terms, a fifth-order determinant has 120 terms).

In practice, determinants are usually evaluated from formulas for the expansion of determinants by rows (columns), which are given below.

Let us consider an nth order determinant

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

Let i be a number from 1, 2, ..., n. Each term of the determinant contains one element of the ith row as a factor. We combine all terms involving a_{i1} (the first element of the ith row), put before the brackets the common factor a_{i1}, and denote the expression in brackets by A_{i1}. Then we combine all terms containing a_{i2} (their sum is $a_{i2}A_{i2}$), and so on. As a result the sum in (1) decomposes into n parts:

$$a_{i1}A_{i1},\ a_{i2}A_{i2},\ \dots,\ a_{in}A_{in}$$

Consequently

$$\Delta = a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in} \tag{2}$$

Equation (2) is the *expansion of a determinant in terms of the elements of the ith row* (or simply the *expansion by the ith row*). Expression A_{ij} is called the *cofactor* or a *signed minor* of the element a_{ij} in the determinant Δ.

Thus, *a determinant is equal to the sum of the products of elements of any row by their cofactors.*

As an example let us expand a third-order determinant by the second row. From formula (7) in Sec. 3.3 we have

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

The expressions in brackets are the cofactors A_{21}, A_{22}, and A_{23}.

Remark. Any term of a determinant having an element a_{ij} as a factor cannot contain any other element from the ith row and the jth column of the determinant as a factor. Whence it follows that the expression for A_{ij} does not involve elements from the ith row and the jth column of the determinant. In other words, the number A_{ij} is completely defined by elements from other (not the ith) rows and other (not the jth) columns of the determinant.

The proof of the statement concerning columns is quite similar. Let j be a number from 1, 2, ..., n. Each term of the determinant contains an element from the jth column as a factor. We combine all the terms involving a_{1j} (the first element of the jth column). Their sum is equal to $a_{1j}A_{1j}$. Then we combine all terms containing a_{2j} (their sum is equal to $a_{2j}A_{2j}$), and so on. The result is

$$\Delta = a_{1j}A_{1j} + a_{2j}A_{2j} + \dots + a_{nj}A_{nj} \qquad (3)$$

which is the *expansion of the determinant Δ in terms of the elements of the jth column* (or *simply the expansion by the jth column*).

Thus, *a determinant is equal to the sum of the products of elements of any column by their cofactors.*

Formulas (2) and (3) can be used for evaluating the determinant Δ. But we must know how to find cofactors. To do this we first deduce some properties of nth-order determinants.

3.6. PROPERTIES OF *n*th-ORDER DETERMINANTS

Properties 1^0-7^0 of a determinant concern rows.

1^0. *If all the elements of a row of a determinant are zero, the determinant is zero.*

In order to prove the statement, it is sufficient to expand a determinant by that row.

2^0. *If any two rows of a determinant are interchanged, the determinant is multiplied by* -1.

□ We have to distinguish between two cases.

Case 1. Interchanging neighboring rows. Suppose that we interchange two neighboring rows, say the first and second, in the determinant:

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

The result is the new determinant

$$\Delta' = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

Let us show that $\Delta' = -\Delta$. We consider some term of Δ:

$$a_{1j_1}a_{2j_2}\dots a_{nj_n} \tag{1}$$

The factors $a_{1j_1}, \dots, a_{nj_n}$ of this product are also the elements of Δ' and they belong to different rows and different columns. Consequently, any term of the determinant Δ is also a term of Δ'. The converse is also true: any term of the determinant Δ' is also a term of Δ. Thus, the determinants Δ and Δ' contain the same terms. In order to prove that $\Delta' = -\Delta$, it is sufficient to verify that the sign (+ and −) of the term of the determinant Δ is opposite to the sign of the term of the determinant Δ'.

The sign of the product in (1) of Δ is defined by the factor $(-1)^{\sigma(J)}$, where $J = (j_1, j_2, \dots, j_n)$. In order to find the sign of the same term of Δ', we should write the product in (1) so that the first be an element from the first row in Δ', the second be an element from the second row, and so on. We know how to obtain the determinant Δ' and so we write (1) in the form $a_{2j_2}a_{1j_1}a_{3j_3}\dots a_{nj_n}$. The sign of this product in Δ' is defined by the factor $(-1)^{\sigma(J')}$, where $J' = (j_2, j_1, j_3, \dots, j_n)$.

When passing from J to J', we have interchanged j_1 and j_2. As a result the number of inversions in a permutation should change (increase or decrease) by 1. Indeed, if $j_1 < j_2$, then a change from J to J' results in a new inversion (j_2, j_1), and if $j_2 < j_1$, then the change eliminated the inversion (j_1, j_2), other inversions being unchanged. Thus, the numbers $\sigma(J)$ and $\sigma(J')$ differ by unity, whence it follows that $(-1)^{\sigma(J)} = -(-1)^{\sigma(J')}$. This means that the sign of the product (1) in the determinant Δ is opposite in sign to that in Δ'.

Case 2. Interchanging the rows which are not neighbors. Suppose that we interchange the ith and the jth rows, with $i < j$. This can be done in several steps. First we interchange the ith and the $(i + 1)$th rows. Then we interchange the $(i + 1)$th and $(i + 2)$th rows in the resultant determinant, and so on until the ith row of the initial determinant is in the jth position. Then we shift the initial jth row "upward" (it is now in the $(j - 1)$th position) until it is in the ith position. All in all, we should accomplish $2(j - i) - 1$ interchanges of neighboring rows ($j - i$ when shifting the ith row "downward" and $j - i - 1$ when shifting the jth row "upward"). Upon each interchange the determinant is multiplied by -1; the total number of interchanges is odd, and therefore the resulting determinant is multiplied by -1. ■

Remark. The proof of property 2^0 shows that the initial determinant Δ and the determinant Δ' obtained from Δ by interchanging two rows consist of the same terms but taken with opposite signs.

3^0. *If a determinant contains two identical rows, it is zero.*

□ Indeed, if we interchange two identical rows in a determinant Δ, we should obtain the same determinant Δ; on the other hand under such an interchange the determinant must be multiplied by -1. Consequently, $\Delta = -\Delta$, whence $2\Delta = 0$ or $\Delta = 0$. ■

Example. The determinant

$$\begin{vmatrix} -1 & 2 & 3 \\ 4 & -5 & 6 \\ 4 & -5 & 6 \end{vmatrix}$$

is zero since its second and third rows are identical.

4^0. *A common factor of the elements of any row can be taken outside the determinant.*

□ For instance, let the elements of the first row have a common factor k, namely $ka_{11}, ka_{12}, \ldots, ka_{1n}$. Then

$$\begin{vmatrix} ka_{11} & ka_{12} & \ldots & ka_{1n} \\ a_{12} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = ka_{11}A_{11} + ka_{12}A_{12} + \ldots + ka_{1n}A_{1n}$$

$$= k(a_{11}A_{11} + a_{12}A_{12} + \dots + A_{1n}A_{1n}) \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \quad \blacksquare$$

5^0. *If each element of a row is a sum of two addends, then the determinant can be represented as a sum of two determinants in each of which all the elements are the same as in the initial determinant except for the elements of the indicated row. The indicated row in the first determinant contains the first addends and that in the second determinant contains the second addends.*

For instance,

$$\begin{vmatrix} b_{11} + c_{11} & b_{12} + c_{12} & \dots & b_{1n} + c_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

In order to prove this, it is sufficient to write

$$(b_{11} + c_{11})\, A_{11} + (b_{12} + c_{12})\, A_{12} + \dots + (b_{1n} + c_{1n})\, A_{1n}$$
$$= (b_{11}A_{11} + b_{12}A_{12} + \dots + b_{1n}A_{1n})$$
$$+ (c_{11}A_{11} + c_{12}A_{12} + \dots + c_{1n}A_{1n})$$

6^0. *The value of a determinant is left unchanged, if the elements of a row are altered by adding to them any constant multiple of the corresponding elements of any other row.*

For instance,

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} + ka_{21} & a_{12} + ka_{22} & \dots & a_{1n} + ka_{2n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

Indeed, according to property 5^0, the determinant on the right-hand side of the equality can be represented as

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} ka_{21} & ka_{22} & \dots & ka_{2n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + k \begin{vmatrix} a_{21} & a_{22} & \dots & a_{2n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

(the determinant containing the factor k is zero since it has two identical rows, the first and the second).

Example. Evaluate the determinant

$$\begin{vmatrix} 2 & 5 & 9 \\ 26 & 57 & 92 \\ 263 & 571 & 920 \end{vmatrix}$$

○ Note that the second row differs "slightly" from the first multiplied by 10, and the third row from the second multiplied by 10. Therefore, we add the second row multiplied by -10 to the third row, and the first row multiplied by -10 to the second row. The resulting determinant is

$$\begin{vmatrix} 2 & 5 & 9 \\ 6 & 7 & 2 \\ 3 & 1 & 0 \end{vmatrix}$$

which, according to property 6^0, must be equal to the initial determinant. By a direct calculation (say, by the triangle rule), we find that the determinant is 109. Consequently, the initial determinant is also 109. ●

A remark must be made before formulating the next property. We know that the sum of the elements of a row multiplied by their cofactors equals the determinant. Suppose we add the elements of a row multiplied by the cofactors of the corresponding elements of another row. For instance, find the sum $a_{21}A_{11} + a_{22}A_{12} + \ldots + a_{2n}A_{1n}$, which contains the elements of the second row multiplied by the cofactors of the corresponding elements of the first row. The result can be deduced due to the next property.

7^0. *The sum of the elements of a row in a determinant multiplied by the cofactors of the corresponding elements of another row is zero.*

□ For instance, let us show that

$$a_{21}A_{11} + a_{22}A_{12} + \ldots + a_{2n}A_{1n} = 0 \tag{2}$$

For arbitrary numbers $b_1, b_2, \ldots, b_n$ the following equality is valid:

$$\begin{vmatrix} b_1 & b_2 & \ldots & b_n \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = b_1A_{11} + b_2A_{12} + \ldots + b_nA_{1n}$$

(this is the expansion of the determinant by the first row). We assume that $b_1 = a_{21}, b_2 = a_{22}, \ldots, b_n = a_{2n}$ and find that

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = a_{21}A_{11} + a_{22}A_{12} + \ldots + a_{2n}A_{1n}$$

The determinant on the left-hand side is zero since it contains two identical rows, the first and the second row. Equality (2) is proved. ■

Example. Given the determinant

$$\begin{vmatrix} 3 & 2 & 1 \\ 1 & -5 & 2 \\ 4 & -3 & 3 \end{vmatrix}$$

the sum of the elements of the first row multiplied by the cofactors of the corresponding elements of the third row is

$$3\cdot\begin{vmatrix} 2 & 1 \\ -5 & 2 \end{vmatrix} - 2\cdot\begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix} + 1\cdot\begin{vmatrix} 3 & 2 \\ 1 & -5 \end{vmatrix} = 0$$

3.7. MINORS. EVALUATION OF nth-ORDER DETERMINANTS

1. Minors.

Let us consider an nth-order determinant

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

We choose one element a_{ij}. If we delete the ith row and the jth column (i.e. the row and the column of the element a_{ij}) from Δ, we obtain a determinant of order $n - 1$. This is called the *minor of the element* a_{ij} in the determinant Δ and is denoted by M_{ij}. For instance, the minor of the element $a_{23} = -2$ in the fourth-order determinant is

$$M_{23} = \begin{vmatrix} 1 & -1 & 7 \\ 0 & 0 & 6 \\ 0 & 0 & -2 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 7 \\ 0 & 0 & 6 \\ 0 & 0 & -2 \end{vmatrix} = 0$$

2. Relation between Minors and Cofactors.

Let us prove the following fundamental theorem.

Theorem. *The cofactor of any element a_{ij} of the determinant Δ is equal to the minor of the element multiplied by $(-1)^{i+j}$, that is,*

$$A_{ij} = (-1)^{i+j}M_{ij} \qquad (1)$$

In other words, the same equality $A_{ij} = \pm M_{ij}$ is always true, the plus sign being valid when the sum $i + j$ is even and the minus sign when it is odd.

□ First we prove formula (1) for the special case when $i = n$, $j = n$. Since the number $n + n$ is even, we need only verify that

$$A_{nn} = M_{nn} \tag{2}$$

According to the definition of a cofactor, in order to find A_{nn} we must combine all the terms of the determinant Δ which involve the element a_{nn} as a factor, that is, the terms of the form

$$a_{1j_1}a_{2j_2}\dots a_{n-1,j_{n-1}}a_{nn} \tag{3}$$

The sign of such a term in Δ is defined by the factor $(-1)^{\sigma(J)}$, where $J = (j_1, j_2, \dots, j_{n-1}, n)$, Eliminating the number n from the permutation J, we obtain a permutation $J' = (j_1, j_2, \dots, j_{n-1})$ of the numbers $1, 2, \dots, n-1$. The number n is the largest of the numbers $1, 2, \dots, n-1$, and therefore it does not form an inversion with either of the numbers $j_1, j_2, \dots, j_{n-1}$ in the permutation J; consequently, the number of inversions in J coincides with that in J', that is,

$$\sigma(J) = \sigma(J') \tag{4}$$

If we factor out a_{nn} from all terms of form (3) (each taken with the respective factor $(-1)^{\sigma(J)}$) the expression in brackets will be equal to A_{nn}. Noting (4), we thus have

$$A_{nn} = \sum_{J'}(-1)^{\sigma(J')}a_{1j_1}a_{2j_2}\dots a_{n-1,j_{n-1}}$$

where J' is any permutation of the numbers $1, 2, \dots, n-1$. The sum on the right-hand side is the determinant

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1,n-1} \\ a_{21} & a_{22} & \dots & a_{2,n-1} \\ \dots & \dots & \dots & \dots \\ a_{n-1,1} & a_{n-1,2} & \dots & a_{n-1,n-1} \end{vmatrix}$$

that is, the minor M_{nn}. Equality (2) is proved.

Let us now consider the general case when i and j are any integers from 1 to n. We reduce this to the special case of $i = n$, $j = n$ discussed above.

We interchange the ith and $(i + 1)$th rows in the determinant Δ. Then we interchange the $(i + 1)$th and $(i + 2)$th rows in the resulting determinant, and so on, until the ith row of Δ changes places with the nth last row. The element a_{ij} in the resulting determinant appears in the nth row and jth column. Then by interchanging neighboring columns (first the jth and the $(j + 1)$th, then the $(j + 1)$th and $(j + 2)$th columns, and so on), we shift the element a_{ij} "to the right" until it appears in the bottom right corner of the determinant, i.e. in the nth row and nth column. We have made $2n - i - j$ interchanges of neighboring rows and columns ($n - i$ when shifting the ith row "downward" and $n - j$ when shifting the jth column "to the right"). These interchanges leave the determinant unchanged, while changing its sign. Given that the number $2n - (i + j)$ is even or odd as $i + j$ is even or odd, we infer that, as a result of the interchanges, the terms in the determinant Δ are multiplied by $(-1)^{i+j}$.

The minor M'_{nn} in the resulting determinant Δ' is the same determinant of order $n - 1$, as the minor M_{nn}. Thus, the cofactor A'_{ij} is equal to $(-1)^{i+j}A_{ij}$. Given that $A'_{nn} = M_{nn}$ (from what we have just proved), we have

$$(-1)^{i+j}A_{ij} = M_{ij}$$

and this is essentially the same as (1). ■

3. Evaluation of nth-order determinants.

In practice, nth-order determinants are evaluated from the formulas for expanding a determinant by a row or column. For instance, expanding a determinant Δ by the ith row we have

$$\Delta = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}$$

or noting formula (1)

$$\Delta = a_{i1}(-1)^{i+1}M_{i1} + a_{i2}(-1)^{i+2}M_{i2} + \ldots + a_{in}(-1)^{i+n}M_{in}$$

This equality reduces the evaluation of the nth-order determinant Δ to the evaluation of a set of $(n-1)$th-order determinants $M_{i1}, M_{i2}, \ldots, M_{in}$.

The formula for the expansion of a determinant by a row (or column) is simpler when all the elements of this row (or column) are zero except for one element, say a_{ij}:

$$\Delta = a_{ij}A_{ij}$$

or

$$\Delta = a_{ij}(-1)^{i+j}M_{ij}$$

Thus to evaluate the nth-order determinant Δ means to evaluate a single $(n-1)$th-order determinant.

Although a given determinant Δ may not contain a row (or column) with the required number of zeros, we can always transform the determinant, leaving its value unchanged, so that the elements of the row (or column) we have chosen are zero except for one element. To do this we use one of the properties of a determinant, namely, the value of a determinant remains unaltered if the elements of one row (or column) are altered by adding a multiple of the corresponding elements of any other row (or column), or in short, if a row (or column) multiplied by an arbitrary number is added to another row (or column). A concrete example illustrates how a determinant can be transformed to the desired form.

Example. Evaluate the fifth-order determinant

$$\Delta = \begin{vmatrix} 1 & 3 & -2 & -2 & 1 \\ 0 & -2 & 2 & 1 & 0 \\ -1 & 2 & 1 & -1 & 3 \\ 1 & -1 & 3 & 0 & 1 \\ 2 & 2 & 3 & -2 & 1 \end{vmatrix}$$

○ We choose the second row since it contains more zeros (there are two). We shall try to transform the determinant without changing its value, so that all the elements of the second row are zero, except for $a_{24} = 1$. Obviously, it is sufficient to add the fourth row multiplied by 2 to the second row and the fourth row multiplied by -2 to the third row. The result is the following determinant

$$\begin{vmatrix} 1 & -1 & 2 & -2 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 3 & -1 & 3 \\ 1 & -1 & 3 & 0 & 1 \\ 2 & -2 & 7 & -2 & 1 \end{vmatrix}$$

which equals the initial one. We expand it by the second row and obtain

$$\begin{vmatrix} 1 & -1 & 2 & 1 \\ -1 & 0 & 3 & 3 \\ 1 & -1 & 3 & 1 \\ 2 & -2 & 7 & 1 \end{vmatrix}$$

We should choose the second row in the last determinant since it contains zero. Let us transform the determinant so that all the elements of the second row are zero, except for $a_{12} = -1$. By subtracting the first row from the third one and the first row multiplied by 2 from the fourth one we obtain the determinant

$$\begin{vmatrix} 1 & -1 & 2 & 1 \\ -1 & 0 & 3 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & -1 \end{vmatrix}$$

which equals the initial determinant Δ. Its expansion by the second row yields

$$\Delta = (-1)(-1)^{1+2} \begin{vmatrix} -1 & 3 & 3 \\ 0 & 1 & 0 \\ 0 & 3 & -1 \end{vmatrix} = \begin{vmatrix} -1 & 3 & 3 \\ 0 & 1 & 0 \\ 0 & 3 & -1 \end{vmatrix}$$

This third-order determinant can be directly evaluated, but to expand it by the first row is much simpler:

$$\Delta = (-1)(-1)^{1+1} \begin{vmatrix} 1 & 0 \\ 3 & -1 \end{vmatrix} = 1 \bullet$$

3.8. CRAMER'S RULE FOR AN $n \times n$ SYSTEM

An important application of nth-order determinants is solving a system of n linear equations in n unknowns. We shall call such a system an $n \times n$ system.

Let us introduce some notation. We shall use the same letter x with different indices to denote different unknowns, viz. x_1 is the first unknown, x_2 is the second one, and so on. The coefficients of the unknowns in an ith equation are: a_{i1} is the coefficient of x_1, a_{i2} is that of x_2, and so on. So the ith equation of the system is written as

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = b_i$$

where n is the number of unknowns in the system.

We shall write an $n \times n$ system as

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases} \tag{1}$$

The determinant

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} \tag{2}$$

formed from the coefficients of the unknowns in system (1) is called the *determinant of this system.*

A *solution of system* (1) is the set of values of the unknowns $x_1 = \alpha_1, x_2 = \alpha_2, \ldots, x_n = \alpha_n$ such that it satisfies every equation of the system.

The following theorem, which is called *Cramer's rule for an* $n \times n$ *system,* is valid and it is similar to the theorem we have proved in Sec. 3.1 (Cramer's rule for a 2×2 system).

Theorem (Cramer's rule for an $n \times n$ system). *If the determinant Δ of an $n \times n$ system is nonzero, then the system has* *a unique solution, which can be found from the formulas*

$$x_1 = \frac{\Delta_1}{\Delta}, \quad x_2 = \frac{\Delta_2}{\Delta}, \quad \ldots, \quad x_n = \frac{\Delta_n}{\Delta} \tag{3}$$

where each determinant $\Delta_1, \Delta_2, \ldots, \Delta_n$ *is obtained from the determinant* Δ *by interchanging the corresponding column (the first, second, ..., nth) by the column consisting of the constant terms of the equations:*

$$\Delta_1 = \begin{vmatrix} b_1 & a_{12} & \ldots & a_{1n} \\ b_2 & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ b_n & a_{n2} & \ldots & a_{nn} \end{vmatrix}, \quad \Delta_2 = \begin{vmatrix} a_{11} & b_1 & \ldots & a_{1n} \\ a_{21} & b_2 & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & b_n & \ldots & a_{nn} \end{vmatrix}, \quad \ldots$$

□ First we prove that if a solution $x_1, x_2, \ldots, x_n$ exists for system (1), then it is satisfied by formulas (3). Then we show that the values of the unknowns defined by (3) are valid for (1).

(I) Suppose that each of the equalities in (1) is satisfied by some numbers $x_1, x_2, \ldots, x_n$. We multiply both sides of the first equality by A_{11}, those of the second by A_{21}, and so on, and both sides of the nth equality by A_{n1}. In other words, we multiply the equalities in (1) by the cofactors of the first column of the determinant Δ. Summing these equations yields

$$\begin{aligned} &A_{11}(a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n) \\ &+ A_{21}(a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &+ A_{n1}(a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n) \\ &= b_1a_{11} + b_2A_{21} + \ldots + b_nA_{n1} \end{aligned}$$

By combining the terms containing first x_1, then x_2, and so on, we rewrite the equation in the form

$$\begin{aligned} &(a_{11}A_{11} + a_{21}A_{21} + \ldots + a_{n1}A_{n1})\, x_1 \\ &+ (a_{12}A_{11} + a_{22}A_{21} + \ldots + a_{n2}A_{n1})\, x_2 \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &+ (a_{1n}A_{11} + a_{2n}A_{21} + \ldots + a_{nn}A_{n1})\, x_n \\ &= b_1A_{11} + b_2A_{21} + \ldots + b_nA_{n1} \end{aligned} \tag{4}$$

The coefficients of $x_1, x_2, \ldots, x_n$ can be easily simplified using the properties of a determinant. For example, the coefficients of x_1 is the sum of the elements of the first column in Δ by their cofactors; consequently, this coefficient is Δ. The coefficient of x_2 is the sum of the products of the elements of the second column in Δ by the cofactors of the corresponding elements of the first column; consequently, this coefficient is zero (property 7^0 of determinants). Similarly, the coefficients of $x_3, \ldots, x_n$ are zero. As a result, the left-hand side of (4) is equal to $\Delta \cdot x_1$. The right-hand side of (4) is the expansion of Δ_1 by the first row; consequently, the right-hand side of (4) is equal to Δ_1. Thus, it follows from (4) that

$$\Delta \cdot x_1 = \Delta_1$$

and noting that $\Delta \neq 0$ we have

$$x_1 = \frac{\Delta_1}{\Delta}$$

We have obtained the first formula in (3).

In order to obtain the second formula from (3) we multiply both sides of the first equality in (1) by A_{12}, those of the second by A_{22}, and so on, and those of the last by A_{n2}. In other words, we multiply the equalities in (1) by the cofactors of the elements of the second row in the determinant Δ. Adding these equalities yields

$$\Delta \cdot x_2 = \Delta_2$$

i.e. the second formula in (3).

The remaining formulas of (3) can be obtained in a similar way.

(II) Let us now check whether the values of the unknowns defined by (3) satisfy (1). We verify this for the first equation. We have

$$a_{11}\frac{\Delta_1}{\Delta} + a_{12}\frac{\Delta_2}{\Delta} + \ldots + a_{1n}\frac{\Delta_n}{\Delta}$$

$$= \frac{1}{\Delta}[a_{11}(b_1A_{11} + b_2A_{21} + \ldots + b_nA_{n1}) + a_{12}(b_1A_{12} + b_2A_{22} + \ldots + b_nA_{n2}) + \ldots + a_{1n}(b_1A_{1n} + b_2A_{2n} + \ldots + b_nA_{nn})]$$

$$= \frac{1}{\Delta}[b_1(a_{11}A_{11} + a_{12}A_{12} + \ldots + a_{1n}A_{1n}) + b_2(a_{11}A_{21} + a_{12}A_{22} + \ldots + a_{1n}A_{2n}) + \ldots + b_n(a_{11}A_{n1} + a_{12}A_{n2} + \ldots + a_{1n}A_{nn})].$$

The factor of b_1 on the right-hand side of this equality is equal to Δ since it is the expansion of Δ by the first row; the factor of b_2 is zero since it is the sum of the products of the elements of the first row in Δ by the cofactors of the corresponding elements of the second row; similarly, the factors of the remaining b_i are also zero. Consequently, we have

$$\frac{1}{\Delta}\cdot b_1\Delta = b_1 \quad \blacksquare$$

Remark. Formulas (3) are seldom used since they require evaluating the determinants $\Delta, \Delta_1, \ldots, \Delta_n$. A practical method for the solution of system (1) is Gaussian elimination (see Chapter 9).

Example 1. Solve the 3×3 system

$$\begin{cases} x + y + z = -2 \\ 4x + 2y + z = -4 \\ 9x + 3y + z = -8 \end{cases} \tag{5}$$

We have

$$\Delta = \begin{vmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{vmatrix} = 2 \neq 0$$

and therefore the system has a unique solution. We find that

$$\Delta_1 = \begin{vmatrix} -2 & 1 & 1 \\ -4 & 2 & 1 \\ -8 & 3 & 1 \end{vmatrix} = -2, \quad \Delta_2 = \begin{vmatrix} 1 & -2 & 1 \\ 4 & -4 & 1 \\ 9 & -8 & 1 \end{vmatrix} = 2$$

$$\Delta_3 = \begin{vmatrix} 1 & 1 & -2 \\ 4 & 2 & -4 \\ 9 & 3 & -8 \end{vmatrix} = -4$$

Consequently, $x = -2/2 = -1$, $y = 2/2 = 1$, $z = -4/2 = -2$.

Example 2. Given three points in the plane: $M_1(1, -2)$, $M_2(2, -4)$, $M_3(3, -8)$, find the coefficients p, q, r such that the graph of the function $y = px^2 + qx + r$ passes through M_1, M_2, and M_3. In other words, draw a parabola through these points.

○ The condition for the parabola to pass through $M_1(1, -2)$ is

$$-2 = p \cdot 1^2 + q \cdot 1 + r$$

The similar condition for $M_2(2, -4)$ is

$$-4 = p \cdot 2^2 + q \cdot 2 + r$$

and for $M_3(3, -8)$ is

$$-8 = p \cdot 3^2 + q \cdot 3 + r$$

We have a system of three equations in three unknowns

$$\begin{cases} p + q + r = -2 \\ 4p + 2q + r = -4 \\ 9p + 3q + r = -8 \end{cases}$$

which only differs from system (5) in Example 1 by the notation for the unknowns. Whence we have $p = -1$, $q = 1$, $r = -2$, and the desired function is $y = -x^2 + x + 2$. ●

3.9. A HOMOGENEOUS $n \times n$ SYSTEM

Definition. A linear equation is said to be *homogeneous* if its constant term is zero. A system composed of homogeneous equations is *homogeneous*.

Let us write a homogeneous system of n equations in n

unknowns in the general form

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = 0 \end{cases} \tag{1}$$

Setting $x_1 = 0, x_2 = 0, \ldots, x_n = 0$ we obtain one of the solutions of a homogeneous system. This solution is called a *zero* (or *trivial*) solution. It is frequenly necessary to know whether a homogeneous system has nontrivial solutions, i.e. such that at least one of the numbers $x_1, x_2, \ldots, x_n$ is nonzero. The following theorem gives an answer.

Theorem. *A system of n homogeneous equations in n unknowns has a nontrivial solution if and only if the determinant is zero.*

□ As usual we use Δ to denote the determinant of a system

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

We must prove two statements.

(1) If system (1) has nontrivial solution, then $\Delta = 0$.

(2) If $\Delta = 0$, then system (1) has a nontrivial solution.

Let us prove the first statement. Suppose that system (1) has a nontrivial solution $\alpha_1, \alpha_2, \ldots, \alpha_n$. Then the system has at least two distinct solutions: $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ and $(0, 0, \ldots, 0)$. Whence it follows that $\Delta = 0$. Indeed, if $\Delta \neq 0$, then as we know the system has a unique solution.

Let us now prove the second statement. We use induction on n. For $n = 2$ we have the system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = 0 \\ a_{21}x_1 + a_{22}x_2 = 0 \end{cases}$$

whose determinant, by the hypothesis, is zero:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} = 0$$

This equation implies that any pair of numbers, $(a_{12}, -a_{11})$ or $(a_{22}, -a_{21})$, is a solution of the system (check it). Therefore, if there are zeros among the numbers a_{11}, a_{12}, a_{21}, a_{22}, then a nontrivial solution exists. Now, if all four numbers are zero, then any pair of numbers x_1, x_2 gives the solution; consequently, nontrivial solutions exist in this case too.

Let us now suppose that the second statement is valid for a system of $n - 1$ equations; we prove the statement for a system of n equations.

We shall not consider a system of the form

$$\begin{cases} 0\cdot x_1 + 0\cdot x_2 + \ldots + 0\cdot x_n = 0 \\ 0\cdot x_1 + 0\cdot x_2 + \ldots + 0\cdot x_n = 0 \\ \ldots\ldots\ldots\ldots\ldots \\ 0\cdot x_1 + 0\cdot x_2 + \ldots + 0\cdot x_n = 0 \end{cases}$$

since any set of the numbers $x_1, x_2, \ldots, x_n$ provides its solution. Consequently, we may assume that at least one of the coefficients of the unknowns in (1) is nonzero. Two cases are possible here.

Case 1 when $a_{11} \neq 0$. We add the corresponding sides of the first equation multiplied by $-a_{21}/a_{11}$ to both sides of the second equation in (1). We obtain

$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n$$
$$- \frac{a_{21}}{a_{11}}(a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n) = 0$$

i.e. an equation of the form

$$a'_{22}x_2 + \ldots + a'_{2n}x_n = 0$$

(the term containing x_1 is absent). Similarly, adding the corresponding sides of the first equation multiplied by $-a_{31}/a_{11}$ to both sides of the third equation in (1) yields

$$a'_{32}x_2 + \ldots + a'_{3n}x_n = 0$$

and so on. As a result system (1) assumes the form

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0 \\ \qquad a'_{22}x_2 + \dots + a'_{2n}x_n = 0 \\ \dots\dots\dots\dots\dots\dots \\ \qquad a'_{n2}x_2 + \dots + a'_{n2}x_n = 0 \end{cases} \qquad (1')$$

System (1′) is equivalent to system (1), that is, it has the same solution as (1). Indeed, since (1′) is derived from (1), any solution $x_1, x_2, \dots, x_n$ of (1) satisfies system (1′). The converse is also true, that is, system (1) can be derived from (1′) (to do this we multiply the first equation in (1′) by a_{21}/a_{11} and add the result to the second equation, then multiply the first equation by a_{31}/a_{11} and add the result to the third equation, and so on), hence it follows that every solution of system (1′) satisfies (1).

Let us consider the determinant of system (1′)

$$\Delta' = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a'_{22} & \dots & a'_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & a'_{n2} & \dots & a'_{nn} \end{vmatrix}$$

Obviously, this can be obtained from the determinant Δ in the same way as system (1′) was obtained from system (1), namely, we multiply the first row by $-a_{21}/a_{11}$ and add the result to the second row, then multiply the first row by $-a_{31}/a_{11}$ and add the result to the third row, and so on. Each of these operations leaves the determinant unchanged (property 6^0 of a determinant). Consequently, $\Delta' = \Delta$. But by the hypothesis $\Delta = 0$ and, consequently, $\Delta' = 0$.

Expansion of Δ' by the first row yields

$$a_{11} \cdot \begin{vmatrix} a'_{22} & \dots & a'_{2n} \\ \dots & \dots & \dots \\ a'_{n2} & \dots & a'_{nn} \end{vmatrix} = 0$$

and since $a_{11} \neq 0$, we have

$$\begin{vmatrix} a'_{22} & \dots & a'_{2n} \\ \dots & \dots & \dots \\ a'_{n2} & \dots & a'_{nn} \end{vmatrix} = 0$$

According to the inductive assumption, this means that an $(n-1) \times (n-1)$ system

$$\begin{cases} a'_{22}x_2 + \ldots + a'_{2n}x_n = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a'_{n2}x_2 + \ldots + a'_{nn}x_n = 0 \end{cases}$$

which is a part of system (1′), has a nontrivial soluton. Let the solution be $x_2 = \alpha_2, \ \ldots, \ x_n = \alpha_n$. Having found x_1 from the first equation in (1′) $\left(\alpha_1 = -\frac{a_{12}}{a_{11}}\alpha_2 - \ldots - \frac{a_{1n}}{a_{11}}\alpha_n\right)$, we get a complete solution $\alpha_1, \alpha_2, \ldots, \alpha_n$ of system (1′) and therefore of system (1).

Case 2 when $a_{11} = 0$. By interchanging, if necessary, the equations in system (1) (as a result the rows in the determinant will interchange and it becomes zero), we get the first equation with at least one of the coefficients of the unknowns to be zero. If it is the coefficient of x_1, we have Case 1. If it is the coefficient of, say, x_2, we must interchange the unknowns, namely, substitute x_2 for x_1 and x_1 for x_2 (the first and second rows interchange in the determinant and it remains zero). As a result, the coefficient of x_1 in the first equation becomes nonzero, i.e. we again arrive at Case 1. ■

Example 1. Find whether the system has a nontrivial solution:

$$\begin{cases} x - 3y + 2z = 0 \\ 2x - y + z = 0 \\ 8x + y - z = 0 \end{cases}$$

○ We find the determinant of the system

$$\Delta = \begin{vmatrix} 1 & -3 & 2 \\ 2 & -1 & 1 \\ 8 & 1 & -1 \end{vmatrix} = 1 \cdot \begin{vmatrix} -1 & 1 \\ 1 & -1 \end{vmatrix}$$

$$-(-3) \cdot \begin{vmatrix} 2 & 1 \\ 8 & -1 \end{vmatrix} + 2 \cdot \begin{vmatrix} 2 & -1 \\ 8 & 1 \end{vmatrix} = 50$$

Since $\Delta \neq 0$, the system has a unique solution, $x = 0$, $y = 0$, $z = 0$. There is no nontrivial solution. ●

Example 2. Find whether the system has a nontrivial solution:

$$\begin{cases} x + 3y - z = 0 \\ 5x + y - 2z = 0 \\ 3x - 5y = 0 \end{cases}$$

○ The determinant is

$$\Delta = \begin{vmatrix} 1 & 3 & -1 \\ 5 & 1 & -2 \\ 3 & -5 & 0 \end{vmatrix} = 0$$

(check this). Thus the system has a nontrivial solution. ●

Example 3. Find the value of a at which the homogeneous system

$$\begin{cases} ax + y + 5z = 0 \\ 7x - ay + z = 0 \\ -3x + 2y + 3z = 0 \end{cases}$$

has a nontrivial solution.

○ The determinant of the system

$$\Delta = \begin{vmatrix} a & 1 & 5 \\ 7 & -a & 1 \\ -3 & 2 & 3 \end{vmatrix} = -3a^2 - 17a + 46$$

depends on a. A nontrivial solution exists if and only if $\Delta = 0$, i.e. if $3a^2 + 17a - 46 = 0$. By solving this quadratic equation, we find its two roots, $a_1 = 2$ and $a_2 = -23/3$. The system has nontrivial solutions at these values of a only. ●

3.10. A CONDITION FOR A DETERMINANT TO BE ZERO

We first consider a second-order determinant

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

Each row of such a determinant is a sequence of two numbers. In a Cartesian coordinate system in the plane, each row is associated with a vector.

Similarly, each row of a third-order determinant

$$\Omega = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

is a sequence of three numbers. In a Cartesian coordinate system in space, each row is again associated with a vector.

Theorem 1. A *second-order determinant is zero if and only if its row vectors are collinear.*

Theorem 2. *A third-order determinant is zero if and only if its row vectors are coplanar.*

Let us first prove the following lemma.

Lemma. *Two vectors* $\mathbf{p}_1$ *and* $\mathbf{p}_2$ *are collinear if and only if one of them is a linear combination of the other. Three vectors* $\mathbf{p}_1$, $\mathbf{p}_2$, *and* $\mathbf{p}_3$ *are coplanar if and only if one of them is a linear combination of the other two.*

□ The first assertion is obvious since it means that $\mathbf{p}_1$ and $\mathbf{p}_2$ are collinear if and only if the equality $\mathbf{p}_1 = \lambda\mathbf{p}_2$ or $\mathbf{p}_2 = \lambda\mathbf{p}_1$ holds true, where λ is a number.

Let us prove the second assertion of the lemma. We first assume that the vectors $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$ are coplanar and reduce them to the origin. If $\mathbf{p}_1 = 0$, then $\mathbf{p}_1$ is a linear combination of $\mathbf{p}_2$ and $\mathbf{p}_3$: $\mathbf{p}_1 = 0 \cdot \mathbf{p}_2 + 0 \cdot \mathbf{p}_3$. Therefore we assume that $\mathbf{p}_1 \neq 0$. Two cases are possible here.

(1) Vector $\mathbf{p}_2$ is collinear with $\mathbf{p}_1$, i.e. $\mathbf{p}_2 = \lambda\mathbf{p}_1$, λ being a scalar. We write this equation in the form $\mathbf{p}_2 = \lambda\mathbf{p}_1 + 0 \cdot \mathbf{p}_3$ and see that one of the three vectors can be represented as a linear combination of the other two.

(2) Vector $\mathbf{p}_2$ is not collinear with $\mathbf{p}_1$. Then $\mathbf{p}_1$ and $\mathbf{p}_2$ represent a plane, and vector $\mathbf{p}_3$, which is coplanar with $\mathbf{p}_1$ and $\mathbf{p}_2$ should lie in this plane, and this means that $\mathbf{p}_3$ can be represented as a linear combination of $\mathbf{p}_1$ and $\mathbf{p}_2$ (see Item 1 in Sec. 1.2).

Conversely, suppose one of the vectors $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$ is a linear combination of the other two, namely $\mathbf{p}_3 = \lambda\mathbf{p}_1 + \mu\mathbf{p}_2$. It is obvious that $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$ are coplanar. ■

□ Let us now prove Theorem 1. We use $\mathbf{p}_1$ and $\mathbf{p}_2$ to denote vectors corresponding to the rows of the determinant ($\mathbf{p}_1$ corresponds to the first row and $\mathbf{p}_2$ to the second). We must prove two statements:

(1) if vectors $\mathbf{p}_1$ and $\mathbf{p}_2$ are collinear, then $\Delta = 0$, and

(2) if $\Delta = 0$, then $\mathbf{p}_1$ and $\mathbf{p}_2$ are collinear.

Let us prove the first statement. Suppose that $\mathbf{p}_1$ and $\mathbf{p}_2$ are collinear, then, according to the lemma, one of them is a linear combination of the other, namely, $\mathbf{p}_2 = \lambda\mathbf{p}_1$, λ being a number. This means that $a_{21} = \lambda a_{11}$, $a_{22} = \lambda a_{12}$. It follows that

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} = \lambda(a_{11}a_{12} - a_{12}a_{11}) = 0$$

the proof is complete.

Let us prove the second statement. Suppose $\Delta = 0$.

Let us consider the following system of equations

$$\begin{cases} a_{11}x_1 + a_{21}x_2 = 0 \\ a_{12}x_1 + a_{22}x_2 = 0 \end{cases} \tag{1}$$

Its determinant

$$\Delta^* = \begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix}$$

is obtained by transposing Δ. Consequently, Δ^* is also zero, and this means, according to the theorem on homogeneous systems, that system (1) has nontrivial solutions, say $x_1 = \alpha_1$, $x_2 = \alpha_2$. Thus, the following equations are valid:

$$a_{11}\alpha_1 + a_{21}\alpha_2 = 0, \quad a_{12}\alpha_1 + a_{22}\alpha_2 = 0$$

where either α_1 or α_2 is nonzero. These equations show that if we multiply the first row of the determinant by α_1 and the second by α_2 and add the results, we obtain the zeroth row. In other words,

$$\alpha_1\mathbf{p}_l + \alpha_2\mathbf{p}_2 = 0 \tag{2}$$

Since either α_1 or α_2 is nonzero, we can represent one of the vectors $\mathbf{p}_1$, $\mathbf{p}_2$ as a linear combination of the other. For instance, if $\alpha_2 \neq 0$, then it follows from (2) that

$$\mathbf{p}_2 = -\frac{\alpha_1}{\alpha_2}\mathbf{p}_1$$

Thus, $\mathbf{p}_1$ and $\mathbf{p}_2$ are collinear. ∎

□ Let us prove Theorem 2. We use $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$ to denote the vectors corresponding to the rows of the determinant Ω. As before, we must prove two assertions:

(1) if vectors $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$ are coplanar, then $\Omega = 0$, and

(2) if $\Omega = 0$, then vectors $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$ are coplanar.

Suppose $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$ are coplanar, then, according to the lemma, one of the vectors is a linear combination of the other two. For instance, $\mathbf{p}_3$ can be represented as a linear combination of $\mathbf{p}_1$ and $\mathbf{p}_2$, i.e. $\mathbf{p}_3 = \lambda\mathbf{p}_1 + \mu\mathbf{p}_2$, where λ and μ are numbers. We write this equation as $\mathbf{p}_3 - \lambda\mathbf{p}_1 - \mu\mathbf{p}_2 = 0$. Hence we can see that if we multiply the first row of the determinant Ω by $-\lambda$ and the second row by $-\mu$ and add them to the third row, then we obtain the determinant

$$\Omega' = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 0 \end{vmatrix}$$

Using the properties of a determinant, we can write $\Omega' = \Omega$ (property 6^0) and $\Omega' = 0$ (property 1^0). Consequently, $\Omega = 0$, which was required to prove.

The second statement can be proved in the same way as the similar statement of Theorem 1 (we leave it to the reader). ∎

Example. Find whether the vectors $\mathbf{p}_1 = \langle -4, 1, -7\rangle$, $\mathbf{p}_2 = \langle 2, 5, 9\rangle$, and $\mathbf{p}_3 = \langle -8, 13, -3\rangle$, defined by their coordinates in a Cartesian system, are coplanar.

○ We use the coordinate of the vectors to form a third-order determinant

$$\Omega = \begin{vmatrix} -4 & 1 & -7 \\ 2 & 5 & 9 \\ -8 & 13 & -3 \end{vmatrix}$$

Direct evaluation shows that it is equal to zero. Hence it follows, according to Theorem 2, that $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$ are coplanar. ●

Exercises to Chapter 3

3.1

1. Evaluate the determinants:

$$\text{(a)} \begin{vmatrix} 7 & -4 \\ 5 & -3 \end{vmatrix}, \quad \text{(b)} \begin{vmatrix} 1 & 5 \\ 11 & 55 \end{vmatrix}, \quad \text{(c)} \begin{vmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & -\cos\alpha \end{vmatrix}$$

2. Solve the equations:

$$\text{(a)} \begin{vmatrix} 0 & x \\ x & 0 \end{vmatrix} = 1, \quad \text{(b)} \begin{vmatrix} x+2 & -3 \\ x-2 & x \end{vmatrix} = 0$$

3. Using Cramer's rule, solve the systems of equations:

$$\text{(a)} \begin{cases} 10x - 3y + 41 = 0 \\ 3x + 2y - 8 = 0, \end{cases} \quad \text{(b)} \begin{cases} 3x + y + 5 = 0 \\ x + 5y - 3 = 0, \end{cases}$$

$$\text{(c)} \begin{cases} x\cos\alpha - y\sin\alpha = a \\ x\sin\alpha + y\cos\alpha = b \end{cases}$$

4. Prove that the square polynomial $ax^2 + bx + c$, with $a \neq 0$, is a perfect square if and only if

$$\begin{vmatrix} a & b \\ b & c \end{vmatrix} = 0$$

3.2

1. Evaluate the determinants:

$$\text{(a)} \begin{vmatrix} 2 & 1 & 3 \\ 5 & 3 & 2 \\ 1 & 4 & 3 \end{vmatrix}, \quad \text{(b)} \begin{vmatrix} -3 & 0 & 1 \\ -5 & 2 & 4 \\ 0 & 3 & 7 \end{vmatrix}, \quad \text{(c)} \begin{vmatrix} 0 & a & 0 \\ b & c & d \\ 0 & c & 0 \end{vmatrix}$$

2. Solving the equations:

$$\text{(a)} \begin{vmatrix} 1 & 2 & 3 \\ -1 & x & 0 \\ 3 & 2 & 4 \end{vmatrix} = -1, \quad \text{(b)} \begin{vmatrix} 1 & x & 2 \\ x & -1 & x \\ -5 & -5 & 4 \end{vmatrix} = 0,$$

$$\text{(c)} \begin{vmatrix} a & a & a \\ -a & a & x \\ -a & -a & x \end{vmatrix} = 0, \quad \text{(d)} \begin{vmatrix} a+x & x & x \\ x & b+x & x \\ x & x & c+x \end{vmatrix} = 0$$

3.3

1. Find the number of inversions in the permutations:
(a) (7, 6, 9, 1, 2, 3, 5, 4, 8), (b) $(n, n-1, n-2, \ldots, 2, 1)$. Are the permutations in (a) and (b) even?

2. Find whether the product (a) $a_{55}a_{44}a_{53}a_{21}a_{32}$, (b) $a_{51}a_{23}a_{34}a_{45}a_{12}$ is a term of a fifth-order determinant. If the answer is affirmative, find the sign of the term.

3. Choose the values of i and k so that the product $a_{47}a_{63}a_{1i}a_{55}a_{7k}a_{24}a_{31}$ is a term of a determinant (find its order) and has the plus sign.

4. Prove that the determinant

$$\begin{vmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 \\ a_{31} & a_{32} & a_{33} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix}$$

all whose elements located "higher" than $a_{11}, a_{22}, \dots, a_{nn}$ are zeros, is equal to the product $a_{11}a_{22}a_{33}\dots a_{nn}$.

5. Prove that the fourth-order determinant

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & 0 & 0 \\ a_{41} & a_{42} & 0 & 0 \end{vmatrix}$$

is zero.

6. Find the terms containing x^4 and x^3 of the determinant

$$\begin{vmatrix} 5x & 1 & 2 & 3 \\ x & x & 1 & 2 \\ 1 & 2 & x & 3 \\ x & 1 & 2 & 2x \end{vmatrix}$$

3.6

1. Using the properties of a determinant, show that
(a) the equation

$$\begin{vmatrix} 1 & x & x^2 \\ 1 & a & a^2 \\ 1 & b & b^2 \end{vmatrix} = 0$$

has the roots $x_1 = a$ and $x_2 = b$,
(b) each of the determinants

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}, \quad \begin{vmatrix} \sin^2\alpha & 1 & \cos^2\alpha \\ \sin^2\beta & 1 & \cos^2\beta \\ \sin^2\gamma & 1 & \cos^2\gamma \end{vmatrix}, \quad \begin{vmatrix} \sin\alpha & \cos\alpha & \sin(\alpha+\delta) \\ \sin\beta & \cos\beta & \sin(\beta+\delta) \\ \sin\gamma & \cos\gamma & \sin(\gamma+\delta) \end{vmatrix}$$

is zero.

2. Using the properties of a determinant, evaluate

$$\begin{vmatrix} 1 & 3 & 6 \\ 12 & 31 & 62 \\ 122 & 315 & 623 \end{vmatrix}$$

3. Prove the equation

$$\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = (b - a)(c - a)(c - b)$$

both by a direct evaluation of the determinant and by using the properties of a determinant.

3.7

1. Find the minor of a_{ij} in the determinant Δ

(a) $\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$, $i = 1, j = 2$, (b) $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$, $i = 2, j = 2$,

(c) $\begin{vmatrix} 5 & -7 & 2 \\ 9 & 1 & 3 \\ 4 & 1 & 6 \end{vmatrix}$, $i = 2, j = 3$, (d) $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}$, $i = 4, j = 4$

2. Expand the determinant

(a) $\begin{vmatrix} 2 & -3 & 4 & 1 \\ 4 & -2 & 3 & 2 \\ a & b & c & d \\ 3 & -1 & 4 & 3 \end{vmatrix}$

by the third row, and the determinant

(b) $\begin{vmatrix} 5 & a & 2 & -1 \\ 4 & b & 4 & -3 \\ 2 & c & 3 & -2 \\ 4 & d & 5 & -4 \end{vmatrix}$

by the second column.

