ALGEBRA and GEOMETRY

Nikolai Yu. Zolotykh

Faculty of Computing Math. and Cybernatics,

Faculty of Foreign Students,

Nizhni Novgorod State University

2006, Fall; 2007, Spring

Contents

Chapter 1

Vector Geometry

A *vector* is a pair (ordered) of two points (in a plane or in space).

A vector v is denoted as follows: $v = \overrightarrow{AB}$, where

- the point A is the *start*, or *initial point*, of the vector v, and
- the point B is the *end*, or *final point*, of the vector v.

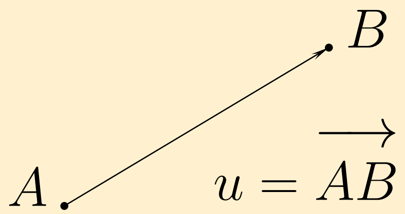
It is represented as an arrow with the *tail* A and the *head* B.

A *vector* is a pair (ordered) of two points (in a plane or in the space).

A vector v is denoted as follows: $v = \overrightarrow{AB}$, where

- the point A is the *start*, or *initial point*, of the vector v, and
- the point B is the *end*, or *final point*, of the vector v.

It is represented as an arrow with the *tail* A and the *head* B.

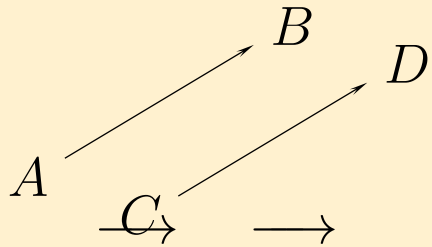


A vector has a *direction* and a *magnitude* (*length* or *modulus*).

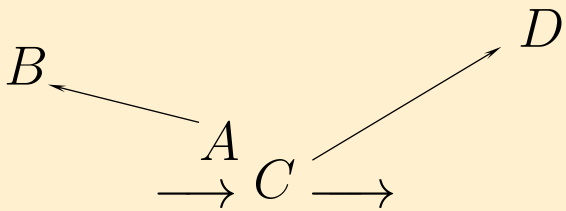
The magnitude of the vector $v = \overrightarrow{AB}$ is the distance between A and B. It is denoted as $|v| = |\overrightarrow{AB}|$ or, simply, $|AB|$.

A *zero* vector, denoted o, is a vector of length 0. Its direction is undefined.

Two vectors u and v are called *equal* if they have the same direction and the same magnitude (regardless of whether they have the same initial points).

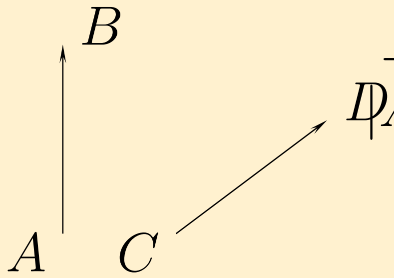


$$\overrightarrow{AB} = \overrightarrow{CD}$$



$$\overrightarrow{AB} \neq \overrightarrow{CD}$$

Vectors can have the same magnitude but different direction. They are not equal.



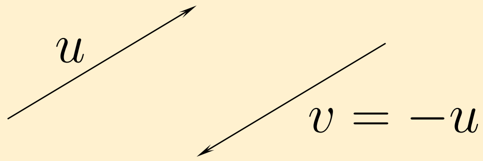
$$|\overrightarrow{AB}| = |\overrightarrow{CD}|$$
$$\overrightarrow{AB} \neq \overrightarrow{CD}$$

Parrallel vectors are also called *collinear*

Zero vector o is the unique vector which is collinear to any vector.

Two vectors u and v are called *opposite* if they are collinear, have opposite direction and have the same magnitude.

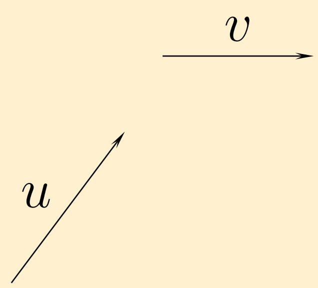
If v is opposite to u then we'll write $v = -u$ (and $u = -v$)



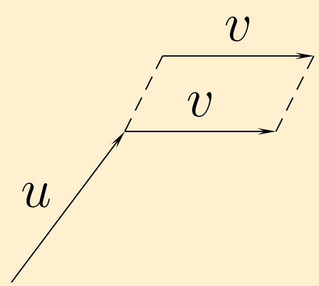
In particular, $\overrightarrow{AB} = -\overrightarrow{BA}$

The opposite to zero vector is zero vector itself, i.e. $-o = o$.

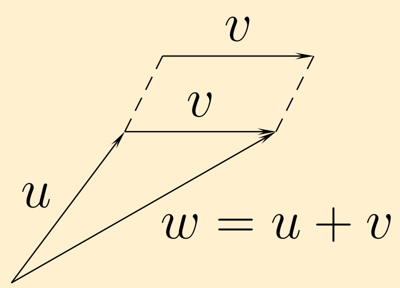
Vector addition. The *sum* of two vectors u and v is a vector w, which can be obtained by placing the initial point of v on the final point of u and then drawing an arrow from the initial point of u to the final point of v (*"tip-to-tail" method*, or *triangle rule*).



Vector addition. The *sum* of two vectors u and v is a vector w, which can be obtained by placing the initial point of v on the final point of u and then drawing an arrow from the initial point of u to the final point of v (*"tip-to-tail" method*, or *triangle rule*).

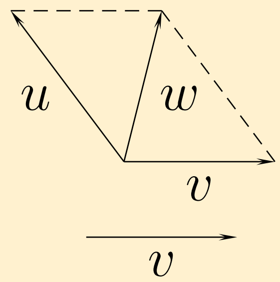


Vector addition. The *sum* of two vectors u and v is a vector w, which can be obtained by placing the initial point of v on the final point of u and then drawing an arrow from the initial point of u to the final point of v (*"tip-to-tail" method*, or *triangle rule*).

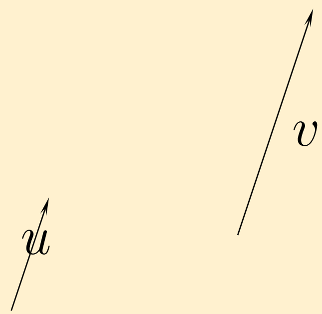


The operation of vector addition is written as $w = u + v$.

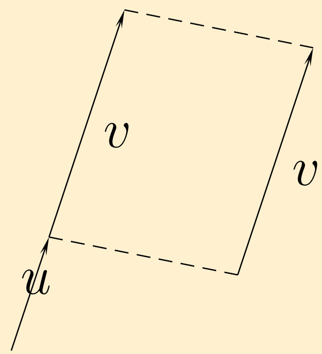
Another way to find the sum of two vectors is to use a *parallelogram rule*: One can place the initial point of v on the initial point u and construct the parallelogram whose two sides are u and v. The diagonal of the parallelogram beginning from the initial points of u and v is obviously the sum w of u and v.



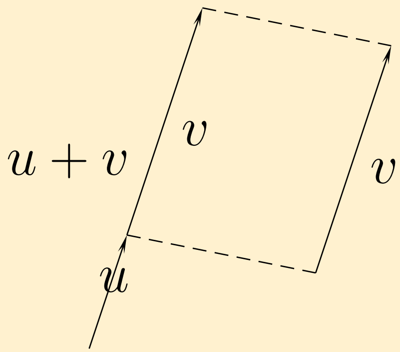
If vectors are collinear...



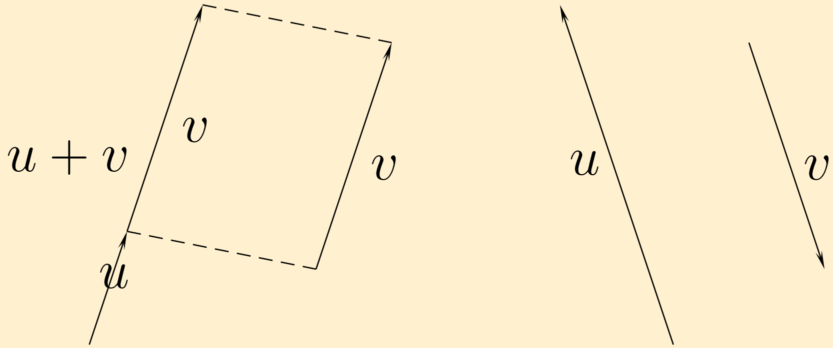
If vectors are collinear...



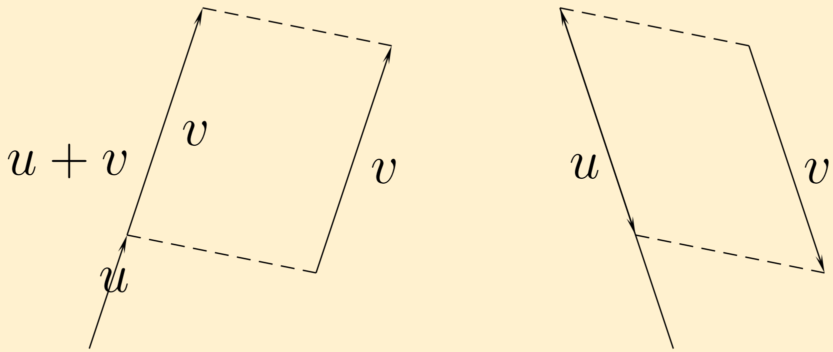
If vectors are collinear...



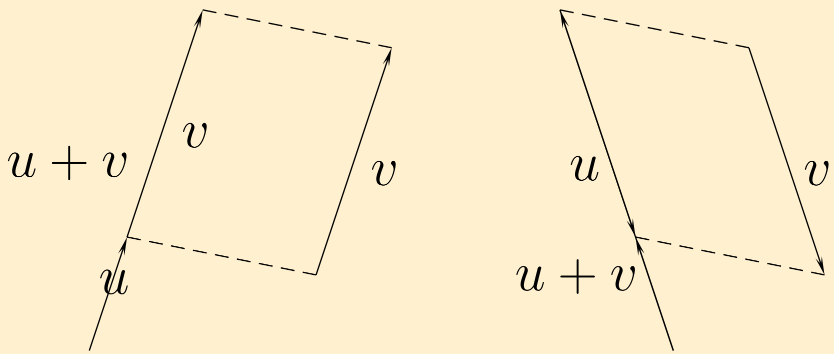
If vectors are collinear...



If vectors are collinear...



If vectors are collinear...



Multiplication of a vector by a scalar. Let v be a vector and α be a real number (scalar).

The product of v by α is a vector whose length is $|\alpha||v|$ and the direction is the same as the direction of v if $\alpha > 0$, and the direction is opposite if $\alpha < 0$.

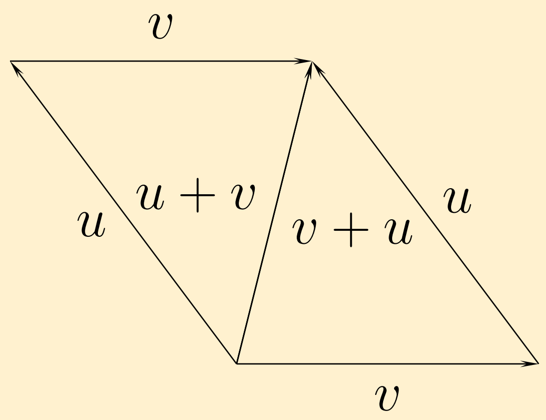
If $v = o$ or $\alpha = 0$ then $\alpha v = o$.

Properties of operations

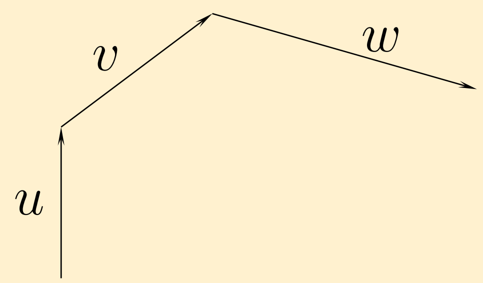
1. For all vectors u, v $\quad u + v = v + u$ (*commutativity*)
2. For all vectors u, v, w $\quad (u + v) + w) = u + (v + w)$ (*associativity*)
3. For any vector v $\quad v + o = v$
4. For any vector v $\quad v + (-v) = o$
5. For any vector v $\quad 1v = v$
6. For all scalars α and β and all vectors v $\quad \alpha(\beta v) = (\alpha\beta)v$
7. For all scalars α and β and all vectors v $\quad (\alpha + \beta)v = \alpha v + \beta v$
8. For all scalars α and all vectors u and v $\quad \alpha(u + v) = \alpha u + \alpha v$

Proofs

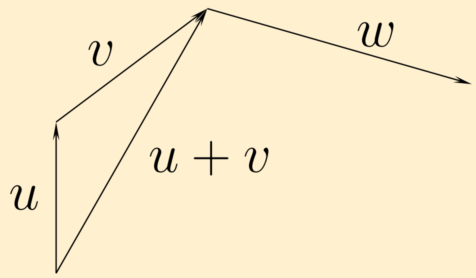
1. For all vectors u, v $\quad u + v = v + u$ (*commutativity*)



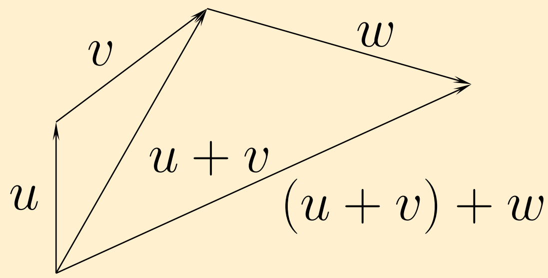
2. For all vectors u, v, w $\quad (u + v) + w = u + (v + w)$ (*associativity*)



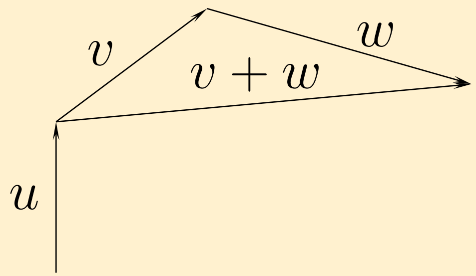
2. For all vectors u, v, w $\quad (u + v) + w = u + (v + w)$ (*associativity*)



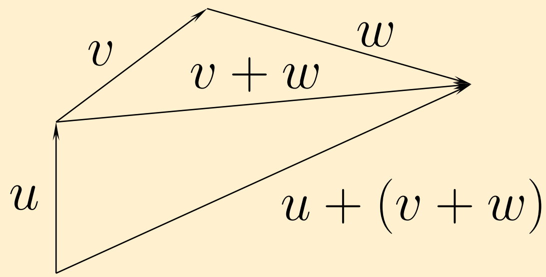
2. For all vectors u, v, w $\quad (u+v)+w = u+(v+w)$ (*associativity*)



2. For all vectors u, v, w $\quad (u+v)+w = u+(v+w)$ (*associativity*)

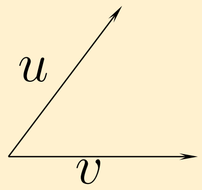


2. For all vectors u, v, w $\quad (u + v) + w = u + (v + w)$ (*associativity*)



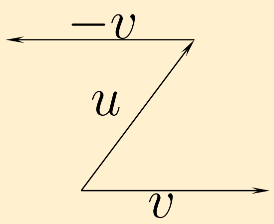
Subtraction

$$u - v = u + (-v)$$



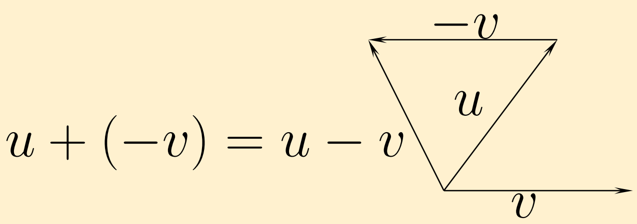
Subtraction

$$u - v = u + (-v)$$



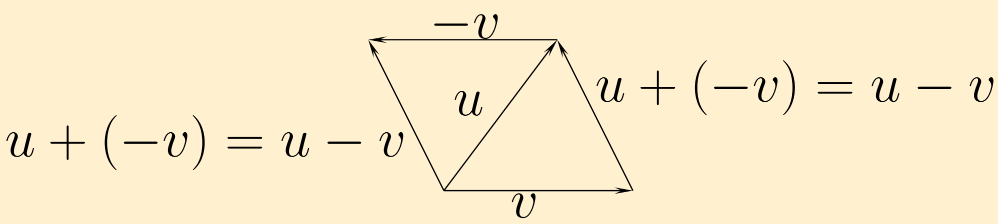
Subtraction

$$u - v = u + (-v)$$



Subtraction

$$u - v = u + (-v)$$



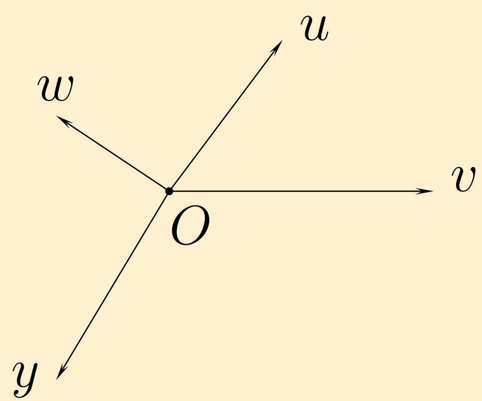
If u and v is placed such that they have the same initial point then $u - v$ connects the end of v with the end of u.

New Properties

1. For all vectors u and v the equation $u + x = v$ has the unique solution $x = v - u$
2. For all vectors u $\quad 0u = o$
3. For all scalars α and vectors v $\quad (-\alpha v) = (-\alpha)v$;
   in particular, $(-1)v = -v$
4. For all vectors u and v and all scalars α $\quad \alpha(u - v) = \alpha u - \alpha v$
5. For all vectors u and all scalars α and β $\quad (\alpha - \beta)u = \alpha u - \beta u$
6. If $\alpha v = o$ then $\alpha = 0$ or $v = 0$
7. For all vectors u and v $\quad |u| - |v| \leq \Big||u| - |v|\Big| \leq |u + v| \leq |u| + |v|$ (*triangle inequality*)

Radius Vectors. Take in the plane a fixed origin O.

As a rule, we will draw vectors as arrows beginning at the origin.



Such vectors are called *radius vectors*.

A vector $\overrightarrow{OA}$ is called the *radius vector of the point A*.

Basis in a Plane. Consider two non-collinear vectors in a plane.

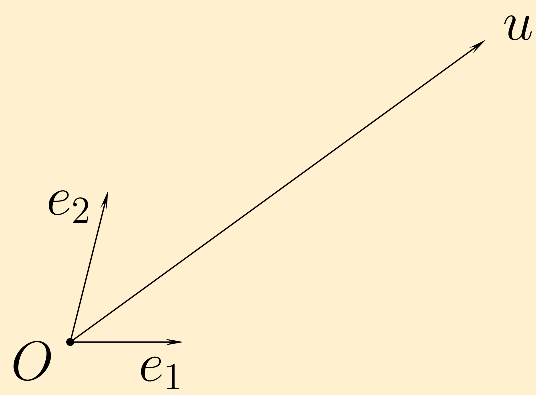
We will say that they form a *basis* in the plane.

Theorem 1.1. *Let e_1, e_2 form a basis in a plane. Then for any vector u in the plane there exist scalars α_1 and α_2 such that*

$$u = \alpha_1 e_1 + \alpha_2 e_2. \tag{1.1}$$

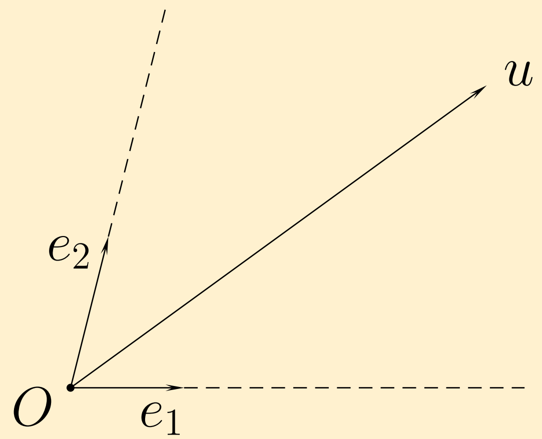
The equality (1.1) is called a *decomposition of a vector u by the basis e_1, e_2.*

Proof.