3. Evaluate the determinants:

(a) $\begin{vmatrix} 3 & 1 & 1 & 1 \\ 1 & 3 & 1 & 1 \\ 1 & 1 & 3 & 1 \\ 1 & 1 & 1 & 3 \end{vmatrix}$, (b) $\begin{vmatrix} 1 & 2 & 3 & 4 \\ -2 & 1 & -4 & 3 \\ 3 & -4 & -1 & 2 \\ 4 & 3 & -2 & -1 \end{vmatrix}$, (c) $\begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{vmatrix}$,

(d) $\begin{vmatrix} 0 & 1 & 1 & 3 \\ 1 & 0 & 1 & 6 \\ 1 & 1 & 0 & -2 \\ 3 & 6 & -2 & 0 \end{vmatrix}$, (e) $\begin{vmatrix} 2 & 0 & -1 & 3 & 4 \\ 0 & 1 & 0 & 2 & 1 \\ -1 & 0 & 1 & 5 & 3 \\ -3 & 1 & -2 & 0 & 2 \\ -1 & -1 & 2 & 5 & 1 \end{vmatrix}$, (f) $\begin{vmatrix} -1 & 2 & 3 & 1 & 1 \\ 0 & -3 & 2 & 1 & 0 \\ 2 & 1 & -1 & 0 & 2 \\ 4 & 1 & -2 & 7 & 0 \\ 0 & 0 & 3 & 5 & 0 \end{vmatrix}$

3.8

1. Using Cramer's rule, solve the systems of equations:

(a) $\begin{cases} x + 2y + 3z = 7 \\ 2x - y + z = 9 \\ x - 4y + 2z = 11, \end{cases}$ (b) $\begin{cases} x + y + z = 1 \\ 2x - 3y + z = 1 \\ 4x + y - 5z = 1, \end{cases}$

(c) $\begin{cases} 2x + y + z + u = 5 \\ x + 2y + z + u = 4 \\ x + y + 2z + u = 7 \\ x + y + z + 2u = 4, \end{cases}$ (d) $\begin{cases} 2x_1 + 2x_2 - x_3 + x_4 = 4 \\ 4x_1 + 3x_2 - x_3 + 2x_4 = 6 \\ 8x_1 + 5x_2 - 3x_3 + 4x_4 = 12 \\ 3x_1 + 3x_2 - 2x_3 + 2x_4 = 6 \end{cases}$

2. Find all the solutions of the system of equations if the determinant of the system is nonzero:

(1) $\begin{cases} a_1x + b_1y + c_1z = a_1 \\ a_2x + b_2y + c_2z = a_2 \\ a_3x + b_3y + c_3z = a_3, \end{cases}$ (2) $\begin{cases} ax_1 + bx_2 + bx_3 + bx_4 = c \\ bx_1 + ax_2 + bx_3 + bx_4 = c \\ bx_1 + bx_2 + ax_3 + bx_4 = c \\ bx_1 + bx_2 + bx_3 + ax_4 = c \end{cases}$

3.9

1. Find nontrivial solutions of the system of equations if it has any:

(a) $\begin{cases} 2x - 4y = 0 \\ 5x - 10y = 0, \end{cases}$ (b) $\begin{cases} -x + y = 0 \\ 2x + 7y = 0, \end{cases}$ (c) $\begin{cases} -2x + y + z = 0 \\ x - 2y + z = 0 \\ x + y - 2z = 0 \end{cases}$

2. Find the values of a at which the system has nontrivial solutions:

(a) $\begin{cases} x + 2y - 3z = 0 \\ 3x - 2y + z = 0 \\ ax - 14y + 15z = 0, \end{cases}$ (b) $\begin{cases} ax + y + z = 0 \\ x + ay + z = 0 \\ x + y + ax = 0 \end{cases}$

3. Prove that if the system

$$\begin{cases} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \\ a_3x + b_3y = c_3 \end{cases}$$

is consistent, then

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0$$

3.10

1. Find whether the vectors are coplanar:

(a) $\mathbf{a} = \langle -1, 4, 2\rangle$, $\mathbf{b} = \langle 3, 1, -5\rangle$, $\mathbf{c} = \langle 1, 9, -1\rangle$,
(b) $\mathbf{a} = \langle 0, 1, 0\rangle$, $\mathbf{b} = \langle \alpha, \beta, \gamma\rangle$, $\mathbf{c} = \langle 0, 2, 0\rangle$

2. Find the values of λ for which the vectors $\mathbf{a} = \langle 3, 4, 2\rangle$, $\mathbf{b} = \langle 6, 8, 7\rangle$, $\mathbf{c} = \langle 9, 12, \lambda\rangle$ are coplanar.

Chapter 4

THE EQUATION OF A LINE IN THE PLANE. A STRAIGHT LINE IN THE PLANE

4.1. THE EQUATION OF A LINE

Suppose that we have an equation

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

in two unknowns x and y. If the numbers x_0 and y_0 are such that their substitution into the equation turns it into an equality then we say that the pair of numbers x_0, y_0 satisfies the equation. For example, the pair $x_0 = 10$, $y_0 = 9$ satisfies the equation $1 - x + y = 0$, while the pair $x_1 = 2$, $y_1 = 3$ does not.

A pair of numbers x, y satisfying a given equation is not arbitrary: if x is defined, then y cannot be arbitrary since the value of y is related to x. It follows that a change in x causes a change in y. A point (x, y) will trace a line in the coordinate plane.

1. A line defined by an equation. An equation corresponding to a line.

Definition 1. Suppose we have an equation $F(x, y) = 0$. Then the set of all points (x, y) in a coordinate plane whose coordinates satisfy the equation is called the *line defined by this equation.*

Definition 2. Suppose we have a line l in a coordinate plane. The *equation corresponding to this line* or simply the equation of the line l is an equation $F(x, y) = 0$ which defines the line l.

In other words, if we have an equation and must find the line defined by this equation, we have to collect all the points (x, y) whose coordinates satisfy the equation. Conversely, if we have a line and must find its equation, we have to choose an equation in two unknowns such that it is satisfied by all the points of the line and only those points.

If we have the equation of a line, then studying the geometric

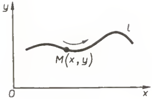


Figure 48

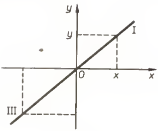


Figure 49

properties of the line reduces to studying its equation and this is the main purpose of analytic geometry. Its importance is that it is easier to study an equation than to investigate a line geometrically, especially since many methods have been developed in algebra and mathematical analysis for studying equations.

The term *running point* of a line is often used when studying lines. This is a variable point $M(x, y)$ which moves along the line (Fig. 48). The coordinates x and y of a running point are called *running coordinates*.

An important remark must be made. An equation $F(x, y) = 0$ relating x and y can sometimes be solved for y, i.e. y can be explicitly expressed in terms of x. In this case the above equation can be replaced by an equation of the form $y = f(x)$. A curve defined by such an equation is called the *graph of the function* $f(x)$. For instance, the equation $1 - x + y = 0$ implies the relationship $y = x - 1$ (or $x = y + 1$).

Finally we note that the equation of a line can be considered in any coordinates in the plane and not only in rectangular Cartesian coordinates. For instance, in polar coordinates the equation of a line is of the form $f(r, \varphi) = 0$.

Example 1. We have an equation $x - y = 0$ or $y = x$. If x assumes an arbitrary value, then y assumes the value of x. Consequently, a line defined by this equation consists of a point equidistant from the x- and y-axes and located in the first or the third quadrant. In other words, the equation defines the bisector of the first and the third quadrant angles (Fig. 49).

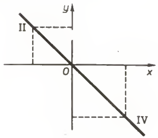


Figure 50

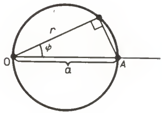


Figure 51

Example 2. We have an equation $x + y = 0$ or $y = -x$. This equation defines the bisector of the second and the fourth quadrant angles (Fig. 50).

Example 3. We have an equation $y = 0$. Although the equation does not involve the x coordinate, we may consider it to be an equation in two unknowns, x and y (in order to show that this assumption is valid, we can write this equation in the form $0 \cdot x + y = 0$; now the left-hand side formally contains x). The line l defined by this equation consists of the points with zero ordinate; the abscissa can be arbitrary since the equation imposes no conditions on x. Consequently l is the x-axis.

Example 4. We have an equation $x^2 + y^2 = 0$. It is satisfied by a single pair of values of x and y, namely $x = 0$ and $y = 0$. Therefore, the equation defines just one point and this is the origin. We may consider this point as a degenerate line.

Example 5. We have an equation $\sin^2(x + y) - 3 = 0$. Obviously, this equation cannot be satisfied by any values of x and y. Consequently, this equation defines an "empty" line (a line containing no points).

Example 6. We have an equation in polar coordinates $r - a\cos\varphi = 0$, where $a > 0$ is a constant number. This equation defines a circle of radius $R = a/2$ which passes through the pole O, the center of the circle lying on a polar ray (Fig. 51).

Let us consider a point A on a polar ray at the distance a from the pole. The condition $r = \cos\varphi$ means that $\angle OMA$ is a right angle. Every point on the indicated circle possesses such a property and only they do.

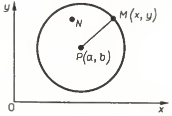


Figure 52

The examples considered above specified an equation and we had to find the line defined by the given equation. We now consider the inverse problem: to form the equation of a given line. Bear in mind that the desired equation can be written in different forms. For instance, $x - y = 0$, $y - x = 0$, $(x - y)^2 = 0$, $|x - y| = 0$ are different forms of the equation of the same line, namely the bisector of the first and third quadrant angles.

Example 7. Set up the equation of the circle with center at $P(a, b)$ and radius R (Fig. 52).

○ We mark an arbitrary running point $M(x, y)$ on the circle. Then the distance from M to P equals R:

$$\sqrt{(x - a)^2 + (y - b)^2} = R \tag{2}$$

and this is the equation of the circle. Indeed,

(1) if the point $M(x, y)$ lies on the circle, then its x, y coordinates satisfy equation (2) since $|MP| = R$,

(2) if a point $N(x, y)$ is not on the circle, then its coordinates do not satisfy equation (2). Indeed, in this case the distance between N and P is not equal to R.

A simpler form of equation (2) can be obtained by squaring its both sides:

$$(x - a)^2 + (y - b)^2 = R^2$$

which is an equation of a circle in rectangular Cartesian coordinates.

If the center P of the circle coincides with the origin, then $a = 0$, $b = 0$ and the equation assumes the form

$$x^2 + y^2 = R^2 \quad \bullet$$

2. A problem on the intersection of two lines.

Suppose that lines l_1 and l_2 are specified by their equations $F(x, y) = 0$ and $\Phi(x, y) = 0$ respectively. Let us find the points at which the lines intersect. From the viewpoint of analytic geometry, finding intersection points means finding the coordinates in a given coordinate system.

If $M(x_0, y_0)$ is an intersection point, then $F(x_0, y_0) = 0$ since M belongs to l_1 and $\Phi(x_0, y_0) = 0$ as well, since M also belongs to l_2. Thus the numbers x_0 and y_0 satisfy the system of equations

$$\begin{cases} F(x, y) = 0 \\ \Phi(x, y) = 0 \end{cases}$$

Conversely, if a pair of numbers x_0, y_0 satisfies the system, then the point $M(x_0, y_0)$ belongs to the two lines simultaneously, that is, it is a point at which they intersect.

Thus, *in order to find the points of intersection of two lines, we must solve the system of the equations of these lines.*

If the system has no solution, then the lines do not intersect; if the system has a unique solution (a single pair of numbers x_0, y_0), then the lines intersect only at one point; if the system has several solutions, then the lines intersect at several points.

Example. Find the point of intersection of the bisectors of the first and third quadrant angles with the circle of radius $R = \sqrt{5}$ centered on the y-axis a unit from the origin (Fig. 53).

○ We know that the equation of this bisector has the form $x - y = 0$. The equation of the circle is $x^2 + (y - 1)^2 = 5$. The system

$$\begin{cases} x - y = 0 \\ x^2 + (y - 1)^2 = 5 \end{cases}$$

has two solutions: $x_1 = 2$, $y_1 = 2$ and $x_2 = -1$, $y_2 = -1$. This means that the lines intersect at two points, $(2, 2)$ and $(-1, -1)$. ●

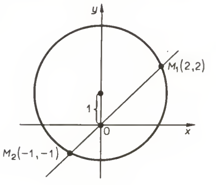


Figure 53

4.2. PARAMETRIC EQUATIONS OF A LINE

Suppose that a point M moves along a line l in the plane. If a coordinate system is specified, the position of M is defined by its coordinates x and y. As the point moves, its coordinates vary, that is they depend on time:

$$x = f(t), \ y = \varphi(t) \tag{1}$$

The symbols $f(t)$ and $\varphi(t)$ mean there is a functional dependence. The variable t (time) is a parameter; given a certain value of t we can find the values of x and y from formulas (1), i.e. the position of M. Therefore we usually say that formulas (1) specify the line l *parametrically*.

Let us digress from physics considerations and suppose that we have a varying quantity t (not necessarily time). We specify two functions

$$x = f(t), \ y = \varphi(t) \tag{2}$$

If we assign arbitrary values to the parameter t (such that $f(t)$ and $\varphi(t)$ are meaningful), then x and y will vary and the point $M(x, y)$ will change its position in the coordinate plane.

Definition. The set of all points $M(x, y)$ whose coordinates are defined by formulas (2) is called the *line represented*

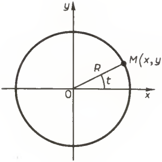


Figure 54

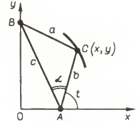


Figure 55

parametrically, and formulas (2) are called the *parametric equations* or parametric representation of this line.

Example 1. Find the line defined by the parametric equations

$$x = R\cos t, \quad y = R\sin t \tag{3}$$

○ Squaring both sides of the equations and adding the results yields $x^2 + y^2 = R^2\cos^2 t + R^2\sin^2 t = R^2$. But $x^2 + y^2 = R^2$ defines a circle of radius R with center at the origin. Thus, any point whose coordinates are defined by formulas (3) lies on this circle.

If t varies from 0 to 2π, then the point $x = R\cos t$, $y = R\sin t$ traverses the entire circle. Therefore formulas (3) with the additional condition $0 \leqslant t \leqslant 2$ consists of the parametric equations of a circle.

The meaning of the parameter t can be understood from Fig. 54: t is the polar angle of the point M belonging to the circle. If we neglect the additional condition ($0 \leqslant t \leqslant 2\pi$), then the point M will traverse the circle many times as t varies from $-\infty$ to $+\infty$. ●

Example 2. A triangle ABC with sides a, b, and c and opposite angles α, β, and γ slides with its vertices A and B along the coordinate axes. Set up the parametric equations of the curve described by the vertex C when it moves. Assume the angle t is the parameter (see Fig. 55).

○ Let the current position of C be defined by the coordinates x and y. We can write $\overrightarrow{OC} = \overrightarrow{OA} + \overrightarrow{AC}$ or in terms of coordinates

$$\langle x, y \rangle = \langle c \cdot \cos(\pi - (t + \alpha), 0 \rangle + \langle b \cos t, b \sin t \rangle$$

Whence

$$x = c \cdot \cos(\pi - (t + \alpha)) + b \cos t, \ y = b \sin t$$

Using reduction formulas we simplify the above expressions:

$$x = -c \cdot \cos(t + \alpha) + b \cos t, \ y = b \sin t. \ \bullet$$

4.3. A STRAIGHT LINE IN THE PLANE AND ITS EQUATION

The position of a straight line in the plane can be defined in different ways. For example, we can specify

(a) two points M_0 and M_1 belonging to the straight line,

(b) one of the points M_0 of the straight line and a vector $\mathbf{a}$ to which the line must be parallel,

(c) one of the points M_0 of the line and a vector $\mathbf{n}$ to which the line must be perpendicular,

(d) one of the points M_0 of the line and an angle φ which the line must form with, say, the x-axis.

We assume in this section that a rectangular Cartesian coordinate system is introduced in the plane.

1. The equation of a straight line which passes through a given point perpendicular to a given vector.

Given a point $M_0(x_0, y_0)$ and a nonzero vector $\mathbf{n} = \langle A, B \rangle$,

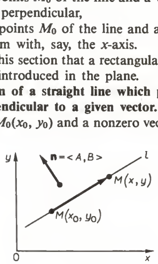


Figure 56

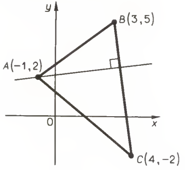


Figure 57

set up the equation of a straight line l passing through M_0 and perpendicular to $\mathbf{n}$ (Fig. 56).

Obviously, $M(x, y)$ belongs to l if and only if the vector $\overrightarrow{M_0M}$ is perpendicular to $\mathbf{n}$. The condition for two vectors to be perpendicular is: the sum of the products of their corresponding coordinates is zero. Noting that the coordinates of the vector $\overrightarrow{M_0M}$ are $x - x_0$, $y - y_0$, we write the condition that $\overrightarrow{M_0M} \perp \mathbf{n}$ in the form

$$A(x - x_0) + B(y - y_0) = 0 \tag{1}$$

Thus, the point $M(x, y)$ belongs to the straight line l if and only if its coordinates satisfy condition (1). Consequently, (1) is the equation of the straight line l.

We arrive at a conclusion: *the equation of the straight line passing through a point* $M_0(x_0, y_0)$ *and perpendicular to the vector* $\mathbf{n} = \langle A, B \rangle$ *has the form in* (1).

Note that a nonzero vector which is perpendicular to a given straight line is called a *normal vector* to the line. There are an infinite number of normal vectors to a straight line, all the vectors being mutually collinear.

Example 1. The equation of a straight line passing through the point $M_0(-1, 2)$ and perpendicular to the vector $\mathbf{n} =$

$\langle 5, -4\rangle$ has the form

$$5(x - (-1)) - 4(y - 2) = 0$$

or, on simplifying, $5x - 4y + 13 = 0$.

Example 2. Three points $A(-1, 2)$, $B(3, 5)$, and $C(4, -2)$ are given in the plane. Ascertain that the figure ABC is a triangle and set up the equation of its altitude drawn from the vertex A (Fig. 57).

○ The points A, B, and C are not collinear since the condition of collinearity is not fulfilled:

$$\frac{3 - (-1)}{1 - (-1)} \neq \frac{5 - 2}{-2 - 2}$$

Consequently, ABC is a triangle.

Before forming the equation of the altitude drawn from the point A, we note that we can take the vector $\vec{BC} = \langle 1, -7\rangle$ as a normal vector to the altitude. Hence it follows that the equation of the altitude has the form $1\cdot(x - (-1)) - 7(y - 2) = 0$ or simpler $x - 7y + 15 = 0$. (There is no need for a drawing since we have solved the problem analytically.) ●

2. The general equation of a straight line.

After removing the brackets equation (1) assumes the form

$$Ax + By + (-Ax_0 - By_0) = 0$$

We use C to denote the number $(-Ax_0 - By_0)$ and get

$$Ax + By + C = 0 \qquad (2)$$

Thus, every straight line is associated with a linear equation in two unknowns x and y, at least one of the coefficients A or B being nonzero (note that A and B are the coordinates of the normal vector to the straight line).

Let us prove the converse: If we have an equation like (2) in which either A or B is nonzero, then this equation defines a straight line.

For example, let $B \neq 0$. Then (2) can be written as

$$A(x - 0) + B\left(y - \left(-\frac{C}{B}\right)\right) = 0$$

According to Item 1, this equation defines a straight line passing through the point M_0 $(0, -C/B)$ and perpendicular to the vector $\mathbf{n} = \langle A, B \rangle$. Now, if $A \neq 0$, the reasoning is similar.

Thus, if a rectangular Cartesian coordinate system is fixed in the plane, we have

(1) *every straight line is defined by a linear equation* $Ax + By + C = 0$ *in which* A *or* B *is nonzero,*

(2) *an equation of the form* $Ax + By + C = 0$ *in which* A *or* B *is nonzero defines a straight line.*

The equation $Ax + By + C = 0$, provided that A or B is nonzero, is called the equation of a straight line in the plane written in a general form or simply the *general equation of a straight line.*

Note an important fact we have established above when considering (2): *the coefficients* A *and* B *in the equation* $Ax + By + C = 0$ *of a straight line are the coordinates of a normal vector to that line.*

3. The equation of a straight line which is parallel to a given straight line and the equation of a straight line which is perpendicular to a given straight line.

Let us consider the general equation of a straight line

$$Ax + By + C = 0 \tag{3}$$

We have noted that the vector $\langle A, B \rangle$ is a normal vector to the straight line.

The following propositions are valid:

(1) *If a straight line* l *is defined by equation* (3), *then any straight line parallel to* l *can be defined by an equation of the form*

$$Ax + By + C' = 0 \tag{4}$$

(2) *If a straight line* l *is defined by equation* (3), *then any straight line which is perpendicular to* l *can be defined by an equation of the form*

$$Bx - Ay + C'' = 0 \tag{5}$$

Let us prove proposition 1. At any value of C' the straight line (4) is parallel to the straight line (3) since it has the same normal vector. On the other hand, whatever the point $M_0(x_0, y_0)$, we can "make" the straight line (4) pass through this point by selecting the value of C' (we must choose $C' = -Ax_0 - By_0$); consequently, by varying C' we can get *any* straight line parallel to l.

Let us now prove proposition 2. Normal vectors to the straight lines (3) and (5) are perpendicular since their scalar product is zero:

$$A \cdot B + B \cdot (-A) = 0$$

Consequently, for any value of C'' the straight line (5) is perpendicular to the straight line (3). On the other hand, whatever the point $M_0(x_0, y_0)$, a suitable selection of C'' can make the straight line (5) pass through this point; this means that by changing C'' we can obtain any straight line perpendicular to l.

Example. A straight line l is defined by the equation $3x - 7y + 12 = 0$. Form the equation of a straight line passing through the point $M_0(5, 1)$ parallel to l and the equation of the straight line passing through M_0 and perpendicular to l.

○ The respective equations of the two lines are $3x - 7y + C'' = 0$ and $-7x - 3y + C'' = 0$. We choose C' and C'' such that the lines both pass through M_0:

$$3 \cdot 5 - 7 \cdot 1 + C' = 0, \quad -7 \cdot 5 - 3 \cdot 1 + C'' = 0$$

Whence $C' = -8$, $C'' = 38$. The desired lines are

$$3x - 7y - 8 = 0, \quad -7x - 3y + 38 = 0 \quad \bullet$$

4. Studying the general equation of a straight line.

Let a straight line in the plane be defined by the equation

$$Ax + By + C = 0$$

where $A \neq 0$ or $B \neq 0$. Let us examine the location of a straight line relative to the coordinate axes depending on whether either of the numbers A, B, C is zero (or nonzero).

(1) If $A = 0$, then the equation assumes the form $By + C = 0$ or $y = -C/B$, which means that all the points of

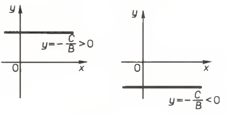


Figure 58

the line have the same ordinate, $-C/B$. Consequently, the straight line is parallel to the x-axis (Fig. 58).

(2) If $B = 0$, then the equation is $Ax + C = 0$ or $x = -C/A$, which means that all the points of the line have the same abscissa, $-C/A$. Thus, the straight line is parallel to the y-axis.

(3) If $C = 0$, then the equation is $Ax + By = 0$. One of the solutions of the equation is the pair of numbers $x = 0$, $y = 0$, which means that the straight line passes through the origin.

(4) If $A = 0$ and $C = 0$, then the equation is $By = 0$ or, since $B \neq 0$, $y = 0$. Thus, the straight line defined by this equation coincides with the x-axis.

(5) If $B = 0$ and $C = 0$, then the equation is $Ax = 0$ or, since $A \neq 0$, $x = 0$. This equation defines the y-axis.

It is easiest to construct a straight line from its equation by finding its two points. When $B \neq 0$, we write the equation in the form

$$y = -\frac{A}{B}x = \frac{C}{B}$$

take two different values of x, and find the corresponding values of y. Now, if $B = 0$, the straight line is parallel to the y-axis; in this case, in order to construct this line, we take any two of its points with abscissa $x = -C/B$.

Example 1. Construct the straight line defined by the equation $3x - 2y + 8 = 0$.

○ We rewrite the equation as $y = (3/2)x + 4$. Then take, for instance, $x_1 = 0$ and $x_2 = -2$ and find that $y_1 = 4$ and $y_2 = 1$.

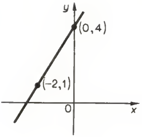


Figure 59

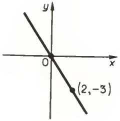


Figure 60

We draw a straight line through the points (0, 4) and $(-2, 1)$ (Fig. 59). ●

Example 2. Construct the straight line defined by the equation $3x + 2y = 0$.

○ Since the constant term is absent, the line passes through the origin. Consequently, we know one point. Let us find another. Setting $x = 2$, we have $y = -3$. We draw a straight line through the origin and the point $(2, -3)$ (Fig. 60). ●

5. The equation of a straight line passing through a given point parallel to a given vector.

Given a point $M_0(x_0, y_0)$ and a nonzero vector $\mathbf{a} = \langle p, q \rangle$, find the equation of the straight line passing through M_0 parallel to $\mathbf{a}$ (Fig. 61).

We restrict our consideration to the case where $p \neq 0$, $q \neq 0$: if $p = 0$ or $q = 0$, then the straight line l is parallel to the y- or x-axis, and its equation has the form $x = x_0$ or $y = y_0$ respectively.

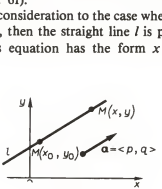


Figure 61

Obviously, a point $M(x, y)$ belongs to l if and only if the vector $\overrightarrow{M_0M}$ is collinear with $\mathbf{a}$, which, in its turn, is the case if and only if the coordinates of $\overrightarrow{M_0M}$ are proportional to the corresponding coordinates of $\mathbf{a}$, i.e.

$$\frac{x - x_0}{p} = \frac{y - y_0}{q} \tag{6}$$

Thus, (6) *is the equation of the straight line passing through* $M_0(x_0, y_0)$ *parallel to* $\mathbf{a} = \langle p, q \rangle$, *where* $p \neq 0$, $q \neq 0$.

A nonzero vector parallel to a straight line is called the *direction vector* of that line. Every straight line has an infinite number of direction vectors, these vectors being mutually collinear.

Proceeding from the foregoing, we can easily solve another problem: to set up the equation of a straight line passing through two given points $M_0(x_0, y_0)$ and $M_1(x_1, y_1)$. We take the vector $\overrightarrow{M_0M_1}$ as a direction vector, whose coordinates are $p = x_1 - x_0$ and $q = y_1 - y_0$ respectively. Assuming that $x_1 \neq x_0$ and $y_1 \neq y_0$, we find from (6) that the equation of l has the form

$$\frac{x - x_0}{x_1 - x_0} = \frac{y - y_0}{y_1 - y_0} \tag{7}$$

Thus, (7) *is the equation of the straight line passing through the two points* $M_0(x_0, y_0)$ *and* $M_1(x_1, y_1)$, *where* $x_1 \neq x_0$, $y_1 \neq y_0$, *or the two-point equation.*

When $x_1 = x_0$, the straight line is parallel to the y-axis and its equation is $x = x_0$. When $y_1 = y_0$, the straight line is parallel to the x-axis and its equation is $y = y_0$.

Example 1. Given three points $A(-1, 7)$, $B(0, 9)$, $C(1, 3)$, set up the equation of the straight line passing through A parallel to the straight line BC.

○ The coordinates of $\overrightarrow{BC}$ are $1 - 0$ and $3 - 9$ respectively, i.e. 1 and -6. We substitute $x_0 = -1$, $y_0 = 7$ (the coordinates of A) and the assumed values $p = 1$, $q = -6$ into equation (6) and obtain

$$\frac{x + 1}{1} = \frac{y - 7}{-6} \quad \text{or} \quad 6x + y - 1 = 0 \; \bullet$$

Example 2. Given a triangle with vertices $A(0, 7)$, $B(2, 5)$, $C(6, -4)$, set up the equation of the median from C.

○ Let P be the midpoint of the side AB whose coordinates are

$$x_0 = \frac{0 + 2}{2} = 1, \quad y_0 = \frac{7 + 5}{2} = 6$$

Since the median passes through C and P, we can write its equation as that of the straight line passing through two points $M_0(1, 6)$ (the point P) and $M_1(6, -4)$ (the point C). We have

$$\frac{x - 1}{6 - 1} = \frac{y - 6}{-4 - 6} \quad \text{or} \quad 2x + y - 8 = 0 \; ●$$

6. An intercept form of the equation of a straight line. Suppose that two points differing from the origin are specified on the coordinate axes: $A(a, 0)$ on the x-axis, with $a \neq 0$, and $B(0, b)$ on the y-axis, with $b \neq 0$. Let us consider the equation

$$\frac{x}{a} + \frac{y}{b} = 1 \tag{8}$$

This equation is satisfied by the coordinates of both A and B. Consequently, (8) defines the straight line AB (Fig. 62).

Equation (8) is called the *intercept equation of a straight line* with the numbers a and b being the intercepts. Any equation of the form

$$Ax + By + C = 0$$

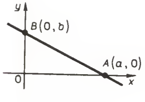


Figure 62

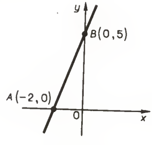


Figure 63

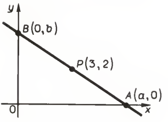


Figure 64

can be reduced to the form of (8) if the three numbers A, B, and C are nonzero. We first write the equation as $Ax + By = -C$, then divide both sides by $-C$ to get

$$\frac{x}{-C/A} + \frac{y}{-C/B} = 1$$

This is the intercept form of the equation of a straight line, where $a = -C/A$ and $b = -C/B$.

For example, the equation $5x - 2y + 10 = 0$ can be transformed as above to get

$$\frac{x}{-2} + \frac{y}{5} = 1$$

This equation defines the straight line passing through the points $(-2, 0)$ and $(0, 5)$ (Fig. 63).

Note that intercept equations are convenient for constructing straight lines.

Example. We draw a straight line through the point $P(3, 2)$ so that P is the midpoint of the line segment AB (Fig. 64).

○ We write the equation of the desired straight line as

$$\frac{x}{a} + \frac{y}{b} = 1$$

Since the point P is the midpoint of AB, where $A(a, 0)$ and $B(0, b)$, we have

$$3 = \frac{a + 0}{2}, \quad 2 = \frac{0 + b}{2}$$

that is, $a = 6$ and $b = 4$. Thus, the equation of our straight line is

$$\frac{x}{6} + \frac{y}{4} = 1 \bullet$$

7. Slope of a straight line.

Consider a straight line l in the plane. The *angle of inclination* of l to the x-axis is φ (Fig. 65), through which we must turn the x-axis for it to coincide with l; such angles are measured counterclockwise.

This definition is not unique. In fact, if we turn the x-axis through an angle of $\varphi + \pi n$, where n is an integer, it will also coincide with l. The angle of inclination φ is the least nonnegative angle of rotation, i.e. such that $0 \leqslant \alpha < \pi$.

The quantity $k = \tan \varphi$ is called the *slope* of the straight line l.

If $\varphi = \pi/2$, then $\tan \varphi = 0$. Hence, a straight line which is parallel to the y-axis has no slope. All other straight lines have slopes.

If $\varphi = 0$, then $\tan \varphi = 0$. Hence the straight line is parallel to the x-axis and its slope is zero. Conversely: if $k = 0$, then $\tan \varphi = 0$ and hence $\varphi = 0$, and this means that the straight line is parallel to the x-axis.

For finding the slope of a straight line, we must know its two points. If $M_0(x_0, y_0)$ and $M_1(x_1, y_1)$ are two points of a

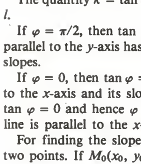


Figure 65

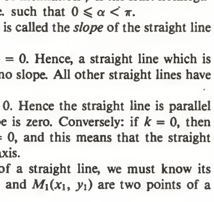


Figure 66

straight line l, then its slope is

$$k = \frac{y_1 - y_0}{x_1 - x_0} \tag{9}$$

The proof follows directly from Fig. 66. If two straight lines are parallel, then their respective slopes are equal.

Example 1. The slope of the straight line passing through points $A(a, 0)$ and $B(0, b)$ is

$$k = \frac{b - 0}{0 - a} = -\frac{b}{a}$$

We assume here that a is nonzero.

Example 2. The slope of the straight line passing through the origin $O(0, 0)$ and a point $A(a, b)$ is

$$k = \frac{b - 0}{a - 0} = \frac{b}{a}$$

We assume here that a is nonzero.

8. The equation of a straight line passing through a given point and having a given slope.

Suppose we have a point $M_0(x_0, y_0)$ in the plane and a slope k. Let us derive the equation of the straight line which passes through M_0 and have the slope k.

If a point $M(x, y)$ belongs to l, then from (9) we have

$$\frac{y - y_0}{x - x_0} = k$$

Now, if M does not lie on l, then the above equation is not satisfied. Thus this equation is the equation of the straight line l. We rewrite it as

$$y - y_0 = k(x - x_0) \tag{10}$$

Thus, (10) is the *point-slope equation of the straight line l that passes through $M_0(x_0, y_0)$ and has slope k.*

If a point M_0 lies on the y-axis, i.e. $M_0(0, b)$, then (10) assumes the form $y - b = k(x - 0)$ or

$$y = kx + b \tag{11}$$

This is the form of the equation of the straight line that has slope k and whose y-intercept equals b.

Note that the equation of any straight line having a slope, i.e. which is not parallel to the y-axis, can be written in the form (10) or (11).

Equation $y = kx + b$ is the *slope-intercept equation of the straight line with slope k and y-intercept b.* If the straight line defined by a general equation

$$Ax + By + C = 0 \tag{12}$$

is not parallel to the y-axis (i.e. $B \neq 0$), then in order to reduce its equation to the form in (11) it is sufficient to write (12) as $By = -Ax - C$ and then divide its both sides by B. We get

$$y = -\frac{A}{B}x - \frac{C}{B}$$

Whence it follows that the slope of the line (12) is

$$k = -\frac{A}{B} \tag{13}$$

Example 1. Find the slope of the straight line $3x + 5y - 7 = 0$.

○ We reduce this equation to the form $y = kx + b$. We have $5y = -3x + 7$ or $y = -(3/5)x + 7/5$. Thus, the slope is $k = -3/5$. ●

Example 2. Derive the equation of the straight line that passes through the point $M_0(2, 1)$ and has the slope $k = -1$.

○ Substituting the data into equation (10) yields $y - 1 = -1 \cdot (x - 2)$ or $x + y - 3 = 0$. ●

Example 3. A light ray is directed along the straight line $y = 3x - 5$. When it reaches the x-axis it is reflected. Find the equation of the reflected ray.

○ We find the point at which the incident ray and the x-axis intersect. Since the ordinate of the point of reflection is zero, we put $y = 0$ in the equation of the ray. We have $3x - 5 = 0$ or $x = 5/3$. We know a point $M_0(5/3, 0)$ on the reflected ray; it remains to find its slope. If the angle of inclination of the incident ray to the x-axis is φ, then the angle of inclination of

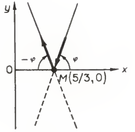


Figure 67

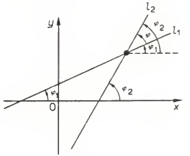


Figure 68

the reflected ray is φ (Fig. 67). We know that $\tan(-\varphi) = -\tan\varphi$. We can see from the equation $k = 3$ for the incident ray. Hence, for the reflected ray $k = -3$. Substituting the coordinates of M_0 and the value of k into (10), we have $y - 0 = -3(x - 5/3)$ or $3x + y - 5 = 0$, which is the equation of the reflected ray.

9. The angle between straight lines.

Let us consider two straight lines l_1 and l_2 in the plane (Fig. 68). One of them we call the first and the other the second. Let k_1 and k_2 be the slopes of these lines.

The angle φ through which we must turn the first straight line for it to coincide with the second is the *slope of the second straight line to the first*. The sign of the angle φ depends on whether such a rotation is counterclockwise or clockwise. Obviously, the angle φ is determined to within the term multiple of π.

If φ_1 and φ_2 are the slopes of the first and second straight lines to the x-axis respectively, then $\varphi = \varphi_2 - \varphi_1$. Whence

$$\tan\varphi = \tan(\varphi_2 - \varphi_1) = \frac{\tan\varphi_2 - \tan\varphi_1}{1 + \tan\varphi_2\tan\varphi_1}$$

But $\tan\varphi_1 = k_1$ and $\tan\varphi_2 = k_2$, thus,

$$\tan\varphi = \frac{k_2 - k_1}{1 + k_2k_1} \tag{14}$$

We have obtained a *formula for finding the angle between two straight lines, provided we know the slopes of the lines.*

Formula (14) is senseless if the denominator on its right-hand side is zero. In this case $\tan\varphi$ does not exist, and this means that $\varphi = \pi/2$. Hence we have the *condition for two straight lines to be perpendicular*: $k_1k_2 + 1 = 0$.

This condition is usually written as

$$k_2 = -\frac{1}{k_1} \quad (\text{if } k_1 \neq 0) \tag{15}$$

Example 1. Find the angle φ between the straight lines $3x - 5y + 7 = 0$ and $2x - 3x + 4 = 0$.

○ We find the slopes of the lines: $k_1 = 3/5$ and $k_2 = 2/3$. Whence

$$\tan\varphi = \frac{\frac{2}{3} - \frac{3}{5}}{1 + \left(\frac{2}{3}\right)\cdot\left(\frac{3}{5}\right)} = \frac{1}{21}$$

From tables of trigonometric functions we find that $\varphi \approx 2°48'$. ●

Example 2. A ray of light is directed along the straight line $y = 2x - 4$. On reaching the straight line $y = x$, the ray is reflected. Find the equation of the reflected ray.

○ Find the point at which the incident ray and the line $y = x$ intersect. Solving the equations $y = 2x - 4$ and $y = x$ simultaneously yields $x = 4$, $y = 4$; the intersection point is $M_0(4, 4)$.

If we denote the angle of inclination of the incident ray to the line $y = x$ by φ, then we have

$$\tan\varphi = \frac{2 - 1}{1 + 2\cdot 1} = \frac{1}{3}$$

The angle of inclination of the reflected ray to the line $y = x$ is obviously $-\varphi$, and its tangent $-\tan\varphi = -1/3$. Then denoting the slope of the reflected ray by k and using (14), we can write

$$-\frac{1}{3} = \frac{k - 1}{1 + k\cdot 1}$$

whence it follows that $k = 1/2$.

Thus, the reflected ray passes through the point $M_0(4, 4)$ and its slope is $k = 1/2$. The equation of this ray is

$$y - 4 = \frac{1}{2}(x - 4)$$

or

$$x - 2y + 4 = 0 \quad \bullet$$

4.4. RELATIVE POSITION OF TWO STRAIGHT LINES IN THE PLANE

Suppose that we have two straight lines l_1 and l_2 in the plane which are defined by the equations

$$A_1x + B_1y + C_1 = 0 \tag{1}$$

for l_1 and

$$A_2x + B_2y + C_2 = 0 \tag{2}$$

for l_2.

Let us show how we can find the relative position of the straight lines l_1 and l_2 from their equations. Three cases are possible.

Case 1. The lines coincide. Whence it follows that their normal vectors, $\mathbf{n}_1 = \langle A_1, B_1 \rangle$ and $\mathbf{n}_2 = \langle A_2, B_2 \rangle$, are collinear. Since $\mathbf{n}_1 \neq 0$, there is a number λ such that $\mathbf{n}_2 = \lambda\mathbf{n}_1$, that is, $A_2 = \lambda A_1$ and $B_2 = \lambda B_1$. We write the equations of the straight lines as

$$A_1x + B_1y + C_1 = 0 \quad \text{(for } l_1\text{)}$$

$$\lambda A_1x + \lambda B_1y + C_2 = 0 \quad \text{(for } l_2\text{)}$$

Any point of l_1 (or of l_2) satisfies each of the above equations and hence the equation

$$0 \cdot x + 0 \cdot y + (C_2 - \lambda C_1) = 0 \tag{3}$$

which results from the subtraction of both sides of the first equation multiplied by λ from the respective sides of the second equation. An equation of form (3) has a solution if its constant term is zero, i.e. $C_2 = \lambda C_1$.

Thus, the *condition for the straight lines l_1 and l_2 to coincide is the existence of a number λ such that $A_2 = \lambda A_1$, $B_2 = \lambda B_1$, $C_2 = \lambda C_1$.*

This condition means that one of equations (1), (2) is obtained from the other by multiplying both sides by the same number λ.

Case 2. The straight lines do not coincide but are parallel. As in case (1) it follows that their normal vectors are collinear, i.e. $A_2 = \lambda A_1$, $B_2 = \lambda B_1$. Since l_1 and l_2 do not coincide, we must have $C_2 \neq \lambda C_1$.

Thus, the *condition for the straight lines l_1 and l_2 to be parallel and not coincident is the existence of a number λ such that $A_2 = \lambda A_1$, $B_2 = \lambda B_1$, $C_2 \neq \lambda C_1$.*

Example. The straight lines $6x - 2y + 4 = 0$ and $9x - 3y + 6 = 0$ coincide since $6 : 9 = (-2):(-3) = 4:6$. These lines are parallel since $6:9 = (-2):(-3) \neq 4:5$.

Case 3. The straight lines l_1 and l_2 are not parallel, they intersect at a single point. In this case their normal vectors, $\mathbf{n}_1$ and $\mathbf{n}_2$, are not collinear. Conversely, if $\mathbf{n}_1$ and $\mathbf{n}_2$ are not collinear, then l_1 and l_2 are not parallel.

Thus, the *condition for the straight lines l_1 and l_2 to be not parallel is that their normal vectors are not collinear.* Noting that vectors $\langle A_1 B_1 \rangle$ and $\langle A_2 B_2 \rangle$ are not collinear if and only if the determinant formed from their coordinates is nonzero (see Sec. 3.10), we can infer that the *condition for the straight lines l_1 and l_2 to be not parallel is*

$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} \neq 0 \tag{4}$$

If l_1 and l_2 are not parallel, we can find their point of intersection by solving the system of equations (1) and (2). According to (4), the solution can also be found using Cramer's rule.

Example. Verify that the straight lines $2x - 3y + 5 = 0$ and $x - 5y + 6 = 0$ are not parallel and find the point at which they intersect.

○ The lines are not parallel since

$$\begin{vmatrix} 2 & -3 \\ 1 & -5 \end{vmatrix} = -7 \neq 0$$

We find the point of their intersection by solving the system

$$\begin{cases} 2x - 3y = -5 \\ x - 5y = -6 \end{cases}$$

According to Cramer's rule,

$$x = \frac{\begin{vmatrix} -5 & -3 \\ -6 & -5 \end{vmatrix}}{-7} = -1, \quad y = \frac{\begin{vmatrix} 2 & -5 \\ 1 & -6 \end{vmatrix}}{-7} = 1.$$

Thus, the lines intersect at the point $(-1, 1)$. ●

4.5. PARAMETRIC EQUATIONS OF A STRAIGHT LINE

Suppose a straight line passes through a point M_0 and parallel to a vector $\mathbf{a}$ (Fig. 69). Then the condition for a point M to belong to this line is the collinearity of the vectors $\overrightarrow{M_0M}$ and $\mathbf{a}$. In other words, the point M lies on this line if and only if the following equation is satisfied for a number t:

$$\overrightarrow{M_0M} = t\mathbf{a} \tag{1}$$

We introduce a rectangular Cartesian coordinate system in the plane, specify a point $M_0(x_0, y_0)$ and a vector $\mathbf{a} = \langle p, q \rangle$ in this system, and denote the coordinates of the running point

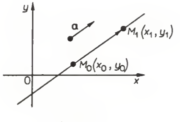


Figure 69

M by x and y. Then the vector equation (1) takes the form $x - x_0 = tp$, $y - y_0 = tq$ or

$$x = x_0 + tp, \quad y = y_0 + tq \tag{2}$$

If the parameter t varies from $-\infty$ to $+\infty$ then the point M, whose coordinates x and y are defined by formulas (2), runs along the entire straight line.

Thus, formulas (2) are *parametric equations of the straight line passing through the point $M_0(x_0, y_0)$ and parallel to the vector $\mathbf{a} = \langle p, q \rangle$.*

Example. The parametric equations of the straight line passing through the point $M_0(2, -1)$ and parallel to the vector $\mathbf{a} = \langle 7, 3 \rangle$ have the form $x = 2 + 7t$, $y = -1 + 3t$.

4.6. DISTANCE BETWEEN A POINT AND A STRAIGHT LINE

Given in the plane a point $M_0(x_0, y_0)$ and a straight line l defined by the equation $Ax + By + C = 0$, find the distance d from M_0 to l.

If $M_1(x_1, y_1)$ is a point of l, then the desired distance (Fig. 70) is

$$d = |\text{proj}_{\mathbf{n}} \overrightarrow{M_0M_1}|$$

where $\text{proj}_{\mathbf{n}}$ is the projection of a vector $\overrightarrow{M_0M_1}$ on $\mathbf{n}$ and $\mathbf{n}$ is

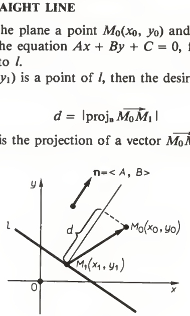


Figure 70

a normal vector to l; or else

$$d = \frac{|\overrightarrow{M_0M_1} \cdot \mathbf{n}|}{|\mathbf{n}|}$$

We can take the vector $\mathbf{n} = \langle A, B \rangle$ as a normal vector to the straight line l. Noting that $\overrightarrow{M_0M_1} = \langle x_1 - x_0, y_1 - y_0 \rangle$, we obtain

$$d = \frac{|A(x_1 - x_0) + B(y_1 - y_0)|}{\sqrt{A^2 + B^2}} \quad (1)$$

Since M_1 belongs to l, we have $Ax_1 + By_1 + C = 0$. Thus we can write formula (1) as

$$d = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}} \quad (2)$$

We have arrived at the following rule: *in order to find the distance from a point M_0 to a straight line l defined by the general equation we have to substitute the coordinates of M_0 for the running coordinates x, y in this equation and divide the absolute value of the result by $\sqrt{A^2 + B^2}$.*

Example 1. Given the equation $3x - 4y + 10 = 0$ of a straight line, find the distance between the point $M(4, 3)$ and this line.

○ From formula (2) we have

$$d = \frac{|3 \cdot 4 - 4 \cdot 3 + 10|}{\sqrt{3^2 + 4^2}} = \frac{10}{5} = 2 \quad \bullet$$

Example 2. Find the distance between the straight lines $3x - 4y + 10 = 0$ and $3x - 4y - 5 = 0$.

○ We choose a point (3, 1) on the second line and find its distance to the first line:

$$d = \frac{|3 \cdot 3 - 4 \cdot 1 + 10|}{\sqrt{3^2 + 4^2}} = \frac{15}{5} = 3 \quad \bullet$$

4.7. HALF-PLANES DEFINED BY A STRAIGHT LINE

Suppose we have a straight line in the plane in rectangular Cartesian coordinates:

$$Ax + By + C = 0 \tag{1}$$

This line divides the plane into two half-planes; the line is their boundary or edge (Fig. 71).

Let us prove the following proposition: *one of the half-planes defined by* (1) *is composed of points* (x, y) *for which*

$$Ax + By + C \geqslant 0 \tag{2}$$

and the other is composed of points (x, y) *for which*

$$Ax + By + C \leqslant 0 \tag{3}$$

□ First we assume that $B \neq 0$. Then (1) is equivalent to the equation $y = kx + b$. Inequalities (2) and (3) reduce to the forms $y \geqslant kx + b$ and $y \leqslant kx + b$ respectively. The first inequality defines the half-plane lying above the straight line $y = kx + b$ (Fig. 72) and the second the half-plane lying below the line. Now, if $B = 0$, then A is necessarily nonzero and (1) is equivalent to the equation $x = c$; (2) and (3) reduce to the forms $x \geqslant c$ and $x \leqslant c$ respectively. The first inequality defines the half-plane lying to the right of the straight line $x = c$ and the second the half-plane lying to the left of the line. ■

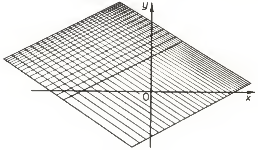


Figure 71

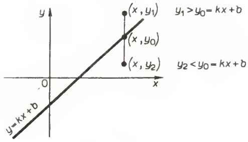


Figure 72

Example 1. Check whether the points $A(-1, -3)$ and $B(2, 2)$ lie on one side or on different sides of the straight line $2x - y - 3 = 0$.

○ We substitute the coordinates of A and B into the left-hand side of the equation. We have $2\cdot(-1) - (-3) - 3 = -2$ and $2\cdot 2 - 2 - 3 = -1$. The resulting numbers have the same sign. Hence the points A and B lie on one side of the line. ●

Exercises to Chapter 4

4.1

1. Find the lines defined by the following equations: (a) $x = 0$, (b) $xy = 0$, (c) $(x - 2)(x + 3) = 0$, (d) $|x| - |y| = 0$, (e) $x^2 - y^2 = 0$.

2. A straight line is displaced so that the area S of a triangle which the line forms with the coordinate axes remains constant. Find the equation of the line described by the midpoint of the segment of the straight line intercepted by the coordinate axes.

3. Describe the line defined in polar coordinates by the equation: (a) $\varphi = 0$, (b) $r = 1$, (c) $r = \varphi$ (the spiral of Archimedes).

4. In order to balance a body of mass m on an inclined plane forming an angle α with the horizontal plane, we must apply a force $Q = mg \sin \alpha$ (Fig. 73). Graph the function $Q(\alpha)$ in polar coordinates assuming that α is the polar angle and Q is the polar radius.

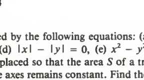


Figure 73

5. Find the points of intersection of two circles. One circle is centered at a point $(a, 0)$ on the x-axis with radius a, and the other at a point $(0, b)$ on the y-axis with radius b.

4.2

1. A body of mass m is thrown with an initial velocity v_0 at an angle α to the horizon. Find the parametric equations of its trajectory with time t as a parameter.

2. A thread is wound onto a fixed disk of radius 1. Then the thread is wound off by pulling its free end so that there is no sag. Compile the parametric equations of the trajectory of its free end. Take the center of the disk as the origin and direct the x-axis toward the free end of the thread (before unwinding). Assume that the length of the unwounded thread is the parameter.

4.3

1. Compile the equation of the indicated straight line:

(a) through the point $(3, -1)$ and parallel to the bisector of the first quadrant angle,

(b) through the point $(3, -1)$ and perpendicular to the bisector of the first quadrant angle.

2. Drop the perpendiculars to the straight line $7x + 3y - 21 = 0$ from the points of intersection with the coordinate axes. Find the equations of the perpendiculars.

3. Find a point B symmetric to $A(-1, 10)$ with respect to a straight line $x - 2y + 6 = 0$.

4. Given the vertices $A(2, -1)$, $B(-1, 3)$, $C(4, 0)$ of the triangle, write the equations of its sides.

5. Find the equations of the straight lines drawn through the vertices of ΔABC (see the previous exercise) and parallel to the opposite sides.

6. Find the points at which the straight line drawn through the points $(-7, 1)$ and $(5, 6)$ intersects the coordinate axes.

7. Given the equations of the sides of ΔABC: $3x + 2y - 8 = 0$ (for AB), $4x - y - 7 = 0$ (for AC), and $10x - 3y + 41 = 0$ (for BC), find the coordinates of its vertices.

8. Compile the equations of the altitudes of ΔABC (see the previous exercise).

9. Find the equation in intercept form of the straight lines: $y = -2x + 3$, $5x + 3y - 15 = 0$, and $3x - 2y + 12 = 0$.

10. Find the area of a triangle formed by the straight line $-3x + 2y + 6 = 0$ and the coordinate axes.

11. Write the equation of a side of a square whose diagonals serve as coordinate axes. The length of a side is a.

12. Draw a straight line through the point $(3, -4)$ so that the area of the triangle it forms with the coordinate axes equals 3.

13. Find the slopes of the straight lines $x/2 + y/7 = 1$ and $9x - 7y + 1 = 0$.

14. Draw a straight line through the point $(-1, 7)$ perpendicular to the straight line $y = (2/5)x + 7$.

15. Form the equations of the legs of an isosceles right triangle if the equation of its base is $y = 3x + 5$ and $(4, -1)$ is the vertex of the right angle.

16. A ray is directed from the point $(6, 9)$ to the straight line $2x - 5y + 6 = 0$ at an angle of 45°. Find the equation of a ray reflected from this straight line.

17. Given in an isosceles triangle the equation of its base $2x + y + 3 = 0$, the equation of one leg $x + 4y + 5 = 0$, and the point $(-28/5, 6/5)$ lying on the other leg, find the equation of the other leg.

4.4

1. Find the relative position of the indicated straight lines:

(a) $3x - 6y - 9 = 0$, $-2x + 4y + 6 = 0$ (b) $3x - 6y - 9 = 0$, $-2x + 4y + 5 = 0$

(c) $y = -(5/3)x + 2$, $5x + 3y - 1 = 0$ (d) $5x - 3y + 11 = 0$, $-4x - 7y + 10 = 0$

2. Find the points on the y-axis which are equidistant from the origin and from the straight line $4x - 3y + 12 = 0$.

3. Find the point on the straight line $x + y = 1$ equidistant from the straight lines $2x - 9y = 1$ and $6x - 7y = 2$.

4. Given the triangle with the vertices $A(2, 1)$, $B(-13, 5)$, $C(7, 3)$, find the altitude from C to the side AB.

5. Find the distance between two straight lines $3x - 4y + 12 = 0$ and $3x - 4y - 7 = 0$.

6. Set up the equations of straight lines which are parallel to the straight line $3x - 4y + 12 = 0$ and a distance 7 units away from it.

7. One side of a square lies on the straight line $x - 3y + 1 = 0$ and one of its vertices is at the point $(3, 0)$. Find the equations of its remaining sides.

8. Through the origin draw tangents to the circle $(x + 4)^2 + (y - 3)^2 = 16$.

4.7

1. Check whether the point $A(5, 2)$ and the origin lie on one side or on different sides of the indicated straight lines: (a) $7x - 12y + 19 = 0$, (b) $2x - 9y + 3 = 0$.

2. Check whether the straight line $5x + 4y - 20 = 0$ intersects the line segment AB, where $A(3, 1)$ and $B(6, -1)$.

3. Given the points $A(3, 2)$, $B(1, 1)$, and $C(-2, 0)$, find whether the entire triangle ABC lies on one side of the straight line $5x - 3y - 6 = 0$.

Chapter 5

CONIC SECTIONS

This chapter deals with *second-degree curves* or *conic sections.* In rectangular Cartesian coordinates in the plane these curves are defined by second-degree algebraic equations whose general form is

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

where at least one of the coefficients A, B, C is nonzero.

The *ellipse*, *hyperbola*, and *parabola* are most interesting conic sections and are often used in mathematics and its applications.

5.1. THE ELLIPSE

1. The definition and canonical equation of an ellipse.

Definition. An *ellipse* is the set of all points in a plane such that the sum of the distances from two fixed points, called the foci, is a constant greater than the distance between the foci.

A *special case of an ellipse* when its foci coincide is the *circle* which is the set of all points in the plane a given distance called the *radius* away from a fixed point.

We can construct an ellipse from its definition (Fig. 74). Take a sheet of plywood, and fix two pins at points F_1 and F_2. Then make a loop of a thread and put it over both pins. If we pull the thread taut with the tip of a pencil and move the pencil over the surface of the plywood so that the thread remains taut, then the line left behind will be an ellipse with foci F_1 and F_2. By changing the distance F_1F_2 between the foci and the length of the thread, we get different ellipses.

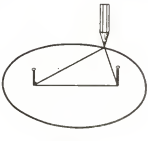


Figure 74

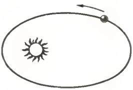


Figure 75

The ellipse is the most common curve in Nature and technology. For instance, the planets orbit the sun along ellipses with the sun at one of the foci (Fig. 75), if a circular cylinder is cut by a plane not parallel to its axis, then the section is an ellipse (Fig. 76).

The notion of an *ellipsoid of inertia* is used in mechanics. This is an imaginary curve whose shape depends on the shape of a plate (Fig. 77).

Let us derive the equation of an ellipse. Suppose that two points F_1 and F_2 are the foci of an ellipse and M its point. We use $2c$, r_1, and r_2 to denote the lengths of the line segments F_1F_2, F_1M, and F_2M (Fig. 78).

The sum of the distances F_1M and F_2M is a constant characterizing the ellipse. We denote this constant by $2a$. By definition $2a > 2c$. Note that if $2a = 2c$, then $F_1M + F_2M = 2a$ is only

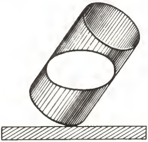


Figure 76

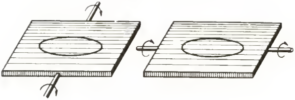


Figure 77

satisfied by the point of the line segment F_1F_2, while for $2a < 2c$ none of the points on the plane satisfy this condition.

Thus, by the definition of an ellipse,

$$r_1 + r_2 = 2a \tag{1}$$

Let us introduce a rectangular Cartesian coordinate system in the plane so that the x-axis passes through the foci and the y-axis divides F_1F_2 in half (Fig. 79). Then (x, y), $(-c, 0)$ and $(c, 0)$ are the coordinates of the points M, F_1, and F_2 respectively. Using the formula for the distance between two points we can write

$$r_1 = \sqrt{(x + c)^2 + y^2}, \quad r_2 = \sqrt{(x - c)^2 + y^2}$$

Substituting these expressions into (1) yields

$$\sqrt{(x + c)^2 + y^2} + \sqrt{(x - c)^2 + y^2} = 2a \tag{2}$$

and this is the equation of an ellipse since it is satisfied by the coordinates of any point M belonging to the ellipse and is not

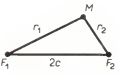


Figure 78

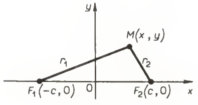


Figure 79

satisfied by the coordinates of points not belonging to it (for such points we have $r_1 + r_2 \neq 2a$).

Equation (2) can be simplified. We first write it as

$$\sqrt{(x + c)^2 + y^2} = 2a - \sqrt{(x - c)^2 + y^2}$$

and square both sides of the last equation:

$$(x + c)^2 + y^2 = 4a^2 - 4a\sqrt{(x - c)^2 + y^2} + (x - c)^2 + y^2$$

By collecting like terms we have

$$4cx = 4a^2 - 4a\sqrt{(x - c)^2 + y^2}$$

or

$$a\sqrt{(x - c)^2 + y^2} = a^2 - cx$$

Then we square both sides of the last equation:

$$a^2x^2 + a^2c^2 - 2a^2cx + a^2y^2 = a^2 - 2a^2cx + c^2x^2$$

The result of a simple transformation is

$$(a^2 - c^2)x^2 + a^2y^2 = a^2(a^2 - c^2)$$

By assumption, the difference $a^2 - c^2$ is positive ($a > c$). We use b to denote the square root of this number: $b = \sqrt{a^2 - c^2}$. Thus, $b^2 = a^2 - c^2$, $b < a$. By dividing both sides of the last equation by a^2b^2, we finally have

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \qquad (3)$$

This equation is obtained from (2). Proceeding in the reverse way we can show (check this) that equation (2) follows from (3). Thus, these equations are equivalent. Equation (3) is called the *canonical equation of an ellipse*. This is a second-degree equation and, hence, an ellipse is a second-degree curve.

Remark. We have noted that in the given coordinate system the parameter a in (3) is greater than b. However, equation (3) can also define an ellipse when a is smaller than b if we change the coordinate axes, thereby reducing the situation to the case when a is greater than b. Hence, (3), when $a < b$, defines an

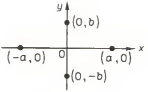


Figure 80

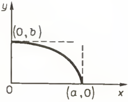


Figure 81

ellipse whose foci lie on the y-axis and the distance between the origin and the foci is $c = \sqrt{b^2 - a^2}$.

2. Studying the shape of an ellipse from its canonical equation.

If an ellipse is defined by a canonical equation and (x, y) is a point on it, then the points $(x, -y)$, $(-x, y)$, and $(-x, -y)$ also belong to the ellipse. Indeed, squaring the numbers $-x$ and $-y$ yields x^2 and y^2 and equation (3) is satisfied for all the indicated points. Hence it follows that the x- and y-axes are the axes of symmetry of an ellipse and the origin is the center of symmetry.

The points at which the ellipse cuts the coordinate axes are *vertices*. Putting $y = 0$ we have from (3) $x^2/a^2 = 1$, i.e. $x = \pm a$. Thus, the vertices lying on the x-axis have coordinates $(-a, 0)$ and $(a, 0)$. Similarly, we find that the coordinates of the vertices lying on the y-axis are $(0, -b)$ and $(0, b)$ (Fig. 80).

Since the ellipse is symmetric with respect to the x- and y-axes, it is sufficient to study its shape within the first quadrant, i.e. for $x \geqslant 0$ and $y \geqslant 0$. In this case we find from (3) that

$$y = \frac{b}{a}\sqrt{a^2 - x^2}$$

This relationship shows that as x increases from 0 to a the coordinate y decreases from b to 0. As x increases further the difference $a^2 - x^2$ becomes negative and y does not exist, hence it follows that the ellipse has no points with abscissas x greater than a. The part of the ellipse we have just considered is depicted in Fig. 81. We can show that the part of the ellipse lying

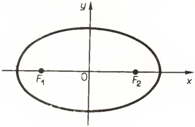


Figure 82

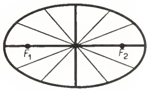


Figure 83

in the first quadrant is convex upward and the tangent at the vertex $(a, 0)$ is parallel to the y-axis, while the tangent at the vertex $(0, b)$ is parallel to the x-axis. The two tangents are shown by dashed lines in Fig. 81.

The parts of the ellipse in the other quadrants are symmetric to the part in the first quadrant with respect to the coordinate axes (Fig. 82).

The center of symmetry of the ellipse is simply called the *center* and it is the midpoint of the line segment F_1F_2 joining its foci. Any chord passing through the center of an ellipse is a *diameter*; Fig. 83 shows several diameters. The diameter passing through the foci of an ellipse is called the *major axis* and the diameter which passes through the center and perpendicular to the major axis is called the *minor axis* of the ellipse.

If, as before, the x-axis passes through the foci F_1 and F_2 of an ellipse and the origin is the midpoint of the line segment F_1F_2, then the major axis coincides with the x-axis and the minor axis with the y-axis; in this case $a > b$ in equation (3) of the ellipse since $b^2 = a^2 - c^2$. In the case where $a < b$ we find that the major and minor axes lie on the y- and x-axes respectively.

The numbers a and b in the canonical equation (3) are the lengths of the corresponding line segments which join the center of an ellipse with its vertices and are called the *semiaxes*. If $a > b$, then a is the *semimajor axis* and b is the *semiminor axis*: if $a < b$, then b is the semimajor axis and a is the semiminor axis. In any case it is customary to denote the semiaxis lying on the x-axis by a and the other semiaxis by b.

Example 1. Find the equation of an ellipse if its major axis is 10 and the distance between the foci is 8. Assume that the major axis coincides with the x-axis and the midpoint of the segment joining the foci coincides with the origin.

○ The equation of the ellipse in the given coordinate system is its canonical equation (3). Here we have $b^2 = a^2 - c^2 = 5^2 - 4^2 = 9$, hence $b = 3$. The desired equation is

$$\frac{x^2}{5^2} + \frac{y^2}{3^2} = 1 \quad \bullet$$

Example 2. Show that the curve $9x^2 + 4y^2 = 36$ is an ellipse. Find the lengths of its axes and the coordinates of its foci.

○ Dividing both sides of the equation by the constant term yields

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

thus the given curve is the ellipse whose semimajor axis is 3 and the semiminor axis is 2. The foci F_1 and F_2 lie on the y-axis $2c$ apart, where $c^2 = 3^2 - 2^2$, i.e. $c = \sqrt{5}$. The coordinates of the foci are $(0, -\sqrt{5})$ and $(0, \sqrt{5})$.

3. The ellipse as a section of a circular cylinder.

We show that *any section of a circular cylinder by a plane not parallel to its axis is an ellipse.*

□ Consider a circular cylinder and a plane α intersecting it at a point O and use l to denote the curve of intersection.

Draw a plane $\tilde{\alpha}$ through O perpendicular to the axis of the cylinder; the section is the circle $\tilde{l}$ with radius r. We choose the coordinate axes in each plane α and $\tilde{\alpha}$ as shown in Fig. 84 (the x-axis is common to both planes and the axes of ordinates are the y-axis in the α plane and the $\tilde{y}$-axis in the $\tilde{\alpha}$ plane).

Let $M(x, y)$ be an arbitrary point on the curve l. Its projection on the plane $\tilde{\alpha}$ is the point $\tilde{M}$ on the circle l. Let x and $\tilde{y}$ be the coordinates of $\tilde{M}$ in the plane $\tilde{\alpha}$. We can see from the figure that $\tilde{y} = y \cos \varphi$, where φ is an (acute) angle between

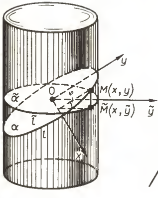


Figure 84

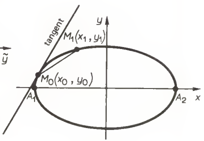


Figure 85

the planes α and $\tilde{\alpha}$. Since $\tilde{M}(x, \tilde{y})$ lies on the circle with radius r and center at the origin, we have $x^2 + \tilde{y}^2 = r^2$, whence

$$x^2 + y^2 \cos^2 \varphi = r^2$$

Thus, the coordinates of any point $M(x, y)$ of the curve l are related as follows

$$\frac{x^2}{r^2} + \frac{y^2}{\left(\frac{r}{\cos \varphi}\right)^2} = 1 \tag{4}$$

Conversely, if point $M(x, y)$ in the plane α satisfies (4), then its projection $\tilde{M}(x, \tilde{y})$, where $\tilde{y} = \tilde{y} \cos \varphi$, satisfies equation $x^2 + \tilde{y}^2 = r^2$, i.e. lies on the circle $\tilde{l}$, hence the point M lies on the curve l.

Thus, the curve of intersection of a cylinder and a plane is defined by equation (4), that is, is an ellipse with the semiminor axis r and the semimajor axis $r/\cos \varphi$, where r is the radius of the section of the cylinder and φ is the angle of inclination of the cutting plane to the perpendicular plane to the axis of the cylinder. ■

4. The equation of a tangent to an ellipse.

Suppose an ellipse is specified by its equation in standard

form

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{5}$$

We choose a point $M_0(x_0, y_0)$ on the ellipse, where $y_0 \neq 0$ (that is, a point which is distinct from the vertices A_1 and A_2 (Fig. 85)). A small portion of the ellipse near M_0 is the graph of a function $y = y(x)$ (we can explicitly express y in terms of x using (5)). Differentiating both sides of the equation

$$\frac{x^2}{a^2} + \frac{y^2(x)}{b^2} = 1$$

with respect to x yields

$$\frac{2x}{a^2} + \frac{2yy'}{b^2} = 1$$

whence it follows that

$$y' = -\frac{b^2}{a^2}\frac{x}{y}$$

This means, for one, that the tangent at M_0 has the slope

$$k = -\frac{b^2}{a^2}\frac{x_0}{y_0}$$

If we know the slope of a tangent and a point $M_0(x_0, y_0)$ on the tangent, then we can write the equation of the tangent as $y - y_0 = k(x - x_0)$ or

$$y - y_0 = -\frac{b^2}{a^2}\frac{x_0}{y_0}(x - x_0)$$

Transposing the terms to the left-hand side yields

$$\frac{x_0 x}{a^2} + \frac{y_0 y}{b^2} - \left(\frac{x_0^2}{a^2} + \frac{y_0^2}{b^2}\right) = 0$$

or, since M_0 lies on the ellipse,

$$\frac{x_0 x}{a^2} + \frac{y_0 y}{b^2} = 1 \tag{6}$$

which is the desired equation of the tangent to an ellipse at a point $M_0(x_0, y_0)$. Equation (6) only differs from equation (5) of an ellipse in that (6) is a linear equation in the running coordinates x and y.

Now, let us take into consideration the vertices A_1 and A_2 of the ellipse. Suppose, for instance, that M_0 coincides with A_1, then we have $x_0 = -a$, $y_0 = 0$, and (6) assumes the form $x = a$, which is, in fact, the equation of the tangent at A_1. Similarly, if M_0 coincides with A_2, then (6) transforms into the equation of the tangent at A_2.

Thus, the *tangent to an ellipse defined by* (5) *at a point* $M_0(x_0, y_0)$ *of the ellipse is defined by* (6).

Example. Find the equation of the tangent at a point M_0 of the ellipse

$$\frac{x^2}{16} + \frac{y^2}{9} = 1$$

the tangent lies on the bisector of the first quadrant angle.

○ We first find the coordinates of M_0. Since $x_0 = y_0$ for this point, we have $\frac{x_0^2}{16} + \frac{x_0^2}{9} = 1$ from the equation of the ellipse, i.e. $x_0^2 = \frac{16 \cdot 9}{25}$, whence $x_0 = \frac{4 \cdot 3}{5} = \frac{12}{5}$ (bear in mind that $x_0 > 0$). According to (6) the equation of the tangent to the ellipse at the point $\left(\frac{12}{5}, \frac{12}{5}\right)$ has the form

$$\frac{12}{5} \cdot \frac{x}{16} + \frac{12}{5} \cdot \frac{y}{9} = 1 \; \bullet$$

5.2. THE HYPERBOLA

1. The definition and canonical equation of a hyperbola.

Definition. The *hyperbola* is the set of all points in the plane such that the absolute value of the difference of the distances from two fixed points called *foci* is a constant smaller than the distance between the foci.

We use F_1 and F_2 to denote the foci of a hyperbola and $2c$

to denote the distance between the foci. Let M be an arbitrary point on the hyperbola and r_1 and r_2 the lengths of the segments F_1M and F_2M. By the definition of a hyperbola,

$$|r_1 - r_2| = 2a \qquad (1)$$

where a is a constant characterizing the hyperbola ($2a < 2c$). In order to derive the equation of a hyperbola, we introduce a rectangular Cartesian coordinate system so that the x-axis passes through the foci and the y-axis divides the line segment F_1F_2 in half (Fig. 86). Then the coordinates of the points F_1, F_2, and M are $(-c, 0)$, $(c, 0)$, and (x, y), respectively. By expressing r_1 and r_2 in terms of the coordinates of F_1, F_2, and M and substituting the results into (1), we obtain

$$\left|\sqrt{(x + c)^2 + y^2} - \sqrt{(x - c)^2 + y^2}\right| = 2a \qquad (2)$$

Perform similar transformations as in the case of the ellipse. Write equation (2) in the form

$$\sqrt{(x + c)^2 + y^2} - \sqrt{(x - c)^2 + y^2} = \pm 2a$$

transfer the second root to the right-hand side, and square both sides. The result is

$$(x + c)^2 + y^2 = 4a^2 \pm 4a\sqrt{(x - c)^2 + y^2} + (x - c)^2 + y^2$$

or

$$cx - a^2 = \pm a\sqrt{(x - c)^2 + y^2}$$

Squaring both sides of the last equation yields

$$c^2x^2 - 2a^2cx + a^4 = a^2x^2 - 2a^2cx + a^2c^2 + a^2y^2$$

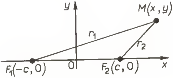


Figure 86

or

$$(c^2 - a^2)x^2 - a^2y^2 = a^2(c^2 - a^2) \tag{2a}$$

We introduce a new number $b = \sqrt{c^2 - a^2}$. Since $c > a$ by definition, we have $c^2 - a^2 > 0$ and the number b exists; with $b^2 = c^2 - a^2$ or, which is the same, $c^2 = a^2 + b^2$. By dividing both sides of (2a) by a^2b^2 we finally have

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \tag{3}$$

Equation (3) is derived from equation (2). If we perform transformations in reverse direction, then we can show that (2) follows from (3) as well. Thus, these equations are equivalent. An equation of form (3) is called the *canonical equation of a hyperbola.*

This is the second-degree equation and thus the hyperbola is a second-degree curve.

Example 1. Find the foci of the hyperbola

$$\frac{x^2}{9} - \frac{y^2}{16} = 1$$

○ We have $a = \sqrt{9} = 3$ and $b = \sqrt{16} = 4$. The distance between the foci is $2c$, where $c^2 = a^2 + b^2 = 9 + 16 = 25$, i.e. $c = 5$. The coordinates of the foci are $(-5, 0)$ and $(5, 0)$. ●

Example 2. Show that the curve

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = -1$$

is a hyperbola and find its foci.

○ The given equation only differs from the canonical equation in that its right-hand side is -1 and not 1. We introduce a new system of coordinates, namely, we consider the x-axis to be the new ordinate denoted by Y-axis and the y-axis to be the new abscissa denoted by X-axis. In new coordinates the given equation takes the form $\frac{Y^2}{a^2} - \frac{X^2}{b^2} = -1$ or

$$\frac{X^2}{a^2} - \frac{Y^2}{b^2} = 1$$

This equation defines a hyperbola whose foci lie on the X-axis at the points $(-c, 0)$ and $(c, 0)$, where $c = \sqrt{a^2 + b^2}$. In the initial xy coordinates the foci lie on the y-axis at the points $(0, -c)$ and $(0, c)$. ●

2. Studying the shape of a hyperbola from its canonical equation.

Since the power of each variable x and y in equation (3) is two, the hyperbola defined by (3) is symmetric with respect to the coordinate axes and the origin O is the center of symmetry of the hyperbola. Thus, to determine the shape of a hyperbola it is sufficient to consider its portion lying in the first quadrant.

We first determine whether a hyperbola intersects the axes of symmetry. It follows from (3) that $|x| \geqslant a$ since for $|x| < a$ the expression $\frac{x^2}{a^2} - \frac{y^2}{b^2}$ is less than unity. Thus, a hyperbola does not intersect the y-axis which is called the *conjugate axis* of the hyperbola. At the same time putting $y = 0$ we find that $x = \pm a$, which means that the hyperbola and the x-axis intersect at two points, $A_1(-a, 0)$ and $A_2(a, 0)$. These points are called the *vertices* and the x-axis the *transverse axis* of the hyperbola.

Thus, the axis of symmetry which does not intersect the hyperbola is its conjugate axis and the axis of symmetry which intersects it at two points (its vertices) is its transverse axis.

Assuming $x \geqslant 0$, $y \geqslant 0$ in the equation of a hyperbola we express y in terms of x

$$y = \frac{b}{a}\sqrt{x^2 - a^2}$$

We can see that as the coordinate x increases y also increases indefinitely. Figure 87 depicts the part of the hyperbola in the first quadrant. We can show that this part is convex upward and the tangent (drawn with a dashed line) to the hyperbola at the vertex A_2 is parallel to the y-axis.

Since a hyperbola is symmetric with respect to the coordinate axes, it can be drawn as illustrated in Fig. 88. The hyperbola consists of two parts called *branches*, one of which corresponds

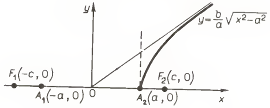


Figure 87

to $x \geqslant a$ and the other to $x \leqslant -a$. The dashed lines in the figure are explained in the next subsection.

3. The asymptotes of a hyperbola.

Let us analyze the straight lines $y = \frac{b}{a}x$ and $y = -\frac{b}{a}x$ which are shown by dashed lines in Fig. 88 and denoted by l_+ and l_- in Fig. 89. These lines possess a remarkable property, namely, if a point M moves along a hyperbola receding indefinitely from the origin, then the hyperbola approaches closer and closer to one of the indicated straight lines.

In order to verify this property we choose the point M in the first quadrant (it is sufficient to consider only this case since a hyperbola is symmetric with respect to the coordinate axes). Let M move along a hyperbola receding indefinitely from the origin O, with $x \to \infty$. The distance $d = MP$ from M to the straight line l_+ is less than the distance MN (MP is the perpen-

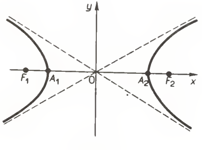


Figure 88

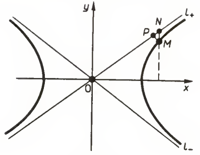


Figure 89

dicular to l_+). But MN is the difference of the ordinates of the hyperbola and the straight line l_+. Since the branch being considered is described by the equation $y = \frac{b}{a}\sqrt{x^2 - a^2}$ and the line l_+ by $y = \frac{b}{a}x$, we have

$$MN = \frac{b}{a}(x - \sqrt{x^2 - a^2})$$

Multiplying and dividing this equation by $x + \sqrt{x^2 - a^2}$ yields

$$MN = \frac{ab}{x + \sqrt{x^2 - a^2}}$$

As the point M moves along the hyperbola receding indefinitely from the origin, its abscissa x becomes as large as desired ($x \to \infty$). We can say the same of the value $x + \sqrt{x^2 - a^2}$. Therefore, the fraction $\frac{ab}{x + \sqrt{x^2 - a^2}}$ becomes arbitrarily small ($MN \to 0$), consequently, the distance d becomes arbitrarily small as well.

Suppose for a curve L there exists a straight line l such that when a point M moves along L to infinity, this point approaches l indefinitely, then the straight line l is called an *asymptote* of the curve L. Thus, the straight lines $y = \frac{b}{a}x$ and $y = -\frac{b}{a}x$ are asymptotes of the hyperbola.

Remark. In order to draw the hyperbola defined by equation (3) we construct a rectangle with center at the origin and sides $2a$ and $2b$, which are parallel to the x- and y-axes respectively (Fig. 90). The straight lines joining opposite vertices of the rectangle are the asymptotes of the hyperbola. Then sketch the branches of the hyperbola: the left branch should touch the rectangle from outside at the point A_1 (one vertex of the hyperbola) and approach the asymptotes with its "ends" and the right branch touches the rectangle at the point A_2 (the other vertex) from the outside and also approaches the asymptotes.

The line segments a and b are *semitransverse* and *semiconjugate* axes respectively.

Example 1. Form the equations of the asymptotes of the hyperbola

$$\frac{x^2}{9} - \frac{y^2}{4} = 1$$

○ We have $a = 3$ and $b = 2$, the equations of the asymptotes are $y = \frac{2}{3}x$ and $y = -\frac{2}{3}x$. ●

Example 2. Set up the equation of the hyperbola whose transverse axis is the x-axis, the equations of whose asymptotes

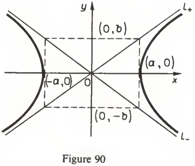


Figure 90

are $5x - 4y = 0$ and $5x + 4y = 0$, and the distance between the vertices is 16.

○ We know that the asymptotes intersect at the center of a hyperbola. We infer from the equations of the asymptotes that the center of a hyperbola is at the origin. Thus, the x-axis is the transverse axis of the hyperbola (by definition) and the y-axis is its conjugate axis. The equation of the hyperbola has the form

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

In order to find the numbers a and b we reduce the equations of the asymptotes to the form $y = \frac{5}{4}x$ and $y = -\frac{5}{4}x$, whence we immediately have $\frac{b}{a} = \frac{5}{4}$. But, by hypothesis, $2a = 16$, i.e. $a = 8$. Consequently, $b = 10$.

The desired equation has the form

$$\frac{x^2}{8} - \frac{y}{10} = 1 \bullet$$

4. The equation of a tangent to a hyperbola.
The equation of a tangent to the hyperbola

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \qquad (4)$$

can be found in the same way as for an ellipse. Therefore, we just give the final result without repeating the argument.

The tangent to the hyperbola defined by (4) *at a point* $M_0(x_0, y_0)$ *on this hyperbola is determined by the equation*

$$\frac{x_0 x}{a^2} - \frac{y_0 y}{b^2} = 1$$

5. The hyperbola as a graph of inverse proportionality.
If in the equation of the hyperbola

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

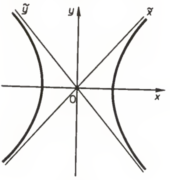


Figure 91

the numbers a and b are equal, we can write

$$x^2 - y^2 = a^2 \tag{5}$$

A hyperbola defined by (5) is *equilateral* and its asymptotes are the straight lines $y = x$ and $y = -x$, i.e. they are the bisectors of the quadrant angles.

Let us turn the coordinate axes through 45° and use $\tilde{x}$ and $\tilde{y}$ to denote the new axes so that they coincide with the asymptotes of the equilateral hyperbola (Fig. 91). We can show that the equation of the hyperbola in the new coordinates has the form

$$\tilde{x}\tilde{y} = \frac{1}{2} a^2$$

We know that two varying quantities $\tilde{x}$ and $\tilde{y}$ related as $\tilde{x}\tilde{y} = c$, where c is a constant, are *inversely proportional*. Thus, we infer that an *equilateral hyperbola is the graph of inverse proportionality.*

5.3. THE PARABOLA

1. The canonical equation of a parabola.

Definition. A *parabola* is the set of all points in the plane equidistant from a fixed point, called the focus, and a fixed line, called the directrix.

We use p to denote the distance between the focus and the directrix, the number $p > 0$ is the *parameter* of the parabola.

In order to derive the equation of a parabola, we take the straight line which is perpendicular to the directrix and passes through the focus F as the x-axis and assume that it is directed from the directrix to the focus (Fig. 92). We place the origin at the midpoint of the line segment joining the focus and the directrix. The coordinates of the focus in our system are $p/2$, 0 and the equation of the directrix is $x = -p/2$.

Suppose $M(x, y)$ is an arbitrary point on the parabola and P is its projection on the directrix, then, by the definition of a parabola, $MP = MF$ or

$$\left|x + \frac{p}{2}\right| = \sqrt{\left(x - \frac{p}{2}\right)^2 + y^2} \tag{1}$$

Squaring both sides of (1) and simplifying yields

$$y^2 = 2px \tag{2}$$

It can easily be shown (we leave it to the reader) that the converse is also true, namely, we can derive (1) from (2). Thus, (1) and (2) are equivalent. Hence it follows that (2) is the equation of the parabola.

Equation (2) is the *canonical equation of a parabola.* We can see from this equation that a parabola is a second-degree curve.

It also follows from the canonical equation that the parabola defined by it is symmetric with respect to the x-axis. In fact,

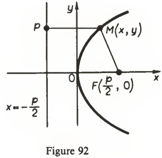


Figure 92

if (2) is satisfied by the coordinates of point (x, y), then it is also satisfied by the coordinates of point $(x, -y)$. In addition, all the points on the parabola have $x \geqslant 0$. The point $(0, 0)$ is the leftmost point of the parabola (in the usual coordinate system) and called the *vertex* of the parabola.

We note once more that if a parabola is defined by the canonical equation $y^2 = 2px$ in the given coordinate system, then its directrix is defined by the equation $x = -p/2$ and its focus is at the point $F(p/2, 0)$.

2. The parabola as the graph of a quadratic equation. School mathematics call the graph of the function

$$y = ax^2 + bx + c$$

a parabola, where a is nonzero. Let us ascertain that this definition is consistent with that given earlier.

We transform the equation as

$$y = a\left(x^2 + \frac{b}{a}x + \frac{c}{a}\right) = a\left(x^2 + 2\cdot\frac{b}{2a}x + \left(\frac{b}{2a}\right)^2 - \left(\frac{b}{2a}\right)^2 + \frac{c}{a}\right) = a\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a}$$

Introducing the notation

$$-\frac{b}{2a} = \alpha, \quad -\frac{b^2 - 4ac}{4a} = \beta$$

we can write

$$y - \beta = a(x - \alpha)^2$$

Now, if we assume the point $\tilde{O}(\alpha, \beta)$ to be the new origin and leave the direction of the coordinate axes unchanged, then the old and new coordinates are related as

$$\tilde{x} = x - \alpha, \ \tilde{y} = y - \beta$$

and the equation of the curve we consider in the new system is

$$\tilde{y} = a\tilde{x}^2 \quad \text{or} \quad \tilde{x}^2 = 2p\tilde{y}$$

with $p = 1/2a$. The last equation only differs from the canonical equation of a parabola in the notation of coordinates.

Thus, a parabola is, in fact, the graph of a quadratic polynomial.

Example. Find the directrix and focus of the parabola defined by the equation

$$y = \frac{1}{2} x^2 - 3x + 6$$

○ We transform the equation as follows

$$y = \frac{1}{2}(x^2 - 6x + 12) = \frac{1}{2}(x^2 - 6x + 9 - 9 + 12)$$

$$= \frac{1}{2}(x - 3)^2 + \frac{3}{2}$$

or

$$y - \frac{3}{2} = \frac{1}{2}(x - 3)^2$$

Then we translate the coordinate axes assuming the point $\tilde{O}(3, 3/2)$ to be the origin. The equation of a parabola in the new coordinate system has the form

$$\tilde{y} = \frac{1}{2} \tilde{x}^2 \quad \text{or} \quad \tilde{x}^2 = 2p\tilde{y}$$

with $p = 1$. This equation only differs from the canonical equation (2) in the notation of the axes. Noting the remark made at the end of subsection 1, we find that the equation of the directrix of the parabola is $y = -p/2 = -1/2$ and the focus is at the point with the coordinates $\tilde{y} = p/2 = 1/2$, $\tilde{x} = 0$. Passing to the old coordinates via the formulas $x = \tilde{x} + 3$ and $y = \tilde{y} + 3/2$ yields the equation of the directrix, viz. $y = 1$, and the coordinates of the focus are $x = 3$ and $y = 2$ (Fig. 93). ●

3. Optical properties of the parabola.

One property of the parabola is widely applied in optics. We first formulate this property geometrically.

Figure 94 shows the parabola $y^2 = 2px$ with focus F, a point $M_0(x_0, y_0)$ on the parabola, the ray l emanating from M_0 and parallel to the axis of the parabola (the x-axis), and the tangent to the parabola at M_0. The property we are interested in is that

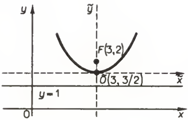


Figure 93

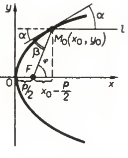


Figure 94

the *tangent makes equal angles with the vector* $\overrightarrow{M_0F}$ *and the ray* l.

This statement is obvious if M_0 is the vertex of the parabola, i.e. $M_0(0, 0)$. Suppose, therefore, that M_0 is not at the point $(0, 0)$. In this case a sufficiently small part of the parabola near M_0 is described by the equation $y = y(x)$. Differentiating the identity $y^2(x) = 2px$ with respect to x yields $2yy' = 2p$, whence $y' = p/y$. This means, for one, that the slope of the tangent to the parabola at M_0 is $k = p/y_0$.

It follows directly from Fig. 94 that

$$\tan \varphi = \frac{y_0}{x_0 - p/2}$$

where φ is the angle the vector $\overrightarrow{FM_0}$ forms with the x-axis. We have at the same time

$$\tan 2\alpha = \frac{2 \tan \alpha}{1 - \tan^2 \alpha} = \frac{2p/y_0}{1 - p^2/y_0^2} = \frac{2py_0}{y_0^2 - p^2}$$

$$= \frac{2py_0}{2px_0 - p^2} = \frac{y_0}{x_0 - p/2}$$

whence it follows that $\varphi = 2\alpha$, and since $\varphi = \alpha + \beta$, we have $\alpha = \beta$ which was desired to obtain.

The property of a parabola we proved can be interpreted as follows. Suppose a source of light is placed at the focus F. Light rays emanating from the focus are reflected from the parabola

Figure 95

governed by the law of reflection, namely, the angle of incidence is equal to the angle of reflection. Since the angles α and β are equal (which we have just proved), the reflected ray is parallel to the x-axis. In other words, all the rays are reflected in a beam parallel to the axis.

The reflection property is exploited in constructing parabolic mirrors which are used in telescopes, searchlights, automobile headlights, solar heating devices, and the like (Fig. 95).

Exercises to Chapter 5

5.1

1. A vertex of a triangle whose base is fixed moves so that the perimeter of the triangle is constant. Find the trajectory of the moving vertex if the base of the triangle is 24 and the perimeter is 50.

2. A line segment of constant length slides with its ends along the sides of a right angle. Prove that any point of the segment moves along the arc of an ellipse.

3. Form the canonical equation of an ellipse if the distance between the foci is 6 and the semimajor axis is 5.

4. Find a point on the ellipse $\frac{x^2}{20} + \frac{y^2}{15} = 1$ at a distance 4 from the minor axis.

5. Given the semiaxes a and b of an ellipse with $a > b$, prove that the distance r from an arbitrary point on the ellipse to its center is such that $b \leqslant r \leqslant a$.

6. Find the side of a square inscribed in the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

7. Given an ellipse passing through the points $(8, 3)$ and $(2\sqrt{5}, 2\sqrt{5})$, find its equation if we know that the coordinate axes are its axes of symmetry.

8. Find the equation of the tangent to the ellipse $\frac{x^2}{3} + \frac{y^2}{12} = 1$ at the point $(1, 3)$.

9. Find the tangents to the ellipse $\frac{x^2}{20} + \frac{y^2}{15} = 1$ with the slope $k = -1$ and determine the distance between these tangents.

10. Prove that the straight line $Ax + By + C = 0$ touches the ellipse if and only if $Aa^2 + Bb^2 = C^2$.

11. Find the tangents common to the two ellipses

$$\frac{x^2}{3} + \frac{y^2}{3} = 1, \quad \frac{x^2}{2} + \frac{y^2}{3} = 1$$

12. Different chords are drawn through the vertex $(a, 0)$ of the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Find the line on which the midpoints of the chords lie.

13. Prove that the tangent to an ellipse at a point M forms equal angles with the line segments MF_1 and MF_2, where F_1 and F_2 are the foci.

14. A cylinder whose section is a circle of radius 10 is cut by a plane which makes an angle of 30° with the axis of the cylinder. The curve of the intersection is the ellipse: find the lengths of its semiaxes.

5.2

1. Form the equation of a hyperbola whose axes coincide with the coordinate axes if the distance between the foci is 20 and the distance between the vertices is 16.

2. Given the foci $F_1(-5, 0)$ and $F_2\ (5, 0)$ and the point $C(4\sqrt{2}, 3)$ on the hyperbola, find its equation.

3. Find a point M on the hyperbola $\frac{x^2}{16} - \frac{y^2}{9} = 1$ for which $MF_1 \perp MF_2$, where F_1 and F_2 are the foci.

4. Find the foci and asymptotes of the hyperbola $\frac{x^2}{64} - \frac{y^2}{36} = 1$.

5. Find the equation of the hyperbola given the equation of its asymptotes $y = \pm(3/2)x$ and the point $(2\sqrt{3}, 3)$ on it.

6. Prove that the product of the distances from any point on a hyperbola to its asymptotes is a constant.

7. Find the points at which the hyperbola $\frac{x^2}{90} - \frac{y^2}{36} = 1$ intersects the following straight lines: (a) $x - 5y = 0$, (b) $2x + y - 18 = 0$, (c) $x - y + 5 = 0$.

8. Find the equation of the tangent to the hyperbola $\frac{x^2}{6} - \frac{y^2}{5} = 1$ at the point $(6, -5)$.

9. Draw the tangent with the slope $k = 2$ to the hyperbola $\frac{x^2}{8} - \frac{y^2}{6} = 1$.

10. Find whether we can draw a tangent to a hyperbola at any angle of inclination to the transverse axis.

11. Find the condition for the straight line $Ax + By + C = 0$ to touch the hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$.

5.3

1. Form the equation of the parabola touching the y-axis if its focus is at the point (6, 0) and the x-axis is its axis of symmetry.

2. Form the equation of the parabola

(a) symmetric with respect to the x-axis, touching the y-axis, and passing through the point (9, 1),

(b) symmetric with respect to the y-axis touching the x-axis, and passing through the point (3, -3).

3. Find the points at which the parabola $y^2 = 3x$ and the following straight lines intersect:

(a) $x - 3y + 6 = 0$, (b) $x - 2y + 3 = 0$, (c) $y + x = -5$.

4. Find the points at which the parabola $y^2 = 12x$ and the ellipse $\frac{x^2}{25} + \frac{y^2}{16} = 1$ intersect.

5. A chord which is perpendicular to the axis of the parabola $y^2 = 2px$ is drawn through its focus. Find the length of the chord.

6. Prove that the tangent to the parabola $y^2 = 2px$ at the point (x_0, y_0) is defined by the equation $y_0 = p(x_0 + x)$.

7. Form the equation of the tangent to the parabola $y^2 = 2px$ at a point with abscissa 3.

8. Draw the tangent to the parabola $x^2 = 5y$ parallel to the straight line $2x - y = 0$.

9. Find the condition that the straight line $y = kx + b$ touches the parabola $y^2 = 2px$.

10. Find the equation of the line formed by the midpoints of the ordinates of the parabola $y^2 = 2px$.

11. An object thrown at an angle to the horizon describes an arc of a parabola and falls 16 meters away. Find the parameter p of the parabolic trajectory if the height reached by the object is 12 meters.

12. Find the equation of the parabola if its minimal ordinate is -4, the focus is at the point (4, 0), and the axis of symmetry is parallel to the y-axis.

Chapter 6

THE PLANE IN SPACE

6.1. THE EQUATION OF A SURFACE IN SPACE. THE EQUATION OF A PLANE

1. The notion of the equation of a plane.
Suppose we have an equation in three variables

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

We introduce a rectangular Cartesian coordinate system in space and consider x, y, z to be the coordinates of a point. Equation (1) is satisfied for some points (x, y, z), while for others it is not. The question arises as what is the set T of the spatial points that satisfy equation (1). Generally, the set T is a *surface* in space. Let us consider the case where we can express coordinate z from (1) as a function of x and y

$$z = f(x, y)$$

When the values $x = x_0$, $y = y_0$ are fixed, we have a single value $z_0 = f(x_0, y_0)$ for z. Geometrically, this means that exactly one point from the set T lies on each straight line parallel to the z-axis. In other words, the set T and every straight line parallel to the z-axis intersect at a unique point. It is natural to consider such a set a surface in space.

Definition 1. The *surface defined by the equation* $F(x, y, z) = 0$ is the set T of all points in space whose coordinates satisfy this equation.

Definition 2. The *equation of a surface* T given in space is the equation $F(x, y, z) = 0$ such that the surface defined by this equation coincides with T.

As an example, we find the equation of a sphere centered at the point $P(a, b, c)$ with radius R.

The point $M(x, y, z)$ belongs to the sphere we consider if

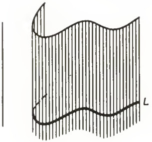


Figure 96

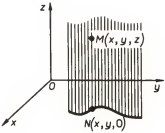


Figure 97

and only if $|PM| = R$ or

$$\sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2} = R$$

This relationship is the equation of the sphere. Squaring both sides we finally obtain

$$(x-a)^2 + (y-b)^2 + (z-c)^2 = R^2$$

2. The equation of a cylindrical surface.

Suppose we have a curve L and a straight line l in space. If we draw a straight line parallel to l, through every point of the curve L, then we obtain a surface (Fig. 96) called *cylindrical*. The curve L is the *directrix* and the lines parallel to l constitute the *generatrix* of the cylindrical surface.

We choose a coordinate system in space so that the z-axis is parallel to the straight line l and assume that the directrix L is in the xy-plane and is defined by the equation

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

We now prove that (2) *is also the equation of the given cylindrical surface in space.*

A point $M(x, y, z)$ belongs to this surface if and only if its projection on the xy-plane, a point $N(x, y, 0)$, lies on the curve L (Fig. 97), and this means that x and y satisfy (2). Thus, M is a point of the cylindrical surface if and only if its x, y coordinates satisfy (2).

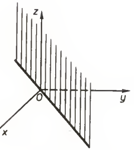


Figure 98

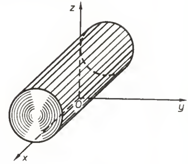


Figure 99

Similarly, we can establish that a cylindrical surface with generatrices parallel to the y-axis is defined by an equation of the form $\Phi(x, z) = 0$. If generatrices are parallel to the x-axis, then the equation of the surface is of the form $\Psi(y, z) = 0$.

Example 1. The equation $y - x = 0$, which does not contain the z coordinate, defines a cylindrical surface with generatrices parallel to the z-axis. The bisector of the first and third quadrant angles in the xy-plane is the directrix (Fig. 98).

Example 2. The equation $(y - 1)^2 + (z + 2)^2 = 1$, which does not involve the x coordinate, defines a cylindrical surface with generatrices parallel to the x-axis. The directrix is the circle centered at the point $(0, 1, -2)$ with radius 1 and lying in the yz-plane. This surface is called the right circular cylinder (Fig. 99).

3. Different ways of defining a plane. The equation of a plane passing through a given point and perpendicular to a given vector.

A plane in space can be specified in several ways. For instance we can specify

(a) a point M_0 of the desired plane and a nonzero normal vector **n** to the plane,

(b) a point M_0 of the desired plane and two noncollinear vectors parallel to the plane,

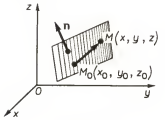


Figure 100

(c) three points of the desired plane M_1, M_2, and M_3 (it is assumed that the points are not collinear).

We introduce a rectangular Cartesian coordinate system in space. Given a point $M_0(x_0, y_0, z_0)$ and a nonzero vector $\mathbf{n} = \langle A, B, C \rangle$ in this system, find the equation of a plane α through M_0 and perpendicular to $\mathbf{n}$.

The point $M(x, y, z)$ lies on the plane α if and only if the vectors $\overrightarrow{M_0M}$ and $\mathbf{n}$ are perpendicular (Fig. 100), i.e.

$$\overrightarrow{M_0M} \cdot \mathbf{n} = 0 \tag{3}$$

Since $\overrightarrow{M_0M} = \langle x - x_0, y - y_0, z - z_0 \rangle$ and $\mathbf{n} = \langle A, B, C \rangle$, condition (3) is equivalent to the equation

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0 \tag{4}$$

which is the equation of the plane α.

Thus, (4) *is the equation of the plane through* $M_0(x_0, y_0, z_0)$ *and perpendicular to* $\mathbf{n} = \langle A, B, C \rangle$.