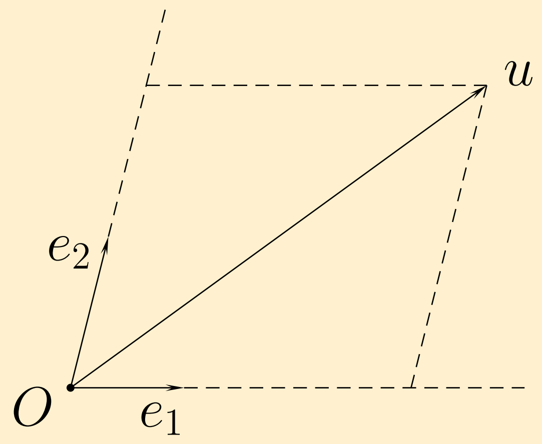
□

Proof.



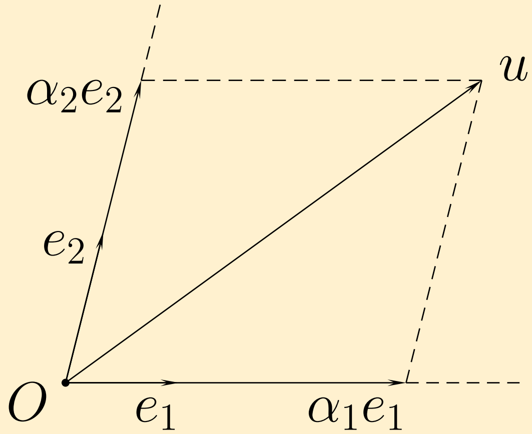
□

Proof.



□

Proof.



$$u = \alpha_1 e_1 + \alpha_2 e_2.$$

□

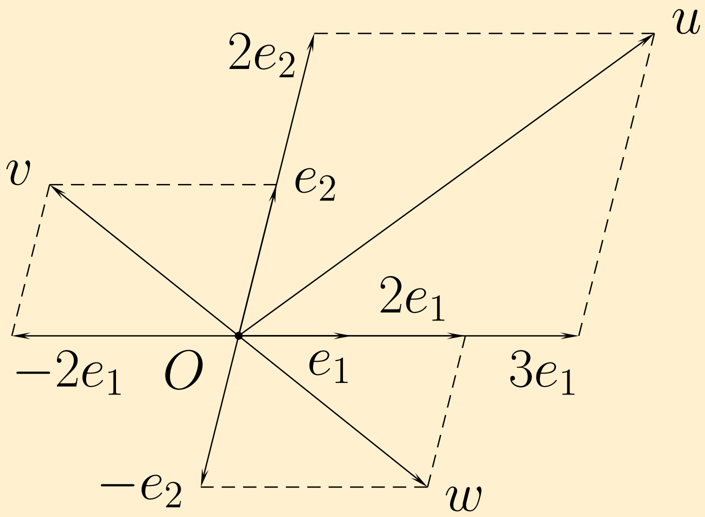
Scalars α_1 and α_2 in the equality $u = \alpha_1 e_1 + \alpha_2 e_2$ are called *coordinates of a vector* u *in the basis* e_1, e_2.

For coordinates of a vector u in the basis e_1, e_2 we will use the following notation:

$$u(\alpha_1, \alpha_2) \quad \text{or} \quad [u] = (\alpha_1, \alpha_2)$$

or (in the case of uncertainty)

$$[u]_{e_1, e_2} = (\alpha_1, \alpha_2).$$



$$u = 3e_1 + 2e_2, \quad v = -2e_1 + e_2, \quad w = 2e_1 - e_2$$

$$[u] = (3, 2), \quad [v] = (-2, 1), \quad [w] = (2, -1)$$

Theorem 1.2. *For any basis e_1, e_2 and any vector u the coordinates of u in the basis e_1, e_2 are unique.*

Proof. By contradiction. Suppose that we have two different decompositions of u:

$$u = \alpha_1 e_1 + \alpha_2 e_2 = \beta_1 e_1 + \beta_2 e_2.$$

Hence, $(\alpha_1 - \beta_1)e_1 + (\alpha_2 - \beta_2)e_2 = o$

Without loss of generality (w.l.g.) we can suppose that $\alpha_1 \neq \beta_1$

Hence

$$e_1 = -\frac{\alpha_2 - \beta_2}{\alpha_1 - \beta_1} e_2$$

This means that e_1 and e_2 are collinear. Contradiction! □

Theorem 1.3. *The coordinates of the sum of two vectors are equal to the sum of their corresponding coordinates:*

if $[u] = (\alpha_1, \alpha_2)$, $[v] = (\beta_1, \beta_2)$ *then* $[u + v] = (\alpha_1 + \beta_1,\ \alpha_2 + \beta_2)$.

The coordinates of the product of a vector by a scalar are equal to the product of its coordinates by the scalar.

if $[u] = (\alpha_1, \alpha_2)$ *then* $[\beta u] = (\beta\alpha_1,\ \beta\alpha_2)$.

Proof. Add vectors $u = \alpha_1 e_1 + \alpha_2$ and $v = \beta_1 e_1 + \beta_2 e_2$. We obtain $u + v = (\alpha_1 + \beta_1)e_1 + (\alpha_2 + \beta_2)e_2$. Hence $[u + v] = (\alpha_1 + \beta_1,\ \alpha_2 + \beta_2)$.

Multiply a vector $u = \alpha_1 e_1 + \alpha_2$ by the scalar β. We obtain $u = \beta\alpha_1 e_1 + \beta\alpha_2 e_2$. Hence $[\beta u] = (\beta\alpha_1,\ \beta\alpha_2)$. □

The aggregate of an origin O and a basis e_1, e_2 is called an *(affine) coordinate system*.

The *coordinates of a point* A in a coordinate system is the coordinates of its radius vector:

$$[A] = [OA]$$

We have $\overrightarrow{AB} = \overrightarrow{OA} - \overrightarrow{OB}$. Hence, *the coordinates of a vector are equal to the coordinates of its head minus coordinates of its tail.*

Example 1.4. Let $A(1,2)$, $B(-3,1)$, $C(7,-4)$, $D(-3,-2)$.

Find coordinates of $\overrightarrow{OA}$, $\overrightarrow{OB}$, $\overrightarrow{OC}$, $\overrightarrow{OD}$, $\overrightarrow{AB}$, $\overrightarrow{CD}$, $2\overrightarrow{AB} - 3\overrightarrow{CD}$.

$\overrightarrow{OA}(1,2)$,

$\overrightarrow{OB}(-3,1)$,

$\overrightarrow{OC}(7,-4)$,

$\overrightarrow{OD}(-3,-2)$,

$\overrightarrow{AB}(-4,-1)$,

$\overrightarrow{CD}(-10,2)$,

$[2\overrightarrow{AB} - 3\overrightarrow{CD}] = (2\cdot(-4) - 3\cdot(-10),\ 2\cdot(-1) - 3\cdot 2) = (22, \ -7)$.

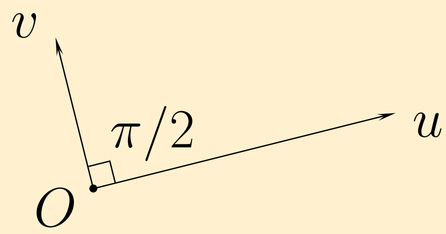
Special case: orthonormal basis and Cartesian coordinate system

A vector of magnitude 1 is called a *unit vector*.

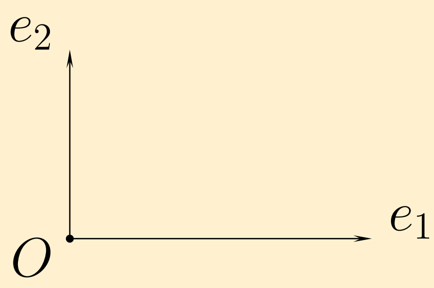
$\dfrac{u}{|u|}$ is a unit vector collinear with u with the same direction.

We say that u is *orthogonal* (or *perpendicular*) with v if the angle between u and v is $\pi/2$.
Notation: $u \perp u$.

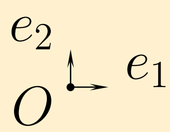
Zero vector is orthogonal to any other vector.



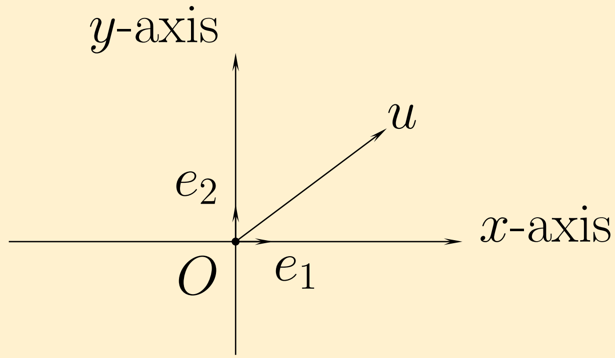
Basis e_1, e_2 is called *orthogonal* if vectors e_1 and e_2 are orthogonal.



Basis e_1, e_2 is called *orthonormal* if vectors e_1 and e_2 are orthogonal unit vectors.



If the basis is orthonormal then a corresponding coordinate system is called *rectangular* (or *Cartesian*) *coordinate system.*



$$u = 4e_1 + 3e_2 \quad \Rightarrow \quad [u] = (4, 3)$$

Vectors in the space

Consider three non-coplanar vectors in the space.

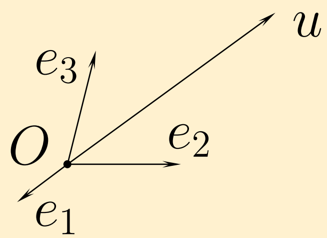
We will say that they form a *basis* in the space.

Theorem 1.5. *Let* e_1, e_2, e_3 *form a basis in the space. Then for any vector* u *in the plane there exist scalars* α_1, α_2 *and* α_3 *such that*

$$u = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3. \tag{1.2}$$

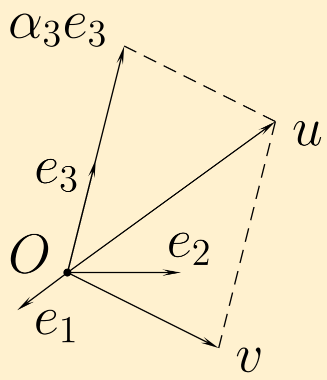
The equality (1.2) is called a *decomposition of a vector* u *by the basis* e_1, e_2, e_3.

Proof.



$\square$

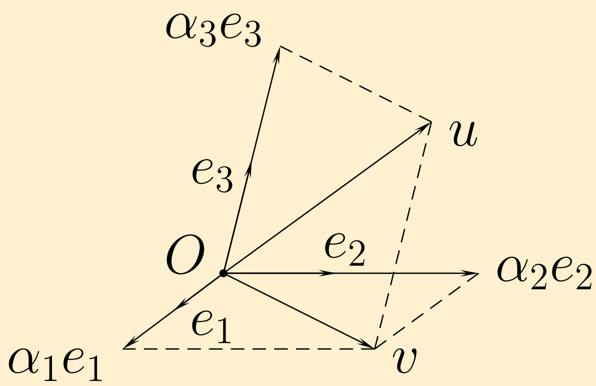
Proof.



$$u = v + \alpha_3 e_3$$

□

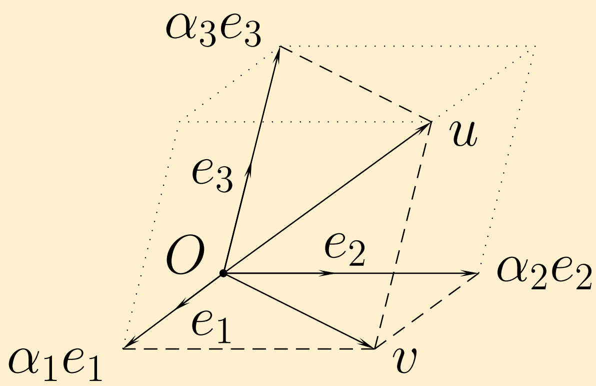
Proof.



$$u = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3$$

□

Proof.



$$u = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3$$

□

Scalars α_1, α_2 and α_3 in the equality $u = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3$ are called *coordinates of a vector* u *in the basis* e_1, e_2, e_3.

For coordinates of a vector u in the basis e_1, e_2, e_3 we will use the following notation:

$$u(\alpha_1, \alpha_2, \alpha_3) \quad \text{or} \quad [u] = (\alpha_1, \alpha_2, \alpha_3)$$

or (in the case of uncertainty)

$$[u]_{e_1, e_2, e_3} = (\alpha_1, \alpha_2, \alpha_3).$$

Theorem 1.6. *For any basis e_1, e_2, e_3 and any vector u the coordinates of u in the basis e_1, e_2, e_3 are unique.*

The aggregate of an origin O and a basis e_1, e_2, e_3 is called an *(affine) coordinate system*.

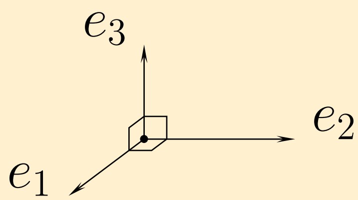
The *coordinates of a point A* in a coordinate system is the coordinates of its radius vector:

$$[A] = [OA]$$

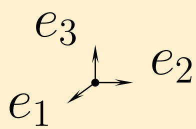
The coordinates of the sum of two vectors are equal to the sum of their corresponding coordinates.

The coordinates of the product of a vector by a scalar are equal to the product of its coordinates by the scalar.

Basis e_1, e_2, e_3 is called *orthogonal* if $e_1 \perp e_2$, $e_1 \perp e_3$, $e_2 \perp e_3$.



Basis e_1, e_2, e_3 is called *orthonormal* if vectors e_1, e_2, e_3 are unit vectors and $e_1 \perp e_2$, $e_1 \perp e_3$, $e_2 \perp e_3$.



If the basis is orthonormal then a corresponding coordinate system is called *rectangular* (or *Cartesian*) *coordinate system*.

Scalar product *Scalar product* (*inner product* or *dot product*) of two vectors:

$$(u, v) = |u| \cdot |v| \cdot \cos\vartheta,$$

where ϑ is the angle between u and v.

Example 1.7. $|u| = 2$, $|v| = 3$, $\vartheta = \dfrac{\pi}{6} \Rightarrow (u, v) = 2 \cdot 3 \cdot \cos\dfrac{\pi}{6} = 3\sqrt{3}$.

$|u| = 1$, $|v| = 3$, $\vartheta = \dfrac{2\pi}{3} \Rightarrow (u, v) = 1 \cdot 3 \cdot \cos\dfrac{2\pi}{3} = -\dfrac{3}{2}$.

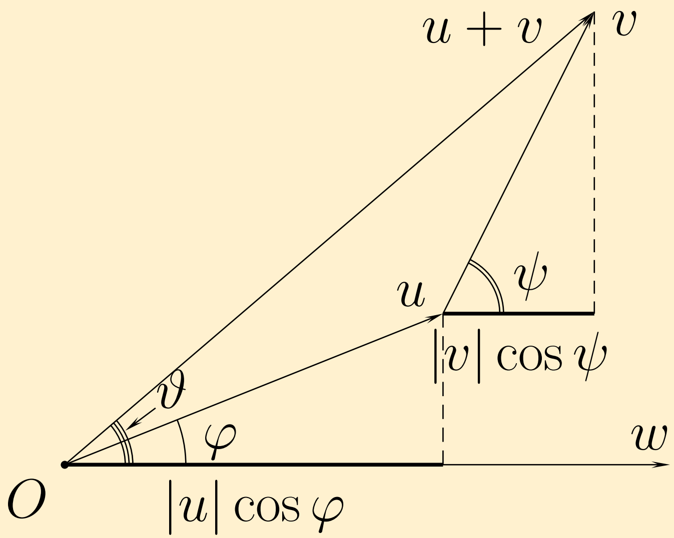
$|u| = 3$, $|v| = 4$, $\vartheta = 0 \Rightarrow (u, v) = 3 \cdot 4 \cdot \cos 0 = 12$.

$|u| = 3$, $|v| = 4$, $\vartheta = \pi \Rightarrow (u, v) = 3 \cdot 4 \cdot \cos\pi = -12$.

$|u| = 3$, $|v| = 4$, $\vartheta = \dfrac{\pi}{2} \Rightarrow (u, v) = 3 \cdot 4 \cdot \cos\dfrac{\pi}{2} = 0$.

$$\text{RHS} = (u, w) + (v, w) = |u| \cdot |w| \cdot \cos\varphi + |v| \cdot |w| \cdot \cos\psi = (|u|\cos\varphi + |v|\cos\psi)|w|$$

LHS and RHS have the same multiplier $|w|$. Let's compare what remains.



$$|u|\cos\varphi + |v|\cos\psi = |u + v|\cos\vartheta$$

(We must consider also the case then angles are not acute) $\square$

Let e_1, e_2, e_3 be orthonormal basis then

$$(e_1, e_1) = (e_2, e_2) = (e_3, e_3) = 1, \qquad (e_1, e_2) = (e_1, e_3) = (e_2, e_3) = 0$$

Theorem 1.9. *Let e_1, e_2 be orthonormal basis of a plane and $[u] = (\alpha_1, \alpha_2)$, $[v] = (\beta_1, \beta_2)$, then*

$$(u, v) = \alpha_1\beta_1 + \alpha_2\beta_2.$$

Let e_1, e_2, e_3 be orthonormal basis of the space and $[u] = (\alpha_1, \alpha_2, \alpha_3)$, $[v] = (\beta_1, \beta_2, \beta_3)$, then

$$(u, v) = \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3.$$

Proof. $(u, v) = (\alpha_1 e_1 + \alpha_2 e_2, \beta_1 e_1 + \beta_2 e_2)$
$= \alpha_1\beta_1(e_1, e_1) + \alpha_1\beta_2(e_1, e_2) + \alpha_2\beta_1(e_2, e_1) + \alpha_2\beta_2(e_2, e_2) = \alpha_1\beta_1 + \alpha_2\beta_2$

The proof for the space is similar. □

Example 1.10. (basis is orthonormal)

$u(3,-2)$, $v(2,6) \Rightarrow (u,v) = 3\cdot 2 + (-2)\cdot 6 = -6$

$u(1,-2)$, $v(2,1) \Rightarrow (u,v) = 1\cdot 2 + (-2)\cdot 1 = 0$ (vectors are orthogonal)

$u(1,-2,3)$, $v(3,5,-1) \Rightarrow (u,v) = 1\cdot 3 + (-2)\cdot 5 + 3\cdot(-1) = -10$

Corollary 1.11. *Let* e_1, e_2 *be orthonormal basis of a plane and* $[u] = (\alpha_1, \alpha_2)$ *then*

$$|u| = \sqrt{\alpha_1^2 + \alpha_2^2}.$$

Let e_1, e_2, e_3 *be orthonormal basis of the space and* $[u] = (\alpha_1, \alpha_2, \alpha_3)$ *then*

$$|u| = \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}.$$

Example 1.12. (basis is orthonormal)