Example. Write the equation of the plane through the point $M_0(1, 2, -3)$ and perpendicular to the vector $\mathbf{n} = \langle 4, 7, 9 \rangle$:

$$4(x - 1) + 7(y - 2) + 9(z + 3) = 0$$

or, after simplifying,

$$4x + 7y + 9z + 9 = 0$$

Theorem. *Any plane in rectangular Cartesian coordinates is defined by the linear algebraic equation, that is, the equation* *of the form*

$$Ax + By + Cz + D = 0 \tag{5}$$

at least one of the coefficients A, B, C being nonzero.

Conversely, *any linear equation in which at least one of the coefficients A, B, C is nonzero defines a plane.*

□ Given a plane α in space. We consider a point $M_0(x_0, y_0, z_0)$ lying on the plane and a nonzero normal vector $\mathbf{n} = \langle A, B, C \rangle$ to the plane. According to what we have proved before, the equation of the plane α has the form in (4). We write (4) as $Ax + By + Cz + D = 0$, where $D = -Ax_0 - By_0 - Cz_0$, and see that the plane α is defined by a linear equation; since $\mathbf{n} \neq 0$, one of the three numbers is nonzero.

Conversely, given equation (5), where one of the numbers A, B, C is nonzero, choose one of the solutions $x = x_0, y = y_0, z = z_0$ of this equation. For instance, if $A \neq 0$, then we can take the solution $x_0 = -D/A, y_0 = 0, z_0 = 0$. Then subtracting the identity

$$Ax_0 + By_0 + Cz_0 + D = 0 \tag{6}$$

from equation (5) yields

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0 \tag{7}$$

which is equivalent to the original ec uation (5). In fact, we obtained (7) from (5) by subtracting (6) from it; similarly, we can derive (5) from (7) by adding (6) to it.

Accordingly equation (7) defines the plane through the point $M_0(x_0, y_0, z_0)$ and perpendicular to the vector $\mathbf{n} = \langle A, B, C \rangle$, hence it follows that (5) also defines the plane. ■

Definition. An equation of the form in (5) where at least one of the coefficients A, B, C is nonzero is called the *general equation of the plane.*

Note that any nonzero vector $\mathbf{n}$ perpendicular to the given plane is called a *normal vector* to this plane. Obviously, a plane has a set of normal vectors, all of them being mutually collinear.

Let us concentrate on a fact we established when proving the second statement of the theorem, namely, *if a plane is defined*

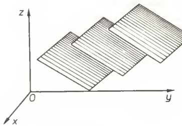


Figure 101

by the equation $Ax + By + Cz + D = 0$, *then* $\mathbf{n} = \langle A, B, C\rangle$ *is a normal vector to the plane.*

4. Analyzing the general equation of a plane.

Let us find whether a change in the coefficients A, B, C and in the number D in the general equation of a plane

$$Ax + By + Cz + D = 0 \qquad (8)$$

changes the position of the plane with respect to the coordinate axes.

(1) If D changes while the coefficients A, B, C do not, then the normal vector $\mathbf{n} = \langle A, B, C\rangle$ is left unchanged: the plane is displaced parallel to itself in space (Fig. 101).

For $D = 0$ the equation assumes the form

$$Ax + By + Cz = 0 \qquad (9)$$

Obviously, the numbers $x = 0$, $y = 0$, $z = 0$ satisfy (9), hence it follows that (9) defines a plane passing through the origin.

(2) A change in A, B, C makes the vector $\mathbf{n}$ rotate and, consequently, the plane itself. If one of the coefficients vanishes, then $\mathbf{n}$ becomes perpendicular to the respective coordinate axis, and therefore the plane is parallel to that axis. For instance, the equation $By + Cz + D = 0$ defines a plane parallel to the x-axis.

If one of the coefficients and the constant term are zero, then the plane passes through the corresponding coordinate axis. For instance, the equation $By + Cz = 0$ defines a plane passing through the x-axis.

(3) If two coordinates of the vector **n** are both zero, then **n** is parallel to that coordinate axis the projection on which is nonzero; the plane is perpendicular to this axis or, in other words, is parallel to the corresponding coordinate plane.

Thus, the planes defined by $Ax + D = 0$, $By + D = 0$, and $Cz + D = 0$ are perpendicular to the x-, y-, and z-axes, respectively or, which is the same, are parallel to the yz, xz, and xy coordinate planes, respectively.

For $D = 0$ the last three equations assume the form $x = 0$, $y = 0$, and $z = 0$. These are the equations of the yz, xz, and xy coordinate planes, respectively.

5. Relative position of two planes.

Suppose two planes α and β are defined by the equations

$$A_1x + B_1y + C_1x + D_1 = 0 \quad \text{for} \quad \alpha$$
$$A_2x + B_2y + C_2z + D_2 = 0 \quad \text{for} \quad \beta$$

Three cases are possible here: (1) the planes coincide, (2) the planes are parallel but not coincident, and (3) the planes are not parallel, i.e. they intersect along a straight line.

How can we determine from the equations which is the case? Case 3 can easily be established, namely for two planes not to be parallel it is necessary and sufficient that their normal vectors $\mathbf{n}_1 = \langle A_1, B_1, C_1 \rangle$ and $\mathbf{n}_2 = \langle A_2, B_2, C_2 \rangle$ be not parallel, i.e. the coordinates of $\mathbf{n}_1$ be not proportional to those of $\mathbf{n}_2$. For instance, the planes defined by $3x - y + z + 4 = 0$ and $x - y + 2z - 5 = 0$ are not parallel since the numbers 3, -1, 1 are not proportional to the numbers 1, -1, 2.

Cases 1 and 2 are common in that $\mathbf{n}_1$ and $\mathbf{n}_2$ are collinear, i.e. $\mathbf{n}_2 = \lambda \mathbf{n}_1$, where λ is a number. The equations of the planes are

$$A_1x + B_1y + C_1z + D_1 = 0 \quad \text{for} \quad \alpha$$
$$\lambda A_1x + \lambda B_1y + \lambda C_1z + D_2 = 0 \quad \text{for} \quad \beta$$

If here the planes have at least one point in common, then the last two equations have the general solution x_0, y_0, z_0. By substituting these coordinates into both equations, we obtain two arithmetic identities. Multiplying the first by λ and subtracting the result from the second yields $D_2 - \lambda D_1 = 0$. Conversely, if $\mathbf{n}_2 = \lambda \mathbf{n}_1$ and $D_2 = \lambda D_1$, then the planes coincide.

Thus, case 1 is characterized by the existence of a number λ such that $A_2 = \lambda A_1$, $B_2 = \lambda B_1$, $C_2 = \lambda C_1$, $D_2 = \lambda D_1$, while for case 2 λ is such that $A_2 = \lambda A_1$, $B_2 = \lambda B_1$, $C_2 \neq \lambda C_1$, $D_2 \neq \lambda D_1$.

Example. The planes $2x + 2y - 4z + 6 = 0$ and $3x + 3y - 6z + 10 = 0$ are parallel but do not coincide since $2/3 = 2/3 = -4/(-6) = 6/10$.

6.2. SPECIAL FORMS OF THE EQUATION OF A PLANE

1. The equation of a plane passing through a given point and parallel to two given vectors.

Given a point $M_1(x_1, y_1, z_1)$ and two noncollinear vectors $\mathbf{a} = \langle p, q, r \rangle$ and $\mathbf{a}' = \langle p', q', r' \rangle$ in a rectangular Cartesian coordinate system. We assume that the vectors originate at M_1 (Fig. 102). Let us form the equation of the plane α through M_1 and parallel to $\mathbf{a}$ and $\mathbf{a}'$.

If a point $M(x, y, z)$ lies in the plane α then the vectors $\overrightarrow{M_1M}$, $\mathbf{a}$, and $\mathbf{a}'$ lie in α and hence are coplanar. Conversely, if M is such that the three vectors are coplanar, then M lies in the plane of $\mathbf{a}$ and $\mathbf{a}'$, i.e. in the plane α.

The condition that three vectors are coplanar is, according to the theorem in Sec. 3.10, that the determinant composed of the coordinates of these vectors is zero. This condition has the

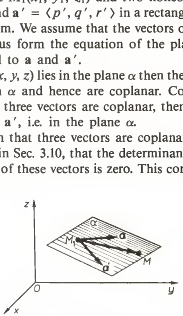


Figure 102

form

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ p & q & r \\ p' & q' & r' \end{vmatrix} = 0 \qquad (1)$$

for the vectors $\overrightarrow{M_1M}$, **a**, and **a**$'$.

Consequently, (1) is the *equation of a plane through the point* $M_1(x_1, y_1, z_1)$ *and parallel to the vectors* $\mathbf{a} = \langle p, q, r\rangle$ and $\mathbf{a}' = \langle p', q', r'\rangle$.

Example 1. Find the equation of the plane through the point $M_1(2, 2, 1)$ and parallel to the vectors $\mathbf{a} = \langle 3, 2, 5\rangle$ and $\mathbf{a}' = \langle 1, -1, 0\rangle$.

○ Substituting the data into (1) yields

$$\begin{vmatrix} x - 2 & y - 2 & z - 1 \\ 3 & 2 & 5 \\ 1 & -1 & 0 \end{vmatrix} = 0$$

To simplify this determinant we expand it by the elements of the first row:

$$(x - 2)\begin{vmatrix} 2 & 5 \\ -1 & 0 \end{vmatrix} - (y - 2)\begin{vmatrix} 3 & 5 \\ 1 & 0 \end{vmatrix} + (z - 1)\begin{vmatrix} 3 & 2 \\ 1 & -1 \end{vmatrix} = 0$$

or

$$5(x - 2) + 5(y - 2) - 5(z - 1) = 0$$

By removing brackets and dividing both sides by 5, we obtain the desired equation $x + y - z - 3 = 0$. ●

Example 2. Find the equation of the plane α through the point $M_1(1, -1, -4)$ and perpendicular to each plane defined by $3x - 2y + 3z + 5 = 0$ and $-5x + 4y - z + 1 = 0$.

○ Since each of the given planes is perpendicular to the plane α its normal vector must be parallel to α. Then we have two vectors $\mathbf{a} = \langle 3, -2, 3\rangle$ and $\mathbf{a}' = \langle -5, 4, -1\rangle$. According to (1) the equation of α has the form

$$\begin{vmatrix} x - 1 & y + 1 & z + 4 \\ 3 & -2 & 3 \\ -5 & 4 & -1 \end{vmatrix} = 0$$

Expanding the determinant yields

$$(x-1)\begin{vmatrix} -2 & 3 \\ 4 & -1 \end{vmatrix} - (y+1)\begin{vmatrix} 3 & 3 \\ -5 & -1 \end{vmatrix} + (z+4)\begin{vmatrix} 3 & -2 \\ -5 & 4 \end{vmatrix} = 0$$

or

$$-10(x-1) - 12(y+1) + 2(z+4) = 0$$

and the final result is the desired equation $5x + 6y - z - 3 = 0$. ●

2. The equation of the plane passing through three given points. Given three noncollinear points $M_1(x_1, y_1, z_1)$, $M_2(x_2, y_2, z_2)$, and $M_3(x_3, y_3, z_3)$, find the equation of the plane passing through these points.

Consider the vectors $\overrightarrow{M_1M_2} = \langle x_2 - x_1, y_2 - y_1, z_2 - z_1 \rangle$ and $\overrightarrow{M_1M_3} = \langle x_3 - x_1, y_3 - y_1, z_3 - z_1 \rangle$ which are, by assumption, noncollinear and lie in the desired plane α (thereby are parallel to α). Thus, we can obtain the equation of the plane α as that of the plane through the point M_1 and parallel to the two vectors $\overrightarrow{M_1M_2}$ and $\overrightarrow{M_1M_3}$, i.e.

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0 \qquad (2)$$

Thus, (2) *is the equation of the plane through three points* M_1, M_2, M_3, or the *three-point form of the equation.*

Example. Find the equation of the plane passing through the three points $M_1(1, -2, -1)$, $M_2(2, 3, 0)$, and $M_3(6, 2, -2)$.

○ Substituting the coordinates of the points into (2) yields

$$\begin{vmatrix} x - 1 & y + 2 & z + 1 \\ 2 - 1 & +2 & +1 \\ 6 - 1 & +2 & +1 \end{vmatrix} = 0$$

or

$$\begin{vmatrix} x - 1 & y + 2 & z + 1 \\ 1 & 5 & 1 \\ 5 & 4 & -1 \end{vmatrix} = 0$$

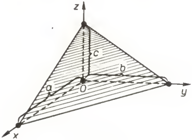


Figure 103

By expanding this determinant by the first row, we obtain

$$(x-1)\begin{vmatrix} 5 & 1 \\ 4 & -1 \end{vmatrix} - (y+2)\begin{vmatrix} 1 & 1 \\ 5 & -1 \end{vmatrix} + (z+1)\begin{vmatrix} 1 & 5 \\ 5 & 4 \end{vmatrix} = 0$$

or

$$-9(x-1) + 6(y+2) - 21(z+1) = 0$$

Removing brackets and dividing both sides of the equation by -3 yields the final result $3x - 2y + 7z = 0$. We can see that the plane passes through the origin.

3. The intercept equation of a plane.

If the plane intersects the three coordinate axes and does not pass through the origin, then it is convenient to write its equation in intercept form.

Given three points on the coordinate axes distinct from the origin, namely, $A(a, 0, 0)$ with $a \neq 0$ on the x-axis, $B(0, b, 0)$ with $b \neq 0$ on the y-axis, and $C(0, 0, c)$ with $c \neq 0$ on the z-axis, find the equation of the plane passing through A, B, and C.

We can take

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1 \qquad (3)$$

as such an equation. In fact, (3) is a linear equation and, therefore, defines a plane. By substituting the coordinates of A into

(3) we ascertain that the plane does pass through this point:

$$\frac{a}{a} + \frac{0}{b} + \frac{0}{c} = 1$$

Similarly, we can show that the plane also passes through B and C.

Equation (3) is called the *intercept equation of a plane* since it defines the plane which cuts the x-intercept a, the y-intercept b, and the z-intercept c.

The intercept form is very convenient and illustrates the position of a plane in space (Fig. 103).

Any general equation of a plane $Ax + By + Cz + D = 0$ can be reduced to form in (3) if A, B, C, and D are all nonzero. We must transfer the constant term D to the right-hand side and then divide both sides by $-D$. The result is

$$\frac{x}{-D/A} + \frac{y}{-D/B} + \frac{z}{-D/C} = 1$$

which is the intercept equation with $a = -D/A$, $b = -D/B$, and $c = -D/C$.

Example 1. Reduce the equation of a plane $4x - 3y + 2z - 12 = 0$ to intercept form.

○ By transforming the equation as above we get $\frac{4x}{12} - \frac{3y}{12} + \frac{2z}{12} = 1$ or $\frac{x}{3} + \frac{y}{-4} + \frac{z}{6} = 1$. ●

Example 2. Find the volume of the pyramid bounded by the plane defined in Example 1 and the coordinate planes.

○ We assume that the base of the pyramid is a triangle with the vertices $O(0, 0, 0)$, $A(3, 0, 0)$, and $B(0, -4, 0)$ in the xy-plane and the altitude of the pyramid is the line segment OC with $C(0, 0, 6)$. Using the formula for the volume of a pyramid and substituting the data yields

$$V = \frac{1}{3} S_{OAB} \cdot h = \frac{1}{3} \cdot \frac{3 \cdot 4}{2} \cdot 6 = 12. \quad ●$$

6.3. DISTANCE BETWEEN A POINT AND A PLANE. ANGLE BETWEEN TWO PLANES

1. Distance from a point to a plane.
Given a plane α defined by the equation

$$Ax + By + Cz + D = 0 \tag{1}$$

and a point $M_0(x_0, y_0, z_0)$ in space, find the distance d from M_0 to α.

This problem is similar to finding the distance between a point and a straight line, which we discussed in Sec. 4.6. Therefore we immediately write

$$d = \frac{|Ax_0 + By_0 + Cz_0 + D|}{\sqrt{A^2 + B^2 + C^2}} \tag{2}$$

We now formulate the following rule: *in order to find the distance from the point M_0 to the plane α defined by the general equation* (1), *we must substitute the coordinates of M_0 for the running coordinates x, y, and z on the left-hand side of* (1) *and divide the absolute value of the result by $\sqrt{A^2 + B^2 + C^2}$.*

Example 1. Find the distance from the point $M_0(1, 1, 1)$ and the plane defined by $2x + 2y - z + 3 = 0$.

○ We have from formula (2)

$$d = \frac{|2\cdot 1 + 2\cdot 1 - 1 + 3|}{\sqrt{2^2 + 2^2 + (-1)^2}} = \frac{6}{\sqrt{9}} = 2 \quad ●$$

Example 2. Find the distance between two parallel planes defined by $2x + 2y - z + 3 = 0$ and $2x + 2y - z - 3 = 0$.

○ The easiest way is to choose a point M_0 on one plane and then find the distance from M_0 to the other plane. For instance, we choose the point $M_0(1, 1, 1)$ on the second plane. According to Example 1, the distance from M_0 to the first plane is 6. ●

6. Angle between two planes.
The *angle between two planes* is the angle between the normal vectors to these planes.

A normal vector to a plane can have either of two opposite

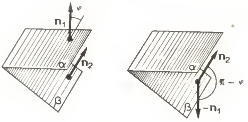


Figure 104

directions and, therefore, the angle between two planes is multivalued (Fig. 104); it may have two values, φ and $\pi - \varphi$. Since $\cos(\pi - \varphi) = -\cos\varphi$, we can find the cosine of the angle between two planes from the formula where $\mathbf{n}_1$ and $\mathbf{n}_2$ are two normal vectors to the planes α and β respectively.

If the planes are defined by the general equations

$$A_1x + B_1y + C_1z + D_1 = 0 \quad \text{for} \quad \alpha$$
$$A_2x + B_2y + C_2z + D_2 = 0 \quad \text{for} \quad \beta$$

then we can choose $\mathbf{n}_1 = \langle A_1, B_1, C_1 \rangle$ and $n_2 = \langle A_2, B_2, C_2 \rangle$ as the normal vectors. Then

$$\cos\varphi = \frac{A_1A_2 + B_1B_2 + C_1C_2}{\sqrt{A_1^2 + B_1^2 + C_1^2}\sqrt{A_2^2 + B_2^2 + C_2^2}}$$

Corollary. *For two planes to be perpendicular it is necessary and sufficient that the condition*

$$A_1A_2 + B_1B_2 + C_1C_2 = 0$$

is satisfied.

In fact, if the planes α and β are perpendicular, then $\mathbf{n}_1 \cdot \mathbf{n}_2 = 0$. The converse is also true.

6.4. HALF-SPACES

We consider a plane α in space. The plane divides the entire space into two half-spaces and is their boundary. Every half-space is indicated by the parallel lines in Fig. 105.

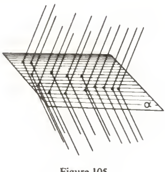


Figure 105

Suppose that the equation of the plane α is

$$Ax + By + Cz + D = 0 \quad (1)$$

We characterize the half-space defined by α in algebraic terms.

One half-space is the set of points $M(x, y, z)$ for which

$$Ax + By + Cz + D \geqslant 0 \quad (2)$$

and the other half-space is the set of points $N(x, y, z)$ for which

$$Ax + By + Cz + D \leqslant 0 \quad (3)$$

□ We first suppose that $C \neq 0$. Then equation (1) reduces to the form

$$z = ax + by + k \quad (4)$$

and (2) is equivalent to the inequality

$$z \geqslant ax + by + k \quad (5)$$

and (3) is equivalent to the inequality

$$z \leqslant ax + by + k \quad (6)$$

Inequality (5) defines the half-space lying above the plane given by (4) (Fig. 106) and (6) defines the half-space lying below the plane.

If $C = 0$, then A or B is nonzero; then reasoning in a similar way (but with respect to another coordinate, x or y) we can obtain the required result. ■

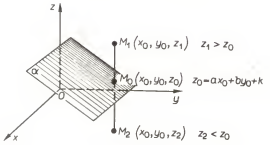


Figure 106

Example 1. Given the plane $2x - 3y + 7z - 5 = 0$ and two points $M(4, -2, -2)$ and $N(1, 3, 1)$, find whether the points lie on one side of the plane.

○ By substituting the coordinates of M and N into the left-hand side of the equation we have $2 \cdot 4 - 3 \cdot (-2) + 7 \cdot (-2) - 5 = -5$ and $2 \cdot 1 - 3 \cdot 3 + 7 \cdot 1 - 5 = -5$. Since the numbers have the same sign, we infer that M and N belong to the same half-space. ●

Example 2. Given the plane α defined by the equation $x + y - 4z + 1 = 0$ and two points $P(1, 1, 1)$ and $Q(2, 2, 1)$, find whether the line segment PQ intersects the plane.

○ Substituting the coordinates of P and Q into the left-hand side of the equation yields $1 + 1 - 4 + 1 = -1$ and $1 + 2 - 4 + 1 = 1$. Since the resulting numbers are different in sign, we infer that P and Q belong to different half-spaces with respect to the plane α. Thus, the line segment PQ intersects the plane. ●

Exercises to Chapter 6

6.1

1. Describe the surfaces defined by the equations:

$$(1)\ x = 0,\ (2)\ y = 0,\ (3)\ z = 0,\ (4)\ xy = 0,$$

$$(5)\ xyz = 0,\ (6)\ x = 2,\ (7)\ xy - y^2 = 0,$$

$$(8)\ \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,\ (9)\ x^2 + y^2 + z^2 = 16$$

2. Form the equation of the plane through the point $P(-1, 2, 3)$ and perpendicular to the radius-vector of the point.

3. Find the equation of the plane through the point $P(1, 3, -4)$ and perpendicular to the radius-vector of the point $Q(-1, 0, 2)$.

4. Given three points $M_1(1, 7, -2)$, $M_2(4, -3, 2)$, and $M_3(3, 4, 2)$, find the equation of the plane through M_1 and perpendicular to the straight line M_2M_3.

5. Find the equation of the plane through the point $A(0, -2, -3)$ and parallel to the plane defined by $x + 5y - 4z + 2 = 0$.

6. Find the equation of the plane through the point $A(1, 1, 1)$ and parallel to the xy-plane and through the point $B(-1, 2, 7)$ and the x-axis.

7. Find the values of p and q such that the following pairs of equations define the parallel planes

(a) $\begin{matrix} x - py + 2z - 3 = 0 \\ 2x - 4y - qz + 7 = 0, \end{matrix}$ (b) $\begin{matrix} (p + 1)x - 6y + (p - 1)z + 2 = 0 \\ 2x - 3y + qz - 1 = 0 \end{matrix}$

8. Find whether the planes defined by $px - y - z = 0$ and $x + y + z = 0$ coincide.

6.2

1. Find the equation of the plane passing through the origin and the points $M_1(1, 2, 3)$ and $M_2(4, 5, 6)$.

2. Find the equation of the plane

(a) through two points $M_1(-1, 1, 4)$ and $M_2(6, 4, -3)$ and parallel to the x-axis.

(b) through two points $M_1(0, -1, 1)$ and $M_2(2, 1, 3)$ and perpendicular to the plane defined by $4x - 3y + 5z = 0$.

3. Find the points at which the plane defined by $7x - 5y + 2z - 14 = 0$ and the coordinate axes intersect.

4. Find the intercept equation of the plane defined by $3x - 5y + 45z + 15 = 0$.

5. Find the equation of the plane cutting the z-intercept $c = -5$ and perpendicular to the vector $\mathbf{n} = \langle 2, -7, 1 \rangle$.

6. Find the equation of the plane through the point $A(-1, 1, 3)$ and cutting equal intercepts on the coordinate axes.

6.3

1. Find the distance from the origin to the plane $2x + 2y - z - 1 = 0$.

2. Find the distance between two parallel planes $6x + 3y + 2z = 5$ and $6x + 3y + 2z = 1$.

3. Find the equation of the planes parallel to the plane $6x + 3y - 2z + 7 = 0$ and a distance 5 away from it.

4. Find the point on the x-axis equidistant from two planes $12x - 16y + 15z + 1 = 0$ and $2x + 2y - z - 1 = 0$.

5. Given two faces of a cube lying in the parallel planes $6x - 3y + 2z + 1 = 0$ and $6x - 3y + 2z + 4 = 0$, find the volume of the cube.

6. Find the distance from the point $P(1, 1, 1)$ to the plane passing through the points $A(4, -1, -1)$, $B(2, 0, -2)$, and $C(3, -1, 2)$.

7. Determine the altitude of the pyramid $SABC$ dropped from the vertex S on the base ABC if $S(0, 6, 4)$, $A(3, 5, 3)$, $B(-2, 11, -5)$, and $C(1, -1, 4)$.

8. The normal vector $\mathbf{n} = \langle 2, -3, 5\rangle$ to the plane α defined by the equation $2x - 3y + 5z - 5 = 0$ originates at a point on the plane. Find whether the terminus of $\mathbf{n}$ lies on the same side of the plane α with the origin of coordinates.

9. Find the angle between two planes:

(a) $4x - 5y + 3z - 1 = 0$, $4x + 5y + 3z - 2 = 0$

(b) $3x - y + 5z + 2 = 0$, $5x + 3y - z + 10 = 0$

10. Given two planes through the point $M(-5, 16, 12)$: one plane contains the x-axis and the other the y-axis, find the cosine of the angle between these planes.

6.4

1. Find whether the point $M(-2, 1, 4)$ and the origin lie on the same side of the plane $3x - 7y + 2z - 1 = 0$.

2. In each of the cases find whether the points $M(2, -1, -1)$ and $N(1, 2, -3)$ lie in the dihedral angle formed by the planes or in adjacent angles or in vertical angles

(a) $3x - y + 2z - 3 = 0$, $x - 2y - z + 4 = 0$

(b) $2x - y + 5z - 1 = 0$, $3x - 2y + 6z - 1 = 0$

Chapter 7

A STRAIGHT LINE IN SPACE

7.1. EQUATIONS OF A LINE IN SPACE. EQUATIONS OF A STRAIGHT LINE

1. A line as the intersection of two surfaces.
A basic method of defining a line in spatial analytic geometry is its representation as the *intersection of two surfaces.*

Suppose we have two surfaces defined by the equations $F(x, y, z) = 0$ and $\Phi(x, y, z) = 0$, respectively. Then, their intersection consists of the points $M(x, y, z)$ such that their coordinates satisfy both equations simultaneously. Naturally, we can adopt the following definitions.

Definition 1. Suppose we have a system

$$\begin{cases} F(x, y, z) = 0 \\ \Phi(x, y, z) = 0 \end{cases}$$

The *line L defined by this system* is the set of all points of space whose coordinates satisfy the system.

Definition 2. The *equations of a line L* in space are any two equations $F(x, y, z) = 0$ and $\Phi(x, y, z) = 0$ such that the line defined by the system of these equations coincides with L.

2. Parametric representation of a line.

Another way of defining a line is to represent it *parametrically.* Suppose we have three arbitrary functions $f(t)$, $\varphi(t)$, and $\psi(t)$ defined on a set T (say, on an interval from a to b). Then the set of points $M(x, y, z)$ whose coordinates are given by the formulas

$$x = f(t),\ y = \varphi(t),\ z = \psi(t) \tag{1}$$

where the parameter t (the argument) runs over the set T, is a *line defined parametrically* and equations (1) are called the *parametric equations of the line.*

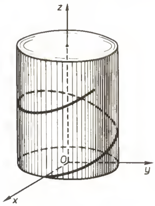


Figure 107

Example. The line defined by the parametric equations $x = R\cos t$, $y = R\sin t$, $z = t$, where t varies from 0 to 2π, is a circular helix whose projection on the xy-plane is a circle centered at the origin (Fig. 107).

3. Parametric equations of a straight line.

The position of a straight line in space can be given in different ways. For instance, a line can be determined by any two points on it, or by two planes which intersect along it, or by a point on it, or by two planes which intersect along it, or by a point on it and a vector parallel to it. Each has a different form for the equations describing the line.

We assume any nonzero vector parallel to a straight line to be the *direction vector of this line.*

Given a point $M_1(x_1, y_1, z_1)$ and a nonzero vector $\mathbf{a} = \langle p, q, r\rangle$ in a rectangular Cartesian coordinate system, find the equation of the straight line l passing through M_1 and whose direction vector is $\mathbf{a}$. Obviously, a point $M(x, y, z)$ lies on l if and only if the vector $\overrightarrow{M_1M}$ is collinear with $\mathbf{a}$, i.e. when there is a number t such that $\overrightarrow{M_1M} = t\mathbf{a}$. We write the last equation in coordinates:

$$x - x_1 = tp,\ y - y_1 = tq,\ z - z_1 = tr$$

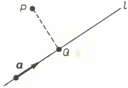


Figure 108

or

$$x = x_1 + tp,\ y = y_1 + tq,\ z = z_1 + tr \tag{2}$$

Thus, relations (2) are the *parametric equations of the straight line passing through the point* $M_1(x_1, y_1, z_1)$ *and whose direction vector is* $\mathbf{a} = \langle p, q, r\rangle$.

When the parameter t varies from $-\infty$ to $+\infty$, the point with the coordinates in (2) runs along the entire straight line.

Example 1. Write the parametric equations of the straight line through the point $M_1(-1, 2, 5)$ and parallel to the vector $\mathbf{a} = \langle 9, -2, 4\rangle$.

○ Using formulas (2) we write the desired equations in the form

$$x = -1 + 9t,\ y = 2 - 2t,\ z = 5 + 4t. \ \bullet$$

Example 2. Find the distance from the point $P(1, 1, 1)$ to the straight line l defined by the parametric equations

$$x = -4t,\ y = 8 + t,\ z = 6 + 3t \tag{3}$$

○ In order to find the distance from the point P to the straight line l it is necessary to find a point $Q(x, y, z)$ such that the vector $\overrightarrow{PQ}$ is perpendicular to l, i.e. $\overrightarrow{PQ} \perp \mathbf{a}$, where $\mathbf{a}$ is a direction vector of the straight line (Fig. 108). We can take the vector $\mathbf{a} = \langle -4, 1, 3\rangle$, whose coordinates equal the coefficients of t in equations (3), as the vector $\mathbf{a}$.

The coordinates of the vector $\overrightarrow{PQ}$ are $x - 1$, $y - 1$, $z - 1$. Using (3) we can write $\overrightarrow{PQ} = \langle -4t - 1, 7 + t, 5 + 3t\rangle$.

The condition for the vectors to be perpendicular is that their scalar product is zero. Therefore, instead of $\overrightarrow{PQ} \perp \mathbf{a}$ we can write

$$(-1 - 4t)\cdot(-4) + (7 + t)\cdot 1 + (5 + 3t)\cdot 3 = 0$$

whence $t = -1$. Thus, the value of the parameter t corresponding to the point Q is -1. Consequently, the coordinates of Q are $x = -4t = 4$, $y = 8 + t = 7$, $z = 6 + 3t = 3$. We can now find the distance from the point P to the straight line l

$$|PQ| = \sqrt{(4-1)^2 + (7-1)^2 + (3-1)^2} = \sqrt{49} = 7$$

4. Canonical equations of a straight line.

The parametric equations of a straight line, i.e. equations (2), express the proportionality between the numbers $x - x_1$, $y - y_1$, $z - z_1$ (the coordinates of $\overrightarrow{M_1M}$) and p, q, r (the coordinates of $\mathbf{a}$). Thus, we can write them as

$$\frac{x - x_1}{p} = \frac{y - y_1}{q} = \frac{z - z_1}{r} \tag{4}$$

Equations (4) are the *canonical (symmetric) equations of a straight line.*

We must clearly understand that (4) is a *system of two equations*

$$\frac{x - x_1}{p} = \frac{y - y_1}{q}$$

$$\frac{x - x_1}{p} = \frac{z - z_1}{r}$$

each of which is of first degree, that is the equation of a plane. Thus, the canonical equations define a straight line as a line along which two planes intersect.

Strictly speaking, representing the equations of a straight line in form (4) is only meaningful when the three numbers p, q, r are all nonzero. Nevertheless, equations (4) can also be used when one or even two of the numbers are zero. Say, for $p = 0$ we write

$$\frac{x - x_1}{0} = \frac{y - y_1}{q} = \frac{z - z_1}{r} \tag{5}$$

Let us show how this notation can be understood. We know that equations (4) express the collinearity of two vectors $\langle x - x_1, y - y_1, z - z_1 \rangle$ and $\langle p, q, r \rangle$. When $p = 0$, collinearity means that

$$x = x_1, \ \frac{y - y_1}{q} = \frac{z - z_1}{r}$$

This is precisely what equations (5) mean.

If $p = 0$ and $q = 0$, then we write (4) as

$$\frac{x - x_1}{0} = \frac{y - y_1}{0} = \frac{z - z_1}{r}$$

which we understand as a system of two equations

$$x = x_1, \ y = y_1$$

(explain why).

Example 1. Find the canonical equations of the straight line through two points $M_1(1, 0, -1)$ and $M_2(-2, 1, 2)$.

○ We take $\overrightarrow{M_1M_2} = \langle -3, 1, 3 \rangle$ as the direction vector, then the canonical equations (4) assume the form

$$\frac{x - 1}{-3} = \frac{y}{1} = \frac{z + 1}{3} \quad \bullet$$

Example 2. Write the canonical equations of the straight line through the point (1, 0, 2) and parallel to the y-axis.

○ We take the unit vector $\mathbf{j} = \langle 0, 1, 0 \rangle$ as the direction vector of the line and write the canonical equations as

$$\frac{x - 1}{0} = \frac{y}{1} = \frac{z - 2}{0}$$

Actually, these equations imply the system $x = 1$, $z = 2$ (y is arbitrary).

7.2. GENERAL EQUATIONS OF A STRAIGHT LINE

1. Finding a vector perpendicular to two given vectors. This subsection is optional. We present a method for finding from two vectors in space $\mathbf{n}_1$ and $\mathbf{n}_2$ a third vector $\mathbf{n}$ perpendicular to each of the vectors. Here if $\mathbf{n}_1$ and $\mathbf{n}_2$ are not collinear, then $\mathbf{n} \neq 0$; otherwise $\mathbf{n} = 0$.

We introduce a rectangular Cartesian coordinate system in space and assume that the vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ are specified by their coordinates: $\mathbf{n}_1 = \langle A_1, B_1, C_1 \rangle$, $\mathbf{n}_2 = \langle A_2, B_2, C_2 \rangle$.

Theorem. *The vector* $\mathbf{n}$ *with coordinates*

$$\begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix}, \quad - \begin{vmatrix} A_1 & C_1 \\ A_2 & C_2 \end{vmatrix}, \quad \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} \tag{1}$$

is perpendicular to each of the vectors $\mathbf{n}_1$ *and* $\mathbf{n}_2$; $\mathbf{n} = 0$ *if and only if* $\mathbf{n}_1$ *and* $\mathbf{n}_2$ *are collinear.*

Remark. Expression (1) for the coordinates of $\mathbf{n}$ can be easily remembered in the following way. We form the array

$$\begin{matrix} A_1 B_1 C_1 \\ A_2 B_2 C_2 \end{matrix}$$

from the coordinates of $\mathbf{n}_1$ and $\mathbf{n}_2$. Eliminating in turn the first, second, and third columns yields three square matrices

$$\begin{pmatrix} B_1 & C_1 \\ B_2 & C_2 \end{pmatrix} \begin{pmatrix} A_1 & C_1 \\ A_2 & C_2 \end{pmatrix} \begin{pmatrix} A_1 & B_1 \\ A_2 & B_2 \end{pmatrix}$$

whose determinants are the coordinates of the vector $\mathbf{n}$; the second determinant should be multiplied by -1.

□ In order to ascertain that $\mathbf{n}$ and $\mathbf{n}_i$ ($i = 1, 2$) are perpendicular, it is sufficient to verify that their scalar product is zero. We have

$$\begin{aligned} \mathbf{n}_1 \cdot \mathbf{n} &= A_1 \cdot \begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix} - B_1 \cdot \begin{vmatrix} A_1 & C_1 \\ A_2 & C_2 \end{vmatrix} + C_1 \cdot \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} \\ &= A_1(B_1C_2 - B_2C_1) - B_1(A_1C_2 - A_2C_1) \\ &\quad + C_1(A_1B_2 - A_2B_1) \end{aligned}$$

Collecting like terms on the right-hand side yields $\mathbf{n}_1 \cdot \mathbf{n} = 0$. Similarly, we can verify that $\mathbf{n}_2 \cdot \mathbf{n} = 0$.

If the vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ are collinear, then their coordinates are proportional and then each determinant in (1) is zero. Hence it follows in this case that $\mathbf{n}$ is zero.

Conversely, let $\mathbf{n} = 0$, i.e. the three determinants in (1) are zero. We prove that in this case $\mathbf{n}_1$ and $\mathbf{n}_2$ are collinear. If $\mathbf{n}_2 = 0$, then the proof is trivial: a zero vector is collinear with any vector. Let $\mathbf{n}_2 \neq 0$, then at least one coordinate of $\mathbf{n}_2$, say A_2, is nonzero. Since the second determinant in (1) is zero, it follows that its rows are collinear (as vectors in the plane). Because the vector corresponding to the second row is nonzero (since $A_2 \neq 0$), we have, by the theorem in Sec. 1.1, subsec. 4,

$$A_1 = \lambda A_2, \quad C_1 = \lambda C_2$$

where λ is a scalar. Similarly, since the third determinant in (1) is zero, we have

$$A_1 = \mu A_2, \quad B_1 = \mu B_2$$

where μ is a scalar. It follows from the last two equations for A_1 that $\lambda = \mu$ (since $A_2 \neq 0$). Hence,

$$A_1 = \lambda A_2, \quad B_1 = \lambda B_2, \quad C_1 = \lambda C_2$$

that is, $\mathbf{n}_1$ is collinear with $\mathbf{n}_2$. ∎

We denote the vector $\mathbf{n}$, whose coordinates are given in (1), by $\mathbf{n}_1 \times \mathbf{n}_2$.

Thus, we introduce a new operation of mutiplication for vectors in space. Whereas the scalar multiplication of two vectors yields a number, the new operation results in a vector and $\mathbf{n}_1 \times \mathbf{n}_2$ is the *vector multiplication* of $\mathbf{n}_1$ and $\mathbf{n}_2$.

Example 1. Show that

$$\mathbf{i} \times \mathbf{j} = \mathbf{k}, \ \mathbf{j} \times \mathbf{k} = \mathbf{i}, \ \mathbf{k} \times \mathbf{i} = \mathbf{j} \tag{2}$$

where **i**, **j**, and **k** are the unit vectors.

○ Since $\mathbf{i} = \langle 1, 0, 0 \rangle$, $\mathbf{j} = \langle 0, 1, 0 \rangle$, and $\mathbf{k} = \langle 0, 0, 1 \rangle$, the vector $\mathbf{i} \times \mathbf{j}$ has the coordinates

$$\begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix}, \quad -\begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix}, \quad \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$$

i.e. $\mathbf{i} \times \mathbf{j} = \langle 0, 0, 1 \rangle = \mathbf{k}$.

The remaining two equations in (2) can be verified in a similar way. ●

Example 2. Find $\mathbf{a} \times \mathbf{b}$, with $\mathbf{a} = \langle 1, 1, 1 \rangle$ and $\mathbf{b} = \langle 1, -1, 1 \rangle$.

○ The coordinates of the vector $\mathbf{a} \times \mathbf{b}$ are

$$\begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix}, \quad -\begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix}, \quad \begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix}$$

i.e. $\mathbf{a} \times \mathbf{b} = \langle 2, 0, 2 \rangle$. ●

2. General equations of a straight line.

Two nonparallel planes define a straight line along which they intersect. Consequently, the system of equations

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0 \\ A_2x + B_2y + C_2z + D_2 = 0 \end{cases} \tag{3}$$

defines a straight line in space provided the vectors $\mathbf{n}_1 = \langle A_1, B_1, C_1 \rangle$ and $\mathbf{n}_2 = \langle A_2, B_2, C_2 \rangle$ are noncollinear.

Equations (3) are called the *general equations of the straight line.*

We can transform the general equations to canonical ones. To do this it is necessary to choose the direction vector of the straight line and a point on that line.

Let us show that we can take the vector $\mathbf{n} = \mathbf{n}_1 \times \mathbf{n}_2$ as direction vector of the line defined by (3), i.e.

$$\mathbf{n} = \left\langle \begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix}, \quad -\begin{vmatrix} A_1 & C_1 \\ A_2 & C_2 \end{vmatrix}, \quad \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} \right\rangle$$

We use α to denote one plane and β to denote the other plane defined by (3). The vector $\mathbf{n}_1 = \langle A_1, B_1, C_1 \rangle$ is perpendicular to α and the vector $\mathbf{n}_2 = \langle A_2, B_2, C_2 \rangle$ is perpendicular to β. Since $\mathbf{n}$ is perpendicular to $\mathbf{n}_1$ and $\mathbf{n}_2$ (according to the theorem on p. 179), it is parallel to the planes α and β, respectively, and consequently, to the straight line in which the planes intersect. In other words, $\mathbf{n}$ is the direction vector of the straight line defined by (3). Note that $\mathbf{n} \neq 0$ since $\mathbf{n}_1$ and $\mathbf{n}_2$ are not collinear by assumption.

We now show how we can find the coordinates of the point through which the line defined by (3) passes. Since $\mathbf{n} \neq 0$, at least one coordinate of **n** is nonzero. Let, for instance,

$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} \neq 0$$

We write system (3) as

$$\begin{cases} A_1x + B_1y = D_1 - C_1z \\ A_2x + B_2y = D_2 - C_2z \end{cases} \tag{4}$$

and setting z equal to a number z_1 (say, zero), we find from (4) the values of x and y, viz., $x = x_1$, $y = y_1$. The three numbers define the desired point. If we know the point $M_1(x_1, y_1, z_1)$ on the straight line (3) and the direction vector n, then we can write the equation of that line

$$\frac{x - x_1}{\begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix}} = \frac{y - y_1}{-\begin{vmatrix} A_1 & C_1 \\ A_2 & C_2 \end{vmatrix}} = \frac{z - z_1}{\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}}$$

Example. Write the canonical equations of the straight line

$$\begin{cases} x + 2y + 3z - 1 = 0 \\ x - y - z + 2 = 0 \end{cases} \tag{5}$$

○ We first find the direction vector. We have

$$\begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix} = \begin{vmatrix} 2 & 3 \\ -1 & -1 \end{vmatrix} = 1, \quad \begin{vmatrix} A_1 & C_1 \\ A_2 & C_2 \end{vmatrix} = \begin{vmatrix} 1 & 3 \\ 1 & -1 \end{vmatrix} = -4$$

$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 1 & -1 \end{vmatrix} = -3$$

thus, the direction vector is $\mathbf{n} = \langle 1, 4, -3 \rangle$. Then we have to find a point on the line, that is, a solution of system (5). Setting $z = 0$ yields the system

$$\begin{cases} x + 2y = 1 \\ x - y = -2 \end{cases}$$

We solve it to find that $x = -1$, $y = 1$. Thus, $(-1, 1, 0)$ is a point on the straight line defined by (5). The canonical equations of this line have the form

$$\frac{x+1}{1} = \frac{y-1}{4} = \frac{z}{-3} \quad \bullet$$

7.3. RELATIVE POSITION OF TWO STRAIGHT LINES

Suppose two straight lines l_1 and l_2 are defined in a rectangular Cartesian coordinate system by canonical equations

$$\frac{x - x_1}{p_1} = \frac{y - y_1}{q_1} = \frac{z - z_1}{r_1}$$

$$\frac{x - x_2}{p_2} = \frac{y - y_2}{q_2} = \frac{z - z_2}{r_2}$$

Two cases are possible here:

(1) l_1 and l_2 lie in the same plane, and

(2) l_1 and l_2 lie in different planes, that is, they are skew.

Let us find when case 1 occurs. We consider three vectors $\mathbf{a}_1 = \langle p_1, q_1, r_1 \rangle$ and $\mathbf{a}_2 = \langle p_2, q_2, r_2 \rangle$ which are the direction vectors of the straight lines l_1 and l_2, and $\overrightarrow{M_1M_2} = \langle x_2 - x_1, y_2 - y_1, z_2 - z_1 \rangle$ (Fig. 109). For the lines l_1 and l_2 to lie in the same plane, it is necessary and sufficient that the three vectors are coplanar.

The condition that three vectors are coplanar (see Sec. 3.1) is that the third-order determinant composed of the coor-

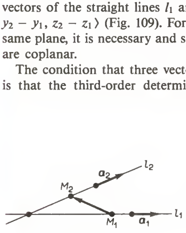


Figure 109

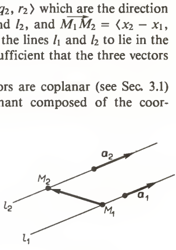


Figure 110

dinates of the vectors is zero:

$$\begin{vmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ p_1 & q_1 & r_1 \\ p_2 & q_2 & r_2 \end{vmatrix} = 0 \qquad (1)$$

Thus, the *straight lines l_1 and l_2 lie in the same plane if and only if condition* (1) *is satisfied.*

Whence the condition that straight lines are skew: the *straight lines are skew if and only if*

$$\begin{vmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ p_1 & q_1 & r_1 \\ p_2 & q_2 & r_2 \end{vmatrix} \neq 0$$

If straight lines lie in the plane, then they intersect, or are parallel and do not coincide, or coincide. When the lines coincide, the three vectors $\overrightarrow{M_1M_2}$, $\mathbf{a}_1$, and $\mathbf{a}_2$ must be collinear. When the lines are parallel, vector $\mathbf{a}_1$ is collinear with $\mathbf{a}_2$ but they are not collinear with $\overrightarrow{M_1M_2}$ (Fig. 110). If the lines intersect, then $\mathbf{a}_1$ and $\mathbf{a}_2$ are not collinear.

Example. Find the relative position of the straight lines

$$\frac{x-1}{1} = \frac{y}{2} = \frac{z-2}{1}, \quad \frac{x-2}{3} = \frac{y-2}{1} = \frac{z-3}{2}$$

○ We have

$$\begin{vmatrix} 2-1 & 2-0 & 3-2 \\ 1 & 2 & 1 \\ 3 & 1 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 1 \\ 1 & 2 & 1 \\ 3 & 1 & 2 \end{vmatrix} = 0$$

thus, the lines are in the same plane. The vectors $\mathbf{a}_1 = \langle 1, 2, 1 \rangle$ and $\mathbf{a}_2 = \langle 3, 2, 1 \rangle$ are not collinear and hence the lines intersect. In order to find the intersection point, we write the equations of the first straight line in parametric form: $x = 1 + t$, $y = 2t$, $z = 2 + t$, and substitute them into the equations of the second line. We have the following system of equations

$$\frac{t-1}{3} = \frac{2t-2}{1} = \frac{t-1}{2}$$

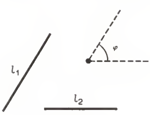


Figure 111

Since the straight lines intersect, the system must be consistent. Solving it yields $t = 1$. Hence the coordinates of the point of intersection are $x = 2$, $y = 2$, $z = 3$. ●

1. Angle between two straight lines.

The angle φ between two straight lines in space is the *smallest of two adjacent angles* formed by two straight lines drawn through an arbitrary point in space and parallel to the two given lines (Fig. 111).

Let $\mathbf{a}_1 = \langle p_1, q_1, r_1 \rangle$ and $\mathbf{a}_2 = \langle p_2, q_2, r_2 \rangle$ be direction vectors of the given lines. The angle between these vectors either equals the angle φ between the lines or is the supplement of φ. Therefore

$$\cos\varphi = \left| \frac{\mathbf{a}_1 \cdot \mathbf{a}_2}{|\mathbf{a}_1| \cdot |\mathbf{a}_2|} \right| = \frac{|p_1p_2 + q_1q_2 + r_2r_2|}{\sqrt{p_1^2 + q_1^2 + r_1^2}\sqrt{p_2^2 + q_2^2 + r_2^2}}$$

If the straight lines are perpendicular, then $\cos\varphi = 0$ (the converse is also true), then the condition that the straight lines are perpendicular has the form

$$p_1p_2 + q_1q_2 + r_1r_2 = 0$$

If the straight lines are parallel, then $\mathbf{a}_1$ and $\mathbf{a}_2$ are collinear (the converse is also true), then the condition that the lines are parallel is

$$\frac{p_1}{p_2} = \frac{q_1}{q_2} = \frac{r_1}{r_2}$$

7.4. RELATIVE POSITION OF A STRAIGHT LINE AND A PLANE

The following cases are possible for a straight line l and a plane α in space:

(1) a straight line l intersects a plane α, i.e. they have a point in common,

(2) l is parallel to α but does not lie in α, in this case they have no points in common,

(3) l lies in α.

Let us find how to distinguish between these cases if the plane α is defined by the equation

$$Ax + By + Cz + D = 0$$

and the straight line l by the canonical equations

$$\frac{x - x_1}{p} = \frac{y - y_1}{q} = \frac{z - z_1}{r}$$

We consider two vectors $\mathbf{n} = \langle A, B, C \rangle$ and $\mathbf{a} = \langle p, q, r \rangle$, one of them being perpendicular to the plane α and the other to the straight line l.