$u(3, -4) \Rightarrow |u| = \sqrt{3^2 + (-4)^2} = \sqrt{25} = 5$

$u(-1, 7) \Rightarrow |u| = \sqrt{(-1)^2 + 7^2} = \sqrt{50} = 5\sqrt{2}$

$u(1, -2, 3) \Rightarrow |u| = \sqrt{1^2 + (-2)^2 + 3^2} = \sqrt{14}$

Corollary 1.13. *Let* e_1*,* e_2 *be orthonormal basis of a plane and* $[A]=(\alpha_1,\alpha_2)$*,* $[B]=(\beta_1,\beta_2)$*, then*

$$|AB|=\sqrt{(\alpha_1-\beta_1)^2+(\alpha_2-\beta_2)^2}.$$

Let e_1*,* e_2*,* e_3 *be orthonormal basis of the space and* $[A]=(\alpha_1,\alpha_2,\alpha_3)$*,* $[B]=(\beta_1,\beta_2,\beta_3)$*, then*

$$(u,v)=\sqrt{(\alpha_1-\beta_1)^2+(\alpha_2-\beta_2)^2+(\alpha_3-\beta_3)^2}.$$

Example 1.14. (basis is orthonormal)

$A(-2,6)$, $B(1,2)\Rightarrow|AB|=\sqrt{(-2-1)^2+(6-2)^2}=\sqrt{25}=5$

$A(1,-2,6)$, $B(0,-1,1)\Rightarrow|AB|=\sqrt{(1-0)^2+(-2+1)^2+(6-1)^2}=\sqrt{27}=3\sqrt{3}$

Angles. Since $(u, v) = |u| \cdot |v| \cdot \cos\vartheta$ then

$$\cos\vartheta = \frac{(u, v)}{|u| \cdot |v|}$$

Example 1.15. Find $\angle ABC$, if $A(2,5)$, $B(1,2)$, $C(3,3)$ (coordinate system is rectangular)

$$[\overrightarrow{BA}] = (1, 3), \qquad [\overrightarrow{BC}] = (2, 1)$$

$$\cos\angle ABC = \frac{(\overrightarrow{BA}, \overrightarrow{BC})}{|BA| \cdot |BC|} = \frac{1 \cdot 2 + 3 \cdot 1}{\sqrt{1^2 + 3^2} \cdot \sqrt{2^2 + 1^2}} = \frac{5}{\sqrt{10}\sqrt{5}} = \frac{\sqrt{2}}{2}$$

hence $\angle ABC = \dfrac{\pi}{4}$

Example 1.16. Find $\angle ABC$, if $A(2,5,1)$, $B(1,2,-1)$, $C(3,3,1)$ (coordinate system is rectangular)

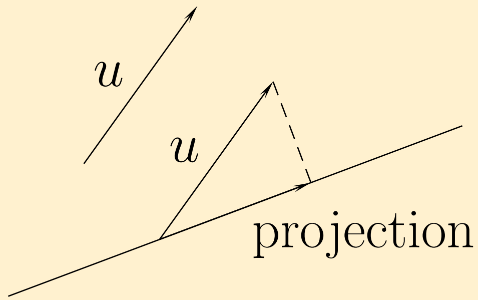
$$[\overrightarrow{BA}] = (1,3,2), \qquad [\overrightarrow{BC}] = (2,1,2)$$

$$\cos\angle ABC = \frac{(\overrightarrow{BA}, \overrightarrow{BC})}{|BA| \cdot |BC|} = \frac{1 \cdot 2 + 3 \cdot 1 + 2 \cdot 2}{\sqrt{1^2+3^2+2^2} \cdot \sqrt{2^2+1^2+2^2}} = \frac{9}{\sqrt{14}\sqrt{9}} = \frac{3}{\sqrt{14}}$$

hence $\angle ABC = \arccos\left(3/\sqrt{14}\right)$

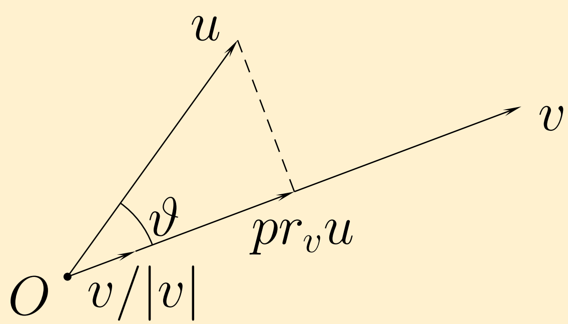
Projection

The orthogonal projection of a vector u along a line is obtained by moving the tail of the vector onto the line and dropping a perpendicular onto the line from the head of the vector.



The resulting vector on the line is the *(orthogonal) projection* of u on the line.

Suppose the line is parallel to non-zero vector v.



Then the projection of u onto the line is

$$\operatorname{pr}_v u = |u| \cdot \cos\vartheta \cdot \frac{v}{|v|} = |u| \cdot |v| \cdot \cos\vartheta \cdot \frac{v}{|v|^2} = \frac{(u,v)}{(v,v)} v$$

Vector product

Let u and v be non-collinear vectors in the space. Let w be a vector such that

1. $|w| = |u| \cdot |v| \cdot \sin\vartheta$, where ϑ is the angle between u and v;
2. $w \perp u$, $w \perp v$;
3. The direction of w follows the *right hand rule*:

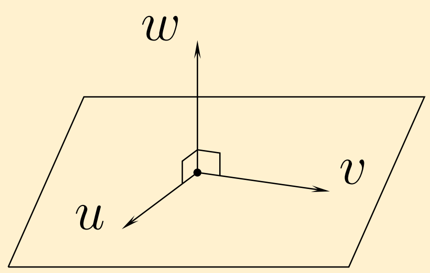
   Point your thumb in the direction of u. Point your fingers in the direction of v. Your palm will face in the direction of w.

   (Use the *right* hand! The left hand will give the wrong answer)

If u and v are collinear then let w be the zero vector.

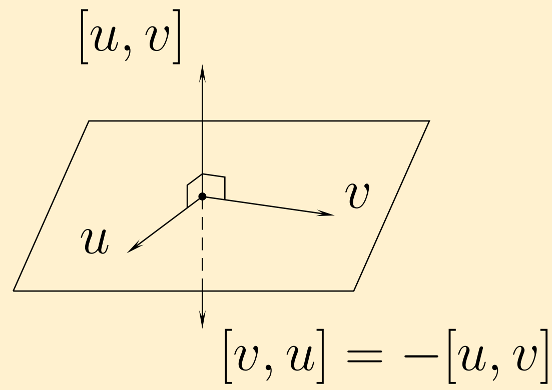
w is called a *vector product* (*outer product* or *cross product*) of u and v. Notation: $w = [u, v]$.

Thus, the vector product involves two vectors and gets as a result another vector which is perpendicular to both vectors; its length is equal to $|u| \cdot |v| \cdot \sin\vartheta$; and direction follows the right hand rule.



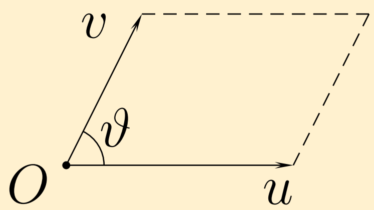
In the vector product order is important.

When we switch the position of our thumb and fingers we have to flip our hand over and our palm then faces in the opposite direction. Hence $[u, v] = -[v, u]$.



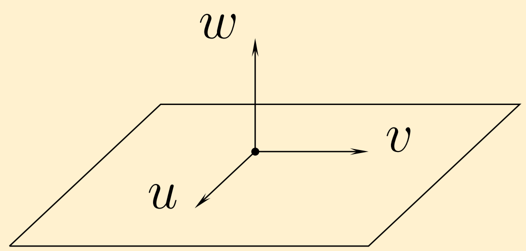
Theorem 1.17. *For all vectors u, v, w in the space and for any scalar α*

1. $[u, v] = -[v, u]$ *(anticommutativity)*
2. $[u + v, w] = [u, w] + [v, w]$
3. $[\alpha u, v] = \alpha[u, v]$
4. $[u, v] = o$ *if and only if* u *and* v *are collinear*
5. $|[u, v]| = |u| \cdot |v| \cdot \sin\vartheta$ *is the area of the parallelogram generated from* u *and* v

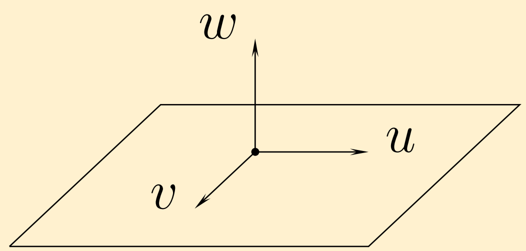


$$\text{Area} = |u| \cdot |v| \cdot \sin\vartheta$$

Right triplet $\langle u, v, w \rangle$



Left triplet $\langle u, v, w \rangle$



Let vectors e_1, e_2, e_3 form an orthonormal basis of the space. Let $\langle e_1, e_2, e_3 \rangle$ be a right triplet (orthonormal *right* basis). Then

$$\begin{aligned} &[e_1, e_2] = e_3, && [e_2, e_3] = e_1, && [e_3, e_1] = e_2 \\ &[e_2, e_1] = -e_3, && [e_3, e_2] = -e_1, && [e_1, e_3] = -e_2 \\ &[e_1, e_1] = o, && [e_2, e_2] = o, && [e_3, e_3] = o \end{aligned}$$

This can be written in a table as follows (the rows correspond to the first operand; the columns correspond to the second one):

	e_1	e_2	e_3
e_1	o	e_3	$-e_2$
e_2	$-e_3$	o	e_3
e_3	e_2	$-e_1$	o

Theorem 1.18. *Let* e_1, e_2, e_3 *be orthonormal* ***right*** *basis of the space and* $[u] = (\alpha_1, \alpha_2, \alpha_3)$, $[v] = (\beta_1, \beta_2, \beta_3)$, *then*

$$[u, v] = (\alpha_2\beta_3 - \alpha_3\beta_2)e_1 + (\alpha_3\beta_1 - \alpha_1\beta_3)e_2 + (\alpha_1\beta_2 - \alpha_2\beta_1)e_3.$$

Proof. Since $u = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3$ and $v = \beta_1 e_1 + \beta_2 e_2 + \beta_3 e_3$ then

$$[u, v] = [\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3, \beta_1 e_1 + \beta_2 e_2 + \beta_3 e_3]$$

$$\begin{aligned} &= [\alpha_1 e_1, \beta_1 e_1] + [\alpha_1 e_1, \beta_2 e_2] + [\alpha_1 e_1, \beta_3 e_3] \\ &+ [\alpha_2 e_2, \beta_1 e_1] + [\alpha_2 e_2, \beta_2 e_2] + [\alpha_2 e_2, \beta_3 e_3] \\ &+ [\alpha_3 e_3, \beta_1 e_1] + [\alpha_3 e_3, \beta_2 e_2] + [\alpha_3 e_3, \beta_3 e_3] \end{aligned}$$

$$\begin{aligned} &= \quad 0 \quad + \alpha_1\beta_2 e_3 - \alpha_1\beta_3 e_2 \\ &- \alpha_2\beta_1 e_3 + \quad 0 \quad + \alpha_2\beta_3 e_1 \\ &+ \alpha_3\beta_1 e_2 - \alpha_3\beta_2 e_1 + \quad 0 \end{aligned}$$

$$= (\alpha_2\beta_3 - \alpha_3\beta_2)e_1 + (\alpha_3\beta_1 - \alpha_1\beta_3)e_2 + (\alpha_1\beta_2 - \alpha_2\beta_1)e_3.$$

□

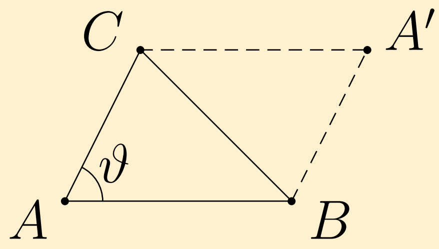
$$[u, v] = (\alpha_2\beta_3 - \alpha_3\beta_2)e_1 + (\alpha_3\beta_1 - \alpha_1\beta_3)e_2 + (\alpha_1\beta_2 - \alpha_2\beta_1)e_3.$$

Example 1.19. Let's find the vector product $w = [u, v]$ if $u(1, 2, 3)$, $v(3, 4, 2)$ (in right orthonormal basis)

$w = [u, v] = (2 \cdot 2 - 3 \cdot 4)e_1 + (3 \cdot 3 - 1 \cdot 2)e_2 + (1 \cdot 4 - 3 \cdot 2)e_3 = -8e_1 + 7e_2 - 2e_3$

Thus the coordinates of $w = [u, v]$ are $(-8, 7, -2)$.

Example 1.20. Let's find the area S of a triangle ABC if $A(0,1,2)$, $B(2,4,3)$, $C(-1,3,0)$ (basis is orthonormal)



Triangles ABC and $A'BC$ are equal. Hence the area of the trangle ABC is the half of the area of the parallelogram $ABA'C$:

$$S = \frac{1}{2} AB \cdot AC \cdot \sin\vartheta = \frac{1}{2}\left|[\overrightarrow{AB}, \overrightarrow{AC}]\right|$$

$$[\overrightarrow{AB}] = (2,3,1),\ [\overrightarrow{AC}] = (-1,2,-2)$$

$$[\overrightarrow{AB}, \overrightarrow{AC}] = (3\cdot(-2) - 1\cdot 2)e_1 + (1\cdot(-1) - 2\cdot(-2))e_2 + (2\cdot 2 - 3\cdot(-1))e_3 = -8e_1 + 3e_2 + 7e_3$$

Thus the coordinates of $[\overrightarrow{AB}, \overrightarrow{AC}]$ are $(-8,3,7)$.

The magnitude of $[\overrightarrow{AB}, \overrightarrow{AC}]$ is $\sqrt{(-8)^2 + 3^2 + 7^2} = \sqrt{122}$. Hence $S = \dfrac{\sqrt{122}}{2}$.

Example 1.21. Let's find the area of a triangle ABC in a plane if $A(\alpha_1, \alpha_2)$, $B(\beta_1, \beta_2)$, $C(\gamma_1, \gamma_2)$ (the basis is orthonormal).

We can suppose that the points are in the space but their third coordinates are 0's:
$A(\alpha_1, \alpha_2, 0)$, $B(\beta_1, \beta_2, 0)$, $C(\gamma_1, \gamma_2, 0)$.
$\overrightarrow{AB}(\beta_1 - \alpha_1, \beta_2 - \alpha_2, 0)$, $\overrightarrow{AC}(\gamma_1 - \alpha_1, \gamma_2 - \alpha_2, 0)$

$$[\overrightarrow{AB}, \overrightarrow{AC}] = 0 \cdot e_1 + 0 \cdot e_2 + \big((\beta_1 - \alpha_1)(\gamma_2 - \alpha_2) - (\beta_2 - \alpha_2)(\gamma_1 - \alpha_1)\big)e_3$$

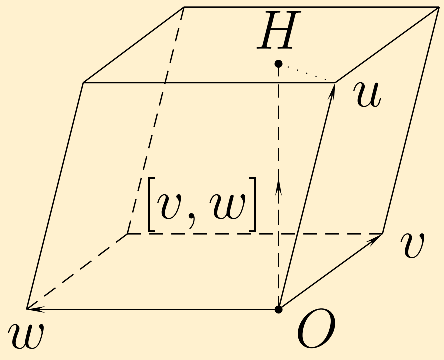
The area of ABC is

$$S = \frac{1}{2}\Big|[\overrightarrow{AB}, \overrightarrow{AC}]\Big| = \frac{1}{2}\Big((\beta_1 - \alpha_1)(\gamma_2 - \alpha_2) - (\beta_2 - \alpha_2)(\gamma_1 - \alpha_1)\Big)$$

Mixed product. Mixed (or triple scalar product) of three vectors u, v, w in the space is

$$(u, v, w) = (u, [v, w]).$$

Theorem 1.22. *Let u, v, w are non-coplanar vectors in the space. The absolute value of their mixed product is the volume of a parallelepiped constructed from u, v, w.*



The volume is $V = Sh$, where
$S = \Big|[v, w]\Big|$ is the area of a parallelogram generated from v and w,
$h = OH = \Big|\mathrm{pr}_{[v,w]}\, u\Big|$ is the height of the parallelepiped

$$V = S \cdot h = \Big|[v, w]\Big| \cdot OH == \Big|[v, w]\Big| \cdot |u| \cdot |\cos\vartheta| = |(u, [v, w])|.$$

Theorem 1.23. *For all vectors u, v, w, y and any scalar α*

1. $(u, v, w) = (v, w, u) = (w, u, v) = -(u, w, v) = -(v, u, w) = -(w, v, u)$

2. $(u + v, w, y) = (u, w, y) + (v, w, y)$

3. $(\alpha u, v, w) = \alpha(u, v, w)$

4a. $(u, v, w) = 0$ *if and only if the vectors u, v, w are co-planar*

4b. $(u, v, w) > 0$ *if and only if the triplet u, v, w is right*

4c. $(u, v, w) < 0$ *if and only if the triplet u, v, w is left*

Proof. The property 4 is verified directly.

If a triplet $\langle u, v, w\rangle$ is right then the triplets $\langle v, w, u\rangle$, $\langle w, u, v\rangle$ are right and the triplets $\langle u, w, v\rangle$, $\langle v, u, w\rangle$ and $\langle w, v, u\rangle$ are left.

If a triplet $\langle u, v, w\rangle$ is left then the triplets $\langle v, w, u\rangle$, $\langle w, u, v\rangle$ are left and the triplets $\langle u, w, v\rangle$, $\langle v, u, w\rangle$ and $\langle w, v, u\rangle$ are right.

Now the property 1 follows from the previous theorem.

Properties 2, 3 follow from corresponding properties of the scalar product:

$$(u+v,w,y)=(u+v,[w,y])=(u,[w,y])+(v,[w,y])=(u,w,y)+(v,w,y),$$

$$(\alpha u,v,w)=(\alpha u,[v,w])=\alpha(u,[v,w])=\alpha(u,v,w).$$

$\square$

Theorem 1.24. *Let e_1, e_2, e_3 be a basis of the space and*

$$[u] = (\alpha_1, \alpha_2, \alpha_3), \quad [v] = (\beta_1, \beta_2, \beta_3), \quad [w] = (\gamma_1, \gamma_2, \gamma_3)$$

then

$$(u, v, w) = (\alpha_1\beta_2\gamma_3 + \alpha_2\beta_3\gamma_1 + \alpha_3\beta_1\gamma_2 - \alpha_3\beta_2\gamma_1 - \alpha_2\beta_1\gamma_3 - \alpha_1\beta_3\gamma_2) \cdot (e_1, e_2, e_3).$$

If basis is right orthonormal then

$$(u, v, w) = \alpha_1\beta_2\gamma_3 + \alpha_2\beta_3\gamma_1 + \alpha_3\beta_1\gamma_2 - \alpha_3\beta_2\gamma_1 - \alpha_1\beta_3\gamma_2 - \alpha_2\beta_1\gamma_3.$$

The proof is similar to the proofs of analogous theorems for scalar product and vector product.

The 3rd order determinant

$$\begin{vmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{vmatrix} = \alpha_1\beta_2\gamma_3 + \alpha_2\beta_3\gamma_1 + \alpha_3\beta_1\gamma_2 - \alpha_3\beta_2\gamma_1 - \alpha_1\beta_3\gamma_2 - \alpha_2\beta_1\gamma_3$$

Corollary 1.25. *Let e_1, e_2, e_3 be a basis of the space and*

$$[u] = (\alpha_1, \alpha_2, \alpha_3), \quad [v] = (\beta_1, \beta_2, \beta_3), \quad [w] = (\gamma_1, \gamma_2, \gamma_3)$$

then

$$(u, v, w) = \begin{vmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{vmatrix} \cdot (e_1, e_2, e_3).$$

Example 1.26. Let's find the triple product of vectors

$$u(1, 2, 3), \quad w(4, 5, 7), \quad v(-2, -3, -4)$$

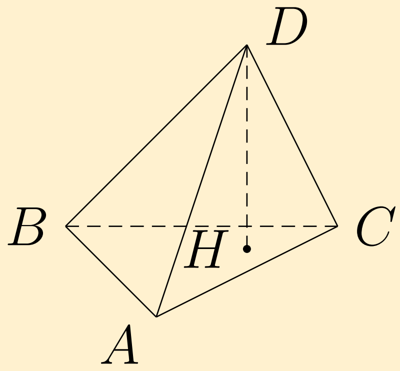
(the basis is right orthonormal)

$$(u, v, w) = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 7 \\ -2 & -3 & -4 \end{vmatrix}$$

$= 1 \cdot 5 \cdot (-4) + 2 \cdot 7 \cdot (-2) + 3 \cdot 4 \cdot (-3) - 3 \cdot 5 \cdot (-2) - 2 \cdot 4 \cdot (-4) - 1 \cdot 7 \cdot (-3) =$
$-20 - 28 - 36 + 30 + 32 + 21 = -1$

What is the volume of a parallelepiped generated from u, v, w?

Example 1.27. Let's find the volume V of a tetrahedron $ABCD$, if $A(1,1,1)$, $B(1,2,3)$, $C(0,2,1)$, $D(3,2,1)$.



The volume is $V = \dfrac{1}{3}Sh$, where $S = \dfrac{1}{2}\left|[\overrightarrow{AB}, \overrightarrow{AC}]\right|$ is the area of the triangle ABC,

$h = DH = \left|\text{pr}_{[\overrightarrow{AB},\overrightarrow{AC}]}\overrightarrow{AD}\right|$ is the height of the tetrahedron

Hence

$$V = \frac{1}{6} \cdot \left|\left(\overrightarrow{AB},\ \overrightarrow{AC},\ \overrightarrow{AD}\right)\right|$$

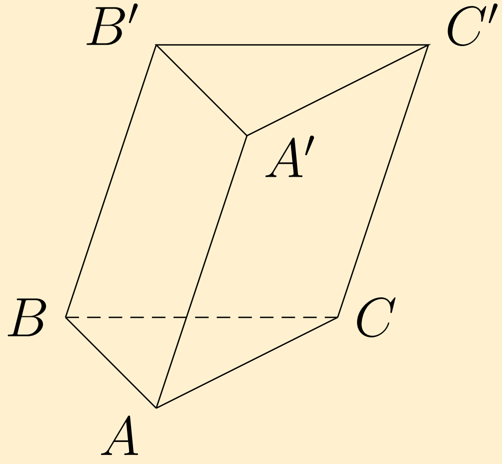
$$\overrightarrow{AB}\,(0,1,2), \quad \overrightarrow{AC}\,(-1,1,0), \quad \overrightarrow{AD}\,(2,1,0)$$

$$(\overrightarrow{AB},\ \overrightarrow{AC},\ \overrightarrow{AD}) = \begin{vmatrix} 0 & 1 & 2 \\ -1 & 1 & 0 \\ 2 & 1 & 0 \end{vmatrix}$$

$$= 0 \cdot 1 \cdot 0 + 1 \cdot 0 \cdot 2 + 2 \cdot (-1) \cdot 1 - 2 \cdot 1 \cdot 2 - 1 \cdot (-1) \cdot 0 - 0 \cdot 0 \cdot 1 = 0 + 0 - 2 - 4 - 0 - 0 = -6.$$

$$V = 1.$$

Example 1.28. Find the volume V of a prism $ABCA'B'C'$, if $A(1,1,1)$, $B(2,4,3)$, $C(3,-2,2)$, $A'(3,4,2)$.



$$V = \frac{1}{2} \cdot \left| \left(\overrightarrow{AB},\ \overrightarrow{AC},\ \overrightarrow{AA'} \right) \right|$$

$$\overrightarrow{AB}\,(1,3,2), \quad \overrightarrow{AC}\,(2,-3,1), \quad \overrightarrow{AA'}\,(2,3,1)$$

$$(\overrightarrow{AB},\ \overrightarrow{AC},\ \overrightarrow{AD}) = \begin{vmatrix} 1 & 3 & 2 \\ 2 & -3 & 1 \\ 2 & 3 & 1 \end{vmatrix}$$

$$= 1 \cdot (-3) \cdot 1 + 3 \cdot 1 \cdot 2 + 2 \cdot 2 \cdot 3 - 2 \cdot (-3) \cdot 2 - 3 \cdot 2 \cdot 1 - 1 \cdot 1 \cdot 3 = -3 + 6 + 12 + 12 - 6 - 3 = 18.$$

$$V = 9$$

Corollary 1.29. *Let e_1, e_2, e_3 be a right orthonormal basis of the space and*

$$[u] = (\alpha_1, \alpha_2, \alpha_3), \quad [v] = (\beta_1, \beta_2, \beta_3),$$

then

$$[u, v] = \begin{vmatrix} e_1 & e_2 & e_3 \\ \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{vmatrix}$$

Example 1.30. Find $[u, v]$ if $u(1, 2, 3)$, $v(3, 2, 1)$ (the basis is right orthonormal).

$$[u, v] = \begin{vmatrix} e_1 & e_2 & e_3 \\ 1 & 2 & 3 \\ 3 & 2 & 1 \end{vmatrix}$$

$$= e_1 \cdot 2 \cdot 1 + e_2 \cdot 3 \cdot 3 + e_3 \cdot 1 \cdot 2 - e_3 \cdot 2 \cdot 3 - e_2 \cdot 1 \cdot 1 - e_1 \cdot 3 \cdot 2 = -4e_1 + 8e_2 - 4e_3.$$

Examples and Exercizes

Exercize 1.31. There are vectors $u(1, 2)$, $v(2, -1)$, $w(-3, 2)$. Find coordinates of $3u + 5v - 3w$ and $-u + 5v + w$. Draw all vectors.

Exercize 1.32. There are vectors $u(0,1,2)$, $v(2,1,1)$, $w(1,-3,-2)$. Find coordinates of $3u+5v-3w$ and $-u+5v+w$.

Example 1.33. Vectors u, v, w have the same tail. Prove that the head of w is on the segment connecting the heads of u and v if and only if

$$w = \alpha u + \beta v \quad \text{with } \alpha \geq 0,\ \beta \geq 0,\ \alpha + \beta = 1$$

What coordinates of w in the basis u, v? In what ratio the head of w divide this segment?

Exercize 1.34. In the triangle ABC the point M is the midpoint of BC (AM is a median). Find coordinates of the vectors $\overrightarrow{AM}$ in basis $\overrightarrow{AB}$, $\overrightarrow{AC}$.

Exercize 1.35. Points E and F are midpoints of edges AB and CD of quadrilateral $ABCD$ respectively. Prove that

$$\overrightarrow{EF} = \frac{1}{2}(\overrightarrow{BC} + \overrightarrow{AD}).$$

Exercize 1.36. Let AD is the bisectix in the triangle ABC. Prove that

$$\overrightarrow{AD} = \frac{|AC|}{|AB| + |AC|} \cdot \overrightarrow{AB} + \frac{|AC|}{|AB| + |AC|} \cdot \overrightarrow{AC}.$$

What coordinates of $\overrightarrow{AD}$ in the basis $\overrightarrow{AB}$, $\overrightarrow{AC}$?

Example 1.37. A quadrilateral $ABCD$ is called *trapezoid* (or *trapezium*) if $AD \parallel BC$ and $AB \nparallel CD$. Sides AD and BC are called *bases*.

Let $AD/BC = 4$. Let A be the origin of a coordinates system and $\overrightarrow{AD}$, $\overrightarrow{AB}$ be a basis. Let M be the intersection of the diagonals of the trapezoid. Find coordinates of A, B, C, D, M.

Example 1.38. Find the coordinates of the intersection of medians in a traingle ABC with $A(x_1, y_1)$, $B(x_2, y_2)$, $C(x_3, y_3)$.

Exercize 1.39. Find $|a|^2 - \sqrt{3}(a,b) + 5|b|^2$ if

1. $|a| = 1$, $|b| = 2$, $\vartheta = \pi/3$;
2. $|a| = 3$, $|b| = 2$, $\vartheta = 5\pi/6$.

(ϑ is the angle between a and b).

Exercize 1.40. Let $a(-1,2)$, $b(2,3)$, $c(1,5)$. Calculate

1. $b\,(a,c) - c\,(a,b)$;
2. $|a|^2 - (b,c)$;
3. $|b|^2 + (b, a+2c)$.

Example 1.41. Prove that vectors a and $b(a, c) - c(a, b)$ are orthogonal.

Exercize 1.42. Prove that $|u - v|^2 = |u|^2 - 2\,|u|\,|v| + |v|^2$.

Exercize 1.43. Find $(\overrightarrow{AC}, \overrightarrow{BC})$ if

1. $|AB| = 5$, $|BC| = 3$, $|AC| = 4$;
2. $|AB| = 7$, $|BC| = 4$, $|AC| = 5$;
3. $|AB| = 3$, $|BC| = 2$, $|AC| = 3$;

(The coordinate system is rectangular)

Exercize 1.44. Find $\mathrm{pr}_v\, u$ if

1. $u(1,1)$, $v(1,-3)$;
2. $u(1,1)$, $v(1,-1)$.

(The basis is orthonormal)

Exercize 1.45. Simplify

1. $[u + v,\ u - v]$;
2. $\left[u + v - \frac{1}{2}w,\ \ -u + 2v - 5w\right]$;

Exercize 1.46. Vectors $u(2,3,1)$ and $v(-1,1,2)$ have the same origin (The basis is orthonormal). Find the area of the triangle generated from u and v and the lengths of all heights in it.

Exercize 1.47. Can the vectors u, v, w form a basis of the space if

1. $u(1, -1, 1)$, $v(1, 2, -5)$, $w(1, 1, 0)$;
2. $u(1, 3, 2)$, $v(3, 2, 5)$, $w(1, -4, 1)$;

Exercize 1.48. Let a, b, c are non-coplanar vectors. Find all λ such that $\lambda a - b + c$, $a + 2b - c$ and $\lambda^2 a + 3c$ are coplanar.

Chapter 2

Complex Numbers

2.1. Numerical sets

We suppose that you know the following notation for numerical sets:

The set of *natural numbers*

$$\mathbb{N} = \{1, 2, 3, \dots\}\,.$$

The set of *integer numbers*

$$\mathbb{Z} = \{\dots,\ \ -3,\ \ -2,\ \ -1,\ 0,\ 1,\ 2,\ 3,\ \ \dots\}\,.$$

The set of *rational numbers*

$$\mathbb{Q} = \left\{\frac{p}{q} :\ p \in \mathbb{Z},\ q \in \mathbb{Z},\ q \neq 0\right\}$$

The set of real numbers $\mathbb{R}$.

Notice that

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}.$$

$$(\text{and}\quad \mathbb{N} \neq \mathbb{Z},\ \mathbb{Z} \neq \mathbb{Q},\ \mathbb{Q} \neq \mathbb{R})$$

Exercize 2.1. Give an example of the number α, such that

$$\alpha \in \mathbb{R}, \quad \text{but} \quad \alpha \notin \mathbb{Q}.$$

2.2. Complex Numbers

A *complex number* is an ordered[1] pair of real numbers (a, b).

For example, $(0, 0)$, $(1, 0)$, $(3/4, 5/14)$.

The set of all complex numbers we denote by $\mathbb{C}$.

Two complex numbers (a, b) and (c, d) are *equal* if (if and only if)

$$a = c \quad \text{and} \quad b = d$$

In this case we'll write $(a, b) = (c, d)$.

[1]The pair is ordered which means that it is important which number is the first, and which one is the second. For instance, $(1, 0)$ and $(0, 1)$ are different pairs.

On the set of all complex numbers let's define the operations *addition* and *multiplication*:[2]:

$$(a,b)+(c,d)=(a+c,b+d), \quad (a,b)\cdot(c,d)=(ac-bd,ad+bc). \tag{2.1}$$

[2]The sign of operation "$\cdot$" will be often omitted as the sign "$\cdot$" or "$\times$" for multiplication of real numbers.

Operations + and $\cdot$ have the following properties:

For all complex numbers (a, b), (c, d), (e, d)

- *associativity*:
$$(a, b) + [(c, d) + (e, f)] = [(a, b) + (c, d)] + (e, f),$$
$$(a, b) \cdot [(c, d) \cdot (e, f)] = [(a, b) \cdot (c, d)] \cdot (e, f);$$
- *commutativity*:
$$(a, b) + (c, d) = (c, d) + (a, b), \quad (a, b) \cdot (c, d) = (c, d) \cdot (a, b);$$
- *distributivity*:
$$(a, b) \cdot [(c, d) + (e, f)] = (a, b) \cdot (c, d) + (a, b) \cdot (e, f).$$

Let's prove the distributivity. Using the definition of addition and multiplication of complex numbers and the properties of real numbers (associativity, commutativity and distributivity) we get:

$$\begin{aligned}(a,b)\cdot\Big[(c,d)+(e,f)\Big] &= (a,b)\cdot(c+e,d+f)\\ &= \Big(a\cdot(c+e)-b\cdot(d+f), a\cdot(d+f)+b\cdot(c+e)\Big)\\ &= (ac+ae-bd-bf, ad+af+bc+be)\end{aligned}$$

On the other hand,

$$\begin{aligned}(a,b)(c,d)+(a,b)(e,f) &= (ac-bd, ad+bc)+(ae-bf, af+be)\\ &= (ac-bd+ae-bf, ad+bc+af+be).\end{aligned}$$

L.h.s. and r.h.s coinside.

Exercize 2.2. Prove the associativity and commutativity of addition and multiplication of complex numbers.

Zero. *Zero* (or *null* or *identity* w.r.t. addition) is a complex number (x, y), such that for any (a, b) it holds that $(a, b) + (x, y) = (a, b)$.

From the definition we get $a + x = a$, $b + y = b$, hence $x = 0$, $y = 0$.

So, zero is the pair $(0, 0)$ and there are no other zeros.

Opposite number. The number *opposite* to (a, b) is a pair (x, y), such that $(a, b) + (x, y) = (0, 0)$.

The opposite number is denoted by $-(a, b)$.

It is not hard to see that $-(a, b) = (-a, -b)$ (prove this).

Difference. The *difference* (or the number obtained as a result of *subtraction*) of two complex numbers (c, d) and (a, b) is the solution (x, y) of the equation

$$(a, b) + (x, y) = (c, d).$$

The difference is denoted by $(c, d) - (a, b)$. It is obvious that $(c, d) - (a, b) = (c - a, d - b)$.

It is not hard to see that the difference $(c, d) - (a, b)$ is a sum of (c, d) and a number opposite to (a, b), i.e.

$$(c, d) - (a, b) = (c, d) + [\,-(a, b)].$$

Unit. *One* (or unit or identity w.r.t multiplication) is a complex number (x, y) such that for any number (a, b) it holds that

$$(a, b)(x, y) = (a, b).$$

From the definition we get

$$\begin{cases} ax - by = a, \\ bx + ay = b. \end{cases}$$

If $a^2 + b^2 \neq 0$, i.e. $(a, b) \neq (0, 0)$, the system hase the only solution:

$$x = 1, \quad y = 0.$$

So, $(a, b)(1, 0) = (a, b)$ for any (a, b) (for $(a, b) = (0, 0)$ this also can be verified).

Thus, $(1, 0)$ is identity w.r.t multiplication.

Inverse number. The number *inverse* to (a, b) is a number (x, y), such that

$$(a, b)(x, y) = (1, 0).$$

The inverse number is denoted by $(a, b)^{-1}$.

If $(a, b) \neq (0, 0)$ the system

$$\begin{cases} ax - by = 1, \\ bx + ay = 0 \end{cases}$$

has the only solution

$$(a, b)^{-1} = \left(\frac{a}{a^2 + b^2}, \;\; -\frac{b}{a^2 + b^2} \right). \tag{2.2}$$

If $(a, b) = (0, 0)$ the system has no solution. Hence there is no inverse number to zero.

Division. The result (the *quotient*) of the *division* (c, d) by (a, b) is the solution (x, y) of the equation $(a, b)(x, y) = (c, d)$. The quotient is denoted by

$$\frac{(c, d)}{(a, b)} \quad \text{or} \quad (c, d)/(a, b).$$

The equation $(a, b)(x, y) = (c, d)$ is equivalent to the system

$$\begin{cases} ax - by = c, \\ bx + ay = d. \end{cases} \tag{2.3}$$

Let's multiply the first equation by a, and the second one by b. Then add the results. We obtain:

$$(a^2 + b^2)x = ac + bd.$$

Now multiply the first equation by b, and the second one by a. Subtract the first from the second. We obtain:

$$(a^2 + b^2)y = ad - cb.$$

If $a^2 + b^2 \neq 0$ we have

$$\frac{(c, d)}{(a, b)} = \left(\frac{ac + bd}{a^2 + b^2}, \frac{ad - cb}{a^2 + b^2} \right). \tag{2.4}$$

If $a^2 + b^2 = 0$, then the result of division is undefined.

Comparing (2.2) and (2.4), we obtain that the quotient $(c, d)/(a, b)$ is the product (c, d) with the number inverse to (a, b):

$$\frac{(c, d)}{(a, b)} = (c, d) \cdot (a, b)^{-1}.$$

Algebraic form of complex numbers. Let's consider the complex numbers kind of $(a, 0)$ as real numbers. More precisely, let

$$(a, 0) = a.$$

It is not hard to verify that

$$(a, 0) + (b, 0) = (a + b, 0), \quad (a, 0)(b, 0) = (ab, 0)$$

for any a, b.

So, *the results of arithmetic operations under real numbers don't depend on the way how they were obtained: using the rules for real numbers or using the rules for complex numbers.*

Now denote $i = (0, 1)$. The number i is called *imaginary unit*.

It is not hard to verify that $(0, b) = (b, 0) \cdot i = bi$, hence

$$(a, b) = a + bi.$$

This is an *algebraic form of a complex number*.

It is not hard to verify that $i^2 = -1$, hence *one can operate with complex numbers as with ordinary algebraic expressions involving "symbol" i substituting i^2 to -1.*

For example, open the parenthesis in

$$(a + bi)(c + di).$$

We get

$$ac + adi + bci + bdi^2.$$

Since $i^2 = -1$, then $(a + bi)(c + di) = ac - bd + (ad - bd)i$.

Now consider the quotient $(c+di)/(a+bi)$. We can multiply the numerator and the denominator by the number $c-di$ (it's called *conjugate* to $c+di$). We get

$$\frac{c+di}{a+bi}=\frac{(c+di)(a-bi)}{(a+bi)(a-bi)}=\frac{ac+bd+i(ad-bc)}{a^2+b^2}=\frac{ac+bd}{a^2+b^2}+\frac{ad-bc}{a^2+b^2}i.$$

The result coincides with (2.4).

So, arithmetic operations are defined in the most obvious way subject to the convention that $i^2=-1$.

Now we have a (complex) number whos square is -1. There is no real number with such property. Indeed, we can indicate another complex number whoes square is -1. It's $-i$.

Example 2.3.

a) $(4+i)(5+3i)+(3+i)(3-2i)=20+12i+5i-3+9-6i+3i+2=28+14i$;

b)

$$\begin{aligned}\frac{(5+i)(7-6i)}{3+i}&=\frac{35-30i+7i+6}{3+i}\\&=\frac{41-23i}{3+i}=\frac{(41-23i)(3-i)}{(3+i)(3-i)}\\&=\frac{123-41i-69i-23}{10}=\frac{100-110i}{10}=10-11i.\end{aligned}$$

Exercize 2.4. Calculate

a) $\dfrac{(5+i)(3+5i)}{2i}$; b) $\dfrac{(2+i)(4+i)}{1+i}$.

Usually complex numbers are denoted by one letter. For example, $z=a+bi$, where a is called the real part of z, and b is called the imaginary part of z.

We also will use the notations: $\operatorname{Re} z=a$, $\operatorname{Im} z=b$.

2.3. Geometric Representation of Complex Numbers

A real number can be considered as a point on the line.