Obviously, the line l is parallel to α if and only if $\mathbf{a}$ is perpendicular to $\mathbf{n}$, that is when the scalar products of these vectors are zero. Consequently,

$$Ap + Bq + Cr = 0 \tag{1}$$

is the *necessary and sufficient condition for the straight line l and the plane α to be parallel.* Now, if this condition is not fulfilled, i.e. *if*

$$Ap + Bq + Cr \neq 0$$

then l and α intersect.

Suppose that the straight line l is parallel to the plane α, i.e. condition (1) is valid.

The straight line lies in the plane α if and only if a point $M_1(x_1, y_1, z_1)$ on l lies in this plane. Consequently, the necessary and sufficient conditions that the straight line l lies in the plane α can be written as

$$Ap + Bq + Cr = 0$$
$$Ax_1 + By_1 + Cz_1 + D = 0$$

Example. Let us consider the straight line

$$\frac{x-3}{5} = \frac{y+2}{3} = \frac{z}{-1}$$

and the three planes

$$3x - 4y + 7z + 4 = 0 \quad (\alpha)$$
$$x - y + 2z + 3 = 0 \quad (\beta)$$
$$x - y + 2z - 5 = 0 \quad (\gamma)$$

The straight line l intersects the plane α since $3\cdot 5 + (-4)\cdot 3 + 7\cdot(-1) \neq 0$. At the same time, the line is parallel to a plane β since $1\cdot 5 + (-1)\cdot 3 + 2\cdot(-1) = 0$, but does not lie in β since the point $M_1(3, -2, 0)$ on this line does not belong to β. Finally, l lies in the plane γ since l is parallel to γ and M_1 belongs to γ.

1. Angle between a straight line and a plane.

The angle θ between a straight line and a plane is the *angle formed by the line and its projection on the plane*. The angle θ varies from 0 to 90°.

Let a plane be defined by the equation

$$Ax + By + Cz + D = 0$$

and a straight line by the canonical equations

$$\frac{x-x_1}{p} = \frac{y-y_1}{q} = \frac{z-z_1}{r}$$

The vector $\mathbf{n} = \langle A, B, C\rangle$ is perpendicular to the plane and the vector $\mathbf{a} = \langle p, q, r\rangle$ is parallel to the straight line.

We denote the angle between the vectors $\mathbf{n}$ and $\mathbf{a}$ by φ. Figure 112 shows that $\varphi = 90° \pm \theta$, hence $\sin\theta = |\cos\varphi|$. Whence we have

$$\sin\theta = \left|\frac{|\mathbf{n}\cdot\mathbf{a}|}{|\mathbf{n}|\cdot|\mathbf{a}|}\right|$$

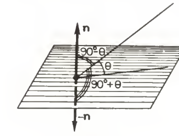


Figure 112

or in coordinate form

$$\sin\theta = \frac{|Ap + Bq + Cr|}{\sqrt{A^2 + B^2 + C^2}\sqrt{p^2 + q^2 + r^2}} \tag{2}$$

We can find from (2) the angle between a straight line and a plane using the coefficients of their equations.

Example. Find the angle between the straight line

$$\begin{cases} 6x - 2y - z - 20 = 0 \\ 15x - 2y - 4z - 8 = 0 \end{cases}$$

and the plane $6x + 15y - 10z + 31 = 0$.

○ We first find the direction vector of the line. To do this, according to Sec. 7.2, we should take the normal vectors of the planes defining the line, i.e. $\mathbf{n}_1 = \langle 6, -2, -1\rangle$ and $\mathbf{n}_2 = \langle 15, -2, -4\rangle$, and then find the vector $\mathbf{a} = \mathbf{n}_1 \times \mathbf{n}_2$. We have

$$\mathbf{a} = \left\langle \begin{vmatrix} -2 & -1 \\ -2 & -4 \end{vmatrix}, -\begin{vmatrix} 6 & -1 \\ 15 & -4 \end{vmatrix}, \begin{vmatrix} 6 & -2 \\ 15 & -2 \end{vmatrix} \right\rangle$$

or $\mathbf{a} = \langle 6, 9, 18\rangle$. To simplify our calculation still further, it is convenient to divide the coordinates of $\mathbf{a}$ by 3; we have $\mathbf{a}' = \langle 2, 3, 6\rangle$. Using (2) we find the sine of the angle between the straight line and the plane:

$$\sin\theta = \frac{|6\cdot 2 + 15\cdot 3 - 10\cdot 6|}{\sqrt{6^2 + 15^2 + (-10)^2}\sqrt{2^2 + 3^2 + 6^2}} = \frac{3}{19\cdot 7} = \frac{3}{133}$$

From trigonometric tables we find that $\theta \approx 1°18'$. ●

Exercises to Chapter 7

7.1

1. Find the equation of the straight line through the origin and the point (a, b, c).

2. Draw a straight line through the point $(3, -2, 7)$ and parallel to the straight line $\frac{x-1}{3} = \frac{y+2}{-3} = \frac{z+4}{5}$.

3. Given the equations of motion of a point $x = 1 - 6t$, $y = 2 + 3t$, $z = 5 - 2t$, find its velocity v.

4. Given the equations of motion of a point $x = 3 - 2t$, $y = -4 + t$, $z = 1 + 2t$, find the distance covered by the point in the time interval from $t_1 = 0$ to $t_2 = 5$.

5. Draw a straight line through the point $(-1, 2, 4)$ and perpendicular to the plane $3x + 2y - 7z + 5 = 0$.

6. Find the projection of the point $(11, -1, 6)$ on the plane $5x - y + 2z - 8 = 0$.

7. Write the equations of the sides of ΔABC with vertices $A(1, 1, 1)$, $B(-2, 3, 4)$, and $C(4, -5, 5)$.

8. Set up the equation of a plane passing through the point $(-1, 0, 1)$ and the straight line $x = 1 + 5t$, $y = -4 + 2t$, $z = -1 - t$.

9. Write the equation of the projection of the straight line $\frac{x+1}{4} = \frac{y-1}{6} = \frac{z-2}{-1}$ on the xy-plane.

10. Write the equations of the perpendicular from the point $(2, 4, 1)$ onto the straight line $\frac{x-1}{5} = \frac{y-1}{-1} = \frac{z-2}{2}$.

11. Find a point symmetric to the point $(6, 5, -4)$ with respect to the straight line $\frac{x+2}{7} = \frac{y-3}{1} = \frac{z+5}{-3}$.

12. Find the distance between two straight lines

$$\frac{x}{5} = \frac{y+3}{-2} = \frac{z-1}{1} \quad \text{and} \quad \frac{x-6}{5} = \frac{y-1}{-2} = \frac{z-9}{1}.$$

13. Given a cube with side equal to 1, find the distance from a vertex to the diagonal not passing through it.

14. Verify whether a plane perpendicular to the diagonal of a cube and passing through its midpoint intersects the cube along a regular hexagon.

7.2

1. Reduce the equations of a straight line

$$\begin{cases} x + y - z + 6 = 0 \\ 4x - 11y + 2z + 27 = 0 \end{cases}$$

to canonical form.

2. Prove that the straight lines

$$\frac{x+2}{3} = \frac{y-1}{-2} = \frac{z}{1} \quad \text{and} \quad \begin{cases} x + y - z = 0 \\ x - y - 5z - 8 = 0 \end{cases}$$

are parallel.

3. Describe relative positions of the straight lines:

$$(1) \begin{cases} 5x + 7 = 0 \\ 3y - 2 = 0 \end{cases} \quad (2) \begin{cases} 3x - 1 = 0 \\ 2y + 5z = 0 \end{cases} \quad (3) \begin{cases} A_1x + B_1y + C_1z = 0 \\ A_2x + B_2y + C_2z = 0 \end{cases}$$

4. Form the equations of the projection of the straight line

$$\begin{cases} x - y - z = 0 \\ 2x + 3y + z - 5 = 0 \end{cases}$$

on the yz-plane.

5. Form the equations of the projection of the straight line

$$\begin{cases} x - 4y + 2z - 5 = 0 \\ 3x + y - z + 2 = 0 \end{cases}$$

on the plane $2x + 3y + z - 6 = 0$.

7.3

1. Prove that the straight lines $\dfrac{x-1}{3} = \dfrac{y+1}{2} = \dfrac{z}{-5}$ and $x = -3 + 7t$, $y = 6 - 5t$, $z = -6 + t$ lie in the same plane and find the equation of the plane. Find the point at which the lines intersect.

2. Given the straight lines

$$\frac{x-6}{7} = \frac{y-1}{-1} = \frac{z-1}{3} \quad \text{and} \quad \frac{x+1}{4} = \frac{y-2}{a} = \frac{z-4}{1}$$

find a such that the lines intersect.

3. Find whether the straight lines

$$\begin{cases} 3x + 5y + 5z - 3 = 0 \\ x + 5y + 14 = 0 \end{cases} \quad \text{and} \quad \frac{x-2}{1} = \frac{y-6}{3} = \frac{z+7}{-4}$$

intersect.

4. Find the angle between the straight lines

$$\frac{x-1}{3} = \frac{y+2}{6} = \frac{z-5}{2} \quad \text{and} \quad \frac{x}{2} = \frac{y-3}{9} = \frac{z+1}{6}$$

7.4

1. Given the straight line through the points $(x_1, y_1, 0)$ and $(x_2, 0, z_2)$, find the point at which this line and the yz-plane intersect.

2. Find the value of A for which the plane $Ax - 3y + 7z - 2 = 0$ is parallel to the straight line $\frac{x+1}{2} = \frac{y}{1} = \frac{z-4}{-3}$.

3. Find the values of A and B such that the plane $Ax + By + z - 9 = 0$ is perpendicular to the straight line $\frac{x}{4} = \frac{y+1}{-5} = \frac{z-3}{-1}$.

4. Find the point at which the straight line $\frac{x}{2} = \frac{y-3}{5} = \frac{z+1}{-3}$ and the plane $3x - 5y - 10z - 6 = 0$.

5. Find the point at which the straight line

$$\begin{cases} 7x + 2y + 3z - 15 = 0 \\ 5x - 3y + 2z - 15 = 0 \end{cases}$$

and the plane $10x - 11y + 5z - 36 = 0$ intersect.

6. Find the sines of the angles formed by the straight line

$$\begin{cases} 4x - 6y + 3z + 18 = 0 \\ x - z + 3 = 0 \end{cases}$$

and the coordinate axes.

Chapter 8

QUADRIC SURFACES

Any surface whose equation in rectilinear Cartesian coordinates has the form

$$\Phi(x, y, z) = 0$$

where $\Phi(x, y, z)$ is the second-degree polynomial in x, y, z, i.e.

$$\Phi(x, y, z) = Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + \\ + Gx + Hy + Kz + L$$

is called a *quadric surface.*

We shall not consider cylindrical surfaces, i.e. those defined in some coordinates by an equation of the form

$$\Phi(x, y) = 0$$

(its left-hand side is independent of z). Then, as can be shown, the equation of a surface in convenient coordinates reduces to one of the two forms:

$$Ax^2 + By^2 + Cz^2 + L = 0 \text{ with } A, B, C, L \text{ nonzero} \quad \text{(I)}$$

$$Ax^2 + By^2 + Kz = 0 \text{ with } A, B, K \text{ nonzero} \quad \text{(II)}$$

Ellipsoids and *hyperboloids* are the most interesting examples of surfaces defined by (I), and the graphs of (II) are *paraboloids.*

8.1. THE ELLIPSOID

Definition. The *ellipsoid* is a surface defined by the equation in convenient rectangular coordinates:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \ a \neq 0, \ b \neq 0, \ c \neq 0 \quad (1)$$

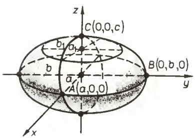


Figure 113

In order to visualize the shape of an ellipsoid and to sketch it in a plane, we use the method of parallel sections, that is examine its intersections with various planes.

Let the cutting planes be parallel to the xy-plane, each of them being defined by the equation $z = h$ (= const). The line of intersection of the plane and an ellipsoid is determined by the system of equations

$$\begin{cases} \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1 \\ z = h \end{cases} \quad \text{or} \quad \begin{cases} \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1 - \dfrac{h^2}{c^2} \\ z = h \end{cases}$$

We rewrite it as

$$\begin{cases} \dfrac{x^2}{a^2(1 - h^2/c^2)} + \dfrac{y^2}{b^2(1 - h^2/c^2)} = 1 \\ z = h \end{cases} \qquad (2)$$

If $|h| < c$, i.e. $-c < h < c$, then $1 - h^2/c^2$ is positive and the section is an ellipse (Fig. 113) with the semiaxes

$$a_1 = a\sqrt{1 - \frac{h^2}{c^2}}, \quad b_1 = b\sqrt{1 - \frac{h^2}{c^2}}$$

The semiaxes have maximal values when $h = 0$, namely $a_1 = a$ and $b_1 = b$.

If h increases from 0 to c or decreases from 0 to $-c$, then a_1 and b_1 decrease from a to 0 and from b to 0, respectively.

When $h = c$ or $h = -c$, the sections of the ellipsoid are the points $(0, 0, c)$ and $(0, 0, -c)$. When $|h| > c$, the first equation in system (2) has no solution since its left-hand side is nonpositive; consequently, there are no points belonging to the ellipsoid outside of the strips between the planes $z = c$ and $z = -c$.

Similar results are obtained when the ellipsoid and the planes parallel to the xz- and yz-axes intersect.

Note once more that the intersection of the ellipsoid defined by (1) and the xy-plane is the ellipse

$$\begin{cases} \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \\ y = 0 \end{cases}$$

and this ellipse has the longest semiaxes (a and b) out of all its sections by the planes parallel to the xy-plane. Similarly, the intersection of the ellipsoid and the xz-plane is the ellipse

$$\begin{cases} \frac{x^2}{a^2} + \frac{z^2}{c^2} = 1 \\ x = 0 \end{cases}$$

and it has the longest semiaxes (a and c) out of all its sections by the planes parallel to xz-plane. Finally, the intersection of the ellipsoid and the yz-plane is the ellipse

$$\begin{cases} \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \\ x = 0 \end{cases}$$

It has the longest semiaxes (b and c) out of all its sections by the planes parallel to the yz-axis.

The quantities a, b, and c are the *semiaxes of the ellipsoid* defined by (1). If $a \neq b$, $b \neq c$, $c \neq a$, the ellipsoid is called *tri-axial* or *symmetric*.

When two of the semiaxes are equal, the ellipsoid is a surface of revolution. For instance, if $a = b$, the equation of the ellipsoid assumes the form

$$\frac{x^2}{a^2} + \frac{y^2}{a^2} + \frac{z^2}{c^2} = 1$$

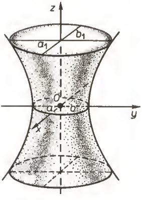


Figure 114

and its section by the plane $z = h$, with $|h| < c$, is a *circle* with center on the z-axis, and the ellipsoid is a surface of revolution whose axis is the z-axis. Such an ellipsoid is called an *ellipsoid of revolution* (or *spheroid*).

Finally, when the three semiaxes of an ellipsoid are equal, the equation of the ellipsoid assumes the form

$$\frac{x^2}{a^2} + \frac{y^2}{a^2} + \frac{z^2}{a^2} = 1$$

i.e. it is a *sphere* with the equation

$$x^2 + y^2 + z^2 = a^2$$

8.2. THE HYPERBOLOID

1. The hyperboloid of one sheet.

Definition. The *hyperboloid of one sheet* is the surface defined by the following equation in convenient coordinates:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1, \ a > 0, \ b > 0, \ c > 0 \tag{1}$$

By cutting the hyperboloid defined by (1) by planes parallel to the coordinate planes, we can find its shape in a similar manner as we did in the case of an ellipsoid.

The section of the hyperboloid in (1) by the plane $z = h$ is a line defined by the equation

$$\begin{cases} \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 + \frac{h^2}{c^2} \\ z = h \end{cases}$$

i.e. is an ellipse (Fig. 114) centered at the point $(0, 0, h)$ and whose semiaxes are $a_1 = a\sqrt{1 + h^2/c^2}$ and $b_1 = b\sqrt{1 + h^2/c^2}$. When $h = 0$, the semiaxes have minimal values $a_1 = a$ and $b_1 = b$; as $|h|$ increases, i.e. as the cutting plane moves away from the xy-plane, the values of the semiaxes increase indefinitely.

The intersection of the surface in (1) and the xz-plane is the hyperboloid defined by the equations

$$\begin{cases} \frac{x^2}{a^2} - \frac{z^2}{c^2} = 1 \\ y = 0 \end{cases}$$

The section of the surface by the yz-axis is the hyperbola with the equations

$$\begin{cases} \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \\ x = 0 \end{cases}$$

It can be shown that any section of the surface in (1) by planes passing through the z-axis is a hyperbola.

The hyperboloid of one sheet given by (1) opens indefinitely along the z-axis as $|z|$ grows; the section by the plane $z = 0$ is the "neck".

When $a = b$, equation (1) becomes

$$\frac{x^2}{a^2} + \frac{y^2}{a^2} - \frac{z^2}{c^2} = 1$$

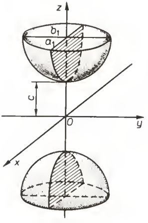


Figure 115

Sections of this surface by the planes $z = h$ are circles with centers on the z-axis. Consequently, in this case the hyperboloid is a surface of revolution about the z-axis. Such a surface is called a *hyperboloid of revolution of one sheet.*

2. Hyperboloid of two sheets.

Definition. The *hyperboloid of two sheets* is the surface defined by the following equation in rectangular coordinates:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1, \quad a > 0, \quad b > 0, \quad c > 0 \tag{2}$$

This equation differs from (1) in the sign of the right-hand side. In order to find the shape of the surface in (2), we consider its sections by the planes $z = h$. These sections are lines lying in the planes $z = h$ and defined by the equations

$$\begin{cases} \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = \dfrac{h^2}{c^2} - 1 \\ z = h \end{cases} \tag{3}$$

When $|h| < c$, the right-hand side of (3) is negative; consequently, the section is "empty". Thus, there are no points belonging to the hyperboloid in (3) in the strip between the planes $z = c$ and $z = -c$. When $|h| = c$, i.e. when $z = c$ or $z = -c$, the only solution of (3) is $x = 0$, $y = 0$. Hence it follows that the sections by the planes $z = c$ and $z = -c$ are the points (0, 0, c) and (0, 0, $-c$) respectively. When $|h| > c$, equation (3) defines an ellipse with the semiaxes $a_1 = a\sqrt{h^2/c^2 - 1}$ and $b_1 = b\sqrt{h^2/c^2 - 1}$ in the plane $z = h$. As $|h|$ increases from c to infinity, a_1 and b_1 increase from zero to infinity.

Thus, the surface (2) consists of two sheets, one lying in the subspace $z > c$ and the other in the subspace $z < -c$ (Fig. 115).

The sections of the surface by the xz- and yz-planes are the hyperbolas

$$\begin{cases} \dfrac{x^2}{a^2} - \dfrac{z^2}{c^2} = -1 \\ y = 0 \end{cases} \quad \text{and} \quad \begin{cases} \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = -1 \\ x = 0 \end{cases}$$

It can be shown that any section by a plane passing through the z-axis is a hyperbola.

For $a = b$ (2) is a surface of revolution about the z-axis, which is called a *hyperboloid of revolution of two sheets.*

8.3. THE PARABOLOID

1. The elliptic paraboloid.

Definition. The *elliptic paraboloid* is the surface defined in rectangular coordinates by the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = z, \; a > 0, \; b > 0 \tag{1}$$

This surface and the plane $z = h$ can only intersect when $h \geqslant 0$. For $h = 0$ the section is a point, the origin. For $h > 0$ the section is the ellipse defined by

$$\begin{cases} z = h \\ \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = h \end{cases}$$

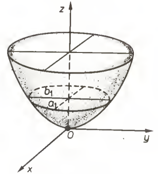


Figure 116

with the semiaxes $a_1 = a\sqrt{h}$ and $b_1 = b\sqrt{h}$. As h increases from zero to infinity, a_1 and b_1 also increase from zero to infinity.

Let us consider the section of the surface (1) by the xz-plane. This section is defined by the equations $y = 0$ and $\frac{1}{a^2} x^2 = z$, which is a parabola with the z-axis and lying in the plane $y = 0$. Similarly, the section by the yz-plane is the parabola $x = 0$, $\frac{1}{b^2} y^2 = z$ with the z-axis and lying in the plane $x = 0$. It can be shown that any section by a plane passing through the z-axis is a parabola with the z-axis.

It is now clear why the surface (1) is called the elliptic paraboloid (Fig. 116). For $a = b$ equation (1) assumes the form

$$x^2 + y^2 = a^2 z \tag{2}$$

In this case the sections by the planes $z = h$, where $h > 0$, are circles with centers on the z-plane. This means that (2) is a surface of revolution with the z-axis and is called an *elliptic paraboloid of revolution*.

2. The hyperbolic paraboloid.

Definition. The *hyperbolic paraboloid* is the surface defined in rectangular coordinates by the equation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = z, \ a > 0, \ b > 0 \tag{3}$$

The section by the plane $y = 0$ is the parabola

$$y = 0,\ z = \frac{1}{a^2} x^2 \tag{4}$$

with the vertex at the origin and the z-axis as the symmetry axis. Since here $1/a^2 > 0$, the parabola lies in the half-plane $y = 0$, $z > 0$, i.e. the parabola opens upward (with respect to the z-axis).

The section by the plane $x = 0$ is a parabola with the equations

$$x = 0,\ z = -\frac{1}{b^2} y^2 \tag{5}$$

i.e. its vertex is at the origin and the parabola is symmetric about the z-axis; it opens downward since here $-1/b^2 < 0$.

In order to visualize the shape of the surface in (3), clearly, we consider the section by the plane $x = h$, where h is a number. This is the parabola

$$x = h,\ z = -\frac{1}{b^2} y^2 + \frac{h^2}{a^2}$$

which results from the parabola in (5) by displacing it upward (i.e. in the positive direction of the z-axis) by h^2/a^2. The vertex of the new parabola has the coordinates

$$x = h,\ y = 0,\ z = \frac{h^2}{a^2}$$

i.e. it lies on the parabola in (4).

Thus, in order to obtain the surface (3) we should use the parabolas (4) and (5), which lie in perpendicular planes and have their vertices and the axes of symmetry in common, one parabola opening upward and the other opening downward. Then one of them should be translated in space so that its vertex moves along the other parabola. The resulting surface is saddle-shaped (Fig. 117).

Let us consider sections of this surface by the planes $z = h$.

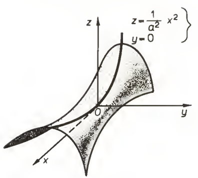


Figure 117

Their equations

$$z = h, \quad \frac{x^2}{a^2} - \frac{y^2}{b^2} = h$$

show that the section by the plane $z = h$ for $h \neq 0$ is a hyperbola. If $h > 0$, then the transverse axis of the parabola is parallel to the x-axis (the straight line $z = h$, $y = 0$), while for $h < 0$ the transverse axis is parallel to the y-axis (the straight line $z = h$, $x = 0$). When $h = 0$, the section is defined by

$$z = 0, \quad \frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$$

i.e. it degenerates into a pair of straight lines $\frac{x}{a} - \frac{y}{b} = 0$ and $\frac{x}{a} + \frac{y}{b} = 0$ lying in the xy-plane.

Part Two

LINEAR ALGEBRA

The first part of this book dealt with the elements of analytic geometry in the plane and in space on the basis of the ideas of *coordinates* and *vectors*. Let us remind that we began manipulating vectors after introducing two operations, namely, vector addition and the multiplication of a vector by a scalar, which follow the laws of arithmetic.

Mathematics and its applications often involve objects which cannot be presented geometrically but admit operations with them like those indicated above for vectors. Such objects are also called *vectors*, but now the term has a wider sense. As in the case of ordinary vectors, these objects can be defined by sets of numbers—coordinates—with the difference that there may be any number of coordinates (not necessarily two or three). The study of such objects is called the *theory of vector spaces* or *linear algebra*.

We begin the second part of the book with a chapter on systems of linear algebraic equations. The chapter is optional, but is necessary for presentation and is practically important.

Chapter 9

SYSTEMS OF LINEAR EQUATIONS

This chapter considers systems of first-degree algebraic equations. Special cases of such systems were discussed in Chapter 3 when studying systems of n linear equations in n unknowns. In what follows the number of equations and the number of unknowns are not necessarily equal; both numbers are arbitrary.

Algebraic first-degree equations are called *linear* since a first-degree equation in two unknowns defines a *straight line* in the plane.

9.1. ELEMENTARY TRANSFORMATIONS OF A SYSTEM OF LINEAR EQUATIONS

Let us revise the notation introduced in Chapter 3. We denote unknowns by $x_1, x_2, \ldots$, the coefficients of the unknowns in the ith equation of the system by $a_{i1}, a_{i2}, \ldots$, and the constant term in the ith equation by b_i. In this notation the general form of the system of m linear equations in n unknowns (an $m \times n$ system) is

$$\begin{array}{ll} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 & \text{(1st equation)} \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 & \text{(2nd equation)} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m & \text{(}m\text{th equation)} \end{array} \qquad (1)$$

Systems can contain equations of the form

$$0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = b$$

Obviously, for $b = 0$ this equation is satisfied by any set of unknowns, while for $b \neq 0$ no set of unknowns satisfies it.

The *solution* of system (1) is a set of n numbers α_1, α_2, ..., α_n which when substituted for the unknowns in each equation of the system (α_1 for x_1, α_2 for x_2, and so on) transform all the equations into valid numerical equalities.

A system is said to be *consistent* if it has at least one solution and *inconsistent* if it has no solution.

We later show that three cases are possible:

(1) the system is inconsistent,

(2) the system has a single solution, and

(3) the system has infinitely many solutions.

It is impossible to have a finite number of solutions greater than one.

To analyze a system of equations means to establish whether it is consistent. If the system is consistent and has a single solution, we should determine the solution; when the number of solutions is infinite, we should find the set of solutions using the most efficient method.

A method for solving systems of linear equations, which is both sufficient and suitable for drawing theoretical conclusions, by a systematic process of elimination of unknowns is called *Gaussian elimination* or the *Gauss algorithm*.

When solving a system of linear equations, we shall always start with a system written in the form in (1), namely, the order of terms in the equations of the initial system is preset: the first term is that with the unknown x_1, the second is with x_2, and so on.

This order can be violated in the process of transforming the system, since the following operations are possible:

(a) *interchanging two terms in all the equations.* For instance, interchanging the terms with x_2 and x_4 in the system

$$\begin{cases} 5x_1 - 7x_2 + 4x_3 - x_4 = 2 \\ 3x_1 + 6x_2 - x_3 + x_4 = 3 \end{cases}$$

yields

$$\begin{cases} 5x_1 - x_4 + 4x_3 - 7x_2 = 2 \\ 3x_1 + x_4 - x_3 + 6x_2 = 3 \end{cases}$$

Thus, a transformation of type in (a) is an *identity transformation*,

(b) *eliminating an equation of the form*

$$0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = 0 \tag{2}$$

from the system,

(c) *adding a constant multiple of one equation to another equation.*

Definition. The operations (a), (b) and (c) are called *elementary operations.*

The following proposition is true.

A system obtained from an initial system by finitely many elementary operations is said to be equivalent to it.

This is obvious for (a) and (b), while (c) requires additional arguments.

Suppose we have the system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \tag{3}$$

We transform the system as follows: add both sides of the second equation multiplied by a number c to the first equation. The result is the new system

$$\begin{cases} (a_{11} + ca_{21})x_1 + (a_{12} + ca_{22})x_2 = b_1 + cb_2 \\ \qquad\qquad\qquad a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \tag{4}$$

Every solution of the initial system (3) satisfies the new system (4). On the other hand, we can obtain (3) from (4) using an operation similar to (c). To do this we add the second equation in (4) multiplied by $-c$ to the first equation. Hence it follows that every solution of the new system satisfies the initial system. Thus, the systems are equivalent.

9.2. GAUSSIAN ELIMINATION

1. Reducing a system to the echelon form.

The essence of Gaussian elimination is to use elementary operations to reduce a system of equations to a form showing all its solutions. Let us assume that *if the algorithm involves* *the equation*

$$0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = 0 \qquad (1)$$

then the equation is discarded, that is, a type (b) elementary operation is applied.

Let us consider the system of linear equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases} \qquad (2)$$

Two cases are possible here.

Case 1. The system involves an equation in which all the coefficients of the unknowns are zero, while the constant term is nonzero:

$$0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = b \qquad (3)$$

where b is nonzero. No set of unknowns can satisfy (3) and therefore a system containing such an equation is inconsistent.

Case 2. The system has no equations like (3) with $b \neq 0$. In other words, each equation has at least one nonzero coefficient. Let us dwell on this case. *Suppose that* $a_{11} \neq 0$. Then we can eliminate x_1 from all the equations beginning from the second one by adding the first equation multiplied by $-a_{21}/a_{11}$ to the second equation, the first equation multiplied by $-a_{31}/a_{11}$ to the third equation and so on. The resulting system is

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ \boxed{\begin{array}{l} a'_{22}x_2 + \ldots + a'_{2n}x_n = b'_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a'_{m2}x_2 + \ldots + a'_{mn}x_n = b'_m \end{array}} \end{cases}$$

The first step is completed. The boxed equations of the system, which remain to be transformed, are the *residual part*.

Let us apply step 1 to the residual part. As a result we either find an equation of type (3) (where $b \neq 0$) thereby establishing that the system is inconsistent, or *assuming* $a'_{22} \neq 0$ perform the next step, i.e. arrive at the system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n = b_1 \\ \qquad a'_{22}x_2 + a'_{23}x_3 + \ldots + a'_{2n}x_n = b'_2 \\ \qquad\qquad \boxed{\begin{array}{l} a''_{33}x_3 + \ldots + a''_{3n}x_n = b''_3 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a''_{m3}x_3 + \ldots + a''_{mn}x_n = b''_m \end{array}} \end{cases}$$

The second step is completed. The residual part contains one equation less.

Continuing the process we arrive at one of two cases. We either obtain a system whose residual part contains an equation of the form $0 \cdot x_1 + \ldots + 0 \cdot x_n = b$, with $b \neq 0$, the system thus having no solution (being inconsistent) and there is no reason to transform it any further. Or, since the number of steps cannot exceed n (the number of unknowns), we finally obtain a system having no residual part, that is, to a system of the form

$$\begin{cases} b_{11}x_1 + b_{12}x_2 + \ldots + b_{1r}x_r + \ldots + b_{1n}x_n = c_1 \\ \qquad b_{22}x_2 + \ldots + b_{2r}x_r + \ldots + b_{2n}x_n = c_2 \\ \qquad\qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad\qquad\qquad b_{rr}x_r + \ldots + b_{rn}x_n = c_r \end{cases} \tag{4}$$

where the coefficients $b_{11}, b_{22}, \ldots, b_{rr}$ are nonzero. The number of equations may be different to system (2) (r equations instead of m initial ones) because equations of form (1) are discarded.

We shall call the system in (4) to be in the *echelon* form due to the appearance of the system. Just look at system (4), the broken line beneath the first terms is like a staircase.

Now, let us see what we should do if the assumptions $a_{11} \neq 0$, $a'_{22} \neq 0$, and so on, required when obtaining (4) are not fulfilled.

At each step the coefficient of the first unknown in the first equation is called the *pivotal* coefficient. In system (2) the pivotal coefficient is a_{11}. If $a_{11} = 0$, then one of the coefficients in the first equation is nonzero, say a_{12}. Then we interchange terms with x_1 and x_2 in all the equations (a type (a) elementary operation) to make the pivotal coefficient nonzero. Then we eliminate x_2, and not x_1, from all the equations of the system. Similarly, we may interchange terms at each step of the algorithm to ensure that the pivotal coefficient is nonzero. The resulting system is not necessarily of form (4); it may differ in the subscripts of the unknowns, but it will still be in the echelon form.

Let us summarize.

Theorem. *If the elementary operations do not result in the equation of the form*

$$0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = b$$

with b nonzero, then the system reduces either to form (4) *or to a form differing from* (4) *in the subscripts of the unknowns.*

Example. Reduce (if possible) the following system of four equations in five unknowns to the echelon form:

$$\left\{\begin{array}{l} -x_1 + 3x_2 + 3x_3 + 2x_4 + 5x_5 = 2 \\ -3x_1 + 5x_2 + 2x_3 + 3x_4 + 4x_5 = 2 \\ -3x_1 + x_2 - 5x_3 \qquad\quad - 7x_5 = -2 \\ -5x_1 + 7x_2 + x_3 + 16x_4 + x_5 = 10 \end{array}\right. \qquad (5)$$

○ The pivotal coefficient is nonzero (it equals -1), thus, we can eliminate x_1 from all the equations of the system starting with the second, namely, we subtract the first equation multiplied by -3, -3, -5 from the second, third, and fourth equations, respectively. The resulting system is

$$\left\{\begin{array}{l} -x_1 + 3x_2 + 3x_3 + 2x_4 + 5x_5 = 2 \\ \boxed{\begin{array}{l} -4x_2 - 7x_3 - 3x_4 - 11x_5 = -4 \\ -8x_2 - 14x_3 - 6x_4 - 22x_5 = -8 \\ -8x_2 - 14x_3 + 6x_4 - 24x_5 = 0 \end{array}} \end{array}\right.$$

The pivotal coefficient in the boxed part of the system is nonzero (it is -4) and we can eliminate x_2 from the second and third equations. We have

$$\begin{cases} -x_1 + 3x_2 + 3x_3 + 2x_4 + 5x_5 = 2 \\ - 4x_2 - 7x_3 - 3x_4 - 11x_5 = -4 \\ \boxed{0\cdot x_3 + 12x_4 - 2x_5 = 8} \end{cases}$$

(the number of equations is one less than we started with since we discarded the equation $0 = 0$). The pivotal coefficient in the residual part is zero, while other coefficients, say of x_5, are nonzero. Interchanging the terms with x_3 and x_5 yields a system in the echelon form

$$\begin{cases} -x_1 + 3x_2 + 5x_5 + 2x_4 + 3x_3 = 2 \\ - 4x_2 - 11x_5 - 3x_4 - 7x_3 = -4 \\ - 2x_5 + 12x_4 = 8 \end{cases} \tag{6}$$

which is equivalent to the initial system (5). ●

2. Solving system in the echelon form.

We now consider the staired system (4). Two cases are possible here: $r = n$ and $r < n$.

Case 1. $r = n$. Then the system has the form

$$\begin{cases} b_{11}x_1 + b_{12}x_2 + \ldots + b_{1n}x_n = c_1 \\ \phantom{b_{11}x_1 +} b_{22}x_2 + \ldots + b_{2n}x_n = c_2 \\ \qquad \cdots\cdots\cdots\cdots\cdots \\ \phantom{b_{11}x_1 + b_{22}x_2 + \ldots +} b_{nn}x_n = c_n \end{cases} \tag{7}$$

and is called *triangular*. Noting that $b_{nn} \neq 0$ we find from the last equation in (7) the unique value of x_n, then substituting it into the preceding equation and noting that $b_{n-1,n-1} \neq 0$ we find a unique value for x_{n-1}, then substituting the values of x_n and x_{n-1} we found into the preceding equation we find a unique value of x_{n-2}, and so on.

Thus, *a system in the echelon form in which the number r of equations equals the number n of unknowns has a unique solution.*

Case 2. $r < n$. We transfer the terms with $x_{r+1}, x_{r+2}, \ldots$

to the right-hand side of all the equations in (4) to obtain

$$\begin{cases} b_{11}x_1 + b_{12}x_2 + \ldots + b_{1r}x_r = c_1 - b_{1,r+1}x_{r+1} - \ldots - b_{1n}x_n \\ \qquad b_{22}x_2 + \ldots + b_{2r}x_r = c_2 - b_{2,r+1}x_{r+1} - \ldots - b_{2n}x_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad\qquad b_{rr}x_r = c_r - b_{r,r+1}x_{r+1} - \ldots - b_{rn}x_n \end{cases} \quad (8)$$

Substituting arbitrary numbers for $x_{r+1}, x_{r+2}, \ldots, x_n$ into (8) yields a system of triangular form for finding the remaining unknowns $x_1, x_2, \ldots, x_r$. Solving this new system gives a unique set of values of $x_1, x_2, \ldots, x_r$. Hence we infer that for $r < n$, system (4) has infinitely many solutions as described by (8) (given arbitrary $x_{r+1}, x_{r+2}, \ldots, x_n$, we can find $x_1, x_2, \ldots, x_r$ from (8).

Unknowns whose values can be arbitrarily chosen are said to be *free*; $x_{r+1}, x_{r+2}, \ldots, x_n$ are free unknowns for system (4).

Thus, a *system in the echelon form of r equations in n unknowns with r smaller than n has infinitely many solutions; here n − r unknowns are free.*

Remark. If when reducing the initial system to the echelon form we interchange the unknowns (type (b) elementary operations), then any $n - r$ unknowns may be free in the resulting system and not necessarily $x_{r+1}, x_{r+2}, \ldots, x_n$.

3. Analyzing a system of linear equations.

Up till now our task was to describe all the solutions of a system of linear equations. In some cases, however, it is sufficient to know *how many* solutions the system has (no solution, or a single solution, or infinitely many solutions) and we say that we analyze the system. If the system has infinitely many solutions, we should also find the free unknowns (the values of the remaining unknowns are uniquely determined once we know the values of free unknowns).

Gaussian elimination is sufficient to analyze a system of linear equations. As a result we either ascertain that the system has no solution (if an equation of the form $0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = b$, where $b \neq 0$, appear) or reduce the system to the echelon form. Suppose we have the second case and r and

n are the number of equations and the number of unknowns in the system in the echelon form, respectively. If $r = n$, then the initial system has a single solution. If $r < n$, then the system has infinitely many solutions, and we can consider the unknowns undefined by the first terms of the equations in the system having the echelon form as free unknowns.

Example 1. Analyze and solve system (5).

○ The system has been already analyzed (reduced to form (6)). It has infinitely many solutions; x_4 and x_3 are free unknowns.

To solve system (5) means to find all its solutions explicitly. We transfer terms with free unknowns x_4 and x_3 to the right-hand side of each equation in (6), i.e. we rewrite the system as

$$\begin{cases} -x_1 + 3x_2 + \;\;5x_5 = \;\;\;2 - \;\;2x_4 - 3x_3 \\ \quad\;\; - 4x_2 - 11x_5 = -4 + \;\;3x_4 + 7x_3 \\ \qquad\qquad\; - \;\;2x_5 = \;\;\;8 - 12x_4 \end{cases} \tag{9}$$

Given x_4 and x_3, we can find x_5, x_2, and x_1.

Thus, the problem is solved. However, we can continue the solution and express x_5, x_2, x_1 in terms of x_4 and x_3. We have from the last equation in (9) $x_5 = -4 + 6x_4$. Substituting this value into the second equation in (9) yields

$$-4x_2 - 11(-4 + 6x_4) = -4 + 3x_4 + 7x_3$$

whence $x_2 = 12 - \dfrac{69}{4} x_4 - \dfrac{7}{4} x_3$.

Finally, substituting the expressions for x_5 and x_2 into the first equation yields

$$-x_1 + 3\left(12 - \frac{69}{4} x_4 - \frac{7}{4} x_5\right) + 5(-4 + 6x_4)$$
$$= 2 - 2x_4 - 3x_3$$

whence $x_1 = 14 - \dfrac{79}{4} x_4 - \dfrac{9}{4} x_3$.

Thus, explicit expressions for x_5, x_2, and x_1 in terms of x_4 and x_3 are $x_5 = -4 + 6x_4$, $x_2 = 12 - \dfrac{69}{4} x_4 - \dfrac{7}{4} x_3$, and

$x_1 = 14 - \frac{79}{4}x_4 - \frac{9}{4}x_3$. The system has infinitely many solutions. For example, putting $x_4 = 0$ and $x_3 = 0$ yields a particular solution $x_1 = 14$, $x_2 = 12$, $x_3 = 0$, $x_4 = 0$, and $x_5 = -4$. ●

Example 2. Solve the system

$$\begin{cases} 2x_1 - x_2 + x_3 - x_4 = 2 \\ 2x_1 - x_2 - 3x_4 = 1 \\ 3x_1 - x_3 + x_4 = 8 \\ 2x_1 + 2x_2 - 2x_3 + 5x_4 = 11 \end{cases} \qquad (10)$$

○ The pivotal coefficient is nonzero, nevertheless it is convenient to begin by interchanging x_1 and x_2 (in each equation). The result is

$$\begin{cases} -x_2 + 2x_1 + x_3 - x_4 = 2 \\ -x_2 + 2x_1 - 3x_4 = 1 \\ 3x_1 - x_3 + x_4 = 8 \\ 2x_2 + 2x_1 - 2x_3 + 5x_4 = 11 \end{cases}$$

We have further

$$\begin{cases} -x_2 + 2x_1 + x_3 - x_4 = 2 \\ \boxed{\begin{array}{r} 0\cdot x_1 - x_3 - 2x_4 = -1 \\ 3x_1 - x_3 + x_4 = 8 \\ 6x_1 - 0\cdot x_3 + 3x_4 = 15 \end{array}} \end{cases} \qquad \begin{cases} -x_2 + x_3 + 2x_1 - x_4 = 2 \\ \boxed{\begin{array}{r} -x_3 + 0\cdot x_1 - 2x_4 = -1 \\ -x_3 + 3x_1 + x_4 = 8 \\ 0\cdot x_3 + 6x_1 + 3x_4 = 15 \end{array}} \end{cases}$$

$$\begin{cases} -x_2 - x_3 + 2x_1 - x_4 = 2 \\ -x_3 + 0\cdot x_1 - 2x_4 = -1 \\ \boxed{\begin{array}{r} 3x_1 + 3x_4 = 9 \\ 6x_1 + 3x_4 = 15 \end{array}} \end{cases} \qquad \begin{cases} -x_2 + x_3 + 2x_1 - x_4 = 2 \\ -x_3 + 0\cdot x_1 - 2x_4 = -1 \\ 3x_1 + 3x_4 = 9 \\ -3x_4 = -3 \end{cases}$$

The last system has a triangular form, by proceeding from the last equation to the first, we obtain a single solution: $x_4 = 1$, $x_1 = 2$, $x_3 = 1$, $x_2 = 0$. ●

4. Practical remark.

The process of elimination can be simplified by considering the matrices of the equations of a system. We take system (10) as an example. We write the coefficients of the unknowns and the constant terms in matrix form. The result is the *augmented matrix of system* (10):

$$\begin{pmatrix} 2 & -1 & 1 & -1 & 2 \\ 2 & -1 & 0 & -3 & 1 \\ 3 & 0 & -1 & 1 & 8 \\ 2 & 2 & -2 & 5 & 11 \end{pmatrix}$$

We assume the following notation: we write the corresponding unknown under the column of its coefficients. Thus, the matrix of system (10) has the form

$$\begin{pmatrix} 2 & -1 & 1 & -1 & 2 \\ 2 & -1 & 0 & -3 & 1 \\ 3 & 0 & -1 & 1 & 8 \\ 2 & 2 & -2 & 5 & 11 \end{pmatrix}$$
$$\begin{matrix} x_1 & x_2 & x_3 & x_4 \end{matrix}$$

Elementary operations for matrices correspond to elementary operations for equations, namely,

(a) an interchange of two columns corresponds to interchange of two unknowns,

(b) discarding a row consisting of zeros corresponds to discarding the equation $0 \cdot x_1 + \ldots + 0 \cdot x_n = 0$,

(c) addition of one row multiplied by a constant k to another row corresponds to the addition of one equation of the system multiplied by a constant k to another equation.

Now we write the solution of system (10):

$$\underset{\begin{matrix} x_1 & x_2 & x_3 & x_4 \end{matrix}}{\begin{pmatrix} 2 & -1 & 1 & -1 & 2 \\ 2 & -1 & 0 & -3 & 1 \\ 3 & 0 & -2 & 5 & 8 \\ 2 & 2 & -2 & 5 & 11 \end{pmatrix}} \sim \underset{\begin{matrix} x_2 & x_1 & x_3 & x_4 \end{matrix}}{\begin{pmatrix} -1 & 2 & 1 & -1 & 2 \\ -1 & 2 & 0 & -3 & 1 \\ 0 & 3 & -2 & 5 & 8 \\ 2 & 2 & -2 & 5 & 11 \end{pmatrix}}$$

$$\sim \begin{pmatrix} -1 & 2 & 1 & -1 & 2 \\ 0 & 0 & -1 & -2 & -1 \\ 0 & 3 & -1 & 1 & 8 \\ 0 & 6 & 0 & 3 & 15 \end{pmatrix} \sim \begin{pmatrix} -1 & 1 & 2 & -1 & 2 \\ 0 & -1 & 0 & -2 & -1 \\ 0 & -1 & 3 & 1 & 8 \\ 0 & 0 & 6 & 3 & 15 \end{pmatrix}$$

$$\begin{matrix} x_2 & x_1 & x_3 & x_4 \end{matrix} \qquad \begin{matrix} x_2 & x_3 & x_1 & x_4 \end{matrix}$$

$$\sim \begin{pmatrix} -1 & 1 & 2 & -1 & 2 \\ 0 & -1 & 0 & -2 & -1 \\ 0 & 0 & 3 & 3 & 9 \\ 0 & 0 & 6 & 3 & 15 \end{pmatrix} \sim \begin{pmatrix} -1 & 1 & 2 & -1 & 2 \\ 0 & -1 & 0 & -2 & -1 \\ 0 & 0 & 3 & 3 & 9 \\ 0 & 0 & 0 & -3 & -3 \end{pmatrix}$$

$$\begin{matrix} x_2 & x_3 & x_1 & x_4 \end{matrix} \qquad \begin{matrix} x_2 & x_3 & x_1 & x_4 \end{matrix}$$

where the $\sim$ sign indicates that an elementary operation is performed on the matrix.

The last matrix corresponds to the triangular system

$$\begin{cases} -x_2 + x_3 + 2x_1 - x_4 = 2 \\ \quad - x_3 \qquad - 2x_4 = -1 \\ \qquad\qquad 3x_1 + 3x_4 = 9 \\ \qquad\qquad\quad - 3x_4 = -3 \end{cases}$$

from which we find the unique solution: $x_4 = 1, x_1 = 2, x_3 = 1, x_2 = 0$.

Example. Solve the system

$$\begin{cases} 2x_1 - x_2 + 3x_3 - 2x_4 + 4x_5 = -1 \\ 4x_1 - 2x_2 + 5x_3 + x_4 + 7x_5 = 2 \\ 2x_1 - x_2 + x_3 + 8x_4 + 2x_5 = 1 \end{cases}$$

○ We have

$$\begin{pmatrix} 2 & -1 & 3 & -2 & 4 & -1 \\ 4 & -2 & 5 & 1 & 7 & 2 \\ 2 & -1 & 1 & 8 & 2 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & -1 & 3 & -2 & 4 & -1 \\ 0 & 0 & -1 & 5 & -1 & 4 \\ 0 & 0 & -2 & 10 & -2 & 2 \end{pmatrix}$$

$$\begin{matrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{matrix} \qquad \begin{matrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{matrix}$$

$$\sim \begin{pmatrix} 2 & 3 & -1 & -2 & 4 & -1 \\ 0 & -1 & 0 & 5 & -1 & 4 \\ 0 & -2 & 0 & 10 & -2 & 2 \end{pmatrix}$$
$$\begin{matrix} x_1 & x_3 & x_2 & x_4 & x_5 \end{matrix}$$

$$\sim \begin{pmatrix} 2 & 3 & -1 & -2 & 4 & -1 \\ 0 & -1 & 0 & 5 & -1 & 4 \\ 0 & 0 & 0 & 0 & 0 & -6 \end{pmatrix}$$
$$\begin{matrix} x_1 & x_3 & x_2 & x_4 & x_5 \end{matrix}$$

The last row corresponds to the equation

$$0 \cdot x_1 + 0 \cdot x_3 + 0 \cdot x_2 + 0 \cdot x_4 + 0 \cdot x_5 = -6$$

Thus, the system has no solution (is inconsistent). ●

5. Consequence of Gaussian elimination for homogeneous systems.

Remember that the system is said to be *homogeneous* if all its constant terms are zero, namely,

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0 \end{cases}$$

A homogeneous system is always consistent: one of its solutions is of the form $x_1 = 0$, $x_2 = 0$, $\ldots$, $x_n = 0$ and is called the *trivial solution*. When considering a homogeneous system, it is important to establish whether it has at least one *nontrivial* (nonvanishing) solution.

Theorem. *A homogeneous system in which the number of equations is smaller than the number of unknowns* $(m < n)$ *always has a nontrivial solution.*

□ When solving the system by Gaussian elimination, we can never obtain an equation of the form $0 \cdot x_1 + \ldots + 0 \cdot x_n = b$ with $b \neq 0$, since all the constant terms in the system are zero. Thus, the process ends with a system in the echelon form in which the number of equations is still smaller than the number of unknowns. But such systems have infinitely many solutions, nontrivial ones included.

Exercises to Chapter 9

1. What is a solution of a system of linear equations in n unknowns $x_1, x_2, \dots, x_n$.

2. Find all the solutions of the equation $2x_1 = 3$ if it is in (a) one unknown x_1 and (2) n unknowns $x_1, x_2, \dots, x_n$.

3. Solve the following systems of equations by Gaussian elimination. Find whether the system is consistent. If the system is consistent, reduce it to the echelon form. If there are free unknowns, use them to express the remaining unknowns.

$$(1)\ \begin{cases} x_1 + 2x_2 + x_3 = 4 \\ 3x_1 - 5x_2 + 3x_3 = 1 \\ 2x_1 + 7x_2 - x_3 = 8 \end{cases} \qquad (2)\ \begin{cases} x + 2y + 2z = 1 \\ 2x + y - 2z = 2 \\ 2x + 2y + z = 3 \\ 5x + 4y + 7z = 6 \end{cases}$$

$$(3)\ \begin{cases} 2x_1 - y + z = -2 \\ x + 2y + 3z = -1 \\ x - 3y - 2z = 3 \end{cases} \qquad (4)\ \begin{cases} x_1 - 2x_2 + 3x_3 - 4x_4 = 0 \\ 2x_1 + x_3 - x_4 = 0 \\ -3x_1 + x_2 + x_3 - 2x_4 = 0 \end{cases}$$

$$(5)\ \begin{cases} x_1 - 2x_2 + x_3 + x_4 = 1 \\ x_1 - 2x_2 + x_3 - x_4 = -1 \\ x_1 - 2x_2 + x_3 + 5x_4 = 5 \end{cases} \qquad (6)\ \begin{cases} x + y + z + t = 1 \\ x + y - z - t = 1 \\ x - y + z - t = 1 \\ x - y - z + t = 1 \end{cases}$$

$$(7)\ \begin{cases} x_1 + x_2 + 2x_3 + 3x_4 = 1 \\ 3x_1 - x_2 - x_3 - 2x_4 = -4 \\ 2x_1 + 3x_2 - x_3 - x_4 = -6 \\ x_1 + 2x_2 + 3x_3 - x_4 = -4 \end{cases} \qquad (8)\ \begin{cases} 2x_1 - x_2 + x_3 - x_4 = 0 \\ 2x_1 - x_2 - 3x_4 = 0 \\ 3x_1 - x_3 + x_4 = 0 \\ 2x_1 + 2x_2 - 2x_3 + 5x_4 = 0 \end{cases}$$

$$(9)\ \begin{cases} x + y - 3z = -1 \\ 2x + y - 2z = 1 \\ x + y + z = 3 \\ x + 2y - 3z = 1 \end{cases} \qquad (10)\ \begin{cases} 2x_1 - x_2 + 3x_3 - 2x_4 + 4x_5 = -1 \\ 4x_1 - 2x_2 + 5x_3 + x_4 + 7x_5 = 2 \\ 2x_1 - x_2 + x_3 + 8x_4 + 2x_5 = 1 \end{cases}$$

4. Find a for which the system is consistent and solve it for those a:

(1) $\begin{cases} x + y + z = 1 \\ x + ay + z = 1 \\ x + y + az = 1 \end{cases}$ (2) $\begin{cases} 2x_1 + x_2 + x_3 + x_4 = 1 \\ x_1 + 2x_2 - x_3 + 4x_4 = 2 \\ x_1 + 7x_2 - 4x_3 + 11x_4 = a \end{cases}$

(3) $\begin{cases} 5x_1 - x_2 + 2x_3 + x_4 = 2a \\ 2x_1 + x_2 + 4x_3 - 2x_4 = 1 \\ x_1 - 3x_2 - 6x_3 + 5x_4 = 0 \end{cases}$

5. Find the equation of a sphere in space passing through the points (a) $M_1(1, 0, 0)$, $M_2(1, 1, 0)$, $M_3(1, 1, 1)$, $M_4(0, 1, 1)$, (b) $M_1(2, 2, 1)$, $M_2(2, 1, 2)$, $M_3(1, 2, 2)$, $M_4(0, 0, 3)$, (c) $M_1(3, 1, 1)$, $M_2(3, 0, 2)$, $M_3(2, 1, 2)$, $M_4(1, -1, 3)$.

6. Find the equation of a third-degree parabola passing through the points $M_1(1, 0)$, $M_2(0, -1)$, $M_3(-1, -2)$, $M_4(2, 7)$.

Remark. A third-order parabola is defined by the equation $y = ax^3 + bx^2 + cx + d$.

Chapter 10

VECTOR SPACES

10.1. ARITHMETIC VECTORS AND OPERATIONS WITH THEM

1. Arithmetic *n*-dimensional vector space.

Definition. Any finite set of numbers

$$a_1, a_2, \ldots, a_n \tag{1}$$

is called an *arithmetic vector*, and the numbers are the *coordinates of this vector.*

We use the notation $\langle a_1, a_2, \ldots, a_n \rangle$ for an arithmetic vector.

For instance, $\langle -1, 2, 0, 7 \rangle$ is an arithmetic vector whose coordinates are $-1, 2, 0, 7$.

We can interpret geometrically only arithmetic vectors having one, two, or three coordinates, namely, each set of the form $\langle a_1 \rangle$, $\langle a_1, a_2 \rangle$, or $\langle a_1, a_2, a_3 \rangle$ is associated with a "true" vector, i.e. a directed line segment on a straight line, in the plane, or in space (in a Cartesian coordinate system). An arithmetic vector having more than three coordinates cannot be interpreted geometrically, since by definition such a vector is a purely arithmetic object (the set in (1)).

Arithmetic vectors are often encountered in mathematics with different numbers of coordinates. Here are several examples:

(a) an ordinary (geometric) vector on a straight line can be considered as an arithmetic vector having one, two, or three coordinates,

(b) each row in an nth-order determinant is an arithmetic vector with n coordinates,

(c) the coefficients of the unknowns in the linear equation

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$$

form an arithmetic vector $\langle a_1, a_2, \ldots, a_n \rangle$ having n coordinates, any solution $\langle x_1^0, x_2^0, \ldots, x_n^0 \rangle$ of this equation is an arithmetic vector with n coordinates.

Definition. Two arithmetic vectors $\langle a_1, a_2, \ldots, a_n \rangle$ and $\langle a_1', a_2', \ldots, a_n' \rangle$ are said to be *equal* if and only if they have the same number of coordinates ($n = m$) and if the respective coordinates are equal, $a_1 = a_1'$, $a_2 = a_2'$, $\ldots$, $a_n = a_n'$.

In what follows we shall omit the word "arithmetic" and simply say "vector", and use lower-case letters in bold face to denote vectors. If two vectors **a** and **b** are equal, then we write $\mathbf{a} = \mathbf{b}$.

Definition. The *sum* of two vectors having the same number of coordinates $\mathbf{a} = \langle a_1, a_2, \ldots, a_n \rangle$ and $\mathbf{b} = \langle b_1, b_2, \ldots, b_n \rangle$ is the vector $\mathbf{a} + \mathbf{b}$ defined by the equation

$$\mathbf{a} + \mathbf{b} = \langle a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n \rangle$$

For instance, $\langle -1, 0, 7, 5 \rangle + \langle 4, 1, -7, 3 \rangle = \langle 3, 1, 0, 8 \rangle$.

Thus, to add vectors means to add their respective coordinates. This definition is in line with that given earlier for the addition of vectors in a three-dimensional space (the coordinates of the sum of two vectors are equal to the sums of the corresponding coordinates of the vectors).

Definition. The *product of the vector* $\mathbf{a} = \langle a_1, a_2, \ldots, a_n \rangle$ *by a scalar* k is the vector $k\mathbf{a}$ defined by the equation

$$k\mathbf{a} = \langle ka_1, ka_2, \ldots, ka_n \rangle$$

Vector addition is commutative and associative:

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

We can also easily verify the following equations:

$$(k + l)\mathbf{a} = k\mathbf{a} + l\mathbf{a}$$
$$k(\mathbf{a} + \mathbf{b}) = k\mathbf{a} + k\mathbf{b}$$
$$k(l\mathbf{a}) = (kl)\mathbf{a}$$

Definition. A vector whose all coordinates are zero is called the *zero vector* and is denoted by $\mathbf{0}$.

Strictly speaking, such a notation is inaccurate since $\mathbf{0}$ means any of the vectors $\langle 0 \rangle$, $\langle 0, 0 \rangle$, $\langle 0, 0, 0 \rangle$, ... However, in what follows it is clear how many coordinates the vector $\mathbf{0}$ has in each specific case.

Obviously, $\mathbf{a} + \mathbf{0} = \mathbf{a}$ whatever the vector $\mathbf{a}$ having the same number of coordinates as the vector $\mathbf{0}$.

The vector $(-1)\mathbf{a}$ is the negative (or the *opposite* vector) of $\mathbf{a}$ and denoted by $-\mathbf{a}$. If $\mathbf{a} = \langle a_1, a_2, \ldots, a_n \rangle$, then $-\mathbf{a} = \langle -a_1, -a_2, \ldots, -a_n \rangle$. Hence

$$\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$$

Note that *if* $k\mathbf{a} = \mathbf{0}$, *then either* $k = 0$ *or* $\mathbf{a} = \mathbf{0}$.

□ Given $\mathbf{a} = \langle a_1, a_2, \ldots, a_n \rangle$, the equation $k\mathbf{a} = \mathbf{0}$ means that

$$ka_1 = 0, \ ka_2 = 0, \ \ldots, \ ka_n = 0 \qquad (2)$$

If $k = 0$, then there is nothing to be proved. If $k \neq 0$, then we have $a_1 = 0$, $a_2 = 0$, ..., $a_n = 0$, whence $\mathbf{a} = \mathbf{0}$. ■

Definition. The set of all arithmetic vectors having the given number n of coordinates in which vector addition and multiplication by a scalar are defined as above is called an *arithmetic n-dimensional vector space* and denoted by $\mathbf{R}^n$.

We shall omit the word "arithmetic" and simply say an *n-dimensional vector space* $\mathbf{R}^n$ or, in short, an $\mathbf{R}^n$ *space*.

We can interpret geometrically only the spaces $\mathbf{R}^1$, $\mathbf{R}^2$, and $\mathbf{R}^3$; $\mathbf{R}^1$ can be associated with the set of vectors on a straight line, $\mathbf{R}^2$ with the set of vectors in the plane, and $\mathbf{R}^3$ with the set of vectors in space.

2. The set of all solutions of a homogeneous system of linear equations as a subspace of $\mathbf{R}^n$.

The theory of arithmetic vectors allows one to interpret a number of facts concerning systems of linear equations. We take a homogeneous system of linear equations as an example.

In the notation adopted in Chapter 9, the general form of a homogeneous system of m linear equations in n unknowns is

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0 \end{cases} \qquad (3)$$

Each solution of (3) is a set of n numbers $x_1, x_2, \ldots, x_n$ and can therefore be considered as a vector $\mathbf{x}$ from $\mathbf{R}^n$: $\mathbf{x} = \langle x_1, x_2, \ldots, x_n \rangle$. In this sense we can speak about the sum of two solutions and the product of a solution and a scalar.

The set of all solutions of a homogeneous system possesses two specific properties.

1°. *The sum of two solutions is a solution.*

2°. *The product of a solution and a scalar is a solution.*

Let us prove the first property. Suppose $\boldsymbol{\alpha} = \langle \alpha_1, \alpha_2, \ldots, \alpha_n \rangle$ and $\boldsymbol{\beta} = \langle \beta_1, \beta_2, \ldots, \beta_n \rangle$ are two solutions of system (3). We have

$$\boldsymbol{\alpha} + \boldsymbol{\beta} = \langle \alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n \rangle$$

We verify whether $\boldsymbol{\alpha} + \boldsymbol{\beta}$ is a solution of (3). We should substitute the coordinates of the vector $\boldsymbol{\alpha} + \boldsymbol{\beta}$ for the unknowns into the left-hand side of each equation of the system to ascertain that the result is zero. Take, for example, the first equation in (3). We have

$$\begin{aligned} \alpha_{11}(\alpha_1 + \beta_1) + a_{12}(\alpha_2 + \beta_2) + \ldots + a_{1n}(\alpha_n + \beta_n) \\ = (a_{11}\alpha_1 + a_{12}\alpha_2 + \ldots + a_{1n}\alpha_n) \\ + (a_{11}\beta_1 + a_{12}\beta_2 + \ldots + a_{1n}\beta_n) \end{aligned}$$

the first sum on the right-hand side is zero since $\boldsymbol{\alpha}$ is a solution and the second sum is zero since $\boldsymbol{\beta}$ is a solution.

Let us prove property 2°. Suppose the vector $\boldsymbol{\alpha} = \langle \alpha_1, \alpha_2, \ldots, \alpha_n \rangle$ is a solution of (3). We must prove that the vector

$$k\boldsymbol{\alpha} = \langle k\alpha_1, k\alpha_2, \ldots, k\alpha_n \rangle$$

where k is a scalar, is also a solution. Substituting the coordinates of this vector for the unknowns into, say, the first equation in (3) yields

$$a_{11}\cdot k\alpha_1 + a_{12}\cdot k\alpha_2 + \ldots + a_{1n}\cdot k\alpha_n$$
$$= k(a_{11}\alpha_1 + a_{12}\alpha_2 + \ldots + a_{1n}\alpha_n) = k\cdot 0 = 0$$

Similarly, we can prove this for the remaining equations.

We denote by M the set of all solutions of system (3). This is a set of vectors in $\mathbf{R}^n$. According to what we have proved, the set M possesses the following properties.

1°. *If vectors* $\mathbf{x}$ *and* $\mathbf{y}$ *belong to* M, *then their sum* $\mathbf{x} + \mathbf{y}$ *also belongs to* M.

2°. *If a vector* $\mathbf{x}$ *belongs to* M, *then the vector* $k\mathbf{x}$, *with* k *a scalar, also belongs to* M.

Any set of vectors in $\mathbf{R}^n$ possessing these properties is called a *subspace*. Thus, we can say that the *set of all solutions of the homogeneous system* (3) *is a subspace of vector space* $\mathbf{R}^n$.

10.2. LINEAR DEPENDENCE OF VECTORS

1. Linear combination of several vectors.

Concrete problems deal, as a rule, with a set or *system* of vectors rather than with a single vector (all vectors in a system have the same number of coordinates). In this case the vectors are denoted by the same letter with different subscripts. For instance,

$$\mathbf{a}_1 = \langle 1, 0, 3, -2, -1\rangle, \mathbf{a}_2 = \langle -1, 1, 4, 3, 0\rangle, \quad \mathbf{a}_3 = \langle -5, 3, 5, 3, 7\rangle \tag{1}$$

is a system of three vectors each of which has five coordinates, that is belongs to the space $\mathbf{R}^5$, while

$$\mathbf{b}_1 = \langle 1, 1, 1, 1\rangle, \mathbf{b}_2 = \langle 1, 1, 1, 1\rangle \tag{2}$$

is the system of two vectors belonging to $\mathbf{R}^4$. Do not be confused to see that $\mathbf{b}_1 = \mathbf{b}_2$ since a *system of vectors may have repetitions*.

Suppose

$$\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p \tag{3}$$

is a system of vectors from $\mathbf{R}^n$. We choose arbitrary numbers $k_1, k_2, \ldots, k_p$ and compose the vector

$$\mathbf{a} = k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + \ldots + k_p\mathbf{a}_p \tag{4}$$

Definition. Any vector $\mathbf{a}$ of form (4) is a *linear combination of vectors* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p$ and $k_1, k_2, \ldots, k_p$ are its *coefficients.*

Example. Find the linear combination $2\mathbf{a}_1 - 3\mathbf{a}_2 + \mathbf{a}_3$ of vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ from (1).

○ Adding the vectors $2\mathbf{a}_1 = \langle 2, 0, 6, -4\rangle$, $-3\mathbf{a}_2 = \langle 3, -3, -12, -9\rangle$, $\mathbf{a}_3 = \langle -5, 3, 5, 3\rangle$ yields

$$2\mathbf{a}_1 - 3\mathbf{a}_2 + \mathbf{a}_3 = \langle 0, 0, -1, 10\rangle$$ ●

Instead of saying "the vector $\mathbf{a}$ is a linear combination of vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p$", we can say that "$\mathbf{a}$ is linearly expressed in terms of the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p$" or "$\mathbf{a}$ decomposes into vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p$".

2. Linear dependence and independence of a system of vectors.

One of the basic concepts of vector space theory is that of *linear dependence.*

Definition. The vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p$ are said to be *linearly dependent* or form a *linearly dependent system* if there are numbers $c_1, c_2, \ldots, c_p$ not all zero such that

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \ldots + c_p\mathbf{a}_p = \mathbf{0} \tag{5}$$

Otherwise, i.e. when (5) is only valid for $c_1 = c_2 = \ldots = c_p = 0$, we say that the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p$ are *linearly independent* or the system of these vectors is *linearly independent.*

Let us stress that the concepts of linear dependence and linear independence only refer to systems with finite numbers of vectors.

Theorem. *A system consisting of one vector is linearly dependent if and only if the vector is zero. A system consisting of several vectors is linearly dependent if and only if at least one of the vectors can be represented as a linear combination of the others.*

□ Suppose a system consists of one vector $\mathbf{a}_1$. If the system is linearly dependent, then there is a nonzero number c_1 such that $c_1\mathbf{a}_1 = \mathbf{0}$. Since $c_1 \neq 0$, this equality is equivalent to $\mathbf{a}_1 = \mathbf{0}$. Thus, the first statement is proved. Let us prove the second statement.

Suppose the system

$$\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p, \; p > 1 \tag{6}$$

is linearly dependent. This means, by definition, that there are numbers $c_1, c_2, \ldots, c_p$, not all zero, such that

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \ldots + c_p\mathbf{a}_p = \mathbf{0}$$

Suppose $c_1 \neq 0$. Adding the vector $-c_1\mathbf{a}_1$ to the last equation yields

$$-c_1\mathbf{a}_1 = c_2\mathbf{a}_2 + \ldots + c_p\mathbf{a}_p$$

and multiplying this equation by $-1/c_1$ gives

$$\mathbf{a}_1 = \left(-\frac{c_2}{c_1}\right)\mathbf{a}_2 + \ldots + \left(-\frac{c_p}{c_1}\right)\mathbf{a}_p$$

Thus, we have shown that if system (6) is linearly dependent, one of its vectors can be expressed as a linear combination of the others. Let us prove the converse.

Suppose one of the vectors of (6) can be represented as a linear combination of the others, for the sake of clarity, let it be the vector $\mathbf{a}_1$:

$$\mathbf{a}_1 = k_2\mathbf{a}_2 + \ldots + k_p\mathbf{a}_p$$

where $k_2, \ldots, k_p$ are numbers some of which or all may be zero. We rewrite the equation as

$$(-1)\mathbf{a}_1 + k_2\mathbf{a}_2 + \ldots + k_p\mathbf{a}_p = \mathbf{0}$$

that is, add the vector $-\mathbf{a}_1$. This equation has the form as in (5), where $c_1 = -1$, $c_2 = k_2, \ldots, c_p = k_p$, with at least one of the numbers $c_1, \ldots, c_p$ nonzero (in our case $c_1 \neq 0$). Thus, system (6) is linearly dependent. ■

Remark. The assertion in the theorem "there is a vector which can be represented as a linear combination of the others" does not mean that *every* vector can be expressed in terms of the others. We illustrate this by way of an example. Let $\mathbf{a}$ be a nonzero vector. We consider a system of two vectors $\mathbf{a}_1 = \mathbf{0}$ and $\mathbf{a}_2 = \mathbf{a}$. The system is linearly dependent since, for instance, $1 \cdot \mathbf{a}_1 + 0 \cdot \mathbf{a}_2 = \mathbf{0}$ ($c_1 = 1$ is thus nonzero). Consequently, one of the vectors must be a linear combination of the other. In this case $\mathbf{a}_1$ can be expressed in terms of $\mathbf{a}_2$ since $\mathbf{a}_1 = 0 \cdot \mathbf{a}_2$, while $\mathbf{a}_2$ cannot be expressed in terms of $\mathbf{a}_1$ since $\mathbf{a}_2 \neq \mathbf{0}$.

Generally speaking, it is not so easy to establish whether a system of vectors is linearly dependent (or independent). However, it is easy for a system of two vectors. Indeed, linear dependence of the system $\mathbf{a}_1, \mathbf{a}_2$ means that one of the vectors, say $\mathbf{a}_1$, can be expressed in terms of the other, i.e. $\mathbf{a}_1 = k\mathbf{a}_2$. Two vectors, say $\mathbf{a}_1$ and $\mathbf{a}_2$, which form a linearly dependent system, are *collinear*. In this case the coordinates of $\mathbf{a}_1$ are proportional to the corresponding coordinates of $\mathbf{a}_2$. For example, the vector $\mathbf{a}_1 = \langle -1, 0, 3, 4 \rangle$ is collinear with the vector $\mathbf{a}_2 = \langle -5, 0, 15, 20 \rangle$.

Thus, a *system of two vectors is linearly dependent if and only if the two vectors are collinear.*

3. Geometric interpretation of linear dependence for a system of three vectors in $\mathbf{R}^3$.

We introduce a Cartesian coordinate system in ordinary space. Then every arithmetic vector from $\mathbf{R}^3$, i.e. every set of three numbers $\langle a_1, a_2, a_3 \rangle$ can be associated with an ordinary vector, i.e. a directed line segment $\mathbf{a}$ in the customary space; for clarity we assume that $\mathbf{a}$ originates from a fixed point O, the origin. This geometric vector can be considered to be the image of the arithmetic vector. We denote both of them by $\mathbf{a}$.

To explain geometrically linear dependence of a system of three vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ in $\mathbf{R}^3$ we use the lemma from Sec. 3.10, which can be formulated thus: *A system of three vectors in $\mathbf{R}^3$ is linearly dependent if and only if the vectors are coplanar.*

Example 1. Find whether the system of the three vectors

$$\mathbf{a} = \langle -1, 0, 7, 2, 3 \rangle$$
$$\mathbf{b} = \langle 2, 0, -14, -4, -6 \rangle$$

in $\mathbf{R}^5$ is linearly dependent.

○ The coordinates of **b** are proportional to the corresponding coordinates of **a**:

$$2:(-1) = 0:0 = (-14):7 = (-4):2 = (-6):3$$

Thus, the system **a**, **b** is linearly dependent (**b** = 2**a**). ●

Example 2. Find whether the system of the three vectors $\mathbf{a}_1 = \langle 2, 3, 5\rangle$, $\mathbf{a}_2 = \langle -4, 5, 7\rangle$, $\mathbf{a}_3 = \langle 10, -7, -9\rangle$ is linearly dependent.

○ A system of three vectors $\mathbf{R}^3$ is linearly dependent if the vectors are coplanar, and in turn, according to the theorem in Sec. 3.10, the vectors are coplanar if and only if the determinant formed from their coordinates is zero. We have the determinant

$$\begin{vmatrix} 2 & 3 & 5 \\ -4 & 5 & 7 \\ 10 & -7 & -9 \end{vmatrix}$$

which is zero (check this). Thus, $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ are linearly dependent. ●

Example 3. Find whether the system of the three vectors $\mathbf{a}_1 = \langle 1, 2, 3, 0\rangle$, $\mathbf{a}_2 = \langle -1, 0, 3, -2\rangle$, $\mathbf{a}_3 = \langle -1, 3, 12, -5\rangle$ is linearly dependent. If the answer is affirmative, set up an expression relating the vectors.

○ According to the definition of linear dependence, we must find whether the equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{0} \tag{7}$$

in three unknowns x_1, x_2, x_3 has at least one *nontrivial* solution.

We write (7) using vector coordinates

$$x_1\langle 1, 2, 3, 0\rangle + x_2\langle -1, 0, 3, -2\rangle + x_3\langle -1, 3, 12, -5\rangle = \langle 0, 0, 0, 0\rangle$$

Reducing to the echelon form yields

$$\langle x_1\cdot 1 + x_2\cdot(-1) + x_3\cdot(-1),\ x_1\cdot 2 + x_2\cdot 0 + x_3\cdot 3,\ x_1\cdot 3 + x_2\cdot 3 + x_3\cdot 12,\ x_1\cdot 0 + x_2\cdot(-2) + x_3\cdot(-5)\rangle = \langle 0, 0, 0, 0\rangle$$

Since the equality of two vectors means the equality of their respective coordinates, the last vector equation is equivalent to the system of four numerical equations

$$\begin{cases} x_1 - x_2 - x_3 = 0 \\ 2x_1 + 3x_3 = 0 \\ 3x_1 + 3x_2 + 12x_3 = 0 \\ - 2x_2 - 5x_3 = 0 \end{cases} \tag{8}$$

We solve (8) using Gaussian elimination:

$$\begin{cases} x_1 - x_2 - x_3 = 0 \\ 2x_1 + 3x_3 = 0 \\ 3x_1 + 3x_2 + 12x_3 = 0 \\ - 2x_2 - 5x_3 = 0 \end{cases} \to \begin{cases} x_1 - x_2 - x_3 = 0 \\ \boxed{\begin{array}{r} 2x_2 + 5x_3 = 0 \\ 6x_2 + 15x_3 = 0 \\ -2x_2 - 5x_3 = 0 \end{array}} \end{cases}$$

$$\to \begin{cases} x_1 - x_2 - x_3 = 0 \\ 2x_2 + 5x_3 = 0 \end{cases}$$

The unknown x_3 in the last system is free (it may take any numerical value). Hence (8) has nontrivial solutions. For instance, putting $x_3 = -2$ yields the solution $x_1 = 3$, $x_2 = 5$, $x_3 = -2$. Consequently, $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ are related as

$$3\mathbf{a}_1 + 5\mathbf{a}_2 - 2\mathbf{a}_3 = 0 \tag{9}$$

Thus, the system $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ is linearly dependent and its vectors are related by (9). ●

10.3. PROPERTIES OF LINEAR DEPENDENCE

We consider the system of vectors

$$\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p \tag{1}$$

from $\mathbf{R}^n$.

1°. *If one of the vectors is the zero vector, then the system is linearly dependent.*

□ Suppose, for instance, that the first vector in (1) is a nonzero vector. Then obviously

$$\mathbf{0} = 0 \cdot \mathbf{a}_2 + 0 \cdot \mathbf{a}_3 + \ldots + 0 \cdot \mathbf{a}_p$$

which shows that $\mathbf{a}_1$ is a linear combination of the other vectors, that is, system (1) is linearly dependent. ■

2°. *If a part of a system is linearly dependent, then the entire system is linearly dependent.*

□ Suppose a part of system (1), say, the first k vectors ($k < p$), is linearly dependent. This means that there are numbers $c_1, c_2, \ldots, c_k$ not all zero such that

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \ldots + c_k\mathbf{a}_k = \mathbf{0}$$

which implies that

$$c_1\mathbf{a}_1 + \ldots + c_k\mathbf{a}_k + 0\cdot\mathbf{a}_{k+1} + \ldots + 0\cdot\mathbf{a}_p = \mathbf{0}$$

i.e. the entire system (1) is linearly dependent. ■

3°. *If the linearly independent system* (1) *becomes linearly dependent, once a vector* **a** *is added, then* **a** *can be represented as a linear combination of the vectors of* (1).

□ The linear dependence of the system $\mathbf{a}_1, \ldots, \mathbf{a}_p, \mathbf{a}$ means there is an equation

$$c_1\mathbf{a}_1 + \ldots + c_p\mathbf{a}_p + c\mathbf{a} = \mathbf{0} \tag{2}$$

where $c_1, c_2, \ldots, c_p, c$ are not all zero. Note that the coefficient of **a** in (2) is necessarily nonzero; otherwise, (2) would imply that the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_p$ are linearly dependent. Since $c \neq 0$, **a** can be represented as a linear combination of $\mathbf{a}_1, \ldots, \mathbf{a}_p$. ■

1. The echelon form of a system of vectors.

Another property of linear dependence concerns a system of vectors which has a special structure. We say that such systems have the *echelon* form.

Definition. A system of vectors

$$\begin{array}{l} \mathbf{a}_1 = \langle a_{11}, a_{12}, \ldots, a_{1p}, \ldots, a_{1n}\rangle \\ \mathbf{a}_2 = \langle 0, \; a_{22}, \ldots, a_{2p}, \ldots, a_{2n}\rangle \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \mathbf{a}_p = \langle 0, \quad 0, \; \ldots, a_{pp}, \ldots, a_{pn}\rangle \end{array} \tag{3}$$

is said to have the *echelon* form.

Here the coordinates below the dashed "diagonal" are all zero, while the coordinates next to the diagonal and above it, i.e. $a_{11}, a_{22}, \ldots, a_{pp}$, are all nonzero.

Obviously, $p \leqslant n$ for the echelon form of a system, that is, the number of vectors does not exceed the number of vector coordinates.

4°. *A system of vectors having the echelon form is linearly independent.*

□ Let us find when a vector of the form

$$\mathbf{a} = c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \ldots + c_p\mathbf{a}_p$$

is the zero vector. Using (3) we find that the first coordinate of $\mathbf{a}$ is c_1a_{11}; if it is zero, then $c_1 = 0$ since $a_{11} \neq 0$. Then the term $c_1\mathbf{a}_1$ in the equation for $\mathbf{a}$ vanishes and the second coordinate of $\mathbf{a}$ is c_2a_{22}; if it is also zero, then $c_2 = 0$ since $a_{22} \neq 0$, and so on. Thus, we infer that

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \ldots + c_p\mathbf{a}_p = \mathbf{0}$$

is valid if $c_1 = c_2 = \ldots = c_p = 0$, which implies that system (3) is linearly independent. ■

Example. Show that the vectors in $\mathbf{R}^5$ $\mathbf{a}_1 = \langle 0, 3, 4, -2, 1\rangle$, $\mathbf{a}_2 = \langle -1, 0, 2, 3, -5\rangle$, and $\mathbf{a}_3 = \langle 0, 0, -2, 1, 1\rangle$ are linearly independent.

○ Arranging the vectors in the following order: $\mathbf{a}_2 = \langle -1, 0, 2, 3, -5\rangle$, $\mathbf{a}_1 = \langle 0, 3, 4, -2, 1\rangle$, $\mathbf{a}_3 = \langle 0, 0, -2, 1, 1\rangle$ yields the echelon form of a system of vectors. ●

2. The number of vectors in a linearly independent system. The corollary follows from property 4°: *there is a linearly independent system in* $\mathbf{R}^n$ *consisting of n vectors exactly.*

An example is the echelon form of a system of vectors

$$\begin{aligned} \mathbf{e}_1 &= \langle 1, 0, \ldots, 0\rangle \\ \mathbf{e}_2 &= \langle 0, 1, \ldots, 0\rangle \\ &\ldots\ldots\ldots\ldots \\ \mathbf{e}_n &= \langle 0, 0, \ldots, 1\rangle \end{aligned} \tag{4}$$

one of whose coordinates is 1 and all the others are 0. Any subsystem of it is also linearly independent, otherwise, the linear dependence of the subsystem would imply that the system is linearly dependent.

Naturally we would like to know whether there is a linearly independent system of *more than n* vectors in $\mathbf{R}^n$. In the next subsection we show that any system in $\mathbf{R}^n$ consisting of more than n vectors is *linearly dependent*.

We mention in conclusion an important property of the system $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ (see (4)), namely, any vector $\mathbf{a}$ from $\mathbf{R}^n$ can be represented as a linear combination of the vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. In fact, suppose we have vector $\mathbf{a} = \langle a_1, a_2, \ldots, a_n \rangle$. Multiplying $\mathbf{e}_1$ by a_1, $\mathbf{e}_2$ by a_2, and so on and adding the resulting vectors yield

$$a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \ldots + a_n\mathbf{e}_n$$
$$= \langle a_1, 0, \ldots, 0 \rangle = \langle 0, a_2, \ldots, 0 \rangle + \ldots + \langle 0, 0, \ldots, a_n \rangle$$
$$= \langle a_1, a_2, \ldots, a_n \rangle = \mathbf{a}$$

Thus, *there is a system of n linearly independent vectors in* $\mathbf{R}^n$ *in terms of which any vector of* $\mathbf{R}^n$ *can be linearly represented.*

10.4. BASES IN SPACE $\mathbf{R}^n$

Definition. A linearly independent system of vectors in a space $\mathbf{R}^n$ is called the *basis* of $\mathbf{R}^n$ if any vector from $\mathbf{R}^n$ can be represented as a linear combination of any vector of the system.

The last proposition formulated in the previous subsection implies that there is a basis composed of n vectors in $\mathbf{R}^n$. Other bases also exist for $\mathbf{R}^n$.

Lemma. *If there is a basis of* $\mathbf{R}^n$ *composed of p vectors, then any* $p + 1$ *vectors in* $\mathbf{R}^n$ *are linearly dependent.*

□ For the sake of simplicity, let $p = 2$. Thus, we have a basis of two vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ in $\mathbf{R}^n$: we prove that any three vectors $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_3$ are linearly dependent.

We denote the basis vectors by $\mathbf{a}$ and $\mathbf{b}$ instead of $\mathbf{a}_1$ and $\mathbf{a}_2$.

Since **a** and **b** form a basis, any vector $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ can be represented in terms of **a** and **b**:

$$\mathbf{b}_1 = \alpha_1 \mathbf{a} + \beta_1 \mathbf{b}, \quad \mathbf{b}_2 = \alpha_2 \mathbf{a} + \beta_2 \mathbf{b}, \quad \mathbf{b}_3 = \alpha_3 \mathbf{a} + \beta_3 \mathbf{b}$$

Multiplying the first equation by a number x_1 (so far arbitrary), the second equation by x_2, and the third equation by x_3 and adding the results yield

$$\begin{aligned} & x_1 \mathbf{b}_1 + x_2 \mathbf{b}_2 + x_3 \mathbf{b}_3 \\ & = x_1(\alpha_1 \mathbf{a} + \beta_1 \mathbf{b}) + x_2(\alpha_2 \mathbf{a} + \beta_2 \mathbf{b}) + x_3(\alpha_3 \mathbf{a} + \beta_3 \mathbf{b}) \\ & = (x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3)\mathbf{a} + (x_1\beta_1 + x_2\beta_2 + x_3\beta_3)\mathbf{b} \end{aligned}$$

If we can choose x_1, x_2, x_3 such that not all are zero and both expressions in parentheses are zero

$$\begin{aligned} \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 = 0 \\ \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 = 0, \end{aligned} \tag{1}$$

then

$$x_1 \mathbf{b}_1 + x_2 \mathbf{b}_2 + x_3 \mathbf{b}_3 = \mathbf{0}$$

is valid, i.e. the system $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ is linearly dependent. Thus, in order to prove the theorem, it is sufficient to show that system (1) has a nontrivial solution. Since system (1) is homogeneous and the number of its equations is smaller than the number of the unknowns (two equations in three unknowns), it has nontrivial solutions according to the theorem in Sec. 9.2, Item 5. Thus, the theorem is proved for the special case of $p = 2$.

A similar argument can be used for the general case (for any p). Instead of system (1), we have a homogeneous system of p equations in $p + 1$ unknowns which, according to the theorem, has a nontrivial solution. ■

Given the lemma, we can easily prove the following theorems.

Theorem 1. *Any two bases of a space* $\mathbf{R}^n$ *have the same number of vectors.*

Note that since one of the bases of $\mathbf{R}^n$, namely, the basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ in the preceding subsection, consists of n vectors, *any basis of* $\mathbf{R}^n$ *has* n *vectors.*

□ Suppose a basis of $\mathbf{R}^n$ has p vectors and another basis has q vectors. We shall prove this indirectly: let $p \neq q$, say $p < q$. Then we can choose $p + 1$ vectors from the second basis. According to the lemma, these vectors must be linearly dependent, and, therefore the second basis is linearly dependent, but this contradicts the definition of a basis. ■

Theorem 2. *Any system in* $\mathbf{R}^n$ *having more than* n *vectors is linearly dependent.*

□ A space $\mathbf{R}^n$ contains a basis of n vectors. According to the lemma, any system of $n + 1$ vectors in $\mathbf{R}^n$ must be linearly dependent. But then, generally, any system of more than n vectors is linearly dependent, since we can always choose a subsystem of $n + 1$ vectors and this is linearly dependent (according to the lemma). ■

In conclusion let us characterize all the bases of $\mathbf{R}^n$.

The bases of $\mathbf{R}^n$ *are different linearly independent systems of* n *vectors.*

Given a basis, its vectors are linearly independent and their number is n. Conversely, any linearly independent system of n vectors

$$\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \tag{2}$$

is a basis of $\mathbf{R}^n$. To make sure, we add a vector $\mathbf{a}$ in $\mathbf{R}^n$ to system (2) and obtain a system of $n + 1$ vectors which according to Theorem 2 must be linearly dependent. By property 3°, Sec. 10.3, of linear dependence we see that $\mathbf{a}$ can be represented as a linear combination of $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$. This proves that the system $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ is a basis of $\mathbf{R}^n$. ■

Example 1. The system of vectors $\mathbf{a}_1 = \langle 1, 2, 3 \rangle$, $\mathbf{a}_2 = \langle -1, 0, 3 \rangle$, $\mathbf{a}_3 = \langle 2, 5, -2 \rangle$, $\mathbf{a}_4 = \langle 4, 12, 2 \rangle$ from $\mathbf{R}^3$ is linearly dependent since the number of vectors is greater than three.

Example 2. Verify whether the vectors $\mathbf{a}_1 = \langle 1, 2, 3 \rangle$, $\mathbf{a}_2 = \langle 0, 1, -1 \rangle$, $\mathbf{a}_3 = \langle 2, 4, 5 \rangle$ form a basis in $\mathbf{R}^3$.

○ Since we have three vectors, it only remains to be established whether they are linearly independent or, which is the same, whether the equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{0}$$

has a trivial solution ($x_1 = x_2 = x_3 = 0$). This equation is equivalent to the system

$$\begin{cases} x_1 + 2x_3 = 0 \\ 2x_1 + x_2 + 4x_3 = 0 \\ 3x_1 - x_2 + 5x_3 = 0 \end{cases}$$

According to the theorem in Sec. 3.9, a homogeneous 3×3 system has a nontrivial solution if and only if its determinant is zero. The determinant of our system

$$\begin{vmatrix} 1 & 0 & 2 \\ 2 & 1 & 4 \\ 3 & -1 & 5 \end{vmatrix} = -1$$

is nonzero, thus, the system has only a trivial solution. Hence it follows that the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ form a basis of $\mathbf{R}^3$. ●

10.5. ABSTRACT VECTOR SPACES

Up till now we studied an arithmetic vector space $\mathbf{R}^n$ which is the set of all arithmetic vectors $\langle a_1, a_2, \ldots, a_n \rangle$ in which two algebraic operations, viz., vector addition and the multiplication of a vector by a scalar, are defined.

A thorough analysis of the basic theorems on arithmetic vectors shows that the notation of a vector in the form of a line of several numbers is not essential. The important thing is that there are two operations: vector addition and multiplication by a scalar, which possess certain properties permitting to operate on expressions of the form $k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + \ldots + k_p\mathbf{a}_p$, where $\mathbf{a}_1, \ldots, \mathbf{a}_p$ are vectors and $k_1, \ldots, k_p$ are scalars and which obey ordinary algebraic laws.

This suggests a principle for generalizing the notion of an arithmetic vector. We should call a vector any object for which two operations, viz., vector addition and multiplication by a scalar, are defined, provided the operations obey some natural requirements. The advantage of such an approach is that there is now no need to consider a vector to be a set of numbers and we may call objects of diverse nature, such as functions and matrices, vectors.

1. Definition of an abstract vector space.

Definition. A set L is called a *vector* (or *linear space*) and its elements are called *vectors* if on this set

(I) the operation of *addition* is defined: every pair of vectors $\mathbf{a}$ and $\mathbf{b}$ from L is associated with a third vector from L calied the *sum* of $\mathbf{a}$ and $\mathbf{b}$ and denoted by $\mathbf{a} + \mathbf{b}$,

(II) the operation of *multiplication of a vector by a scalar*: each pair $\mathbf{a}$, k (where $\mathbf{a}$ is a vector and k is a scalar) is associated in L with a vector called the *product of k and* $\mathbf{a}$ and denoted by $k\mathbf{a}$,

(III) these operations possess the following properties:

(1) $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ for any two vectors $\mathbf{a}$ and $\mathbf{b}$ in L,

(2) $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$ for any three vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$, in L,

(3) there is a unique vector $\mathbf{0}$ in L such that $\mathbf{a} + \mathbf{0} = \mathbf{a}$ for any $\mathbf{a}$ in L,

(4) for any vector $\mathbf{a}$ there is a unique vector $\mathbf{a}'$ such that $\mathbf{a} + \mathbf{a}' = \mathbf{0}$ in L,

(5) $1 \cdot \mathbf{a} = \mathbf{a}$ for any $\mathbf{a}$ in L,

(6) $k_1(k_2\mathbf{a}) = (k_1k_2)\mathbf{a}$ for any scalars k_1, k_2 and any $\mathbf{a}$ in L,

(7) $(k_1 + k_2)\mathbf{a} = k_1\mathbf{a} + k_2\mathbf{a}$ for any scalars k_1, k_2 and any $\mathbf{a}$ in L,

(8) $k(\mathbf{a} + \mathbf{b}) = k\mathbf{a} + k\mathbf{b}$ for any scalar k and any vectors $\mathbf{a}$ and $\mathbf{b}$.

The vector 0 in (3) is called the *zero vector*; the vector $\mathbf{a}'$ in (4) is called the *negative* of $\mathbf{a}$ and is denoted by $-\mathbf{a}$.

The branch of mathematics that studies vector spaces is called *linear* algebra.

Note that we have considered two sets of two different operations but used the same notation for each set, viz., vector addition and scalar addition, and multiplication of a vector by a scalar and multiplication of scalars. However, this is not umbiguous since it is clear from the notation which operation is in use. For instance, given the equation

$$(k_1 + k_2)\mathbf{a} = k_1\mathbf{a} + k_2\mathbf{a}$$

the + sign on the left-hand side means addition of scalars, while the + sign on the right-hand side means vector addition.

2. Examples of vector spaces.

Space $\mathbf{R}^n$. We considered this space earlier. Its elements (arithmetic vectors) are the sets of n numbers whose addition and multiplication by scalars are defined by the rules

$$\langle a_1, a_2, \ldots, a_n\rangle + \langle b_1, b_2, \ldots, b_n\rangle$$
$$= \langle a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n\rangle, \; k\langle a_1, a_2, \ldots, a_n\rangle$$
$$= \langle ka_1, ka_2, \ldots, ka_n\rangle .$$

Since n can assume infinitely many values 1, 2, 3, ..., we have an infinite series of vector spaces: $\mathbf{R}^1$, $\mathbf{R}^2$, $\mathbf{R}^3$, ...

Function space. We can consider various mappings of the set $\mathbf{R}$ of real numbers into itself or, which is the same, various functions

$$y = f(x)$$

defined on $\mathbf{R}$ and varying in $\mathbf{R}$. Let us introduce the operations of addition and multiplication for functions. To add two functions $f_1(x)$ and $f_2(x)$ together is to construct a new function $f_1(x) + f_2(x)$ whose value for any $x = a$ is $f_1(a) + f_2(a)$; multiplication of $f(x)$ by a scalar k yields $kf(x)$ whose value for any $x = a$ is $kf(a)$.

The set of all functions on which the operations of addition and multiplication by a scalar are defined is a vector space.

Space of matrices. We consider various matrices of the form

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

having two rows and two columns. Addition of matrices and multiplication of a matrix by a scalar are defined as follows:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix}$$

$$k\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} ka_{11} & ka_{12} \\ ka_{21} & ka_{22} \end{pmatrix}$$

We can easily verify that all the conditions in the definition of a vector space are satisfied here. Thus, we have a vector space which is the *space of 2 × 2 matrices* (i.e. of matrices having two rows and two columns).

Vector spaces of 3 × 3, 4 × 4, etc. matrices can be constructed in a similar way. Vector spaces of $m \times n$ matrices, where $m \neq n$, also exist.

3. Basis.

We can introduce a number of concepts for an abstract vector space in the same way as we did for the space $\mathbf{R}^n$. Specifically, we can speak of a *linear combination*

$$k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + \dots + k_p\mathbf{a}_p$$

of several vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p$ and of *linearly dependent* and *linearly independent* systems of vectors. However, the properties of L may differ from those of $\mathbf{R}^n$ in one respect. In order to formulate this (possible) difference, we introduce the important concept of a *basis* of the space L, which actually repeats the definition of a basis of $\mathbf{R}^n$.

Definition. A *basis* of a vector space L is a linearly independent system of vectors from L such that any vector from L can be represented as a linear combination of the system's vectors.

A basis exists in $\mathbf{R}^n$, an example being the system of n vectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ constructed in Sec. 10.3, Item 2. There are vector spaces L which *have no basis*. In order to exclude such spaces from our consideration, we must add one more condition to (I), (II), and (III) in the definition of a vector space given in Item 1:

(IV) *the space L has at least one basis.*

4. The coordinates of a vector in terms of a basis.

Let us choose a basis of L consisting of n vectors

$$\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n \tag{1}$$

Any vector $\mathbf{a}$ from L can be expressed in terms of the basis, i.e. represented as a linear combination of the basis vectors

$$\mathbf{a} = k_1\mathbf{e}_1 + k_2\mathbf{e}_2 + \dots + k_n\mathbf{e}_n \tag{2}$$

Theorem. *Any vector* **a** *can be expressed in terms of the basis in a unique manner.*

□ Suppose we have another representation of **a** in the basis besides (2):

$$\mathbf{a} = k_1'\mathbf{e}_1 + k_2'\mathbf{e}_2 + \dots + k_n'\mathbf{e}_n \qquad (3)$$

Subtracting (2) from (3) yields

$$(k_1' - k_1)\mathbf{e}_1 + (k_2' - k_2)\mathbf{e}_2 + \dots + (k_n' - k_n)\mathbf{e}_n = \mathbf{0}$$

But by assumption the vectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ are linearly independent since they form the basis. Consequently, the last equation can only be valid if all the coefficients $k_1' - k_1$, $k_2' - k_2, \dots, k_n' - k_n$ are zero, that is,

$$k_1 = k_1',\ k_2 = k_2',\ \dots,\ k_n = k_n'$$

Hence, representations (2) and (3) are the same. ■

Definition. The coefficients $k_1, k_2, \dots, k_n$ in the representation of the vector **a** in terms of the basis (1) are called the *coordinates of* **a** *in the basis.*

Equations

$$(k_1\mathbf{e}_1 + k_2\mathbf{e}_2 + \dots + k_n\mathbf{e}_n) + (l_1\mathbf{e}_1 + l_2\mathbf{e}_2 + \dots + l_n\mathbf{e}_n)$$
$$= (k_1 + l_1)\mathbf{e}_1 + (k_2 + l_2)\mathbf{e}_2 + \dots + (k_n + l_n)\mathbf{e}_n$$

and

$$c(k_1\mathbf{e}_1 + k_2\mathbf{e}_2 + \dots + k_n\mathbf{e}_n)$$
$$= (ck_1)\mathbf{e}_1 + (ck_2)\mathbf{e}_2 + \dots + (ck_n)\mathbf{e}_n$$

following from the properties of vector addition and multiplication of a vector by a scalar show that *in order to add vectors, we add their respective coordinates in the given basis*, and *when multiplying a vector by a scalar we multiply its coordinates in the given basis by that scalar.*

Example 1. The vectors **a** and **b** have the coordinates in a basis of L $-1, 2, 3, 0, 6$ and $1, 1, -3, 0, -4$, respectively. Find the coordinates of the vector $2\mathbf{a} - 3\mathbf{b}$ in the same basis.

○ The coordinates of the vector $2\mathbf{a}$ are $-2, 4, 6, 0, 12$ and those of $-3b$ are $-3, -3, 9, 0, 12$. Hence it follows that the coordinates of $2\mathbf{a} - 3\mathbf{b}$ are $-5, 1, 15, 0, 24$. ●

Example 2. Verify that the arithmetic vectors $\mathbf{a}_1 = \langle 1, 2, 3 \rangle$, $\mathbf{a}_2 = \langle 0, 1, -1 \rangle$, $\mathbf{a}_3 = \langle 2, 4, 5 \rangle$ form a basis of $\mathbf{R}^3$ and find the coordinates of the vector $\mathbf{a} = \langle 5, 13, 9 \rangle$ in that basis.

○ We showed in Sec. 9.4 that the system $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ is a basis of $\mathbf{R}^3$. Thus it remains for us to find the coordinates of $\mathbf{a}$ in terms of the basis $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$.

By denoting the desired coordinates by x_1, x_2, x_3, we can write

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{a} \tag{4}$$

By assumption we have

$$\begin{aligned} & x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 \\ & = \langle x_1 + 2x_3,\ 2x_1 + x_2 + 4x_3,\ 3x_1 - x_2 + 5x_3 \rangle \end{aligned}$$

and therefore (4) is equivalent to the system

$$\begin{cases} x_1 \qquad + 2x_3 = 5 \\ 2x_1 + x_2 + 4x_3 = 13 \\ 3x_1 - x_2 + 5x_3 = 9 \end{cases}$$

Gaussian elimination yields

$$\begin{cases} x_1 + 2x_3 = 5 \\ \boxed{\begin{array}{l} x_2 = 3 \\ -x_2 - x_3' = -6 \end{array}} \end{cases}$$

whence it follows that $x_2 = 3$, $x_3 = 3$, $x_1 = -1$. Thus, $\mathbf{a} = -\mathbf{a}_1 + 3\mathbf{a}_2 + 3\mathbf{a}_3$ whose coordinates in the basis $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ are $-1, 3, 3$.