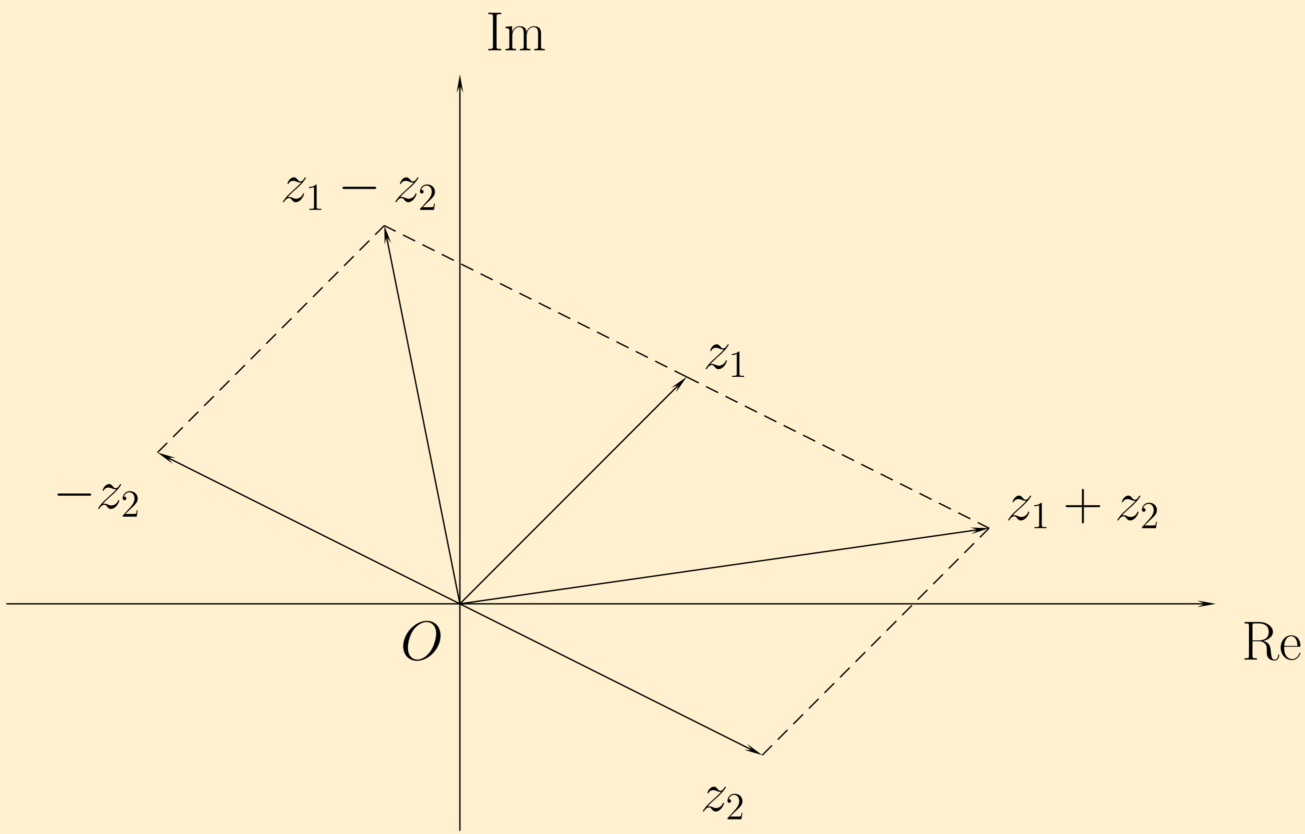
A complex number can be considered as a point in the plane. The point can be identified using its Cartesian coordinates.

Thus, complex number $(a, b) = a + bi$ identifies a point (and corresponding radius–vector) whose x coordinate is a and y coordinate is b.

In this case the x axes is called the *real* axes, and the y axes is called the *imaginary* axes.

The real axes corresponds to the set of all real numbers. The imaginary axes corresponds to the set of imaginary numbers, i.e. to the numbers kind of bi.

The sum and difference of complex numbers are represented by the sum and difference of their radius-vectors.



The absolute value of a complex number $z = a + bi$ is the distance r between the point z and the origin (Fig. **??**).

The absolute value is denoted by $|z|$.

It is obvious that the old definition of absolute value for real numbers remains correct.

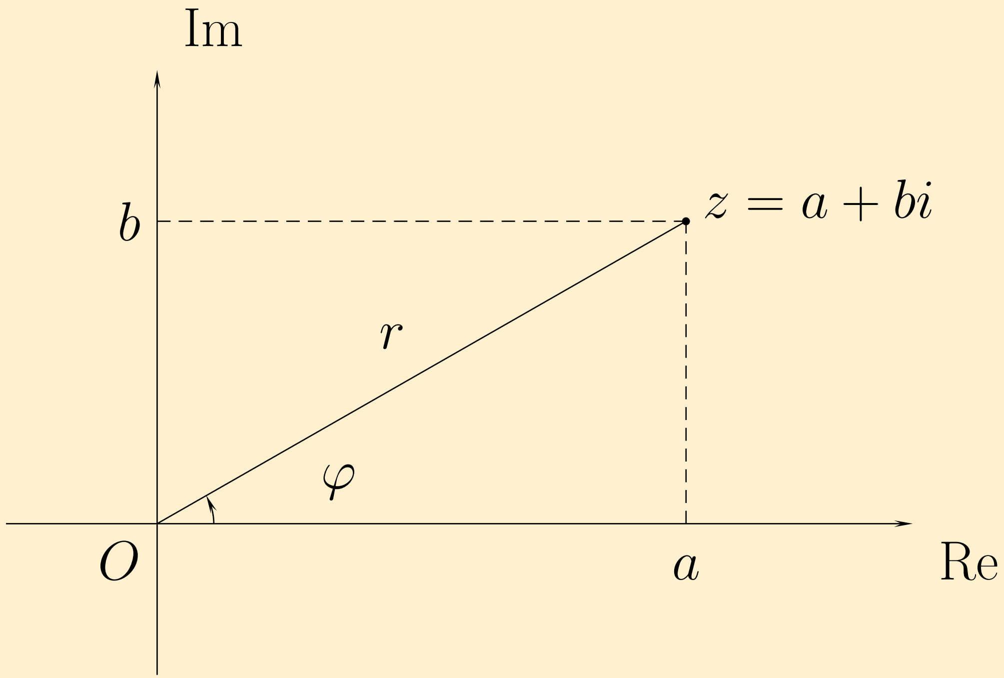
It is obvious that $|z| = 0$, iff (if and only if) $z = 0$.

The *argument* of z is the angle φ, in positive (anti-clock wise) direction from the Re to the radius-vector of z.

The argument is denoted by $\arg z$.

The argument is unique up to $2\pi k$, where $k \in \mathbb{Z}$.

The number 0 doesn't have an argument.



Im
$z = a + bi$
b
r
φ
O
a
Re

Using the Pythagorean theorem and trigonometric formulas we get

$$|a + bi| = \sqrt{a^2 + b^2}, \quad a = r\cos\varphi, \quad b = r\sin\varphi,$$

hence

$$a + bi = r(\cos\varphi + i\sin\varphi).$$

The notion $r(\cos\varphi + i\sin\varphi)$ is called the *trigonometric form* of the complex number $a + bi$.

0 has not a trigonometric form.

Example 2.5. Find the trigonometric form:

a) $8 = 8(\cos 0 + i \sin 0)$;

b) $i = \cos \frac{\pi}{2} + i \sin \frac{\pi}{2}$;

c) $-8 = 8(\cos \pi + i \sin \pi)$;

d) $1 - i = \sqrt{2} \left(\cos \left(-\frac{\pi}{4} \right) + i \sin \left(-\frac{\pi}{4} \right) \right)$, indeed, $|1 - i| = \sqrt{2}$, one of the solution to the system

$$\begin{cases} 1 = \sqrt{2} \cos \varphi, \\ -1 = \sqrt{2} \sin \varphi \end{cases}$$

is $\varphi = -\pi/4$;

e) $-\sqrt{3} - i = 2 \left(\cos \left(-\frac{5\pi}{6} \right) + i \sin \left(-\frac{5\pi}{6} \right) \right)$;

f) $\cos \alpha - i \sin \alpha = \cos(-\alpha) + i \sin(-\alpha)$;

g) $\sin \alpha + i \cos \alpha = \cos \left(\frac{\pi}{2} - \alpha \right) + i \sin \left(\frac{\pi}{2} - \alpha \right)$;

h) $1+\cos\varphi+i\sin\varphi$, if $\varphi\in[-\pi,\pi]$: using the formula for double angle we get:

$$2\cos^2\frac{\varphi}{2}+i2\cos\frac{\varphi}{2}\sin\frac{\varphi}{2}=2\cos\frac{\varphi}{2}\left(\cos\frac{\varphi}{2}+i\sin\frac{\varphi}{2}\right).$$

Exercize 2.6. Find the trigonometric form:

a) $1 - i$; b) $1 - i\sqrt{3}$; c) $\dfrac{1 + i \operatorname{tg} \alpha}{1 - i \operatorname{tg} \alpha}$.

Example 2.7. Let's describe the set of points corresponding to the complex numbers z such that $|z - z_0| = r$, where $z \in \mathbb{R}, z_0 \in \mathbb{R}, r \in \mathbb{R}, r > 0$.

1st method. Let $z = x + yi$, $z_0 = x_0 + y_0 i$, thet the initial equation is equivalent to $(x - x_0)^2 + (y - y_0)^2 = r^2$. This describes a circle with radius r and with center $z_0 = x_0 + iy_0$.

2nd method. $|z - z_0|$ is a distance between points z and z_0. Hence the initial equation describes the set of all points z, such that the distance between them and the fixed point z_0 is constant equaled to r. This is a circle.

Exercize 2.8. Describe the set of all points z, such that $|z - z_1| = |z - z_2|$.

Exercize 2.9. Plot the set of all points z such that

a) $1 \leq |z - 2 + i| < 2$; b) $|\arg z| < \pi/4$; c) $\operatorname{Im} z = 3$.

2.3.1. Multiplication and division of complex numbers in trigonometric form

Let

$$z_1 = r(\cos\varphi + i\sin\varphi)$$

$$z_2 = \rho(\cos\psi + i\sin\psi).$$

Open parenthesis in

$$z_1 z_2 = r(\cos\varphi + i\sin\varphi)\cdot\rho(\cos\psi + i\sin\psi).$$

and use the trigonometric formulas for sum of angles:

$$\begin{aligned} z_1 z_2 &= r\rho(\cos\varphi\cos\psi - \sin\varphi\sin\psi + i(\sin\varphi\sin\psi + \cos\varphi\cos\psi)) \\ &= r\rho\Big[\cos(\varphi+\psi) + i\sin(\varphi+\psi)\Big]. \end{aligned}$$

We have the complex number written in trigonometric form. Its absolute value is $r\rho$ and argument is $\varphi + \psi$:

$$|z_1 z_2| = |z_1|\cdot|z_2|, \quad \arg(z_1 z_2) = \arg z_1 + \arg z_2 + 2\pi k \quad (k \in \mathbb{Z}).$$

Thus to obtain the radius-vector of the product $z_1 z_2$ we should rotate the radius-vector of z_1 to the angle $\arg(z_2)$ and stretch it $|z_2|$ times.

If $\rho \neq 0$ then

$$\Big[\rho(\cos\psi + i\sin\psi)\Big]^{-1} = \rho^{-1}\Big[\cos(-\psi) + i\sin(-\psi)\Big],$$

hence

$$\frac{r(\cos\varphi + i\sin\varphi)}{\rho(\cos\psi + i\sin\psi)} = \frac{r}{\rho}\Big(\cos(\varphi - \psi) + i\sin(\varphi - \psi)\Big),$$

i.e.

$$\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}, \qquad \arg\frac{z_1}{z_2} = \arg z_1 - \arg z_2.$$

Thus, to obtain the radius-vector of the result of division z_1/z_2 we should rotate the radius-vector of z_1 to the angle $-\arg(z_2)$ and compress it $|z_2|$ times.

Example 2.10. Let's perform the operations:

a)

$$\begin{aligned}
&(1+i\sqrt{3})(1+i)(\cos\varphi + i\sin\varphi)\\
&\quad = 2\left(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}\right)\sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right)(\cos\varphi + i\sin\varphi)\\
&\quad = 2\sqrt{2}\left(\cos\left(\frac{7\pi}{12} + \varphi\right) + i\sin\left(\frac{7\pi}{12} + \varphi\right)\right);
\end{aligned}$$

b)

$$\frac{\cos\varphi + i\sin\varphi}{\cos\psi - i\sin\psi} = \frac{\cos\varphi + i\sin\varphi}{\cos(-\psi) + i\sin(-\psi)} = \cos(\varphi+\psi) + i\sin(\varphi+\psi).$$

Example 2.11. Let's find the trigonometric form of the number $1+\cos\varphi+i\sin\varphi$, where $\varphi\in[-\pi,\pi]$.

This example has been considered (Example 2.5 d).

Let's consider another method.

$$\begin{aligned}
&1+\cos\varphi+i\sin\varphi\\
&\quad=(\cos 0+i\sin 0)+(\cos\varphi+i\sin\varphi)\\
&\quad=\left(\cos(-\frac{\varphi}{2})+i\sin(-\frac{\varphi}{2})\right)\left(\cos\frac{\varphi}{2}+i\sin\frac{\varphi}{2}\right)+\left(\cos\frac{\varphi}{2}+i\sin\frac{\varphi}{2}\right)^2\\
&\quad=2\cos\frac{\varphi}{2}\left(\cos\frac{\varphi}{2}+i\sin\frac{\varphi}{2}\right).
\end{aligned}$$

2.3.2. Powers

Let n be a natural number.

As for real numbers we'll say that ζ is the n-th power of z, $\zeta = z^n$, iff

$$\zeta = \underbrace{z \cdot z \cdot \ldots \cdot z}_{n \text{ multipliers}}.$$

We have defined z^{-1} as a number inverse to $z \neq 0$. For any $z \neq 0$ we define $z^0 = 1$ and $z^{-n} = (z^{-1})^n$. Let's prove that

$$(z^{-1})^n = (z^n)^{-1}.$$

We have

$$(z^{-1})^n z^n = z^{-1} \cdot \ldots \cdot z^{-1} z \cdot \ldots \cdot z = (z^{-1} \cdot \ldots \cdot (z^{-1} z) \cdot \ldots \cdot z) = 1.$$

Exercize 2.12. Prove that

$$(z^m)^n = z^{mn}, \quad z^m z^n = z^{m+n}, \quad z_1^n z_2^n = (z_1 z_2)^n$$

for any integers m, n.

Using mathematical induction we get *Moivre's formulas* (A. de Moivre, 1736):

$$\Big[r(\cos\varphi + i\sin\varphi)\Big]^n = r^n\Big[\cos(n\varphi) + i\sin(n\varphi)\Big],$$

where n is any integer number.

Example 2.13. Calculate:

$$\begin{aligned}
(1+i\sqrt{3})^{150} &= 2\left(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}\right)^{150} \\
&= 2^{150}\left(\cos\frac{150\pi}{3} + i\sin\frac{150\pi}{3}\right) \\
&= 2^{150}\Big[\cos(50\pi) + i\sin(50\pi)\Big] \\
&= 2^{150}\left(\cos 0 + i\sin 0\right) = 2^{150}.
\end{aligned}$$

Example 2.14. Let's prove that if $z + \dfrac{1}{z} = 2\cos\vartheta$, then

$$z^n + \frac{1}{z^n} = 2\cos n\vartheta. \tag{2.5}$$

The equation $z + \dfrac{1}{z} = 2\cos\vartheta$ is equavalent to the square one: $z^2 - 2z\cos\vartheta + 1 = 0$.

Its roots are

$$z_{1,2} = \cos\vartheta \pm \sqrt{\cos^2\vartheta - 1} = \cos\vartheta \pm i\sin\vartheta.$$

Hence $z_{1,2}^n = \cos n\vartheta \pm i\sin n\vartheta$, $z_{1,2}^{-n} = \cos n\vartheta \mp i\sin n\vartheta$. The equality to prove follows immediately.

Exercize 2.15. Calculate:

a) $\dfrac{(-1+i\sqrt{3})^{15}}{(1-i)^{20}}+\dfrac{(-1-i\sqrt{3})^{15}}{(1+i)^{20}}$;

b) $(1+\cos\varphi+i\sin\varphi)^n$.

2.4. Complex conjugate numbers

Recall, that $\overline{z} = a - bi$ is called *conjugate* of $z = a + bi$.

It is not hard to see that $\overline{\overline{z}} = z$.

If z is in trigonometric form

$$z = r(\cos\varphi + i\sin\varphi),$$

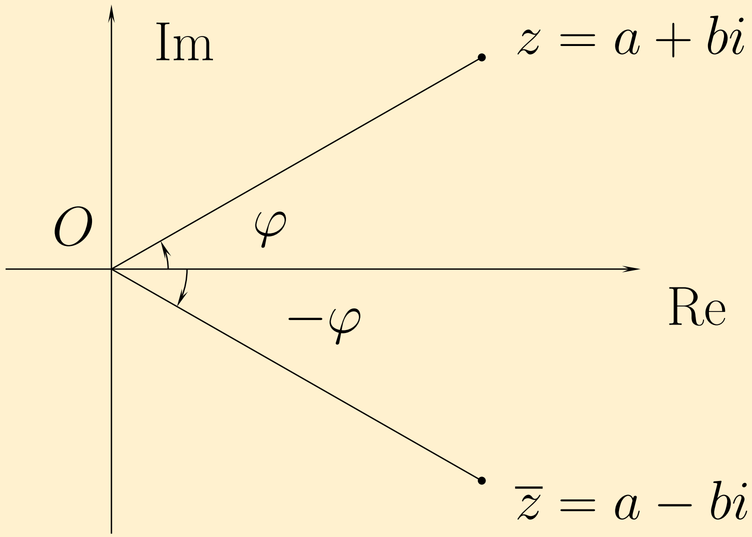
then

$$\bar{z} = r(\cos\varphi - i\sin\varphi) = r(\cos(-\varphi) + i\sin(-\varphi))$$

(see Fig. **??**).

Thus, the arguments of two conjugate numbers are different in sign. The moduluses are equal:

$$|\bar{z}| = |z|, \quad \arg\bar{z} = -\arg z.$$



Im
$z = a + bi$
O
φ
Re
$-\varphi$
$\overline{z} = a - bi$

Properties:

$$\begin{aligned} \overline{z_1 + z_2} &= \bar{z}_1 + \bar{z}_2, \qquad \overline{z_1 - z_2} = \bar{z}_1 - \bar{z}_2, \\ \overline{z_1 \cdot z_2} &= \bar{z}_1 \cdot \bar{z}_2, \qquad \overline{z_1 / z_2} = \bar{z}_1 / \bar{z}_2; \\ z\bar{z} &= |z|^2, \qquad z + \bar{z} = 2 \operatorname{Re} z. \end{aligned} \tag{2.6}$$

Let's prove the first one: $z_1 = a + bi$, $z_2 = c + di$, then

$\overline{z_1 + z_2} = \overline{(a + c) + (b + d)i} = (a + c) - (b + d)i$

On the other hand, $\bar{z}_1 + \bar{z}_2 = a - bi + c - di = (a + c) - (b + d)i$.

LHS = RHS

Exercize 2.16. Prove remain properties (2.6).

Example 2.17. Find the solution $\bar{z} = z^3$.