5. Representation of an abstract vector space L as an arithmetic vector space $\mathbf{R}^n$.

Suppose a vector space L has a basis consisting of n vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. By representing a vector $\mathbf{a}$ from L as a linear combination of the vectors of the basis:

$$\mathbf{a} = k_1\mathbf{e}_1 + k_2\mathbf{e}_2 + \ldots + k_n\mathbf{e}_n$$

we can associate $\mathbf{a}$ with a line of numbers $k_1, k_2, \ldots, k_n$ which are the coordinates of $\mathbf{a}$ in the basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$, thereby, associating $\mathbf{a}$ with an arithmetic vector from $\mathbf{R}^n$.

The properties of the coordinates of a vector in terms of the basis imply that when vectors from L are added, their corresponding vectors from $\mathbf{R}^n$ are also added, and when a vector from L is multiplied by a scalar, its corresponding arithmetic vector is also multiplied by that number. Thus, we can say that *under juxtaposition*

$$\mathbf{a} \to \langle k_1, k_2, \ldots, k_n \rangle$$

the space L transforms into an arithmetic vector space $\mathbf{R}^n$, $\mathbf{R}^n$ being a "numerical model" of L. Thus, everything we proved earlier for $\mathbf{R}^n$ must be valid for L. Specifically,

(1) *any two bases of L contain the same number of vectors,*

(2) *if a basis of L consists of n vectors, then any system of more than n vectors in L is linearly dependent.*

6. Dimensions of a vector space.

Definition. The number of vectors of a basis of a vector space L is called the *dimension* of L. If the number of vectors is n, then L is said to have *dimension* n (or to be *n-dimensional*).

If $\mathbf{R}^n$ contains a basis of n vectors, then $\mathbf{R}^n$ has dimension n or is said to be an arithmetic n-space. We know that any system of $n + 1$ or more vectors in $\mathbf{R}^n$ is linearly dependent and we can give another definition of dimension: the *dimension of a space L is the maximal number of linearly independent vectors in L.*

Exercises to Chapter 10

1. Find the linear combination $2\mathbf{a}_1 - \mathbf{a}_2 + 3\mathbf{a}_3$ of the vectors $\mathbf{a}_1 = \langle 1, -1, 0, 4 \rangle$, $\mathbf{a}_2 = \langle 16, 4, 7, -2 \rangle$, and $\mathbf{a}_3 = \langle 5, 2, 2, -3 \rangle$.

2. Find the vector $5\mathbf{a}_1 - 2\mathbf{a}_2 + 7\mathbf{a}_3$ if $\mathbf{a}_1 = \mathbf{b}_1 - 2\mathbf{b}_2$, $\mathbf{a}_2 = 3\mathbf{b}_1 + \mathbf{b}_2$, $\mathbf{a}_3 = -\mathbf{b} + \mathbf{b}_2 + \mathbf{b}_3$ and $\mathbf{b}_1 = \langle 0, 0, 1 \rangle$, $\mathbf{b}_2 = \langle -1, 2, 3 \rangle$, $\mathbf{b}_3 = \langle -1, 1, 0 \rangle$.

3. Find the vector $\mathbf{x}$ from the equation $\mathbf{a} + 2\mathbf{b} + 3\mathbf{c} + 4\mathbf{x} = 0$, where $\mathbf{a} = \langle 5, -8, -1 \rangle$, $\mathbf{b} = \langle 2, -1, -4 \rangle$, $\mathbf{c} = \langle -3, 2, -5 \rangle$.

4. Find the vector $\mathbf{x}$ from the equation $3(\mathbf{a} - \mathbf{x}) + 2(\mathbf{b} + \mathbf{x}) = 5(\mathbf{c} + \mathbf{x})$, where $\mathbf{a} = \langle 2, 5, 1, 3 \rangle$, $\mathbf{b} = \langle 10, 1, 5, 10 \rangle$, $\mathbf{c} = \langle 4, 1, -1, 1 \rangle$.

5. Show without calculation that each system of vectors is linearly dependent:

(a) $\mathbf{a}_1 = \langle 1, 1, 0, 1 \rangle$, $\mathbf{a}_2 = \langle -3, -3, 0, -3 \rangle$,

(b) $\mathbf{a}_1 = \langle 1, 2, 3, 4 \rangle$, $\mathbf{a}_2 = \langle 4, 3, 2, 1 \rangle$, $\mathbf{a}_3 = \langle 5, 5, 5, 5 \rangle$,

(c) $\mathbf{a}_1 = \langle 1, 2, 3\rangle$, $\mathbf{a}_2 = \langle 3, 2, 1\rangle$, $\mathbf{a}_3 = \langle 1, 1, 1\rangle$.

6. Show without calculation that the system of vectors is linearly independent:

(a) $\mathbf{a}_1 = \langle 1, 0, 0, 0\rangle$, $\mathbf{a}_2 = \langle 0, 1, 0, 0\rangle$, $\mathbf{a}_3 = \langle 0, 0, 1, 0\rangle$, $\mathbf{a}_4 = \langle 2, 3, 4, 5\rangle$.

(b) $\mathbf{a}_1 = \langle 0, 2, 5\rangle$, $\mathbf{a}_2 = \langle 0, 0, 3\rangle$, $\mathbf{a}_3 = \langle -1, 4, 7\rangle$.

7. Given the two vectors $\mathbf{a}_1 = \langle 1, 2, 3, 4\rangle$, and $\mathbf{a}_2 = \langle 0, 0, 0, 1\rangle$, choose, without calculating, two more vectors $\mathbf{a}_3$ and $\mathbf{a}_4$ such that the system $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, $\mathbf{a}_4$ is linearly independent.

8. Find all the values of λ such that the vector $\mathbf{b}$ can be represented as a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$:

$\mathbf{a}_1 = \langle 3, -1\rangle$	$\mathbf{a}_1 = \langle 1, 0, 0\rangle$	$\mathbf{a}_1 = \langle 3, 4, 2\rangle$
(1) $\mathbf{a}_2 = \langle 5, -2\rangle$	(2) $\mathbf{a}_2 = \langle 0, 1, 0\rangle$	(3) $\mathbf{a}_2 = \langle 6, 8, 7\rangle$
$\mathbf{b} = \langle 2, \lambda\rangle$	$\mathbf{b} = \langle 0, 0, \lambda\rangle$	$\mathbf{b} = \langle 9, 12, \lambda\rangle$

9. Find whether each system is linearly dependent:

(a) $\mathbf{a}_1 = \langle 1, 1, 1\rangle$, $\mathbf{a}_2 = \langle 1, 2, 3\rangle$, $\mathbf{a}_3 = \langle 1, 3, 3\rangle$,

(b) $\mathbf{a}_1 = \langle 1, 1, 1, 1\rangle$, $\mathbf{a}_2 = \langle 1, -1, 1, -1\rangle$, $\mathbf{a}_3 = \langle 2, 3, 1, 4\rangle$, $\mathbf{a}_4 = \langle 2, 1, 1, 3\rangle$,

(c) $\mathbf{a}_1 = \langle 2, 3, -4, 1\rangle$, $\mathbf{a}_2 = \langle 1, -2, 4, 0\rangle$, $\mathbf{a}_3 = \langle 0, 0, 5, 1\rangle$.

10. Two vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ have the coordinates $-1, 2$ and $3, -5$, respectively, in a basis of a two-dimensional vector space L, and the coordinates of vector $\mathbf{b}$ are 1, 1. Show that $\mathbf{a}_1$ and $\mathbf{a}_2$ constitute a basis and find the coordinates of $\mathbf{b}$ in this basis.

11. Three vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ have the coordinates $1, -1, 0$; $1, 2, 3$; and $0, 1, -1$ respectively in a basis of a three-dimensional vector space L, and the coordinates of vector $\mathbf{b}$ are 6, 6, 6. Show that $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ constitute a basis and find the coordinates of $\mathbf{b}$ in this basis.

Chapter 11

MATRICES

We already used matrices in Chapter 3 when considering the determinants of square matrices and Chapter 9 when solving systems of linear equations by Gaussian elimination. In fact, matrices are extensively applied in science and engineering.

We recall what a matrix is.

Definition. An $m \times n$ *matrix* is a set of numbers arranged in a rectangular array of m rows and n columns and usually enclosed in parentheses.

For instance,

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

is the 2×3 matrix.

An $n \times n$ matrix, which has the same number of rows and columns is called a *square matrix of order n*.

The numbers constituting a matrix are called the *elements* of the matrix.

We use the usual notation and denote the element in the ith row and jth column by a_{ij}. Thus, the general form of an $m \times n$ matrix is

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

We shall denote a matrix by a single upper-case letter in bold face, **A**, **B**, **C**, and so on.

11.1. RANK OF A MATRIX

1. Definition of the rank of a matrix.
Given an $m \times n$ matrix **A**

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Each row of **A** can be considered as an arithmetic vector having n coordinates, i.e. a vector from $\mathbf{R}^n$. We denote these vectors by $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_m$:

$$\mathbf{A}_1 = \langle a_{11}, a_{12}, \dots, a_{1n} \rangle$$
$$\mathbf{A}_2 = \langle a_{21}, a_{22}, \dots, a_{2n} \rangle$$
$$\dots\dots\dots\dots\dots\dots$$
$$\mathbf{A}_m = \langle a_{m1}, a_{m2}, \dots, a_{mn} \rangle$$

Definition. The *rank* of the matrix **A** is the maximum number of linearly independent vectors in the system $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_m$ of its rows.

When all the elements of a matrix **A** are zero, the order of **A** is zero.

Thus, the rank of a matrix is a number r such that
(1) the matrix contains r linearly independent rows,
(2) any $r + 1$ rows of the matrix are linearly dependent.

Example. Consider the 3×4 matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 3 & -2 \\ -1 & 1 & 4 & 3 \\ 2 & 1 & 13 & -3 \end{pmatrix}$$

The first two rows $\mathbf{A}_1 = \langle 1, 0, 3, -2 \rangle$ and $\mathbf{A}_2 = \langle -1, 1, 4, 3 \rangle$ are linearly independent since the coordinates of $\mathbf{A}_1$ are not proportional to those of $\mathbf{A}_2$. At the same time the system of the rows $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ is linearly dependent since $\mathbf{A}_3 = 3\mathbf{A}_1 + \mathbf{A}_2$ (check this). Hence, the rank of **A** is two.

2. Two propositions on the rank.

(1) *Suppose any r rows of* **A** *are linearly independent. If the*

rank of **A** *is r, then all the remaining rows can be represented as a linear combination of the r rows.*

□ Adding any of the remaining rows to the r rows yields a system of $r + 1$ rows which, by the definition of rank, must be linearly dependent. According to property 3^0, Sec. 10.2, of linear dependence, it follows that the added row can be expressed as a linear combination of the r rows. ■

(2) *Suppose any r rows of* **A** *are linearly independent. If the remaining rows can be represented as a linear combination of the r rows, the rank of* **A** *is r.*

□ The proof is based on the concept of a *basis* of a finite system of vectors.

A basis of a system of vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p$ is a linearly independent part of the system (a subsystem) such that any vector of the system can be represented as a linear combination of the subsystem. The following theorem is valid, and it can be proved in the same way as the theorem on bases (Sec. 10.4).

Theorem. *Any two bases of a given system of vectors consist of the same number of vectors.*

This theorem helps us prove proposition (2). Suppose we have r rows of **A**, say $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_r$, such that the assumptions in (2) are satisfied, i.e. they are linearly independent and any row of **A** can be represented as linear combination of the r rows. Then, these r rows constitute a basis of the system of all the rows of **A**. Let us consider any $r + 1$ rows of **A**. If they were linearly independent, this would mean that there is a basis of more than r vectors in the system of all rows of **A**, which, according to the above theorem, is impossible. Hence it follows that any $r + 1$ rows of **A** are linearly dependent, and the rank of **A** is r. ■

3. Properties of the rank of a matrix.

1^0. *The rank of a matrix* **A** *is not changed, if we add a row (say, next to the last row) which is a linear combination of the rows of* **A**.

□ We denote the new matrix by $\tilde{\mathbf{A}}$. Let us prove that the ranks of **A** and $\tilde{\mathbf{A}}$ are the same.

Suppose that the rank of **A** is r. This means that **A** must contain r linearly independent rows, say, the first r rows $\mathbf{A}_1$, $\mathbf{A}_2$, ..., $\mathbf{A}_r$. The remaining rows of $\mathbf{A}$ can be represented as a linear combination of $\mathbf{A}_1$, $\mathbf{A}_2$, ..., $\mathbf{A}_r$. This implies that the added row can also be expressed as a linear combination of $\mathbf{A}_1$, $\mathbf{A}_2$, ..., $\mathbf{A}_r$; by assumption, it is a linear combination of the rows of $\mathbf{A}$. In other words, all the rows following the first r rows in $\mathbf{A}$ can be expressed as a linear combination of $\mathbf{A}_1$, $\mathbf{A}_2$, ..., $\mathbf{A}_r$. Thus, the rank of $\tilde{\mathbf{A}}$ is also r. ■

Property 1 can be formulated in a different way: *the rank of a matrix does not change if we delete a row which can be represented as a linear combination of the remaining rows of the matrix.*

□ In fact, if we denote the initial matrix by $\tilde{\mathbf{A}}$ and the matrix obtained by deleting the row by $\mathbf{A}$, we can see that $\tilde{\mathbf{A}}$ can be obtained from $\mathbf{A}$ by adding a row which is a linear combination of the rows of $\mathbf{A}$. But in this case, according to property 1^0, the ranks of $\mathbf{A}$ and $\tilde{\mathbf{A}}$ must be the same. ■

2^0. *The rank of a matrix is left unchanged if we add a linear combination of the other rows to any row.*

□ Suppose we add the second row of $\mathbf{A}$ multiplied by a number k to the first row. The result is the matrix $\tilde{\mathbf{A}}$ whose rows are

$$\mathbf{A}_1 + k\mathbf{A}_2, \quad \mathbf{A}_2, \; \ldots, \; \mathbf{A}_m \tag{1}$$

The rank of $\tilde{\mathbf{A}}$ is not changed if we add the row $\mathbf{A}_1$; indeed, $\mathbf{A}_1$ is a linear combination of rows (1), $\mathbf{A}_1 = (\mathbf{A}_1 + k\mathbf{A}_2) - k\mathbf{A}_2)$. We obtain a matrix with rows $\mathbf{A}_1$, $\mathbf{A}_1 + k\mathbf{A}_2$, $\mathbf{A}_2$, ..., $\mathbf{A}_m$. The rank of this matrix is again left unchanged if we delete its row $\mathbf{A}_1 + k\mathbf{A}_2$ (since this row is a linear combination of the other rows). The result is a matrix with rows $\mathbf{A}_1$, $\mathbf{A}_2$, ..., $\mathbf{A}_m$, that is, the initial matrix $\mathbf{A}$. Thus, the ranks of $\tilde{\mathbf{A}}$ and $\mathbf{A}$ are the same. ■

3^0. *The rank of* $\mathbf{A}$ *is not changed if we add a zero row.*

□ A zero row can always be considered as a linear combination of the rows of $\mathbf{A}$ ($\mathbf{0} = 0 \cdot \mathbf{A}_1 + 0 \cdot \mathbf{A}_2 + 0 \cdot \mathbf{A}_m$), and therefore property 3^0 follows from property 1^0. ■

We can give another formulation of property 3^0: *the rank of* $\mathbf{A}$ *is unaltered if we delete a zero row* (of course, provided there is such a row).

11.2. PRACTICAL METHOD FOR FINDING THE RANK OF A MATRIX

In order to find the rank of a matrix, we can apply Gaussian elimination to a given matrix **A** and not to a system of equations.

Definition. The operations in Gaussian elimination are called *elementary operations.* The elementary operations for matrices are

(1) interchanging two columns,

(b) deleting a zero row,

(c) adding one row multiplied by a number to another row.

The following theorem is valid.

Theorem. *The rank of a matrix is not changed during elementary operations.*

□ The statement is obvious for elementary operations of type (a); it follows from properties 2^0 and 3^0 (see Sec. 11.1, Item 3) for elementary operations of types (b) and (c). ■

The main idea of the practical method for computing the rank of a matrix is that the elementary operations can reduce matrix **A** to the form

$$\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1r} b_{1,r+1} & \dots & b_{1n} \\ 0 & b_{22} & \dots & b_{2r} b_{2,r+1} & \dots & b_{2n} \\ 0 & 0 & \dots & b_{rr} b_{r,r+1} & \dots & b_{rn} \end{pmatrix}$$

in which the diagonal elements $b_{11}, b_{22}, \dots, b_{rr}$ are nonzero, while the elements below the diagonal are zero. This form of a matrix is called the *echelon* form. Then we can immediately deduce from the echelon form of matrix that the rank of **A** is r.

In fact, the rows of **B** are linearly independent since they form the echelon form of vectors (see Sec. 10.3), and thus the rank of **B** equals the number of its row, i.e. r. Since the rank remains unchanged by elementary operations, we know that the rank of **A** is also r.

Rule. *For finding the rank of a matrix* **A**, *we should reduce it to the echelon form* **B** *using elementary operations. The rank of* **A** *equals the number of rows in* **B**.

Example 1. Find the rank of the matrix

$$\mathbf{A} = \begin{pmatrix} -1 & 3 & 3 & 2 & 5 \\ -3 & 5 & 2 & 3 & 4 \\ -3 & 1 & -5 & 0 & -7 \\ -5 & 7 & 1 & 4 & 1 \end{pmatrix}$$

○ Since $a_{11} \neq 0$, we can make all elements of the first column below a_{11} vanish. To do this, we multiply the first row by -3 and add the result to the second and third rows, and then we add the first row multiplied by -5 to the fourth row. The result is the matrix

$$\begin{pmatrix} -1 & 3 & 3 & 2 & 5 \\ 0 & -4 & -7 & -3 & -11 \\ 0 & -8 & -14 & -6 & -22 \\ 0 & -8 & -14 & -6 & -24 \end{pmatrix}$$

Noting that $a_{22} \neq 0$, we make all elements of the second column below a_{22} vanish. To this end, we add the second row multiplied by -2 to the third and fourth rows. We have the matrix

$$\begin{pmatrix} -1 & 3 & 3 & 2 & 5 \\ 0 & -4 & -7 & -3 & -11 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 \end{pmatrix}$$

from which we delete the zero row. The third element in the last (now the third) row is zero, however, we can make it non-zero by interchanging the third and fifth columns. Then we get the matrix

$$\mathbf{B} = \begin{pmatrix} -1 & 3 & 5 & 2 & 3 \\ 0 & -4 & -11 & -3 & -7 \\ 0 & 4 & -2 & 0 & 0 \end{pmatrix}$$

which is in echelon form. The rank of **B** is three; consequently the rank of the initial matrix **A** is also three. ●

Example 2. Find the rank of the matrix

$$\mathbf{A} = \begin{pmatrix} -1 & 4 & 3 & a & 2 \\ -2 & 7 & 5 & 1 & 3 \\ -1 & 2 & 1 & 8 & a \end{pmatrix}$$

for different values of the parameter a.

○ We use the elementary operations

$$\mathbf{A} \sim \begin{pmatrix} -1 & 4 & 3 & a & 2 \\ 0 & -1 & -1 & 1-2a & -1 \\ 0 & -2 & -2 & 8-a & a-2 \end{pmatrix}$$

$$\sim \begin{pmatrix} -1 & 4 & 3 & a & 2 \\ 0 & -1 & -1 & 1-2a & -1 \\ 0 & 0 & 0 & 6+3a & a \end{pmatrix}$$

hence it follows that if at least one of the numbers $6 + 3a$ or a is nonzero, then the rank of **A** is three, and if both numbers are zero, then the rank is two. But a and $6 + 3a$ cannot both be zero simultaneously. Thus, the rank of **A** is three for any a. ●

11.3. THEOREM ON THE RANK OF A MATRIX

Suppose we omit some k rows and k columns, $k \leqslant n$ and $k \leqslant m$, from the matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

The elements at the intersection of these rows and columns form a square matrix $\mathbf{A}'$ of order k (a submatrix of **A**). Its determinant $|\mathbf{A}'|$ is called a *kth-order minor of* **A**.

Nonzero minors of a matrix are especially interesting. The theorem given below relates the rank of a matrix and a maximal-order nonzero minor.

Theorem (on the rank of a matrix). *The rank of a matrix equals the maximal order of the nonzero minors of the matrix.*

In other words, if the rank of a matrix **A** is r, then **A** necessarily contains an rth-order nonzero minor, while all the minors of order $r + 1$ or more are zero. We omit the proof of the theorem.

Corollary 1. *The maximum number of linearly independent rows equals the maximum number of linearly independent columns.*

□ In fact, we consider a matrix **A** and its transpose A^*

$$\mathbf{A}^* = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}$$

($\mathbf{A}^*$ is obtained by interchanging the rows and columns in **A**). In order to prove the proposition it is sufficient to verify that

$$\text{rank } \mathbf{A} = \text{rank } \mathbf{A}^* \tag{1}$$

In fact, the left-hand side of (1) expresses the maximum number of linearly independent rows in **A**, while the right-hand side is the maximum number of rows in $\mathbf{A}^*$, which is, obviously, the maximum number of columns in **A**.

Thus, it remains for us to prove (1) which by the theorem on the rank of a matrix is equivalent to the following problem: prove that the maximal orders of the nonzero minors in the matrices **A** and $\mathbf{A}^*$ are the same. This last proposition follows directly from the properties of determinants (i.e. the transposition of a matrix leaves its determinant unchanged), since by transposing **A** we exchange each minor of **A** by the "transposed" minor and hence the minor is unaltered. ■

Corollary 2. *If the determinant is zero, then its rows are linearly dependent.*

□ We consider a square matrix **A** such that $|\mathbf{A}| = 0$. Then the maximal order of the nonzero minors must be less than n (n being the number of rows), hence, rank $\mathbf{A} < n$. This means that the rows of the matrix are linearly dependent. ■

11.4. RANK OF A MATRIX AND SYSTEMS OF LINEAR EQUATIONS

Suppose we have a system of linear equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases} \quad (1)$$

We arrange the coefficients of the unknowns in matrix form

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

and this is called the *coefficient matrix* of (1).

Suppose system (1) is consistent, then Gaussian elimination yields the system in the echelon form

$$\begin{cases} b_{11}x_1 + b_{12}x_2 + \ldots + b_{1r}x_r + b_{1,r+1}x_{r+1} + \ldots + b_{1n}x_n = c_1 \\ \qquad b_{22}x_2 + \ldots + b_{2r}x_r + b_{2,r+1}x_{r+1} + \ldots + b_{2n}x_n = c_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad\qquad b_{rr}x_r + b_{r,r+1}x_{r+1} + \ldots + b_nx_n = c_r \end{cases}$$

Its coefficient matrix is

$$\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1r}b_{1,r+1} & \ldots & b_{1n} \\ 0 & b_{22} & \ldots & b_{2r}b_{2,r+1} & \ldots & b_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & b_{rr}b_{r,r+1} & \ldots & b_{rn} \end{pmatrix}$$

System (1) is obtained from (2) by elementary operations. Now, applying the same elementary operations to the matrix **A** of (1) yields the matrix **B**. We know that elementary operations do not change the rank of a matrix, thus, the rank of **A** is equal to the rank of **B**. Since the rank of **B** equals the number of its rows, i.e. r we can write

$$\text{rank } \mathbf{A} = r$$

Reading this equation "from right to left" (r rank $\mathbf{A}$) we may conclude that *if system* (1) *is consistent, then the number r of the equations in an equivalent system* (2) *obtained from* (1) *by Gaussian elimination, equals the rank of the matrix* $\mathbf{A}$ *of* (1). Or alternatively, *if the system is consistent, then the number of free unknowns is $n - r$, where r is the rank of the matrix* $\mathbf{A}$ *of the system.*

11.5. OPERATIONS WITH MATRICES

The study of *operations* with matrices is an important part of matrix theory. *Multiplication of matrices*, which is similar in certain respects to multiplication of numbers, is especially interesting and underlies *matrix algebra* which is frequently used in mathematics and its applications.

1. Addition of matrices.

Suppose $\mathbf{A}$ and $\mathbf{B}$ are two matrices of the same dimension, i.e. having the same number of rows and the same number of columns,

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix}$$

Their sum is the matrix

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{pmatrix}$$

In other words, in order to add two $m \times n$ matrices we should add their corresponding elements. For instance,

$$\begin{pmatrix} 1 & 2 \\ -1 & 3 \\ 4 & -2 \end{pmatrix} + \begin{pmatrix} -1 & -2 \\ 1 & -2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 5 & 1 \end{pmatrix}$$

An $m \times n$ matrix with all its elements being zeros is called an $m \times n$ *zero matrix* and denoted by $\mathbf{0}$ (strictly speaking, we must denote it by $\mathbf{0}_{m,n}$ to show its dimension). Obviously,

$$\mathbf{A} + \mathbf{0} = \mathbf{A}$$

for any matrix $\mathbf{A}$ of the same dimension as $\mathbf{0}$.

2. Multiplication of a matrix by a number.

To multiply a matrix $\mathbf{A}$ by a number k is, by definition, to multiply all its elements by that number:

$$k\mathbf{A} = \begin{pmatrix} ka_{11} & ka_{12} & \dots & ka_{1n} \\ ka_{21} & ka_{22} & \dots & ka_{2n} \\ \dots & \dots & \dots & \dots \\ ka_{m1} & ka_{m2} & \dots & ka_{mn} \end{pmatrix}$$

For instance,

$$3 \cdot \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 3 & -3 \\ 0 & 6 \end{pmatrix}$$

It can easily be proved that

$$k(\mathbf{A} + \mathbf{B}) = k\mathbf{A} + k\mathbf{B}$$

for any two matrices $\mathbf{A}$ and $\mathbf{B}$ of the same dimension.

3. Matrix multiplication.

This is a special operation in which any two matrices $\mathbf{A}$ and $\mathbf{B}$ such that the number of columns in $\mathbf{A}$ equals the number of rows in $\mathbf{B}$ yield a third matrix $\mathbf{C}$.

Suppose we have an $m \times n$ matrix $\mathbf{A}$

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

and an $n \times k$ matrix $\mathbf{B}$

$$\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1k} \\ b_{21} & b_{22} & \dots & b_{2k} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nk} \end{pmatrix}$$

We note that, by definition, the number n of columns in **A** equals the number n of rows in **B**, or in other words, the *length of a row in* **A** *equals the height of a column in* **B**.

Definition. The *product AB* of two matrices **A** and **B** is a matrix **C** whose elements are such that

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$$

The rule for finding c_{ij} can easily be remembered: every element of the ith row of **A** and the corresponding element of the jth column of **B**

$$a_{i1}a_{i2}\ldots a_{in} \quad \begin{matrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{matrix}$$

are multiplied and then the results are added together. Since we have m rows in **A** and k columns in **B**, the double subscripts of c_{ij} will run over i from 1 to m and over j from 1 to k, whence it follows that **C** is an $m \times k$ matrix.

To illustrate, we compile a table showing the numbers of rows and columns in the matrices **A**, **B**, and **AB**:

	A	B	AB
Number of rows	m	n	m
Number of columns	n	k	k

$\left(\text{You can remember this table thus: } \dfrac{m}{n}\cdot\dfrac{n}{k} = \dfrac{m}{k}.\right)$

Example 1. $\begin{pmatrix} 3 & -2 \\ 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 \\ 0 & 2 \end{pmatrix}$

$$= \begin{pmatrix} 3\cdot 0 + (-2)\cdot 0 & 3\cdot(-1) + (-2)\cdot 2 \\ 1\cdot 0 + 4\cdot 0 & 1\cdot(-1) + 4\cdot 2 \end{pmatrix} = \begin{pmatrix} 0 & -7 \\ 0 & 7 \end{pmatrix}.$$

Example 2. $\begin{pmatrix} 1 & 1 & -3 \\ 1 & 2 & -4 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1\cdot 2 + 1\cdot 1 + (-3)\cdot 1 \\ 1\cdot 2 + 2\cdot 1 + (-4)\cdot 1 \end{pmatrix}$

$$= \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Example 3. $\begin{pmatrix} -1 & 0 \\ 2 & 7 \\ 3 & 5 \end{pmatrix}\cdot\begin{pmatrix} 6 & 2 \\ -4 & 2 \end{pmatrix}$

$$= \begin{pmatrix} -1\cdot 6 + 0\cdot(-4) & (-1)\cdot 2 + 0\cdot 2 \\ 2\cdot 6 + 7\cdot(-4) & 2\cdot 2 + 7\cdot 2 \\ 3\cdot 6 + 5\cdot(-4) & 3\cdot 2 + 5\cdot 2 \end{pmatrix} = \begin{pmatrix} -6 & -2 \\ -16 & 18 \\ -2 & 16 \end{pmatrix}$$

Example 4. $\begin{pmatrix} 1 & 1 & -3 \\ 1 & 2 & 4 \end{pmatrix}\cdot\begin{pmatrix} 5 & 1 \\ 7 & 2 \end{pmatrix}$ the product is meaningless since the length of the row in the first matrix differs from the length of the column in the second one.

11.6. PROPERTIES OF MATRIX MULTIPLICATION

1. Noncommutativity.

It is advisable to compare matrix multiplication and number multiplication.

Matrix multiplication is not commutative in contrast to number multiplication. Generally, this means that $\mathbf{AB} \neq \mathbf{BA}$.

To illustrate, we consider the two matrices

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Since the length of a row in **A** and the height of a column in **B** are the same, the product **AB** is true, but the product **BA** is undefined.

There are pairs if matrices **A** and **B** for which both products **AB** and **BA** are defined, but $\mathbf{AB} \neq \mathbf{BA}$. For instance,

$$\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}\cdot\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

while

$$\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}\cdot\begin{pmatrix}0 & 1\\ 0 & 1\end{pmatrix} = \begin{pmatrix}0 & 2\\ 0 & 0\end{pmatrix}$$

2. Associativity.

Matrix multiplication is associative. This means that given three matrices **A**, **B**, *and* **C**, *if one of the products* (**AB**)**C** *and* **A**(**BC**) *exists, then the other also exists and*

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}) \tag{1}$$

□ Suppose **A** and **B** are two matrices with the product **AB**, and **C** is a matrix for which (**AB**)**C** is defined:

	A	B	AB	C	(AB)C
Number of rows	m	n	m	k	m
Number of columns	n	k	k	l	l

Then the product **A**(**BC**) is also defined. Let us prove equation (1).

We denote the matrix **AB** by **P** and **BC** by **Q**. Let us prove that **PC** = **AQ**.

The element of **PC** in the ith row and the jth column is

$$\sum_{\beta=1}^{k} p_{i\beta}c_{\beta j} = \sum_{\beta=1}^{k}\left(\sum_{\alpha=1}^{n} a_{i\alpha}b_{\alpha\beta}\right)c_{\beta j} = \sum a_{i\alpha}b_{\alpha\beta}c_{\beta j} \tag{2}$$

The subscript α on the right-hand side of (2) varies from 1 to n and β varies from 1 to k.

The element of **AQ** in the same row and column is

$$\sum_{\alpha=1}^{n} a_{i\alpha}q_{\alpha j} = \sum_{\alpha=1}^{n} a_{i\alpha}\left(\sum_{\beta=1}^{k} b_{\alpha\beta}c_{\beta j}\right) = \sum a_{i\alpha}b_{\alpha\beta}c_{\beta j}$$

The right-hand side of this equation is the same as that in (2). Thus, the elements at the same places in the matrices **PC** and **AQ** are the same. Hence, **PC** = **AQ**. ■

Because matrix multiplication is associative, we can write **ABC** without specifying which product **(AB)C** or **A(BC)** is meant.

3. Distributivity.

Let us prove that matrix multiplication is distributive with respect to matrix addition. Since matrix multiplication is not commutative, we have

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$$
$$(\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}$$

□ We only need to prove the first equation since the second can be proved by analogy. We assume that **B** and **C** have the same dimension and the number of columns in **A** and the number of rows in **B** are the same.

The element of $\mathbf{A}(\mathbf{B} + \mathbf{C})$ in the ith row and the jth column is

$$a_{i1}(b_{1j} + c_{1j}) + a_{i2}(b_{2j} + c_{2j}) + \ldots + a_{in}(b_{nj} + c_{nj})$$

where n is the number of columns in **A** (which equals, by assumption, the number of rows in **B** and **C**). Removing brackets yields

$$(a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}) + (a_{i1}c_{1j} + a_{i2}c_{2j} + \ldots + a_{in}c_{nj})$$

The first sum is the element in the kth row and jth column of **AB** and the second sum is the analogous element of **AC**. ■

11.7. INVERSE OF A MATRIX

A more profound analogy between matrix multiplication and number multiplication is observed when considering *square* matrices. We assume that all the matrices discussed below are square having the same order n (the number of rows or columns), and this means that **AB** is always defined.

1. Identity matrix.

We know that the number 1 is such that when multiplied by a number a gives a again:

$$a \cdot 1 = 1 \cdot a = a$$

An $n \times n$ matrix called the *identity matrix* with the form

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

has a similar property. It is easy to prove that

$$\mathbf{AI} = \mathbf{IA} = \mathbf{A}$$

whatever the matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

□ In fact, the element of **AI** in the ith row and jth column is

$$a_{i1} \cdot 0 + \dots + a_{i,j-1} \cdot 0 + a_{ij} \cdot 1 + a_{i,j+1} \cdot 0 + \dots + a_{in} \cdot 0 = a_{ij}$$

whence it follows that $\mathbf{AI} = \mathbf{A}$. Similarly, we can prove that $\mathbf{IA} = \mathbf{A}$. ■

2. Inverse matrix.

An important property of the multiplication of numbers is that any nonzero number a has an *inverse number* b such that

$$ab = 1, \quad ba = 1$$

Matrices have a similar property when the condition that the determinant of the matrix **A** is nonzero is substituted for the condition that $a \neq 0$.

Definition 1. A square matrix **A** is said to be *nonsingular* if its determinant is nonzero:

$$|\mathbf{A}| \neq 0$$

and the matrix is said to be *singular* if $|\mathbf{A}| = 0$.

Definition 2. Given an $n \times n$ square matrix **A**. An $n \times n$ matrix **B** is called the *inverse* of **A** if

$$\mathbf{AB} = \mathbf{I}, \quad \mathbf{BA} = \mathbf{I}$$

The inverse of **A** is usually denoted by $\mathbf{A}^{-1}$. Thus,

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}, \quad \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

The following proposition is valid (we shall accept it without proof): *a singular matrix has no inverse.*

3. Finding the inverse of a nonsingular matrix.
Suppose

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

is a nonsingular matrix, i.e.

$$|\mathbf{A}| \neq 0$$

Theorem. *If* **A** *is a nonsingular matrix, then*

$$\mathbf{B} = \frac{1}{|\mathbf{A}|} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}$$

is the inverse of A.

Here A_{ij} is the cofactor of a_{ij} in the determinant $|\mathbf{A}|$.

Note the special arrangement of A_{ij} in the matrix **B**: A_{ij} is in the jth row and ith column and not in the ith row and jth column.

□ We should prove the equations $\mathbf{AB} = \mathbf{I}$ and $\mathbf{BA} = \mathbf{I}$. We start with the product **AB** and putting $\mathbf{AB} = \mathbf{C}$ we write

$$\mathbf{C} = \mathbf{AB} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

$$\times \frac{1}{|\mathbf{A}|} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}$$

By the rule of matrix multiplication, the element in the ith row and jth column in $\mathbf{C}$ is

$$c_{ij} = \frac{1}{|\mathbf{A}|}(a_{i1}A_{j1} + a_{i2}A_{j2} + \ldots + a_{in}A_{jn})$$

The expression equals 1 for $i = j$, since in this case the expression in the parentheses is the expansion of the determinant $|\mathbf{A}|$ by the ith row; when $i \neq j$, the equation is zero, since the expression in parentheses is the sum of the elements of the ith row in $|\mathbf{A}|$ multiplied by the cofactors of the corresponding elements of the jth row (see property 7^0, Sec. 3.6). Thus, c_{ij} is unity if $i = j$ and is zero if $i \neq j$, which implies $\mathbf{C} = \mathbf{I}$. ■

Example 1. Given the matrix

$$\mathbf{A} = \begin{pmatrix} -7 & 2 \\ 4 & -1 \end{pmatrix}$$

verify whether it has the inverse and find $\mathbf{A}^{-1}$.

○ We have

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

consequently, the matrix $\mathbf{A}^{-1}$ exists. In order to determine it, we first find the numbers $A_{11} = -1$, $A_{12} = -4$, $A_{21} = -2$, $A_{22} = -7$. Then using (1) we can write

$$\mathbf{A}^{-1} = -\begin{pmatrix} -1 & -2 \\ -4 & -7 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 4 & 7 \end{pmatrix} \quad \bullet$$

Example 2. Given the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{pmatrix}$$

verify whether it has the inverse and find it.

○ We have

$$\mathbf{A} = \begin{vmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{vmatrix} = -1$$

Since $|\mathbf{A}| \neq 0$, $\mathbf{A}$ has the inverse $\mathbf{A}^{-1}$.

We have

$$A_{11} = \begin{vmatrix} -1 & 0 \\ 2 & 1 \end{vmatrix} = -1, \quad A_{12} = -\begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = -1$$

$$A_{13} = \begin{vmatrix} 1 & -1 \\ -1 & 2 \end{vmatrix} = 1, \quad A_{21} = -\begin{vmatrix} 2 & 3 \\ 2 & 1 \end{vmatrix} = 4$$

$$A_{22} = \begin{vmatrix} 2 & 3 \\ -1 & 1 \end{vmatrix} = 5, \quad A_{23} = -\begin{vmatrix} 2 & 2 \\ -1 & 2 \end{vmatrix} = -6$$

$$A_{31} = \begin{vmatrix} 2 & 3 \\ -1 & 0 \end{vmatrix} = 3, \quad A_{32} = -\begin{vmatrix} 2 & 3 \\ 1 & 0 \end{vmatrix} = 3$$

$$A_{33} = \begin{vmatrix} 2 & 2 \\ 1 & -1 \end{vmatrix} = -4$$

Consequently, the inverse is

$$\mathbf{A}^{-1} = -\begin{pmatrix} -1 & 4 & 3 \\ -1 & 5 & 3 \\ 1 & -6 & -4 \end{pmatrix} = \begin{pmatrix} 1 & -4 & -4 \\ 1 & -5 & -3 \\ -1 & 6 & 4 \end{pmatrix}$$

We advise the reader to prove the equalities $\mathbf{AA}^{-1} = \mathbf{I}$ and $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. ●

11.8. SYSTEMS OF LINEAR EQUATIONS IN MATRIX FORM

The operation of matrix multiplication opens a new insight into a simple, but nevertheless important, problem of linear algebra, namely, solving a system of linear equations.

Suppose we have a system of m linear equations in n unknowns (an $m \times n$ system). For the sake of simplicity, we consider the special case when $m = 2$, $n = 3$, that is, a system of the form

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \end{cases} \tag{1}$$

By introducing the matrices

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

we can write *system* (1) *as a single matrix equation*

$$\mathbf{AX} = \mathbf{B} \tag{2}$$

Indeed, the product **AX** is the matrix with a single column:

$$\mathbf{AX} = \begin{pmatrix} \boxed{a_{11}x_1 + a_{12}x_2 + a_{13}x_3} \\ \boxed{a_{21}x_1 + a_{22}x_2 + a_{23}x_3} \end{pmatrix}$$

whose elements are the right-hand sides of (1). Equating them to the corresponding elements of the matrix **B** yields system (1).

Equation (2) is the *matrix notation of system* (1).

For instance, the matrix notation of the system

$$\begin{cases} -2x_1 + x_2 + 5x_3 = 4 \\ x_1 - 7x_3 = 1 \end{cases}$$

is

$$\begin{pmatrix} -2 & 1 & 5 \\ 1 & 0 & -7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$$

The matrix notation of an $m \times n$ system is still of the form in (2), except that **A** is an $m \times n$ matrix, **B** is the column of m elements, and **X** is the column of n elements.

1. An $n \times n$ system. Notation of a solution in terms of an inverse matrix.

Systems in which the number of equations and unknowns is the same ($n \times n$ systems) are of particular importance. In this case **A** is an $n \times n$ *square matrix*. Suppose **A** is nonsingular, i.e. its determinant is nonzero. Then **A** has the inverse $\mathbf{A}^{-1}$. By using the inverse, we can solve equation (2), namely, multiplying (2) from the left by $\mathbf{A}^{-1}$ yields

$$\mathbf{A}^{-1}(\mathbf{AX}) = \mathbf{A}^{-1}\mathbf{B}$$

or, because matrix multiplication is associative

$$(\mathbf{A}^{-1}\mathbf{A})\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$$

But since $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ and $\mathbf{IX} = \mathbf{X}$, we have

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{B} \tag{3}$$

We cannot yet say that (3) is *equivalent* to (2); we can only assert that any solution $\mathbf{X}$ of (2) also satisfies (3).

In fact, the matrix $\mathbf{X}$ defined by (3) satisfies (2) since

$$\mathbf{A}(\mathbf{A}^{-1}\mathbf{B}) = (\mathbf{A}\mathbf{A}^{-1})\mathbf{B} = \mathbf{IB} = \mathbf{B}$$

Thus, (3) *is the matrix notation of the solution for an* $n \times n$ *system having a nonsingular matrix* $\mathbf{A}$.

Remark. Formula (3) is not so convenient that it greatly simplifies the solution of an $n \times n$ system with a nonsingular matrix $\mathbf{A}$ since we first have to find the matrix $\mathbf{A}^{-1}$, which is in itself difficult. However, (3) is interesting from the viewpoint of theory. Gaussian elimination is still the best method for solving an $n \times n$ system.

Nevertheless, we can illustrate an application of (3).

Example. Solve the system of equations

$$\begin{cases} x_1 \qquad + 3x_3 = 1 \\ 5x_1 + 3x_2 + 7x_3 = 1 \\ 3x_1 + 2x_2 + 5x_3 = 1 \end{cases}$$

○ We have

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 3 \\ 5 & 3 & 7 \\ 3 & 2 & 5 \end{pmatrix}$$

The determinant $|\mathbf{A}|$ is 4, hence $\mathbf{A}$ is nonsingular.

We calculate the inverse $\mathbf{A}^{-1}$ using formula (1) from Sec. 11.6. We have

$$A_{11} = \begin{vmatrix} 3 & 7 \\ 2 & 5 \end{vmatrix} = 1, \quad A_{21} = -\begin{vmatrix} 0 & 3 \\ 2 & 5 \end{vmatrix} = 6$$

$$A_{31} = \begin{vmatrix} 0 & 3 \\ 3 & 7 \end{vmatrix} = -9$$

$$A_{12} = -\begin{vmatrix} 5 & 7 \\ 3 & 5 \end{vmatrix} = -4, \quad A_{22} = \begin{vmatrix} 1 & 3 \\ 3 & 5 \end{vmatrix} = -4$$

$$A_{23} = -\begin{vmatrix} 1 & 0 \\ 3 & 2 \end{vmatrix} = -2$$

$$A_{13} = \begin{vmatrix} 0 & 3 \\ 3 & 7 \end{vmatrix} = -9, \quad A_{32} = -\begin{vmatrix} 1 & 3 \\ 5 & 7 \end{vmatrix} = 8$$

$$A_{33} = \begin{vmatrix} 1 & 0 \\ 5 & 3 \end{vmatrix} = 3$$

$$\mathbf{A}^{-1} = \begin{pmatrix} 1/4 & 6/4 & -9/4 \\ -4/4 & -4/4 & 8/4 \\ 1/4 & -2/4 & 3/4 \end{pmatrix}$$

Then we find the matrix **X** from (3)

$$\mathbf{X} = \begin{pmatrix} 1/4 & 6/4 & -9/4 \\ -1 & -1 & 2 \\ 1/4 & -2/4 & 3/4 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2/4 \\ 0 \\ 2/4 \end{pmatrix}$$

Thus, the solution of the system is $x_1 = -1/2$, $x_2 = 0$, $x_3 = 1/2$. ●

2. Matrix equations.

The unknown matrix **X** in equations of form (2) consists of a single column, however, there are equations of the same form in which **X** consists of several columns. We consider the case when **A**, **B**, and **X** are $n \times n$ square matrices, if **A** is nonsingular, then in order to find **X** we may use the technique applied earlier, namely, multiplying the equation from the left by the matrix $\mathbf{A}^{-1}$. Then we obtain

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$$

The problem is solved.

Note that if the equation is of the form

$$\mathbf{YA} = \mathbf{B}$$

(the unknown matrix $\mathbf{Y}$ is on the left of $\mathbf{A}$), then we obtain

$$\mathbf{Y} = \mathbf{BA}^{-1}$$

Example. Solve the equation

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \mathbf{X} = \begin{pmatrix} 3 & 5 \\ 5 & 9 \end{pmatrix}$$

○ The matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

is nonsingular (its determinant is -2). We find the inverse

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

and then the unknown matrix

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{B} = -\frac{1}{2} \begin{pmatrix} 4 & -2 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} 3 & 5 \\ 5 & 9 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 2 & 3 \end{pmatrix} \quad \bullet$$

Exercises to Chapter 11

1. What can we say about the rows of a matrix if its rank is one or two.
2. Find the ranks of each matrix:

$$(1) \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad (2) \begin{pmatrix} 1 & 0 \\ -1 & 2 \\ 3 & 5 \end{pmatrix}, \quad (3) \begin{pmatrix} 2 & 3 & 4 \\ -1 & 2 & 0 \\ 1 & 1 & 3 \end{pmatrix}, \quad (4) \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 3 & 1 & 1 \\ 1 & 1 & 4 & 1 \\ 1 & 1 & 1 & 5 \end{pmatrix},$$

$$(5) \begin{pmatrix} 2 & -1 & -3 & -2 & 4 \\ 4 & -2 & 2 & 1 & 7 \\ 2 & -1 & 1 & 8 & 2 \end{pmatrix}, \quad (6) \begin{pmatrix} 1 & 2 & 3 & 1 \\ 2 & 3 & 1 & 2 \\ 3 & 1 & 2 & -2 \\ 0 & 4 & 2 & 5 \end{pmatrix},$$

$$(7)\ \begin{pmatrix} 4 & 3 & -1 & 1 & -1 \\ 2 & 1 & -3 & 2 & -5 \\ 1 & -3 & 0 & 1 & -2 \\ 1 & 5 & 2 & -2 & 6 \end{pmatrix}$$

3. Find the rank of the matrix depending on the value of a:

$$(1)\ \begin{pmatrix} 1 & a & -1 & 2 \\ 2 & -1 & a & 5 \\ 1 & 10 & -6 & 1 \end{pmatrix},\quad (2)\ \begin{pmatrix} a & 1 & 1 \\ 1 & a & 1 \\ 1 & 1 & a \end{pmatrix}$$

4. Find the condition under which both products **AB** and **BA** and the product **AA** (which is the square of the matrix **A**) are defined.