1 method. Let $z = x + iy$. Then

$$x - iy = x^3 + 3x^2yi - 3xy^2 - y^3i.$$

Let's equate real and imaginary parts. We obtain the aggregate of systems:

$$\begin{cases} x = x^3 - 3xy^2, \\ -y = 3x^2y - y^3 \end{cases} \quad \text{OR} \quad \begin{cases} x(x^2 - 3y^2 - 1) = 0, \\ y(y^2 - 3x^2 - 1) = 0, \end{cases}$$

Hence

$$\begin{cases} x = 0, \\ y = 0 \end{cases} \text{or} \begin{cases} x = 0, \\ y^2 - 3x^2 = 1 \end{cases}$$
$$\text{or} \begin{cases} x^2 - 3y^2 = 1, \\ y = 0 \end{cases} \text{or} \begin{cases} x^2 - 3y^2 = 1, \\ y^2 - 3x^2 = 1, \end{cases}$$

Then

$$\begin{cases} x = 0, \\ y = 0 \end{cases} \text{or} \begin{cases} x = 0, \\ y^2 = 1 \end{cases} \text{or} \begin{cases} x^2 = 1, \\ y = 0 \end{cases} \text{or} \begin{cases} x^2 - 3y^2 = 1, \\ y^2 - 3x^2 = 1. \end{cases}$$

The last system is inconsistent in $\mathbb{R}$. Indeed, multiplying the first equation by 3 and adding them with the second one we obtain $8y^2 = -4$, that's impossible for real y.

From the 1st, 2nd and 3rd system we get the ANSWER: 0, $\pm i$, ± 1.

2 method. Let's write z in trigonometric form: $z = r(\cos\varphi + i\sin\varphi)$.

The equation has the form:

$$r(\cos(-\varphi) + i\sin(-\varphi)) = r^3(\cos 3\varphi + i\sin 3\varphi),$$

Hence

$$\begin{cases} r = r^3, \\ -\varphi = 3\varphi + 2\pi k,\ \ k \in \mathbb{Z}; \end{cases} \quad \text{hence} \quad \begin{cases} r = 0 \quad \text{or} \quad r = 1, \\ \varphi = \dfrac{\pi k}{2},\ \ k \in \mathbb{Z}. \end{cases}$$

The *solutions* to the system: 0, $\pm i$, ± 1.

Exercize 2.18. Solve equations:

a) $|z| + z = 8 + 4i$;

b) $\bar{z} = z^2$;

c) $|z| - iz = 1 - 2i$;

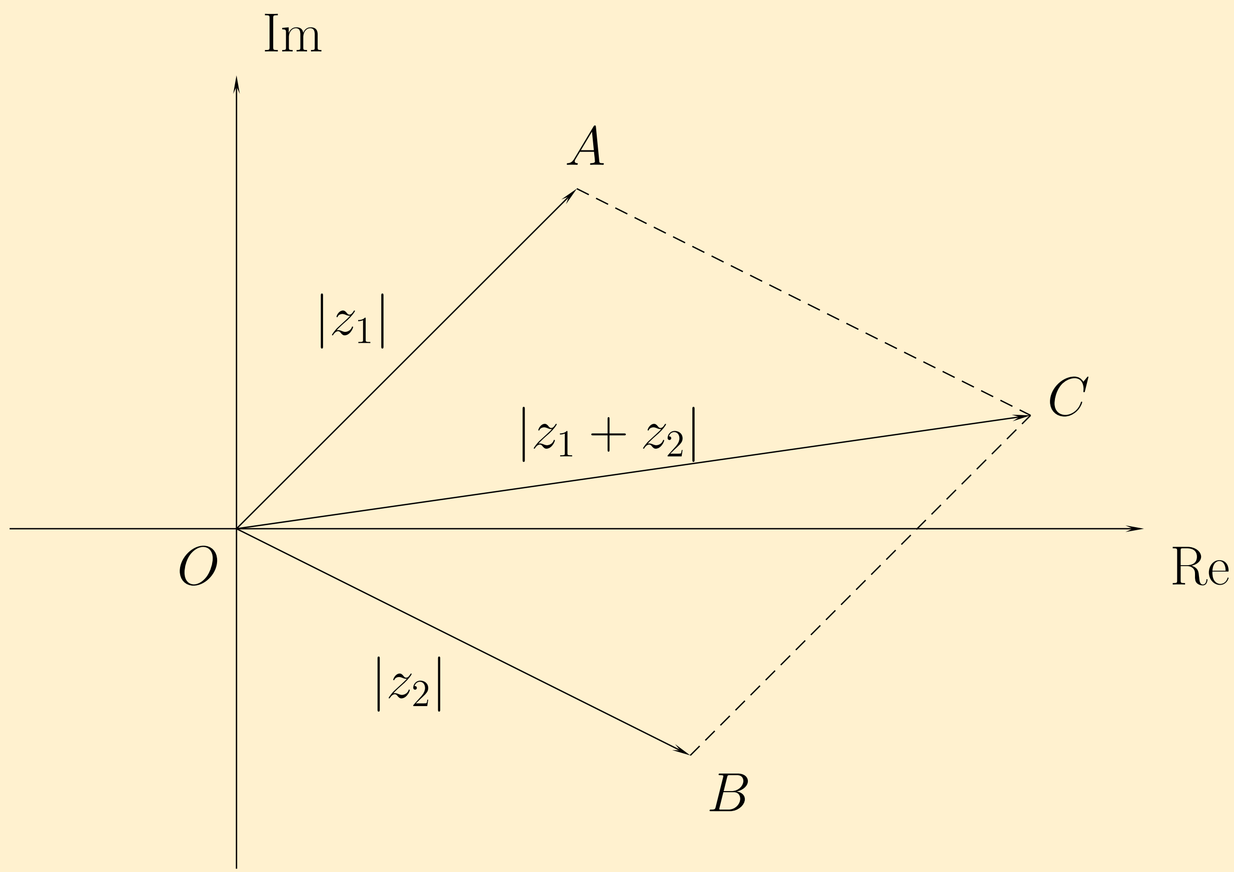
d) $z^2 = \bar{z}^3$;

e) $z^2 + z|z| + |z^2| = 0$.

2.5. Triangle inequality

Proposition 2.19 (Triangle inequality)**.** *For any complex numbers z_1, z_2 the following inequality holds*

$$|z_1| - |z_2| \leq ||z_1| - |z_2|| \leq |z_1 + z_2| \leq |z_1| + |z_2|.$$



Example 2.20. Prove that if $|z| < \frac{1}{2}$ then

$$|(1+i)z^3 + iz| < \frac{3}{4}.$$

From the triangle inequality we get:

$$|(1+i)z^3 + iz| \leq |(1+i)z^3| + |iz|.$$

Then

$$|(1+i)z^3| + |iz||1+i||z|^3 + |i||z| = \sqrt{2}|z|^3 + |z|.$$

Since

$$|z| < \frac{1}{2},$$

then

$$\sqrt{2}|z|^3 + |z| < \sqrt{2}\frac{1}{2^3} + \frac{1}{2} < \frac{3}{4}.$$

Exercize 2.21. Prove that if $|z| \leq 2$, then $1 \leq |z^2 - 5| \leq 9$.

2.6. Roots of complex numbers

Let $n \in \mathbb{N}$, $\zeta \in \mathbb{C}$.

The number z is called a value of *the root of the n-th degree* from ζ, if $z^n = \zeta$.

It is not hard to see that $\zeta = 0$ has the only value (0) of the root of any natural degree.

Let $\zeta \neq 0$, then $z \neq 0$.

Let represent ζ in the trigonomtric form:

$$\zeta = \rho(\cos\psi + i\sin\psi).$$

We want to find $z = r(\cos\varphi + i\sin\varphi)$ such that $z^n = \zeta$.

Using Moivre's formulas, we obtain

$$r^n\Big[\cos(n\varphi) + i\sin(n\varphi)\Big] = \rho(\cos\psi + i\sin\psi).$$

Since for any non-zero complex number the modulus is unique and the argument is unique up to $2\pi k$, where $k \in \mathbb{Z}$, then

$$r^n = \rho, \qquad n\varphi = \psi + 2\pi k.$$

We obtain

$$r = \sqrt[n]{\rho}, \qquad \varphi = \frac{\psi + 2\pi k}{n}$$

Thus, for any $k \in \mathbb{Z}$ every number

$$z = \sqrt[n]{\rho}\left(\cos\frac{\psi + 2\pi k}{n} + i\sin\frac{\psi + 2\pi k}{n}\right) \tag{2.7}$$

is a value of the root of n-th degree from

$$\zeta = \rho(\cos\psi + i\sin\psi).$$

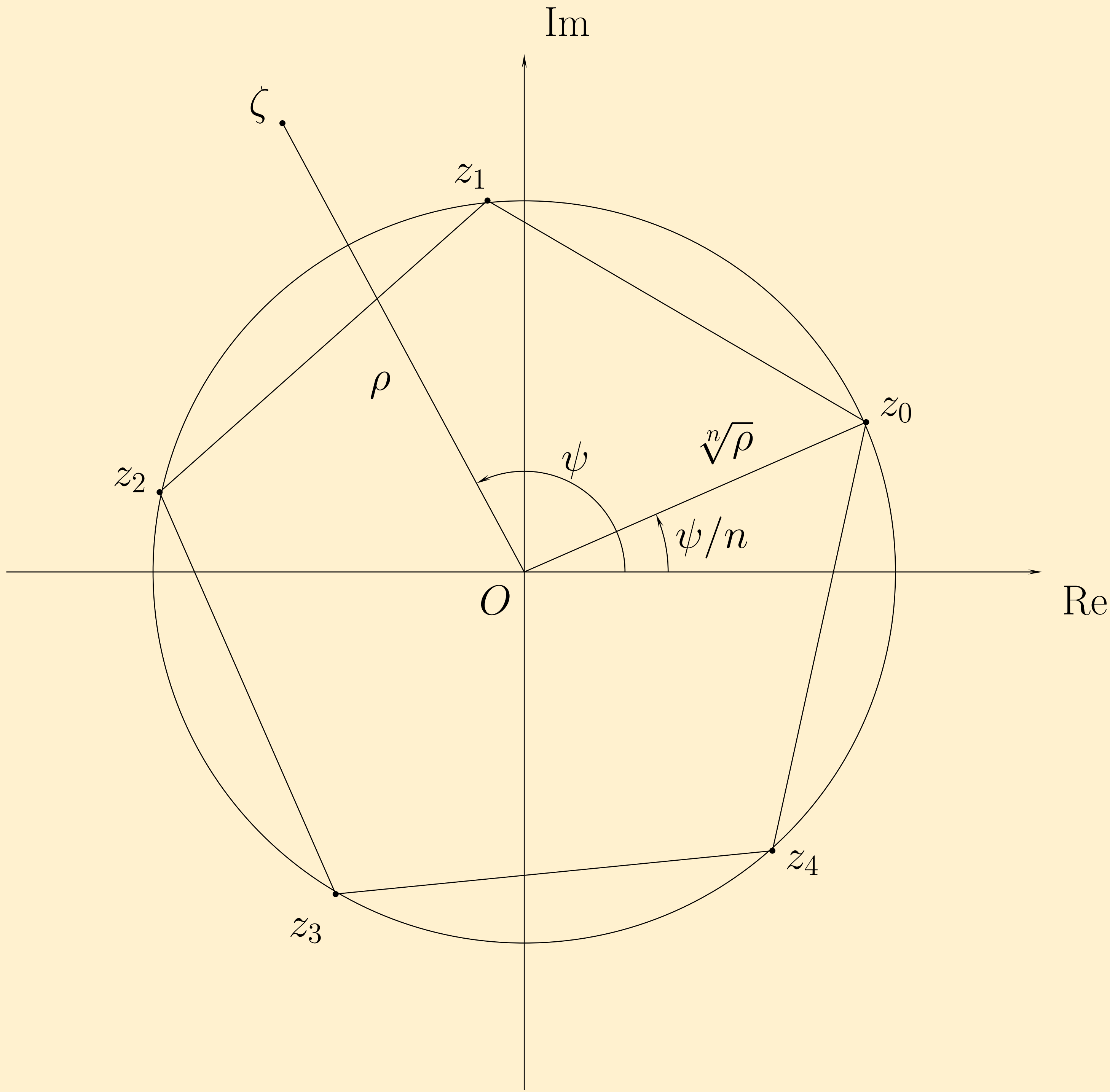
It is enough to consider only $k = 0, 1, \ldots, n-1$. Thus, *any non-zero complex number* $\zeta = \rho(\cos\psi + i\sin\psi)$ *has* n *different values of* n*-th degree:*

$$\sqrt[n]{\zeta} = \sqrt[n]{\rho}\left(\cos\frac{\psi + 2\pi k}{n} + i\sin\frac{\psi + 2\pi k}{n}\right) \quad (k = 0, 1, \ldots, n-1). \tag{2.8}$$

In the complex plane all values of a root is on the equal distance ($\sqrt[n]{\rho}$) from origin.

The angle between two adjacent radious-vectors corresponding to two values of the root is equal to $2\pi/n$.

Thus, points correspoding to all values of the root of degree $n \geq 3$ are the vertices of a regular n-polygon.



Im
Re
ζ
z_1
z_0
z_2
z_3
z_4
ρ
ψ
$\sqrt[n]{\rho}$
ψ/n
O

Example 2.22. Find all values of $\sqrt[3]{-8}$.

What value of $\sqrt[3]{-8}$ do you just know?

$-8 = 8(\cos\pi + i\sin\pi)$. Using (2.8) we obtain:

$$\sqrt[3]{-8} = \sqrt[3]{8}\left(\cos\frac{\pi+2\pi k}{3} + i\sin\frac{\pi+2\pi k}{3}\right) \quad (k = 0, 1, 2).$$

So, $\sqrt[3]{-8}$ has the following values:

$$\begin{aligned} z_0 &= 2\left(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}\right) = 1 + i\sqrt{3}, \\ z_1 &= 2\left(\cos\pi + i\sin\pi\right) = -2, \\ z_2 &= 2\left(\cos\frac{5\pi}{3} + i\sin\frac{5\pi}{3}\right) = 1 - i\sqrt{3}. \end{aligned}$$

Exercize 2.23. Calculate

a) $\sqrt[4]{-\dfrac{18}{1+i\sqrt{3}}}$;

b) $\sqrt[3]{\dfrac{1-5i}{1+i} - 5\dfrac{1+2i}{2-i} + 2}$.

Exercize 2.24. Find an error in the following "proof":

Since

$$\sqrt{a} \cdot \sqrt{b} = \sqrt{ab};$$

then

$$\sqrt{-1} \cdot \sqrt{-1} = \sqrt{(-1) \cdot (-1)};$$

hence

$$(\sqrt{-1})^2 = \sqrt{1}.$$

Thus $-1 = 1$.

Example 2.25.

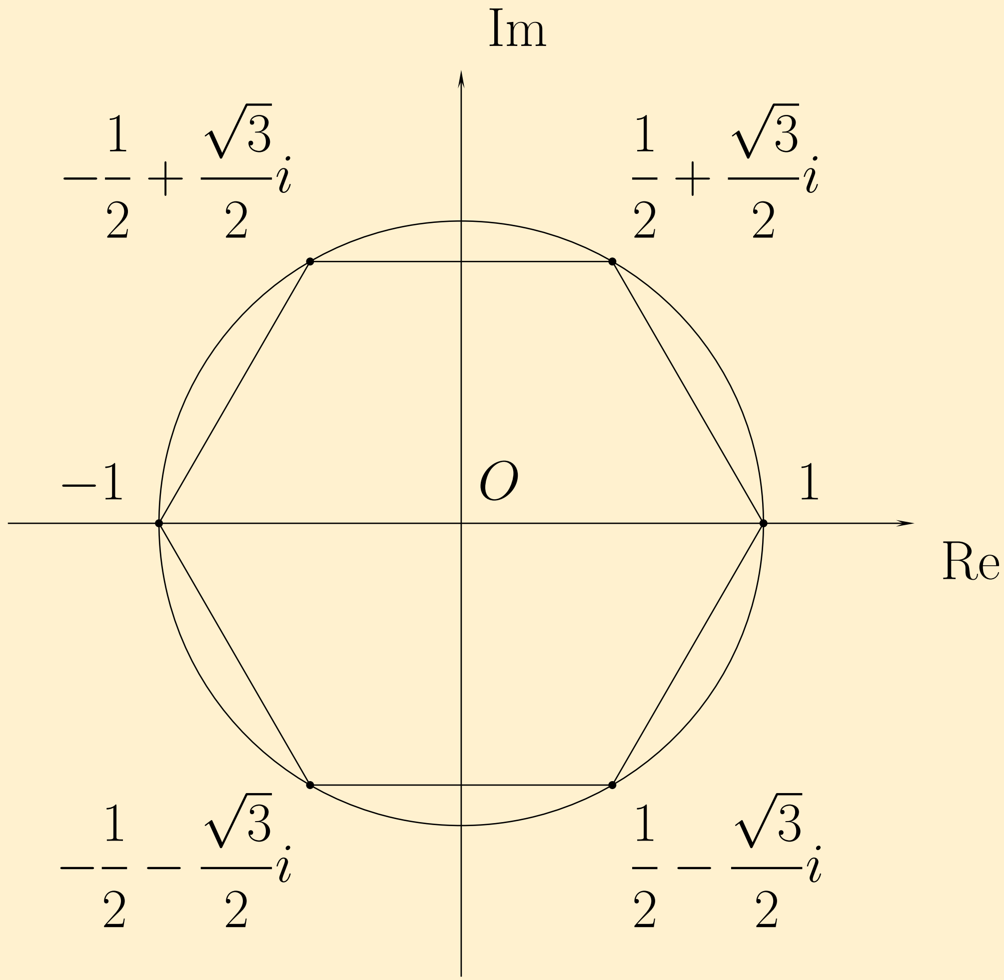
a) $\sqrt[1]{1} = 1$ (one value).

b) $\sqrt[2]{1} = \{1, -1\}$ (two values).

c) $\sqrt[3]{1}$. $\varepsilon_0 = 1$, $\varepsilon_1 = \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3} = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$, $\varepsilon_2 = \cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3} = -\frac{1}{2} - i\frac{\sqrt{3}}{2}$.

d) $\sqrt[4]{1} = \{\pm 1, \pm i\}$.

e) $\sqrt[6]{1} = \left\{\pm 1, \pm\frac{1}{2} \pm i\frac{\sqrt{3}}{2}\right\}$



Im
$-\frac{1}{2}+\frac{\sqrt{3}}{2}i$
$\frac{1}{2}+\frac{\sqrt{3}}{2}i$
-1
O
1
Re
$-\frac{1}{2}-\frac{\sqrt{3}}{2}i$
$\frac{1}{2}-\frac{\sqrt{3}}{2}i$

Exercize 2.26. Find all values of $\sqrt[12]{1}$.

Example 2.27. Find the sum of all values of n-th degree from 1.

If $n = 1$ the sum is obviously is 1.

If $n > 1$, then $\varepsilon_k = (\varepsilon_1)^k$, and:

$$1 + \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_{n-1} = 1 + \varepsilon_1 + \varepsilon_1^2 + \ldots + \varepsilon_1^{n-1} = \frac{1 - \varepsilon_1^n}{1 - \varepsilon_1}.$$

(This is the sum of a *geometric progression*) Since $(\varepsilon_1)^n = 1$, the sum is 0.

Exercize 2.28.

a) Find the product of all values of the root of n-th degree from 1.

2.6.1. Applications

Let's derive the formulae

$$\cos 2\varphi = \cos^2\varphi - \sin^2\varphi, \quad \sin 2\varphi = 2\sin\varphi\cos\varphi \quad (\varphi \in \mathbb{R})$$

Consider the complex number $z = \cos\varphi + i\sin\varphi$

Using the formula for the square of a sum we obtain

$$z^2 = (\cos\varphi + i\sin\varphi)^2 = \cos^2\varphi + 2i\cos\varphi\sin\varphi - \sin^2\varphi.$$

From another hand, using Moivre's fromula we obtain

$$z^2 = (\cos\varphi + i\sin\varphi)^2 = \cos 2\varphi + \sin 2\varphi.$$

Thus,

$$\operatorname{Re} z^2 = \cos 2\varphi = \cos^2\varphi - \sin^2\varphi, \quad \operatorname{Im} z^2 = \sin 2\varphi = 2\cos\varphi\sin\varphi$$

Exercize 2.29. Find formulae for $\cos 3\varphi$, $\sin 3\varphi$, $\cos 4\varphi$, $\cos 4\varphi$.

2.7. Square Roots of Complex Numbers

The square root of $z = a + bi$ is $x + iy$, such that

$$(x + iy)^2 = a + bi.$$

Let $z \neq 0$ then $a + bi = x^2 + 2xyi - y^2$, hence

$$\begin{cases} a = x^2 - y^2, \\ b = 2xy. \end{cases} \tag{2.9}$$

Raise the equations to the square and add them:

$$a^2 + b^2 = x^4 + 2x^2y^2 + y^2,$$

hence

$$a^2 + b^2 = (x^2 + y^2)^2.$$

Since $x^2 + y^2 > 0$, then $x^2 + y^2 = \sqrt{a^2 + b^2}$.

We have

$$\begin{cases} x^2 + y^2 = \sqrt{a^2 + b^2}, \\ x^2 - y^2 = a. \end{cases}$$

Then

$$
\begin{aligned}
x^2 &= \frac{1}{2}(\quad a + \sqrt{a^2 + b^2}), \\
y^2 &= \frac{1}{2}(-a + \sqrt{a^2 + b^2}).
\end{aligned}
\tag{2.10}
$$

Some of the roots of the system can be irrelevant!

Each of these relations gives us two values for x and y.

Combining them we obtain 4 different complex numbers, but only two of them satisfy the original system (2.9):

it is not hard to see from the second equation that *if* $b > 0$ *then the signes of* x *and* y *coincide, if* $b < 0$ *then the signes of* x *and* y *differ.*

Example 2.30. Let's find all values of $\sqrt[4]{2 - i\sqrt{12}} = \sqrt{\sqrt{2 - i\sqrt{12}}}$.

First find all values of $\sqrt{2 - i\sqrt{12}}$.

From (2.10) we have $x^2 = 3$, $y^2 = 1$.

Since the $\operatorname{Im}(2 - i\sqrt{12}) < 0$ then $\sqrt{2 - i\sqrt{12}} = \pm(\sqrt{3} - i)$.

Now compute $\sqrt{\sqrt{3} - i}$.

We obtain

$$x^2 = \frac{1}{2}(\sqrt{3} + 2), \quad y^2 = \frac{1}{2}(-\sqrt{3} + 2).$$

Thus

$$\sqrt{\sqrt{3} - i} = \pm\left(\sqrt{\frac{\sqrt{3} + 2}{2}} - i\sqrt{\frac{-\sqrt{3} + 2}{2}}\right) = \pm\left(\frac{1 + \sqrt{3}}{2} + \frac{1 - \sqrt{3}}{2}i\right).$$

Now let's find $\sqrt{-\sqrt{3} + i}$.

$\langle$*Skipped*$\rangle$

Answer:

$$\pm\left(\frac{1+\sqrt{3}}{2}+\frac{1-\sqrt{3}}{2}i\right), \qquad \pm\left(\frac{-1+\sqrt{3}}{2}+\frac{1+\sqrt{3}}{2}i\right).$$

Exercize 2.31. Find

a) $\sqrt{-8i}$, b) $\sqrt{3-4i}$.

2.8. Equations of the 2nd, 3rd and 4th degree

2.8.1. Quadratic equation

$$ax^2 + bx + c = 0, \qquad a \neq 0$$

Divide the both sides by a:

$$x^2 + px + q = 0. \tag{2.11}$$

Let's solve it by extracting a *perfect square*. In LHS:

$$x^2 + 2x\frac{p}{2} + \left(\frac{p}{2}\right)^2 - \left(\frac{p}{2}\right)^2 + q = \left(x + \frac{p}{2}\right)^2 - \left(\frac{p}{2}\right)^2 + q.$$

$$\left(x + \frac{p}{2}\right)^2 = \left(\frac{p}{2}\right)^2 - q, \quad \text{hence} \quad x + \frac{p}{2} = \sqrt{\left(\frac{p}{2}\right)^2 - q}.$$

The square root in the last formula has two values with opposite signes.

Introduce for them denotion $\pm\sqrt{\ }$.

Thus, Equation (2.11) has two roots:

$$x_{1,2} = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}.$$

Discriminant of the Equation (2.11) is

$$D = \sqrt{\left(\frac{p}{2}\right)^2 - q}.$$

If $D = 0$ then (2.11) has one complex root.

If $D \neq 0$ then (2.11) has two different complex roots.

Example 2.32. $x^2 + x + 1 = 0$

Example 2.33. $x^2 - (2+i)x + (-1+7i) = 0$.

$$x_{1,2} = \frac{2+i}{2} \pm \sqrt{\frac{3+4i}{4} + 1 - 7i} = \frac{2+i}{2} \pm \frac{\sqrt{7-24i}}{2}.$$

$$x^2 = \frac{1}{2}(\quad 7 + \sqrt{7^2 + 24^2}) = 16,$$

$$y^2 = \frac{1}{2}(-7 + \sqrt{7^2 + 24^2}) = 9.$$

Thus, $\sqrt{7-24i} = \pm(4-3i)$,

hence $x_{1,2} = \dfrac{2+i}{2} \pm \dfrac{4-3i}{2}$.

Answer: $x_1 = 3 - i$, $x_2 = -1 + 2i$.

Exercize 2.34. Solve equations:

a) $x^2 - (3 - 2i)x + (5 - 5i) = 0$;

b) $(2 + i)x^2 - (5 - i)x + (2 - 2i) = 0$.

Exercize 2.35. Solve *biquadratic equations*:

a) $x^4 - 3x^2 + 4 = 0$;

b) $x^4 - 30x^2 + 289 = 0$.

Hint: Denote $t = x^2$.

2.8.2. Cube equation

$$y^3 + ay^2 + by + c = 0$$

Let $y = x - \frac{a}{3}$ then the term with the square of the unknown vanishes.

We obtain

$$x^3 + px + q = 0. \tag{2.12}$$

The roots of the equation can be found using the *Cardano formulas* (1545):

$$x = \alpha + \beta, \tag{2.13}$$

where

$$\alpha = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}, \quad \beta = \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}. \tag{2.14}$$

Example 2.36. $x^3 - 6x + 4 = 0$. Using (2.14) we obtain

$$\alpha = \sqrt[3]{-2 + \sqrt{4-8}} = \sqrt[3]{-2+2i},$$
$$\beta = \sqrt[3]{-2 - \sqrt{4-8}} = \sqrt[3]{-2-2i}.$$

The cube root can be found "exactly". For instance, $\alpha_1 = 1 + i$, $\beta_1 = 1 - i$.

$$x_1 = (1+i) + (1-i) = 2,$$
$$x_2 = (1+i)(-\frac{1}{2} + \frac{\sqrt{3}}{2}i) + (1-i)(-\frac{1}{2} - \frac{\sqrt{3}}{2}i) = -1 - \sqrt{3},$$
$$x_3 = (1+i)(-\frac{1}{2} - \frac{\sqrt{3}}{2}i) + (1-i)(-\frac{1}{2} + \frac{\sqrt{3}}{2}i) = -1 + \sqrt{3}.$$

Exercize 2.37. $x^3 - 19x + 30 = 0$.

2.8.3. Equations of the 4th degree

There is a formula for finding all roots of an algebraic equation of the 4-th degree (Ferrari formula, 1545).

⟨*Skipped*⟩

Let $n \geq 5$.

There is no *general* formula involving addition, subtraction, multiplication, division and rooting for finding all roots of an algebraic equation of the n-th degree (E. Galois, 1830).

2.9. Extra examples

Example 2.38. Find algebraic expression for $\cos\dfrac{2\pi}{5}$, $\sin\dfrac{2\pi}{5}$.

$$z = \cos\frac{2\pi}{5} + i\sin\frac{2\pi}{5}$$

is one of the root of equation

$$x^5 - 1 = 0,$$

that is equavalent to

$$(x-1)(x^4 + x^3 + x^2 + x + 1) = 0.$$

Hence z is a root of equation

$$x^4 + x^3 + x^2 + x + 1 = 0$$

This is a *reciprocal* equation.

Divide its both part over x^2:

$$x^2 + x + 1 + \frac{1}{x} + \frac{1}{x^2} = 0, \tag{2.15}$$

or

$$\left(x^2 + \frac{1}{x^2}\right) + \left(x + \frac{1}{x}\right) + 1 = 0.$$

Let

$$u = x + \frac{1}{x}, \tag{2.16}$$

Then $u^2 = x^2 + 2 + \dfrac{1}{x^2}$ and (2.15) has the form

$$u^2 + u - 1 = 0.$$

Its roots are $u_{1,2} = -\dfrac{1}{2} \pm \dfrac{\sqrt{5}}{2}$.

From (2.16) we have

$$x_{1,2} = \quad \frac{\sqrt{5}}{4} - \frac{1}{4} \pm i\frac{1}{4}\sqrt{2}\sqrt{5+\sqrt{5}},$$

$$x_{3,4} = -\frac{\sqrt{5}}{4} - \frac{1}{4} \pm i\frac{1}{4}\sqrt{2}\sqrt{5-\sqrt{5}}.$$

Obviously, $z = \cos\frac{2\pi}{5} + i\sin\frac{2\pi}{5} = x_1$.

Hence

$$\cos\frac{2\pi}{5} = \frac{\sqrt{5}}{4} - \frac{1}{4}, \qquad \sin\frac{2\pi}{5} = \frac{1}{4}\sqrt{2}\sqrt{5+\sqrt{5}}.$$

2.10. Numerical rings and fields

We will say that the numerical set $A \subseteq \mathbb{C}$ is *closed* under addition if for all a and b in A their sum $a + b$ also is in A.

$$\forall\ a \in A,\ \forall\ b \in A\ a + b \in A$$

Similarly the closeness under other arithmetic operations is defined.

For example, the sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are closed under addition and multiplication.

The sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are closed under subtraction.

The set U_n of all values of the root of the n-th degree is closed under multiplication and division.

Non-empty numerical set K is called (*numerical*) *ring* if it is closed under addition, subtraction and multiplication.

The set $\mathbb{N}$ is closed under addition and multiplication, but it is not closed under subtraction, hence $\mathbb{N}$ *is not a ring*.

It is not hard to see that the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are ring. The set $\{0\}$ also is a ring. It is called *null-ring* or *trivial ring*.

Proposition 2.39. *If K is a ring then $0 \in K$.*

Proof. Let $a \in K$. Since K is closed under subtraction then $0 = a - a \in K$. □

Exercize 2.40. Give an example of a ring which doesn't contain 1.

Let $n \in \mathbb{Z}$ then

$$n\mathbb{Z} = \{nk : \; k \in \mathbb{Z}\}$$

(all numbers divisible by n) is a ring.

Exercize 2.41. Does the set of all even numbers form a ring?

Does the set of all odd numbers form a ring?

A numerical ring F is called (*numerical*) *field*, if the set of their non-zero elements is not empty and closed relatively to division.

The set $\mathbb{Z}$ is not field, while $\mathbb{Q}$ and $\mathbb{R}$ are fields: in $\mathbb{Q}$ and $\mathbb{R}$ the division by non-zero elemet is possible and doesn't lead outside the set.

In fact, any numerical field contains $\mathbb{Q}$.

Proposition 2.42. *If F is numerical field then $\mathbb{Q} \subseteq F$.*

Proof. By definition of a field, $0 \in F$ and $a \in F$ where $a \neq 0$.

Since $F \subset \{0\}$ is closed under division then $1 = a/a \in F$.

Further, using that F is closed under addition we obtain that numbers $2 = 1 + 1$, $3 = 1 + 1 + 1$ and so on belongs to F, hence $\mathbb{N} \subseteq F$.

Also, it is true that $-1 = 0 - 1 \in F$, $-2 = 0 - 2 \in F$, $\ldots$.

Thus, $\mathbb{Z} \subseteq F$.

Finally, numbers $\dfrac{p}{q}$ are in F too for all $p \in \mathbb{N}$, $q \in \mathbb{Z}$, hence $\mathbb{Q} \subseteq F$. $\square$

Example 2.43. Consider the set of numbers of the form $\alpha + \beta\sqrt{3}$, where $\alpha \in \mathbb{Z}$, $\beta \in \mathbb{Z}$.

The set is a ring but it is not a field.

Now let $\alpha \in \mathbb{Q}$, $\beta \in \mathbb{Q}$, then the set is a field: it is closed under division since

$$\frac{\alpha + \beta\sqrt{3}}{\gamma + \delta\sqrt{3}} = \frac{(\alpha + \beta\sqrt{3})(\gamma - \delta\sqrt{3})}{(\gamma + \delta\sqrt{3})(\gamma - \delta\sqrt{3})} = \frac{\alpha\gamma - 3\beta\delta}{\gamma^2 - 3\delta^2} + \frac{-\alpha\delta + \beta\gamma}{\gamma^2 - 3\delta^2}\sqrt{3}. \tag{2.17}$$

Notice that the equality $\gamma + \delta\sqrt{3} = 0$ is possible only if $\gamma = 0$ and $\delta = 0$.

Indeed, if $\gamma + \delta\sqrt{3} = 0$ and $\gamma = 0$, then $\delta = 0$. If $\gamma + \delta\sqrt{3} = 0$ and $\delta \neq 0$, then $\sqrt{3} = -\dfrac{\gamma}{\delta}$, that is impossible since $\sqrt{3}$ is irrational number.

Finally, $\gamma + \delta\sqrt{3} \neq 0$, analogously, $\gamma - \delta\sqrt{3} \neq 0$, and consequantly multiplying the numerator and the denominator of the fraction by $\gamma - \delta\sqrt{3}$ in (2.17) is possible.

These calculations are similar to the procedure of division of complex numbers (multiplication by conjugate number).

Chapter 3

Systems of Linear Equations

3.1. Examples and definitions

$$\begin{cases} 2x + 3y + z = 1, \\ 4x + 6y + z = 0, \\ 2x - 3y + 2z = 2 \end{cases}$$

— 3 equations and 3 *unknowns* x, y, z.

$$\begin{cases} x_1 + 2x_2 - x_3 + x_5 = 1; \\ 3x_1 + 4x_2 - x_3 + 2x_4 - x_5 = 1; \\ 4x_1 + 6x_2 - 2x_3 + 2x_4 = 2; \\ x_1 + 2x_2 - x_3 + x_4 + x_5 = 0. \end{cases}$$

— 4 equations and 5 unknowns x_1, x_2, x_3, x_4, x_5.

Notice that the number of equations in this system is not equal to the number of unknowns.

$$x_1 + 2x_2 - 3x_3 + x_4 = 1$$

— the system consisting of the only equation.

Let F be a field.

A *system of* m *linear equations in* n *unknowns* $x_1, x_2, \ldots, x_n$ is a family of linear equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases}$$

where $a_{ij} \in F$, $b_i \in F$ $(i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n)$.

a_{ij}, b_i are called *coefficients* of the system,

x_j are called *unknowns* or *variables*.

(*Partial*) *solution* of the system is an ordered collection of numbers $(x_1, x_2, \ldots, x_n)$, $x_j \in F$ $(j = 1, 2, \ldots, n)$ which satisfy each of the equations simultaneously.

We say that the system is *consistent* if it has a solution. Otherwise the system is called *inconsistent*.

Example. The system

$$\begin{cases} 5x + y - z = 14, \\ 2x + y + 3z = 4, \end{cases}$$

has a partial solution $(2, 3, -1)$, i. e. $x = 2$, $y = 3$, $z = -1$. So the system is consistent.

Another solution is $(1, 29/4, -7/4)$

Are there other solutions to the system?

(Yes! It has infinitely many solutions.)

Example. The system

$$\begin{cases} 2x + y - z = 1, \\ 2x + y - z = -1, \end{cases}$$

has no solution. It is inconsistent.

The problem is

1. to determine if such a system has a solution,
2. to find all solutions of the system.

The matrix

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

is called the *coefficient matrix* of the system.

The matrix

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}$$

is called the *augmented* (or *enlarged*) *matrix* of the system.

3.2. Equivalent systems of linear equations

Two systems with the same number of equations are *equivalent* if and only if they have the same set of solutions.

Let's define some basic actions by wich a system is transformed in an equivalent system.

Elementary operations on the system:

1. Interchange two equations.
2. Multiply one of the equations with a number $\alpha \in F$, $\alpha \neq 0$.
3. Multiply one of the equations with a number α, and add the result to another equation, leaving the original equation unchanged.

3.3. The Gauss–Jordan method

The main idea is consecutive eliminating of unknowns by means of the elementary operations on the system.

Example

$$\begin{cases} 2x + 3y + z = 0, \\ 4x + 6y + z = -2, \\ 2x - 3y + 2z = 2 \end{cases}$$

Multiply the first equation by 1/2 (i. e. divide it by 2):

$$\begin{cases} x + \dfrac{3}{2}y + \dfrac{1}{2}z = 0, \\ 4x + \phantom{\dfrac{3}{2}} 6y + \phantom{\dfrac{1}{2}} z = -2, \\ 2x - \phantom{\dfrac{3}{2}} 3y + \phantom{\dfrac{1}{2}} 2z = 2 \end{cases}$$

Now add the first equation multiplied by -4 to the second one (i. e. subtract the first equation multiplied by 4 from the second one):

$$\begin{cases} \quad x + \dfrac{3}{2}y + \dfrac{1}{2}z = 0, \\ \qquad\qquad\quad - \quad z = -2, \\ 2x - \ 3y + \ 2z = \quad 2 \end{cases}$$

We eliminated the unknown x (and occasionaly y) from the second equation.

Now add the first equation multiplied by -2 to the third one: (i. e. subtract the first equation multiplied by 2 from the third one):

$$\begin{cases} x + \dfrac{3}{2}y + \dfrac{1}{2}z = 0, \\ \qquad\quad - \quad z = -2, \\ \quad - 6y + \quad z = \quad 2 \end{cases}$$

We eliminated the unknown x (and occasionaly y) from the third equation.

Interchange the second and the third equations:

$$\begin{cases} x + \dfrac{3}{2}y + \dfrac{1}{2}z = 0, \\ \quad - 6y + \quad z = \quad 2, \\ \qquad\qquad - \quad z = -2 \end{cases}$$

Multiply the second equation by $-1/6$:

$$\begin{cases} x + \dfrac{3}{2}y + \dfrac{1}{2}z = 0, \\ \qquad y - \dfrac{1}{6}z = -\dfrac{1}{3}, \\ \qquad\qquad - \quad z = -2 \end{cases}$$

Add the second equation multiplied by $-3/2$ to the first one:

$$\begin{cases} x \quad + \frac{3}{4}z = \frac{1}{2}, \\ \quad y - \frac{1}{6}z = -\frac{1}{3}, \\ \quad\quad - z = -2 \end{cases}$$

We eliminated y from the first equation.

Multiply the third equation by -1:

$$\begin{cases} x \quad + \frac{3}{4}z = \frac{1}{2}, \\ \quad y - \frac{1}{6}z = -\frac{1}{3}, \\ \quad\quad z = 2 \end{cases}$$

Add the third equation multiplied by $-3/4$ to the first one:

$$\begin{cases} x \qquad\qquad = -1, \\ \quad y - \dfrac{1}{6}z = -\dfrac{1}{3}, \\ \qquad\qquad z = \quad 2 \end{cases}$$

We eliminated z from the first equation.

Add the third equation multiplied by $1/6$ to the second one:

$$\begin{cases} x \qquad = -1, \\ \quad y \quad = \quad 0, \\ \qquad z = \quad 2. \end{cases}$$

We eliminated z from the first equation.

This system is equivalent to the final system. So the (unique) solution is $(-1, 0, 2)$. It can be verified by substituting it into the original system.

It is convinient to perform all operations with augmented matrix.