5. Find the matrix $\mathbf{A}^2 - 6\mathbf{A}$, where $\mathbf{A} = \begin{pmatrix} 5 & 3 \\ 1 & 2 \end{pmatrix}$.

6. Multiply the matrices in the indicated order:

$$(1)\ \begin{pmatrix} 1 & -2 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix},\quad (2)\ \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 1 & -2 \end{pmatrix},$$

$$(3)\ \begin{pmatrix} -3 & 2 & 0 \\ -4 & 5 & -2 \\ -5 & 3 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & -1 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix},\quad (4)\ \begin{pmatrix} 1 & -2 \\ 4 & 6 \end{pmatrix}\begin{pmatrix} 0 & 7 & 8 \\ 2 & 6 & -3 \end{pmatrix},$$

$$(5)\ (0,\ -1)\begin{pmatrix} 2 \\ 0 \end{pmatrix},\quad (6)\ (1\ \ 3\ \ -2)\begin{pmatrix} 5 & 2 \\ 1 & 7 \\ 1 & 0 \end{pmatrix},\quad (7)\ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}(-3\ \ 4)$$

7. The notation **ABC** means **(AB)C** or , which is the same, **A(BC)**. Find

$$(a)\ \begin{pmatrix} -1 & 3 \\ 4 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -2 & 5 \end{pmatrix}\begin{pmatrix} -1 & 3 \\ 4 & 1 \end{pmatrix},\quad (b)\ (1\ \ 2\ \ -3)\begin{pmatrix} 3 \\ -2 \\ -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix},$$

$$(c)\ \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}(-3\ \ 4)\begin{pmatrix} 0 & -1 \\ 3 & 1 \end{pmatrix}$$

8. The notation $\mathbf{A}^n$ means $\mathbf{A}\cdot\mathbf{A}\ldots\mathbf{A}$ (n times). Find the matrices

$$(1)\ \begin{pmatrix} 4 & -1 \\ 5 & -2 \end{pmatrix}^2,\quad (2)\ \begin{pmatrix} 4 & -1 \\ 5 & -2 \end{pmatrix}^3,\quad (3)\ \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix}^2,$$

$$(4)\ \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix}^3,\quad (5)\ \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix}^n,\quad (6)\ \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n$$

9. Find the change in the matrix

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

when we multiply it from the left by each matrix

$$\text{(a)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \text{(b)} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \text{(c)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \alpha \\ 0 & 1 & 1 \end{pmatrix}, \quad \text{(d)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

10. Find the matrix $\mathbf{A}^{-1}$ if

$$\text{(a) } \mathbf{A} = \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix}, \quad \text{(b) } \mathbf{A} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}, \quad \text{(c) } \mathbf{A} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

11. Find $\mathbf{A}^{-1}$ if

$$\text{(a) } \mathbf{A} = \begin{pmatrix} 0 & 1 & 3 \\ 2 & 3 & 5 \\ 3 & 5 & 7 \end{pmatrix}, \quad \text{(b) } \mathbf{A} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \text{(c) } \mathbf{A} = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

12. Using the inverse matrix solve each system of equations:

$$\text{(a) } \begin{cases} x + 2y + 3z = 7 \\ 2x - y + z = 9 \\ x - 4y + 2z = 11 \end{cases} \qquad \text{(b) } \begin{cases} 2x - y + 3z = 1 \\ 3x - 5y + z = 1 \\ 4x - 7y = 1 \end{cases}$$

13. Solve each matrix equation

$$\text{(1) } \mathbf{X} \cdot \begin{pmatrix} 3 & -2 \\ 5 & -4 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ -5 & 6 \end{pmatrix},$$

$$\text{(2) } \begin{pmatrix} 1 & 2 & -3 \\ 3 & 2 & -4 \\ 2 & -1 & 0 \end{pmatrix} \cdot \mathbf{X} = \begin{pmatrix} 1 & -3 & 0 \\ 10 & 2 & 7 \\ 10 & 7 & 8 \end{pmatrix},$$

$$\text{(3) } \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix} \cdot \mathbf{X} \cdot \begin{pmatrix} -3 & 2 \\ 5 & -3 \end{pmatrix} = \begin{pmatrix} -2 & 4 \\ 3 & -1 \end{pmatrix}$$

Chapter 12

EUCLIDEAN VECTOR SPACES

The concepts of a basis, dimension, and coordinate in terms of a basis were introduced earlier, and are all related to the concept of *linear dependence*. By considering two- and three-dimensional vector spaces (i.e. the sets of all vectors in the plane or in space), we can see that the concept of a vector is considerably extended. For instance, each vector has a *length*, two non-zero vectors form a definite *angle*, and *perpendicular vectors* exist in spaces. This chapter generalizes these concepts (usually called *metric*) to the case of a vector space of any dimension.

12.1. SCALAR PRODUCT. EUCLIDEAN VECTOR SPACES

1. Definition of a scalar product.

We introduced the operation of scalar multiplication of vectors in the ordinary space in Sec. 2.3 by the formula

$$\mathbf{x}\cdot\mathbf{y} = |\mathbf{x}|\cdot|\mathbf{y}|\cdot\cos\varphi$$

where φ is the angle between $\mathbf{x}$ and $\mathbf{y}$. The consequences of this definition are

$$|\mathbf{x}| = \sqrt{\mathbf{x}\cdot\mathbf{x}} \tag{1}$$

$$\cos\varphi = \frac{\mathbf{x}\cdot\mathbf{y}}{\sqrt{\mathbf{x}\cdot\mathbf{x}}\sqrt{\mathbf{y}\cdot\mathbf{y}}} \tag{2}$$

from which we can see that the length of a vector and the angle between two vectors can be expressed via the scalar product. Formulas (1) and (2) suggest a way of generalizing (to an n-dimensional case) the *scalar product*, and so the length of a vector and the angle between two vectors can then be found from formulas (1) and (2).

The scalar product of two vectors in three-dimensional space possesses the following properties (see Sec. 2.3):

1^0. $\mathbf{x}\cdot\mathbf{x} \geqslant 0$, with $\mathbf{x}\cdot\mathbf{x} = 0$ only if $\mathbf{x} = 0$,

2^0. $\mathbf{x}\cdot\mathbf{y} = \mathbf{y}\cdot\mathbf{x}$,

3^0. $(k\mathbf{x})\cdot\mathbf{y} = k(\mathbf{x}\cdot\mathbf{y})$, where k is a scalar,

4^0. $(\mathbf{x} + \mathbf{y})\cdot\mathbf{z} = \mathbf{x}\cdot\mathbf{z} + \mathbf{y}\cdot\mathbf{z}$.

It is advisable that properties 1^0-4^0 remain valid for any generalization of the scalar product.

Definition 1. We say that we can define the *operation of scalar multiplication* (or simply, *scalar multiplication*) in a vector space L if any pair of vectors $\mathbf{x}$ and $\mathbf{y}$ from L can be associated with a number denoted by $\mathbf{x}\cdot\mathbf{y}$ and called the *scalar product* of $\mathbf{x}$ and $\mathbf{y}$, properties 1^0-4^0 being valid.

Definition 2. A vector space L in which scalar multiplication can be defined is called a *Euclidean vector space.*

Properties 1^0-4^0 are often called the *axioms of the scalar product.*

2. Method of a scalar product.

We cannot say from the definition of a scalar product whether we can introduce a scalar product in an n-dimensional vector space. The formula

$$\mathbf{x}\cdot\mathbf{y} = x_1y_1 + x_2y_2 + x_3y_3$$

from Sec. 2.3 suggests an answer; the formula expresses the scalar product of two vectors $\mathbf{x}$ and $\mathbf{y}$ in a three-dimensional space in terms of the Cartesian coordinates of these vectors. Using a similar formula, we can define a scalar multiplication in an n-dimensional vector space L. To do this, we choose a basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ in L and associate a pair of vectors $\mathbf{x}$, $\mathbf{y}$ with a number

$$\mathbf{x}\cdot\mathbf{y} = x_1y_1 + x_2y_2 + \ldots + x_ny_n \tag{3}$$

where $x_1, x_2, \ldots, x_n$ are the coordinates of $\mathbf{x}$ and $y_1, y_2, \ldots, y_n$ are the coordinates of $\mathbf{y}$ in the chosen basis. Then the properties 1^0-4^0 are satisfied.

□ (1) The expression $\mathbf{x}\cdot\mathbf{x} = x_1^2 + x_2^2 + \ldots + x_n^2$ is nonnegative for all values of $x_1, x_2, \ldots, x_n$ and is zero only if $x_1 = 0$, $x_2 = 0, \ldots, x_n = 0$, that is, when $\mathbf{x} = 0$.

(2) That property 2^0 is satisfied is obvious.

(3) $(k\mathbf{x})\cdot\mathbf{y} = (kx_1)\cdot y_1 + (kx_2)\cdot y_2 + \ldots + (kx_n)\cdot y_n = k(x_1y_1 + x_2y_2 + \ldots + x_ny_n) = k(\mathbf{x}\cdot\mathbf{y})$.

(4) $(\mathbf{x} + \mathbf{y})\cdot\mathbf{z} = (x_1 + y_1) z_1 + (x_2 + y_2) z_2 + \ldots + (x_n + y_n) z_n = (x_1z_1 + x_2z_2 + \ldots + x_nz_n) + (y_1z_1 + y_2z_2 + \ldots + y_nz_n) = \mathbf{x}\cdot\mathbf{z} + \mathbf{y}\cdot\mathbf{z}$. ■

Thus, (3) *defines the operation of scalar multiplication in a space L.*

If we choose another basis $\mathbf{e}_1^*, \mathbf{e}_2^*, \ldots, \mathbf{e}_n^*$ and associate a pair of vectors $\mathbf{x}$, $\mathbf{y}$ with a number $(\mathbf{x}\cdot\mathbf{y})^* = x_1^*y_1^* + x_2^*y_2^* + \ldots + x_n^*y_n^*$ (where $x_1^*, \ldots, x_n^*$ and $y_1^*, \ldots, y_n^*$ are the coordinates of $\mathbf{x}$ and $\mathbf{y}$ in the new basis), then, generally speaking, $\mathbf{x}\cdot\mathbf{y} = (\mathbf{x}\cdot\mathbf{y})^*$ is not satisfied. Hence it is clear that the operation of scalar multiplication can be defined in L in many ways.

We shall show later that however *the operation of scalar multiplication is defined in a vector space L, there is always at least one basis such that* (3) *is satisfied.*

3. Properties following from the axioms 1^0-4^0 of a scalar product.

To begin with, we show that two additional properties of axioms 3^0 and 4^0 are valid:

3_1^0. $\mathbf{x}\cdot(k\mathbf{y}) = k(\mathbf{x}\cdot\mathbf{y})$,

4_1^0. $\mathbf{x}\cdot(\mathbf{y} + \mathbf{z}) = \mathbf{x}\cdot\mathbf{y} + \mathbf{x}\cdot\mathbf{z}$.

Property 3_1^0 follows from the chain of equalities

$$\mathbf{x}\cdot(k\mathbf{y}) = (k\mathbf{y})\cdot\mathbf{x} = k(\mathbf{y}\cdot\mathbf{x}) = k(\mathbf{x}\cdot\mathbf{y})$$

each of which uses the axioms of a scalar product. Property 4_1^0 can be proved in a similar manner, viz.,

$$\mathbf{x}\cdot(\mathbf{y} + \mathbf{z}) = (\mathbf{y} + \mathbf{z})\cdot\mathbf{x} = \mathbf{y}\cdot\mathbf{x} + \mathbf{z}\cdot\mathbf{x} = \mathbf{x}\cdot\mathbf{y} + \mathbf{x}\cdot\mathbf{z}$$

By using properties 3^0 and 3_1^0 we can deduce one more property:

$$3_1^0\cdot(k\mathbf{x})\cdot(l\mathbf{y}) = (kl)(\mathbf{x}\cdot\mathbf{y})$$

Further, it follows from properties 4^0 and 4_1^0 that

$$(\mathbf{x}_1 + \mathbf{x}_2 + \ldots + \mathbf{x}_p)\cdot(\mathbf{y}_1 + \mathbf{y}_2 + \ldots + \mathbf{y}_q) = (\mathbf{x}_1, \mathbf{y}_1) + (\mathbf{x}_1, \mathbf{y}_2) + \ldots + (\mathbf{x}_p, \mathbf{y}_q)$$

i.e. the scalar multiplication of a sum by a sum is distributive: each term of the first sum must be multiplied by each term of the second and the results added. This and property 3_2^0 yield a rule for multiplying one linear combination by another:

$$(k_1\mathbf{x}_1 + k_2\mathbf{x}_2 + \ldots + k_p\mathbf{x}_p, l_1\mathbf{y}_1 + l_2\mathbf{y}_2 + \ldots + l_q\mathbf{y}_q) = k_1l_1(\mathbf{x}_1\cdot\mathbf{y}_1) + (k_1l_2)(\mathbf{x}_1\cdot\mathbf{y}_2) + \ldots + k_ql_q(\mathbf{x}_p\cdot\mathbf{y}_q) \qquad (4)$$

For instance

$$(3\mathbf{x}_1 - 2\mathbf{x}_2)\cdot(5\mathbf{y}_1 + 4\mathbf{y}_2) = (3\cdot5)(\mathbf{x}_1\cdot\mathbf{y}_1) + (3\cdot4)(\mathbf{x}_1\cdot\mathbf{y}_2) + ((-2)\cdot5)(\mathbf{x}_2\cdot\mathbf{y}_1) + ((-2)\cdot4)(\mathbf{x}_2\cdot\mathbf{y}_2) = 15(\mathbf{x}_1\cdot\mathbf{y}_1) + 12(\mathbf{x}_1\cdot\mathbf{y}_2) - 10(\mathbf{x}_2\cdot\mathbf{y}_1) - 8(\mathbf{x}_2\cdot\mathbf{y}_2)$$

12.2. SIMPLE METRIC CONCEPTS IN EUCLIDEAN VECTOR SPACES

Having introduced scalar multiplication in a vector space L, we can define such concepts as the *length* or *absolute value* of a vector, the *angle* between two vectors, and the *perpendicularity* of vectors.

Definition 1. The *absolute value* of a vector $\mathbf{x}$ is the number

$$|\mathbf{x}| = \sqrt{\mathbf{x}\cdot\mathbf{x}} \qquad (1)$$

Since the radicand is nonnegative by virtue of axiom 1^0, the root exists.

Definition 2. Suppose $\mathbf{x}$ and $\mathbf{y}$ are two nonzero vectors. The *angle* between $\mathbf{x}$ and $\mathbf{y}$ is the number φ defined by

$$\cos\varphi = \frac{\mathbf{x}\cdot\mathbf{y}}{\sqrt{\mathbf{x}\cdot\mathbf{x}}\sqrt{\mathbf{y}\cdot\mathbf{y}}} \qquad (2)$$

and such that $0 \leqslant \varphi \leqslant \pi$.

The existence of φ will be proved later.

Definition 3. Two vectors **x** and **y** are said to be *perpendicular* or *orthogonal* if their scalar product is zero.

By viritue of (2), two nonzero vectors **x** and **y** are orthogonal if and only if the angle φ between them is $\pi/2$ ($\cos\varphi = 0$).

The *zero vector is orthogonal to any other vector* and this follows from the chain of equalities

$$\mathbf{0}\cdot\mathbf{x} = (0\mathbf{x})\cdot\mathbf{x} = 0(\mathbf{x}\cdot\mathbf{x}) = 0$$

1. Cauchy-Buniakovski's inequality.

Formula (2) for finding the angle between two nonzero vectors **x** and **y** requires additional explanation. The point is that $\cos\varphi$ cannot be any arbitrary number since $\cos\varphi = a$ is only valid if $-1 \leqslant a \leqslant 1$. Thus, the angle φ between two nonzero vectors exists if

$$-1 \leqslant \frac{\mathbf{x}\cdot\mathbf{y}}{|\mathbf{x}|\,|\mathbf{y}|} \leqslant 1$$

or, which is the same,

$$\frac{(\mathbf{x}\cdot\mathbf{y})^2}{|\mathbf{x}|^2|\mathbf{y}|^2} \leqslant 1 \tag{3}$$

Theorem. *The inequality*

$$(\mathbf{x}\cdot\mathbf{y})^2 \leqslant |\mathbf{x}|^2|\mathbf{y}|^2 \tag{4}$$

is valid for any two vectors in a Euclidean vector space L.

Note that (4) is known as *Cauchy-Buniakovski's inequality**. When $\mathbf{x} \neq 0$ and $\mathbf{y} \neq 0$ (4) is equivalent to (3).

□ Consider the scalar product of the vector $\mathbf{y} - t\mathbf{x}$ with itself, where t is any number:

$$(\mathbf{y} - t\mathbf{x})\cdot(\mathbf{y} - t\mathbf{x}) = y\cdot\mathbf{y} - t(\mathbf{x}\cdot\mathbf{y}) - t(\mathbf{y}\cdot\mathbf{x}) + t^2(\mathbf{x}\cdot\mathbf{x})$$

---

* Cauchy A. L. (1789-1857), French mathematician, the author of fundamental works on mathematical analysis

Buniakovski V. J. (1804-1889), Russian mathematician, the author of many works on statistics including the first Russian textbook on probability theory.

We obtain an equation of the form

$$(\mathbf{y} - t\mathbf{x})\cdot(\mathbf{y} - t\mathbf{x}) = \alpha t^2 + 2\beta t + \gamma$$

where $\alpha = \mathbf{x}\cdot\mathbf{x}$, $\beta = -(\mathbf{x}\cdot\mathbf{y})$, $\gamma = \mathbf{y}\cdot\mathbf{y}$. The quadratic polynomial in t on the right-hand side of the equation is nonnegative for any t, since it is the scalar product of a vector with itself. Hence it follows that the discriminant of the quadratic, i.e. $\beta^2 - \alpha\gamma$, is not positive. Consequently,

$$(\mathbf{x}\cdot\mathbf{y})^2 - (\mathbf{x}\cdot\mathbf{x})(\mathbf{y}\cdot\mathbf{y}) \leqslant 0$$

which is (4) in different notation. ■

Thus, we have proved Cauchy-Buniakovski's inequality and thereby validated the definition of the angle between two vectors.

If the scalar product is defined in terms of a basis by $\mathbf{x}\cdot\mathbf{y} = x_1y_1 + x_2y_2 + \ldots + x_ny_n$, then Cauchy-Buniakovski's inequality assumes the form

$$(x_1y_1 + x_2y_2 + \ldots + x_ny_n)^2 \leqslant (x_1^2 + x_2^2 + \ldots + x_n^2)(y_1^2 + y_2^2 + \ldots + y_n^2)$$

which is valid for all x and y.

12.3. ORTHOGONAL SYSTEM OF VECTORS. ORTHOGONAL BASIS

1. Orthogonal system of vectors.

Definition. The system of vectors

$$\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p \tag{1}$$

in a Euclidean vector space L is said to be *orthogonal* if any two vectors of the system are orthogonal, that is,

$$\mathbf{a}_i\cdot\mathbf{a}_j = 0$$

where i and j are any numbers from the set $1, 2, \ldots, p$ and $i \neq j$.

Obviously, any subsystem of an orthogonal system is also orthogonal.

What is then the number of vectors in an orthogonal system? If we do not require that each vector of the system be nonzero, then the system may consist of *any* number of vectors. For instance, the system of p equal vectors $\mathbf{0}, \mathbf{0}, \ldots, \mathbf{0}$ is orthogonal for any p. Therefore, we assume in addition, that all the vectors of the system are nonzero. Given this assumption, the number of vectors must not exceed the dimension of the space L, i.e. n.

This follows from the next Lemma.

Lemma. *An orthogonal system of nonzero vectors is always linearly independent.*

□ Given an orthogonal system of nonzero vectors (1). Suppose they are linearly dependent, i.e. the following equality is true,

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \ldots + c_p\mathbf{a}_p = \mathbf{0}$$

where $c_1, c_2, \ldots, c_p$ are not all zero, say $c_1 \neq 0$. Then scalar multiplication of the last equation by $\mathbf{a}_1$ yields

$$c_1(\mathbf{a}_1 \cdot \mathbf{a}_1) + c_2(\mathbf{a}_2 \cdot \mathbf{a}_1) + \ldots + c_p(\mathbf{a}_p \cdot \mathbf{a}_1) = 0$$

or, since $\mathbf{a}_1$ is orthogonal to every vector $\mathbf{a}_2, \ldots, \mathbf{a}_p$,

$$c_1(\mathbf{a}_1 \cdot \mathbf{a}_1) = 0$$

But $\mathbf{a}_1 \cdot \mathbf{a}_1 \neq 0$ since $\mathbf{a}_1 \neq 0$. Hence $c_1 = 0$. This contradiction proves that system (1) is linearly independent. ■

Since a linearly independent system in an n-dimensional vector space consists of not more than n vectors, the lemma yields the corollary: *A system of mutually orthogonal nonzero vectors in an n-dimensional Euclidean vector space contains no more than n vectors.*

This proposition is obvious for $n = 3$, since we cannot "put" more than three orthogonal nonzero vectors into an ordinary three-dimensional space (Fig. 118).

Figure 118

Is there an orthogonal system of *exactly n* nonzero vectors in an n-dimensional Euclidean vector space L? Since, according to the lemma such a system is necessarily a basis of the space, we may put an alternative question: Is there an orthogonal basis in L? The answer is given in the next subsection.

2. Existence of an orthogonal basis.

Lemma. *Suppose we have a system of k vectors in an n-dimensional Euclidean vector space L, $k < n$, then there exists a nonzero vector orthogonal to all the vectors of the system.*

□ We denote the vectors of the system by $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ and the unknown nonzero vector by $\mathbf{x}$. The vector $\mathbf{x} \neq 0$ and is such that

$$\mathbf{a}_1 \cdot \mathbf{x} = 0, \quad \mathbf{a}_2 \cdot \mathbf{x} = 0, \quad \ldots, \quad \mathbf{a}_k \cdot \mathbf{x} = 0 \tag{2}$$

We choose a basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ in L and assume that

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n$$

Then conditions in (2) take the form

$$\begin{gathered} \mathbf{a}_1 \cdot (x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n) = 0 \\ \mathbf{a}_2 \cdot (x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n) = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \mathbf{a}_k \cdot (x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n) = 0 \end{gathered}$$

Expanding each scalar product and designating $\alpha_{ij} = \mathbf{a}_i \cdot \mathbf{e}_j$ yields

$$\begin{cases} \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1n}x_n = 0 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{2n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{k1}x_1 + \alpha_{k2}x_2 + \ldots + \alpha_{kn}x_n = 0 \end{cases}$$

This is a system of k homogeneous equations in n unknowns $x_1, x_2, \ldots, x_n$. Since $k < n$, the system must have a nontrivial solution (see the theorem in Sec. 3.9). Consequently, there is a nonzero vector $\mathbf{x}$ satisfying (2). ■

Theorem. *There is always a basis of mutually orthogonal vectors in a Euclidean vector space L* (an orthogonal basis).

□ Let L be n-dimensional. In order to construct an orthogonal basis, we take an arbitrary nonzero vector $\mathbf{a}_1$, then a nonzero vector $\mathbf{a}_2$ orthogonal to $\mathbf{a}_1$, then a nonzero vector $\mathbf{a}_3$ orthogonal to all the constructed vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{n-1}$. The existence of each next vector is ensured by the previous lemma. The resulting system $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ is orthogonal consisting of n nonzero vectors and, consequently, is a basis of L. ■

12.4. ORTHONORMAL BASIS

1. Definition of an orthonormal basis.

Definition 1. A vector $\mathbf{e}$ is said to be *normalized* or a *unit* vector if its length is 1.

Multiplying a nonnormalized vector $\mathbf{a}$ ($\mathbf{a} \neq 0$) by a number $\frac{1}{|\mathbf{a}|}$ yields the normalized vector; indeed, a scalar product of the vector $\frac{1}{|\mathbf{a}|}\mathbf{a}$ with itself is

$$\left(\frac{1}{|\mathbf{a}|}\mathbf{a}\right) \cdot \left(\frac{1}{|\mathbf{a}|}\mathbf{a}\right) = \frac{1}{|\mathbf{a}|^2}(\mathbf{a} \cdot \mathbf{a}) = 1$$

Transition from the vector $\mathbf{a}$ to $\frac{1}{|\mathbf{a}|}\mathbf{a}$ is called the *normalization* of $\mathbf{a}$.

Definition 2. Given an n-dimensional Euclidean vector space L, the basis

$$\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n \tag{1}$$

of L is said to be *orthonormal* if

(1) the basis is orthogonal, that is,

$$\mathbf{e}_i \cdot \mathbf{e}_j = 0 \quad \text{for} \quad i \neq j \tag{2}$$

(2) each basis vector is normalized, that is,

$$\mathbf{e}_i \cdot \mathbf{e}_j = 1 \tag{3}$$

An orthonormal basis always exists. In order to form such a basis, it is sufficient to take any orthogonal basis $\mathbf{a}_1, \mathbf{a}_2, \ldots,$

$\mathbf{a}_n$ (the existence of an orthogonal basis was proved in the previous section) and then normalize each vector of the basis

$$\mathbf{e}_1 = \frac{1}{|\mathbf{a}_1|}\mathbf{a}_1, \quad \mathbf{e}_2 = \frac{1}{|\mathbf{a}_2|}\mathbf{a}_2, \quad \ldots, \quad \mathbf{e}_n = \frac{1}{|\mathbf{a}_n|}\mathbf{a}_n$$

2. Scalar product in an orthonormal basis.

Suppose we have an orthonormal basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. Performing scalar multiplication of two arbitrary vectors $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n$ and $\mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + \ldots + y_n\mathbf{e}_n$ yields

$$\mathbf{x}\cdot\mathbf{y} = (x_1\mathbf{e}_1)\cdot\mathbf{y} + (x_2\mathbf{e}_2)\cdot\mathbf{y} + \ldots + (x_n\mathbf{e}_n)\cdot\mathbf{y} \qquad (4)$$

The first term of the right-hand side equals

$$(x_1\mathbf{e}_1)\cdot(y_1\mathbf{e}_1) + (x_1\mathbf{e}_1)\cdot(y_2\mathbf{e}_2) + \ldots + (x_1\mathbf{e}_1)\cdot(y_n\mathbf{e}_n)$$
$$= (x_1y_1)(\mathbf{e}_1\cdot\mathbf{e}_1) + (x_1y_2)(\mathbf{e}_1\cdot\mathbf{e}_2) + \ldots + (x_1y_n)(\mathbf{e}_1\cdot\mathbf{e}_n)$$
$$= x_1y_1\cdot 1 + x_1y_2\cdot 0 + \ldots + x_1y_n\cdot 0 = x_1y_1$$

Similarly, we find that the second term is x_2y_2, and so on. The result is

$$\mathbf{x}\cdot\mathbf{y} = x_1y_1 + x_2y_2 + \ldots + x_ny_n$$

Thus, the *scalar product of two vectors in an orthonormal basis equals the sum of the products of their respective coordinates.*

Exercises to Chapter 12

1. Explain why it is impossible to introduce a scalar product in an n-dimensional vector space, with $n > 1$, using the formula

$$\mathbf{x}\cdot\mathbf{y} = x_1y_1$$

where x_1 and y_1 are the first coordinates of $\mathbf{x}$ and $\mathbf{y}$ in the basis. Which axiom of 1^0-4^0 of the scalar product appears to be violated?

2. Prove the inequality

$$(x_1^2 + x_2^2)(y_1^2 + y_2^2) \geqslant (x_1y_1 + x_2y_2)^2$$

(a particular case of Cauchy-Buniakovski's inequality for $n = 2$) indirectly by considering the difference between the left-hand and right-hand sides.

3. Prove that the previous inequality becomes an equality if and only if $x_1 : y_1 = x_2 : y_2$, that is, when the vectors $\langle \mathbf{x}_1, \mathbf{x}_2 \rangle$ and $\langle \mathbf{y}_1, \mathbf{y}_2 \rangle$ are collinear.

4. Let a scalar product be defined in $\mathbf{R}^3$ by

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + x_3 y_3$$

Employing the method given in Sec. 12.3, form an orthogonal basis with the first basis vector $\langle 1, 1, 1 \rangle$.

4. Suppose a scalar product in $\mathbf{R}^4$ is defined by

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + x_3 y_3 + x_4 y_4$$

(a) Verify the orthogonality of the vectors $\mathbf{a}_1 = \langle 1, 1, 1, 1 \rangle$ and $\mathbf{a}_2 = \langle 1, 1, -1, -1 \rangle$. Construct an orthogonal basis which includes $\mathbf{a}_1$ and $\mathbf{a}_2$. Can we choose the other two vectors so that their coordinates are $+1$ or -1?

(b) Find and normalize a nonzero vector orthogonal to each of the vectors $\langle 1, 1, 1, 1 \rangle$, $\langle 1, -1, -1, 1 \rangle$, $\langle 2, 1, 1, 3 \rangle$.

Chapter 13

AFFINE SPACES. CONVEX SETS AND POLYHEDRONS

13.1 THE AFFINE SPACE $\mathbf{A}^n$

We noted that an arithmetic vector space $\mathbf{R}^n$ is an important object studied in linear algebra. An *arithmetic affine space* $\mathbf{A}^n$ is closely related to $\mathbf{R}^n$.

The relation between $\mathbf{A}^n$ and $\mathbf{R}^n$ is the same as between the set of all points and the set of all vectors in an ordinary three-dimensional space. Each point in an ordinary space is defined (in a given Cartesian coordinate system) by a triple of numbers and a pair of points defines a vector. By analogy, the points of $\mathbf{A}^n$ are defined by sets of n numbers and each pair of points is associated with a vector from $\mathbf{R}^n$.

Definition 1. Any set of n numbers

$$a_1, a_2, \ldots, a_n \tag{1}$$

is called an *arithmetic point* and the numbers $a_1, a_2, \ldots, a_n$ are called the *coordinates* of that point.

To stress that (1) is considered to be an arithmetic point (and not an arithmetic vector) we write it in parentheses:

$$(a_1, a_2, \ldots, a_n) \tag{2}$$

We denote arithmetic points by upper-case letters A, B, C, and so on. For instance, we denote point (2) by A and write $A = (a_1, a_2, \ldots, a_n)$.

Definition 2. Suppose A and B are two arithmetic points having the same number of coordinates n, i.e. $A = (a_1, a_2, \ldots, a_n)$ and $B = (b_1, b_2, \ldots, b_n)$. The arithmetic vector $\langle b_1 - a_1, b_2 - a_2, \ldots, b_n - a_n\rangle$ from $\mathbf{R}^n$ is called the *vector* $\overrightarrow{AB}$ and A is its *initial point* and B is its *terminal point.*

This definition implies that the *coordinates of* $\overrightarrow{AB}$ *are the* *differences of the corresponding coordinates of the initial and terminal points.*

The point (0, 0, 0) is called the *origin* and denoted by O. Obviously, whatever the point A, the coordinates of $\overrightarrow{OA}$ coincide with the coordinates of A. The vector $\overrightarrow{OA}$ is the *radius vector* of A.

Definition 3. The set of all arithmetic points having n coordinates in which each pair of points A and B is associated (by the technique discussed) with a vector $\overrightarrow{AB}$ from $\mathbf{R}^n$ is called an *n-dimensional arithmetic affine space* and denoted by $\mathbf{A}^n$.

We shall frequently omit the word "arithmetic" in our later discussion.

Theorem. *The equality*

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

is valid whatever the points A, B, C from $\mathbf{A}^n$.

□ Let $A = (a_1, a_2, \ldots, a_n)$, $B = (b_1, b_2, \ldots, b_n)$, and $C = (c_1, c_2, \ldots, c_n)$. Then

$$\overrightarrow{AB} = \langle b_1 - a_1, b_2 - a_2, \ldots, b_n - a_n \rangle$$
$$\overrightarrow{BC} = \langle c_1 - b_1, c_2 - b_2, \ldots, c_n - b_n \rangle$$

whence it follows that

$$\begin{aligned}\overrightarrow{AB} + \overrightarrow{BC} &= \langle (b_1 - a_1) + (c_1 - b_1), (b_2 - a_2) + (c_2 - b_2), \\ &\quad \ldots, (b_n - a_n) + (c_n - b_n) \rangle \\ &= \langle c_1 - a_1, c_2 - a_2, \ldots, c_n - a_n \rangle = \overrightarrow{AC} \quad \blacksquare\end{aligned}$$

Laying off a vector. This is an important concept in the geometry of $\mathbf{A}^n$ or *affine* geometry.

Definition 4. Let $A = (a_1, a_2, \ldots, a_n)$ be a point from $\mathbf{A}^n$ and $\mathbf{p} = \langle p_1, p_2, \ldots, p_n \rangle$ be a vector from $\mathbf{R}^n$. *To lay off a vector* $\mathbf{p}$ *from a point A is to find a point B such that*

$$\overrightarrow{AB} = \mathbf{p} \tag{3}$$

Since $\overrightarrow{OB} = \overrightarrow{OA} + \overrightarrow{AB}$, (3) is equivalent to the equation $\overrightarrow{OB} = \overrightarrow{OA} + \mathbf{p}$. Consequently, denoting the coordinates of B by $b_1, b_2, \ldots, b_n$ yields

$$b_1 = a_1 + p_1, \quad b_2 = a_2 + p_2, \quad \ldots, \quad b_n = a_n + p_n \tag{4}$$

which means that the *coordinates of B are the sums of the coordinates of A and the corresponding coordinates of* **p**.

Notation. We write the point B obtained by laying off a vector **p** from the point **A** as $A + \mathbf{p}$.

According to (4),

$$A + \mathbf{p} = (a_1 + p_1,\ a_2 + p_2,\ \ldots,\ a_n + p_n)$$

13.2. SIMPLE GEOMETRIC FIGURES IN $\mathbf{A}^n$

1. Straight line

A straight line in an ordinary space can be defined by a point and a direction vector. The following definition is the generalization of this fact.

Definition. Let X_0 be a fixed point in $\mathbf{A}^n$ and **p** be a fixed vector from $\mathbf{R}^n$. The set of points X of the form

$$X = X_0 + t\mathbf{p} \qquad (1)$$

with t a number, is called a *straight line through* X_0 *in the direction of* **p**, or simply, a *straight line.*

If $X_0 = (x_1^0,\ x_2^0,\ \ldots,\ x_n^0)$ and $\mathbf{p} = \langle p_1,\ p_2,\ \ldots,\ p_n \rangle$, then $x_1^0 + tp_1,\ x_2^0 + tp_2,\ \ldots,\ x_n^0 + tp_n$ are the coordinates of point (1). Therefore, we can say that the *parametric equations of a straight line through* X_0 *in the direction of* **p** *have the form*

$$x_1 = x_1^0 + tp_1,\ x_2 = x_2^0 + tp_2,\ \ldots,\ x_n = x_n^0 + tp_n$$

2. Line segment.

Definition. Let X_0 and X_1 be two points in $\mathbf{A}^n$. The set of points X of the form

$$X = X_0 + t\overrightarrow{X_0X_1} \qquad (2)$$

with t varying over the interval $0 \leqslant t \leqslant 1$ is called a *line segment* X_0X_1.

The points X_0 and X_1 are called the *end points* of the line segment X_0X_1. Obviously X_0 corresponds to $t = 0$ and X_1 to $t = 1$.

Theorem (on a line segment). *A line segment* $\overrightarrow{X_0X_1}$ *in* $\mathbf{A}^n$

consists of the points X such that

$$\overrightarrow{OX} = s\overrightarrow{OX_0} + (1 - s)\ OX_1$$

where s varies within the limits $0 \leqslant s \leqslant 1$.

□ It follows from (2) that

$$\overrightarrow{OX} = OX_0 + t\overrightarrow{X_0X_1} = \overrightarrow{OX_0} + t(\overrightarrow{OX_1} - \overrightarrow{OX_0})$$
$$= (1 - t)\ \overrightarrow{OX_0} + t\overrightarrow{OX_1}$$

Putting $1 - t = s$ yields the statement of the theorem. ■

3. Different kinds of plane in $\mathbf{A}^n$.

We can introduce the concept of a *plane*, and not only a straight line, in the space $\mathbf{A}^n$. However, the dimension n may be larger than three. The point is that there are different kinds of plane in $\mathbf{A}^n$, where $n > 3$: two-dimensional, three-dimensional, and finally $(n - 1)$-dimensional.

Definition. Let k be a real number. *A plane of dimension k* (or a *k-dimensional plane*) in $\mathbf{A}^n$ is the set of points X of the form

$$X = X_0 + t_1\mathbf{p}_1 + t_2\mathbf{p}_2 + \ldots + t_k\mathbf{p}_k$$

where X_0 is a fixed point, $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_k$ are fixed linearly independent vectors from $\mathbf{R}^n$, and $t_1, t_2, \ldots, t_k$ are arbitrary numbers.

The definition may include the condition that k must be *smaller than n*. Indeed, there are no linearly independent systems of vectors whose number exceeds n in $\mathbf{R}^n$. The case when $k = n$ is of no interest since a linearly independent system of n vectors is a basis of $\mathbf{R}^n$, and, therefore, an n-dimensional plane is the set of points of the form

$$X_0 + \mathbf{p}$$

where $\mathbf{p}$ is any vector from $\mathbf{R}^n$; in other words, an n-dimensional plane is the entire space $\mathbf{A}^n$.

Two kinds of plane in $\mathbf{A}^n$ are most interesting, namely, planes of minimal possible dimension 1 and planes of maximal possible dimension $n - 1$. One-dimensional planes are, essentially, *straight lines* defined earlier. Planes of dimension $n - 1$ are

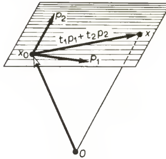


Figure 119

called *hyperplanes*. When $n = 3$, a hyperplane is the set of points of the form

$$X_0 + t_1\mathbf{p}_1 + t_2\mathbf{p}_2$$

where $\mathbf{p}_1$ and $\mathbf{p}_2$ are linearly independent vectors from $\mathbf{R}^3$. Whence it follows that for $n = 3$ the concept of a hyperplane coincides with that of a plane in a three-dimensional space (Fig. 119).

An ordinary plane is defined by an equation of the form

$$ax + by + cz + d = 0$$

Generalizing this fact, we shall give, in the next subsection, another, independent definition of a hyperplane and use it in what follows.

4. Hyperplanes.

Definition. A *hyperplane* in $\mathbf{A}^n$ is the set of numbers $X = (x_1, x_2, \ldots, x_n)$ whose coordinates satisfy the linear equation

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n + b = 0 \tag{3}$$

where $a_1, a_2, \ldots, a_n$, b are fixed numbers and $a_1, a_2, \ldots, a_n$ are not all zero.

Both definitions of a hyperplane are equivalent in that we define it as a set determined by an equation of form (3) and as an $(n - 1)$-dimensional plane, i.e. the set of points X of the

form

$$X = X_0 + t_1\mathbf{p}_1 + t_2\mathbf{p}_2 + \ldots + t_{n-1}\mathbf{p}_{n-1}$$

where $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_{n-1}$ are fixed linearly independent vectors from $\mathbf{R}^n$, and the equivalence can be proved.

5. Half-spaces.

If a plane in an ordinary space is defined by the equation

$$ax + by + cz + d = 0$$

then each of two half-spaces into which the plane divides the space is defined by the respective inequality, namely,

$$ax + by + cz = d \geqslant 0$$

and

$$ax + by + cz + d \leqslant 0$$

Generalizing this fact we make the following definition.

Definition. Given a hyperplane

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n + b = 0$$

in $\mathbf{A}^n$, the following two sets are called *half-spaces* into which this hyperplane divides $\mathbf{A}^n$:

(1) the set of all points $X = (x_1, x_2, \ldots, x_n)$ such that

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n + b \geqslant 0$$

(2) the set of all points $X = (x_1, x_2, \ldots, x_n)$ such that

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n + b \leqslant 0$$

13.3. CONVEX SETS OF POINTS IN $\mathbf{A}^n$. CONVEX POLYHEDRONS

1. Convex sets in the plane and in three-dimensional space.

Definition. The set M of points in the plane and in space is said to be *convex* if it contains the line segment AB joining its two points A and B.

Figure 120 illustrates convex and concave sets.

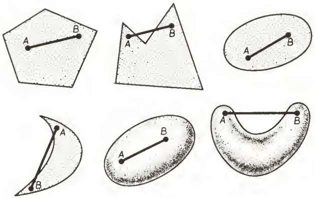


Figure 120

2. Convex sets in $\mathbf{A}^n$.

Definition. The set M of points in $\mathbf{A}^n$ is said to be *convex* if it contains all the points of the line segment AB joining its points A and B.

Theorem. *The hyperplane as well as any of its half-spaces into which it divides* $\mathbf{A}^n$ *are convex sets.*

□ We consider a half-space defined by

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n + b \geqslant 0 \quad (1)$$

We shall prove that it is convex. Suppose the points $X' = (x_1', x_2', \ldots, x_n')$ and $X'' = (x_1'', x_2'', \ldots, x_n'')$ belong to the indicated half-space, i.e.

$$\begin{aligned} a_1x_1' + a_2x_2' + \ldots + a_nx_n' + b \geqslant 0 \\ a_1x_1'' + a_2x_2'' + \ldots + a_nx_n'' + b \geqslant 0 \end{aligned} \quad (2)$$

Let us verify that any point X of the line segment $X'X''$ also belongs to that half-space.

According to the theorem on a line segment (see Sec. 13.2), the coordinates of X can be represented as

$$\begin{aligned} x_1 &= sx_1' + (1 - s)\, x_1'' \\ x_2 &= sx_2' + (1 - s)\, x_2'' \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n &= sx_n' + (1 - s)\, x_n'' \end{aligned}$$

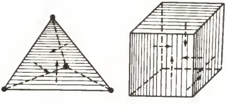


Figure 121

where $0 \leqslant s \leqslant 1$. Substituting these expressions into the left-hand side of (1) yields

$$a_1(sx_1' + (1 - s)\,x_1'') + a_2(sx_2' + (1 - s)\,x_2'') + \ldots + a_n(sx_n' + (1 - s)\,x_n'') + b$$
$$= s(a_1x_1' + a_2x_2' + \ldots + a_nx_n' + b) + (1 - s)(a_1x_1'' + a_2x_2'' + \ldots + a_nx_n'' + b)$$

(note that $sb + (1 - s)\,b = b$). The expression on the right-hand side of this equation is nonnegative on the basis of (2) and since $s \geqslant 0$, $1 - s \geqslant 0$.

Thus, we proved that the half-space is convex. The same can be proved by analogy for a hyperplane. ■

3. Convex polyhedral domains. Convex polyhedrons.

A convex polyhedron in an ordinary space can be visualized as an intersection of several half-spaces. For instance, we can see from Figure 121 that a tetrahedron is the intersection of four half-spaces and a cube is the intersection of six half-spaces.

However, the converse is not always true: the intersection of several half-spaces can be a convex polyhedron. First, the intersection may be empty (Fig. 122) or it can be an unbounded figure (the infinite trihedral angle in Fig. 123). Given the last remark, we introduce the following definitions.

Definition 1. The set of points in $\mathbf{A}^n$ is called a *convex polyhedral domain* if it can be represented as the intersection of several half-spaces.

Definition 2. A convex polyhedral domain in $\mathbf{A}^n$, which is

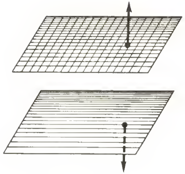


Figure 122

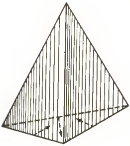


Figure 123

at the same time a bounded set in $\mathbf{A}^n$, is called a *convex polyhedron*.

In a bounded set the coordinates of all its points are less in absolute value than a constant C: $|x_1| \leqslant C$, $|x_2| \leqslant C$, ..., $|x_n| \leqslant C$.

Thus, a convex polyhedral domain is defined by a system of linear inequalities

$$\begin{cases} a_1x_1 + a_2x_2 + \ldots + a_nx_n + b \geqslant 0 \\ a_1'x_1 + a_2'x_2 + \ldots + a_n'x_n + b' \geqslant 0 \end{cases}$$

The number of inequalities in the system can be arbitrary.

Example. Find the convex polyhedral domain in the plane defined by the system

$$\begin{cases} x + y + 1 \geqslant 0 \\ x - 2y - 2 \geqslant 0 \\ 2x - y - 4 \geqslant 0 \end{cases}$$

(In the case of a plane, we should say a polygonal and not "polyhedral" domain.)

● We can write the first inequality of the system as $y \geqslant -x - 1$; it defines the half-plane above its boundary line l_1 (Fig. 124). The second inequality is equivalent to $y \leqslant (1/2)x - 1$ and defines the half-plane below the boundary

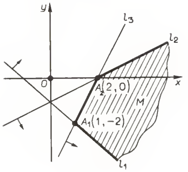


Figure 124

line l_2. The third inequality $y \leqslant 2x - 4$ defines the half-plane below its boundary line l_3. The intersection of these three half-planes is a convex polyhedral domain M shown in Fig. 124; it is unbounded and has two vertices: $A_1(1, -2)$ and $A_2(2, 0)$. ●

Exercises to Chapter 13

1. Prove that if A, B, C, D are four points in a space $\mathbf{A}^n$, with $\overrightarrow{AB} = \overrightarrow{CD}$, then $\overrightarrow{AC} = \overrightarrow{BD}$.

2. Prove that $(A + \mathbf{p}) + \mathbf{q} = A + (\mathbf{p} + \mathbf{q})$, where A is any point from $\mathbf{A}^n$ and $\mathbf{p}$ and $\mathbf{q}$ are vectors from $\mathbf{R}^n$.

3. Prove that two distinct straight lines in space $\mathbf{A}^n$ can only have a single point in common.

4. Prove that three points A, B, C in a space $\mathbf{A}^n$ lie on the same straight line if and only if the vectors $\overrightarrow{AB}$ and $\overrightarrow{AC}$ are collinear.

5. A point X in $\mathbf{A}^n$ divides the line segment X_1X_2 in the ratio $\alpha:\beta$ if $\overrightarrow{X_1X} = (\alpha/\beta)\overrightarrow{XX_2}$. Deduce the formula

$$\overrightarrow{OX} = \frac{\beta}{\alpha + \beta}\,\overrightarrow{OX_1} + \frac{\alpha}{\alpha + \beta}\,\overrightarrow{OX_2}$$

for the radius vector of point X dividing the line segment in the ratio $\alpha:\beta$.

6. Prove that any three noncollinear points A, B, C in space $\mathbf{A}^n$ belong to a two-dimensional plane and that this plane is unique.

7. Find which of the indicated domains on the plane are convex: (1) $y \geqslant x^2$, (2) $y < x^2$, (3) $x^2 + y^2 < 1$, (4) $x^2 + 2y^2 \leqslant 3$, (5) $\begin{cases} x^2 - y^2 \geqslant 1, \\ x \geqslant 1, \end{cases}$ (6) $x^2 - y^2 \leqslant 1$.

8. Let M be a domain in the plane and $\bar{M}$ be its complement (to a complete plane). Can it be that the domains M and $\bar{M}$ are convex?

9. Find the convex polygonal domains in the plane defined by each system of inequalities:

$$\text{(a)} \begin{cases} x - y + 1 \geqslant 0 \\ 3x + 2y - 6 \geqslant 0 \\ -3x - y + 9 \geqslant 0 \end{cases} \quad \text{(b)} \begin{cases} x + y - 1 \geqslant 0 \\ x - y + 1 \geqslant 0 \\ -x + y + 1 \geqslant 0 \\ 2x + y \geqslant 0 \end{cases}$$

$$\text{(c)} \begin{cases} 2x - y + 7 \leqslant 0 \\ -4x + y + 11 \geqslant 0 \\ -x + 3y - 7 \leqslant 0 \\ x - 2 \leqslant 0 \\ x - y - 5 \leqslant 0 \end{cases}$$

Which of the domains are convex polyhedrons?

10. Suppose a convex polyhedral domain M is defined in the plane by

$$\begin{cases} a_1x + b_1y + c_1 \geqslant 0 \\ a_2x + b_2y + c_2 \geqslant 0 \\ \cdots\cdots\cdots\cdots \end{cases}$$

Prove that if M contains a ray with the direction vector $\bar{p} = \langle X, Y\rangle$, then

$$\begin{cases} a_1X + b_1Y \geqslant 0 \\ a_2X + b_2Y \geqslant 0 \\ \cdots\cdots\cdots \end{cases}$$

Formulate a similar problem for a three-dimensional space.

INDEX

TO THE READER

Mir Publishers would be grateful for your comments on the content, translation and design of this book.

We would also be pleased to receive any other suggestions you may wish to make.

Our address is:
Mir Publishers
2 Pervy Rizhsky Pereulok
I-110, GSP, Moscow, 129820
USSR

Other Mir Titles

MATHEMATICS PROBLEMS

ALGEBRA

V. V. Vavilov, I. I. Melnikov,
S. N. Olekhnik, P. I. Pasichenko

This is a study aid which contains the fundamental methods of solving problems in algebra. The exposition of the methods is followed by theoretical data and worked examples. Every theoretical subject is supplied with assignments and exercises for a thorough mastering of the material.

The book is founded on the course of mathematics taught at the preparatory faculty of Moscow University and is closely associated with *Mathematics Problems. Equations and Inequalities* by the same authors.

It is intended for high-school students and for those entering colleges.

MATHEMATICS PROBLEMS

EQUATIONS AND INEQUALITIES

V. V. Vavilov, I. I. Melnikov,
S. N. Olekhnik, P. I. Pasichenko

The book contains information concerning the methods of solving equations and inequalities in one unknown, those containing the sign of the absolute value, and also irrational, exponential and logarithmic equations and inequalities. It includes problems often set at entrance examinations. The methods are illustrated with examples.

It is closely associated with *Mathematics Problems. Algebra* by the same authors.

This text is intended for high-school students and for those entering colleges.

FUNCTIONS OF SEVERAL VARIABLES

V. F. Butuzov, N. Ch. Krutitskaya,
G. N. Medvedev, A. A. Shishkin

This study aid is the continuation of *Mathematical Analysis in Questions and Problems* by the same authors (Mir Publishers, Moscow, 1989), which was devoted to principal problems of the analysis of functions of one variable. This book deals with the differential and integral calculus of functions of several variables. Each section of the book is supplied with problems and exercises for independent work. The book is intended for college students.

GEOMETRY

A. Aleksandrov and N. Netsvetaev

This study aid presents the main sections of the geometry course: analytic geometry, elementary geometry based on axiomatics including geometric transformations and constructions, elements of multidimensional and projective geometry and topology.
It is intended for mathematics students of teacher-training colleges.

Linear Algebra with Elements of Analytic Geometry

This study aid follows the course on linear algebra with elementary analytic geometry and is intended for technical school students specializing in applied mathematics. The text deals with the elements of analytic geometry, the theory of matrices and determinants, systems of linear equations, vectors, and Euclidean spaces. The material is presented in an informal manner. Many interesting examples will help the reader to grasp the material easily.

ISBN 5-03-001789-5