Elementary operation on the system are equivalent to the following operations on the augmented matrix:

1. Interchanging two rows: $R_i \leftrightarrow R_j$ interchanges i-th and j-th rows.
2. Multiplying a row by a non-zero number. $R_i \leftarrow \alpha R_i$ multiplies i-th row by the non-zero number α.
3. Adding a multiple of one row to another row. $R_j \leftarrow R_j + \alpha R_i$ adds α times i-th row to j-th row.

$$\left(\begin{array}{ccc|c} 2 & 3 & 1 & 0 \\ 4 & 6 & 1 & -2 \\ 2 & -3 & 2 & 2 \end{array}\right) \quad \begin{array}{l} R_1 \leftarrow \frac{1}{2} R_1 \\ R_2 \leftarrow R_2 - 4 \cdot R_1 \\ R_3 \leftarrow R_3 - 2 \cdot R_1 \end{array} \quad \left(\begin{array}{ccc|c} 1 & 3/2 & 1/2 & 0 \\ & & -1 & -2 \\ & -6 & 1 & 2 \end{array}\right)$$

$$R_2 \leftrightarrow R_3 \quad \left(\begin{array}{ccc|c} 1 & 3/2 & 1/2 & 0 \\ & -6 & 1 & 2 \\ 0 & 0 & -1 & -2 \end{array}\right) \quad \begin{array}{l} R_2 \leftarrow -\frac{1}{6} R_2 \\ R_1 \leftarrow R_1 - \frac{3}{2} \cdot R_2 \end{array}$$

$$\left(\begin{array}{ccc|c} 1 & 0 & 3/4 & 1/2 \\ 0 & 1 & -1/6 & -1/3 \\ 0 & 0 & -1 & -2 \end{array}\right) \quad \begin{array}{l} R_3 \leftarrow -R_3 \\ R_1 \leftarrow R_1 - \frac{3}{4} \cdot R_3 \\ R_2 \leftarrow R_2 + \frac{1}{6} \cdot R_3 \end{array} \quad \left(\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{array}\right)$$

Example 2.

$$\begin{cases} x_1 + 2x_2 - x_3 + x_5 = 1; \\ 3x_1 + 4x_2 - x_3 + 2x_4 - x_5 = 1; \\ 4x_1 + 6x_2 - 2x_3 + 2x_4 = 2; \\ x_1 + 2x_2 - x_3 + x_4 + x_5 = 0. \end{cases} \tag{3.1}$$

Augmented matrix:

$$\left(\begin{array}{ccccc|c} 1 & 2 & -1 & 0 & 1 & 1 \\ 3 & 4 & -1 & 2 & -1 & 1 \\ 4 & 6 & -2 & 2 & 0 & 2 \\ 1 & 2 & -1 & 1 & 1 & 0 \end{array}\right).$$

$$R_2 = R_2 - 3R_1$$
$$R_3 = R_3 - 4R_1$$
$$R_4 = R_4 - R_1$$

$$\left(\begin{array}{ccccc|c} 1 & 2 & -1 & 0 & 1 & 1 \\ 0 & -2 & 2 & 2 & -4 & -2 \\ 0 & -2 & 2 & 2 & -4 & -2 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right).$$

$$R_2 = -\frac{1}{2}R_2$$

$$R_1 = R_1 - 2R_2$$

$$R_3 = R_3 + 2R_2$$

$$\left(\begin{array}{ccccc|c} 1 & 0 & 1 & 2 & -3 & -1 \\ 0 & 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right).$$

$$R_1 = R_1 - 2R_4$$

$$R_2 = R_2 + R_4$$

$$R_3 \leftrightarrow R_4$$

$$\left(\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & -3 & 1 \\ 0 & 1 & -1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Thus the original system of linear equations is equvalent to the following system:

$$\begin{cases} x_1 \phantom{{}+x_2} + x_3 \phantom{{}+x_4} - 3x_5 = 1; \\ \phantom{x_1 + {}} x_2 - x_3 \phantom{{}+x_4} + 2x_5 = 0; \\ x_4 = -1; \\ 0 = 0. \end{cases}$$

Express x_1, x_2, x_4 in terms of x_3, x_5:

$$\begin{aligned} x_1 &= 1 - x_3 + 3x_5; \\ x_2 &= 0 + x_3 - 2x_5; \\ x_4 &= -1 \end{aligned}$$

It is not hard to see that for any values of x_3, x_5 (*independent unknowns*) there exist values for x_1, x_2, x_4 (*dependent unknowns*), such that all equations of the final and original systems turn into valid equalities. So the general solution of the system can be written as

$$\begin{aligned} x_1 &= 1 - t_1 + 3t_2; \\ x_2 &= \phantom{-1 - {}} t_1 - 2t_2; \\ x_3 &= \phantom{-1 - {}} t_1; \qquad\qquad \text{where } t_1,\ t_2 \text{ are arbitrary numbers.} \\ x_4 &= -1 \\ x_5 &= t_2, \end{aligned}$$

Notice that in the general case the form of the answer is not unique.

Exercize 3.1. Solve the following systems:

$$\begin{cases} x + y = 2, \\ x - y = 3, \end{cases} \qquad \begin{cases} 2x + y - 3z = -9, \\ x - y = 3, \\ x + z = 4, \end{cases}$$

$$\begin{cases} 2x_1 + 3x_2 + 2x_3 + 3x_4 = 10, \\ -2x_1 + 2x_2 + 3x_3 + x_4 = 4, \\ 5x_2 + 5x_3 - 2x_4 = 8, \end{cases} \qquad \begin{cases} 2x + 3y - z = 1, \end{cases}$$

$$\begin{cases} 2x + y = 1, \\ -2x + 2y = 3, \\ 3y = 2. \end{cases}$$

3.4. Reduced row-echelon form

We say that a matrix is in *row-echelon form* (r.e.f.) if

1. all zero rows (if any) are at the bottom of the matrix and

2. if two successive rows are non-zero, the second row starts with more zeros than the first (moving from left to right).

$$\begin{pmatrix} 0 & 0 & 3 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \text{ is in row-echelon form}$$

$$\begin{pmatrix} 0 & 0 & 3 & 0 & 2 & 1 \\ 0 & 0 & 2 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \text{ is not in row-echelon form}$$

We say that a matrix is in *reduced row-echelon form* (r.r.e.f.) if

1. it is in row-echelon form,
2. the leading (leftmost non-zero) entry in each non-zero row is 1,
3. all other elements of the column in which the leading entry 1 occurs are zeros.

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \text{ is in reduced row-echelon form}$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \text{ is not in reduced row-echelon form}$$

If a matrix is in reduced row-echelon form, we denote the column numbers in which the leading entries 1 occur, by $j_1, j_2, \ldots, j_r$, where r is the number of non-zero rows.

For example, in the 4×6 matrix

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

we have

$$r = 3;\ c_1 = 3,\ c_2 = 4,\ c_3 = 6.$$

Gauss–Jordan algorithm starts with a given matrix A and produces a matrix B in reduced row-echelon form by means of elementary row operations.

If A is the augmented matrix of a system of linear equations, then B will be a much simpler matrix than A

From A the consistency or inconsistency of the corresponding system is apparent and the complete solution of the system can be read off.

Case 1: For the last non-zero row of B is $(0, 0, \ldots, 0, 1)$ and the corresponding equation is $0x_1 + 0x_2 + \ldots + 0x_n = 1$ which has no solutions. Consequently the original system has no solutions.

Case 2: The leading entry of the last non-zero row is not in the last column. The system of equations is consistent.

$$B = \begin{pmatrix} 0 & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots & * & 0 & * & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & * & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Express unknowns $x_{j_1}, \dots, x_{j_r}$ in terms of remaining unknowns:

$$\begin{aligned} x_{j_1} &= c_{11}x_{k_1} + \dots + c_{1n-r}x_{k_{n-r}}, \\ x_{j_2} &= c_{21}x_{k_1} + \dots + c_{2n-r}x_{k_{n-r}}, \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ x_{j_r} &= c_{r1}x_{k_1} + \dots + c_{1n-r}x_{k_{n-r}}, \end{aligned} \tag{3.2}$$

where $\{j_1, \dots, j_r\} \cup \{k_1, \dots, k_{n-r}\} = \{1, 2, \dots, n\}$.

$x_{k_1}, \dots, x_{k_{n-r}}$ are called *independent unknowns*,

$x_{j_1}, \dots, x_{j_r}$ are called *dependent unknowns*. which can take on arbitrary values:

$$
\begin{aligned}
x_{j_1} &= c_{10} + c_{11}t_1 + \ldots + c_{1n-r}t_{n-r},\\
x_{j_2} &= c_{20} + c_{21}t_1 + \ldots + c_{2n-r}t_{n-r},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
x_{j_r} &= c_{r0} + c_{r1}t_1 + \ldots + c_{1n-r}t_{n-r},\\
x_{k_1} &= \qquad\quad t_1,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
x_{k_{n-r}} &= \qquad\qquad\qquad\qquad t_{n-r},
\end{aligned}
\tag{3.3}
$$

where $t_1, \ldots, t_{n-r}$ are arbitrary numbers in F.

If $n = r$ then the solution is unique.

3.5. Arithmetic space

Let's consider an ordered collection of n numbers in F written in a column:

$$a = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

a is called a *vector* (or *column-vector*) of height n.

The set of columns of height n is called the n-dimensional arithmetic space and denoted as F^n.

$\mathbb{Q}^n$ is the rational arithmetic space, $\mathbb{R}^n$ is the real arithmetic space
$\mathbb{Z}^n$ is the complex arithmetic space

Let's introduce the following operations:

Summation $a + b$ of two vectors a and b

$$a + b = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix}.$$

Multiplication a vector a with a scalar α:

$$\alpha a = \alpha \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \alpha\alpha_1 \\ \alpha\alpha_2 \\ \vdots \\ \alpha\alpha_n \end{pmatrix}$$

For example,

$$\begin{pmatrix} 1 \\ 2 \\ -3 \\ 2 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 0 \\ 3 \end{pmatrix}.$$

Multiplication a vector a with a scalar α:

$$3 \cdot \begin{pmatrix} 3 \\ 1 \\ 4 \\ -2 \end{pmatrix} = \begin{pmatrix} 9 \\ 3 \\ 12 \\ -6 \end{pmatrix}$$

Now the formulae (3.3) can be written in *column-wise* notation.

For simplicity let's assume that $j_1 = 1, \ldots, j_r = r$ (this can be attained by renumbering of unknowns).

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \\ x_{r+1} \\ x_{r+2} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_{10} \\ c_{20} \\ \vdots \\ c_{r0} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + t_1 \cdot \begin{pmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{r1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots + t_{n-r} \cdot \begin{pmatrix} c_{1n-r} \\ c_{2n-r} \\ \vdots \\ c_{rn-r} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \tag{3.4}$$

For example, formulae

$$\begin{aligned} x_1 &= \ \ 1 - t_1 + 3t_2; \\ x_2 &= \qquad\ t_1 - 2t_2; \\ x_3 &= \qquad\ t_1; \qquad\qquad \text{where } t_1,\ t_2 \text{ are arbitrary numbers.} \\ x_4 &= -1 \\ x_5 &= \qquad\qquad\quad t_2, \end{aligned}$$

can be written as

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} + t_1 \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 3 \\ -2 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

3.6. Homogeneous systems

The system of linear equations with zero right-hand side is called *homogeneous system*:

$$\begin{cases} a_{11}x_1 + \ a_{12}x_2 + \ldots + \ a_{1n}x_n = 0, \\ a_{21}x_1 + \ a_{22}x_2 + \ldots + \ a_{2n}x_n = 0, \\ \dotfill \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0. \end{cases}$$

Such a system always has solution

$$x_1 = 0, x_2 = 0, \ldots, x_n = 0$$

(*trivial* solution) so a homogenius system is always consistent.

If in a solution at least one unknown is not 0 then the solution is called *non-trivial*.

For example, the system

$$\begin{cases} x + y = 0, \\ x - y = 0 \end{cases}$$

has only the trivial solution.

Wherease the system

$$\begin{cases} x + y + z = 0, \\ x - y + z = 0 \end{cases}$$

has complete solution $x = t$, $y = 0$, $z = -t$, where t is arbitrary. In particular, taking $t = 1$ we get the non-trivial solution $x = 1$, $y = 0$, $z = -1$.

Theorem 3.2. *A homogeneous system of m linear equations in n un- knowns always has a non-trivial solution if $m < n$.*

Proof. Suppose that $m < n$ and that the coeffcient matrix of the system is equivalent to B, a matrix in reduced row-echelon form.

Let r be the number of non-zero rows in B.

Then $r \leq m < n$ and hence $n - r > 0$ and so the number $n - r$ of independent unknowns (arbitrary unknowns) is in fact positive.

Taking one of these unknowns to be 1 (or any other non-zero value) gives a non-trivial solution. □

Chapter 4

Vector Spaces

4.1. Definition of Vector Space

Let V be a non-empty set, F be a field.

Let's call the objects in V *vectors*, the objects in F *scalars* (or *numbers*).

V is called the *linear space* (or *vector space*) over the filed F if

I. There is a rule (or operation), called *vector addition*, which associates with each pair of vectors a and b in V a vector c in V, called the *sum* of a and b and denoted as $c = a + b$.

II. There is a rule (or operation), called *scalar multiplication*, which associates with each scalar α in F and each vector a in V a vector b in V, called the *product of* α *and* a and denoted as $b = \alpha \cdot a$, or, simply, $b = \alpha a$.

III. The following conditions holds (*axioms of linear space*):

(1) $\forall a, b, c \in V \;\; a + (b + c) = (a + b) + c$ (addition is *associative*),

(2) $\forall a, b \in V \;\; a + b = b + a$ (addition is *commutative*),

(3) $\exists o \; \forall a \in V \;\; a + o = a$ (o is called the *zero vector*) (The zero vector $o \in V$ and the scalar $0 \in F$ should be distinguished,

(4) $\forall a \in V \ \exists b \in V \ a + b = o$ (b is called the *opposite* to v vector and it's denoted as $-a$),

(5) $\forall a \in V \ 1 \cdot a = a$,

(6) $\forall a \in V \ \forall \alpha, \beta \in F \ \alpha(\beta a) = (\alpha\beta a)$,

(7) $\forall a, b \in V \ \forall \alpha \in F \ \alpha(a + b) = \alpha a + \alpha b$ (distibutivity I),

(8) $\forall a \in V \ \forall \alpha, \beta \in F \ (\alpha + \beta)a = \alpha a + \beta a$ (distributivity II).

Further, as a rule, vectors are denoted by latin letters, whereas scalars are denoted by greek letters.

The filed of scalars is usually $\mathbb{R}$ or $\mathbb{C}$ and the linear space is called *real* or *complex* depending on whether the filed is $\mathbb{R}$ or $\mathbb{C}$. However, over fileds are also possible.

Let's consider a few examples of linear spaces

The origin of the name "vector" is from the following example.

EXAMPLE 1. "Geometric" spaces:

$\mathbf{V}_1$ is a set of all radious-vectors on a line

$\mathbf{V}_2$ is a set of all radious-vectors in the plane

$\mathbf{V}_3$ is a set of all radious-vectors in space

$\mathbf{V}_1$, $\mathbf{V}_2$, $\mathbf{V}_3$ are *real* linear spaces (in respect to usual operations of vector addition and scalar multiplication introduced in chapter 1)

For this spaces we proved all conditions in the definition of vector space (axioms)

EXAMPLE 2. “Arithmetic” spaces:

F^n is a set of all n-tuples (ordered collections) of numbers in F

F^n is the linear space (in respect to usual operations vector addition and scalar multiplication introduced in Section 3.5).

All axioms of linear space can be easily verified.

EXAMPLE 3. $F[x]$, the space of polynomials over F.

A *polynomial* is an expression of the form

$$a_0 + a_1x + a_2x^2 + \ldots + a_nx^n,$$

where $a_j \in F$ $(j = 0, 1, \ldots, n)$.

$F[x]$ is the set of all polynomials over F.

Examples:

$$2x, \qquad 1 + 2x - x^2, \qquad \frac{1}{219} + \frac{3}{311}x^3 + \frac{6}{171}x^4 + \frac{9}{113}x^5.$$

The *sum* of two polynomials

$$f = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

$$g = b_0 + b_1 x + b_2 x^2 + \ldots + b_n x^n$$

is

$$f + g = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \ldots + (a_n + b_n)x^n$$

The product of a polynomial f and a scalar α is

$$\alpha f = \alpha a_0 + \alpha a_1 x + \alpha a_2 x^2 + \ldots + \alpha a_n x^n$$

It is easy to verify that $F[x]$ is a linear space over F.

Example 4. The space of real functions.

Let V be a set of all functions $f\colon \mathbb{R} \to \mathbb{R}$.

Let $F = \mathbb{R}$.

So, a vector in the space is a function. A saclar is a real number.

Addition of two vectors f and g is defined as $(f + g)(x) = f(x) + g(x)$.

Multiplication of a scalar α and a vector f is defined as $(\alpha f)(x) = \alpha\big(f(x)\big)$.

All axioms of linear space can be easily verified.

In particular, the zero vector o is the function $f(x) = 0$.

4.2. Simple corollaries from axioms

Proposition 4.1. *In each linear space the zero vector is unique.*

Proof. Let's consider that there exist two zero vectors $o_1, o_2 \in V$. Then

$$o_1 = o_1 + o_2 = o_2.$$

□

Proposition 4.2. *For each vector $a \in V$ there exists unique opposite vector.*

Proof. Let b_1, b_2 be vectors, opposite to a.

$$b_1 + (a + b_2) = b_1 + o = b_1.$$

$$(b_1 + a) + b_2 = o + b_2 = b_2.$$

But the left hand sides are the same, so $b_1 = b_2$. □

Proposition 4.3. *For any a, b in V the equation $a + x = b$ has unique solution.*

Proof. By substitution we can verify that $x = (-a) + b$ is a solution of the equation.

Now let's prove that the solution is unique.

Suppose that there are two solutions x_1, x_2.

Then $a + x_1 = a + x_2 = b$.

Adding $(-a)$ to both part of the equation, we get $(-a) + a + x_1 = (-a) + a + x_2$, hence $x_1 = x_2$. □

$(-a) + b = b + (-a)$ is called the *difference* of b and a and denoted as $b - a$.

Proposition 4.4. $0a = o$ *for any* $a \in V$.

Proof. We have $a + 0a = 1a + 0a = (1 + 0)a = 1a = a$.

So $a + 0a = a$, hence $0a = a - a = o$. $\square$

Proposition 4.5. *$\alpha o = o$ for each $\alpha \in F$.*

Proof. Let $a \in V$.

We have $\alpha a + \alpha o = \alpha(a + o) = \alpha a$.

Thus, $\alpha a + \alpha o = \alpha a$, hence $\alpha o = \alpha a - \alpha a = o$. □

Proposition 4.6. *For all vectors a, b in V and all scalars α, β in F the following equalities hold:*

1) $(-\alpha)a = -(\alpha a)$;

2) $(-1)a = -a$;

3) $(\alpha - \beta)a = \alpha a - \beta a$;

4) $\alpha(a - b) = \alpha a - \alpha b$.

Proof. 1) Let's show that $(-\alpha)a$ is opposite to αa.

Really, $\alpha a + (-\alpha)a = (\alpha - \alpha)a = 0a = o$.

2)–4) follow from 1). □

Proposition 4.7. *Let $a \in V$ and $\alpha \in F$. If $\alpha a = 0$, then $\alpha = 0$ or $a = 0$.*

Proof. If $\alpha = 0$, then the statement holds.

If $\alpha \neq 0$, then $a = \dfrac{1}{\alpha}(\alpha a) = \dfrac{1}{\alpha} o = o$. □

4.3. Subspaces

Let V be a linear space over a field F.

Let W be a non-empty subset of V.

W is called *subspace* of V if W itself is the space over F.

Theorem 4.8 (Criterion of subspace)**.** *Let V be a linear space over F and $W \neq \emptyset$, $W \subseteq V$. Then W is subspace iff*

1. *$\forall\ a, b \in W\ a + b \in W$ (W is said to be closed under vector addition),*
2. *$\forall\ \alpha \in F$ and $\forall\ a \in W\ \alpha a \in W$ (W is closed under scalar multiplication).*

Proof. Necessity. Conditions 1–2 are necessary because they are in the definition of linear space.

Sufficiency. Verify that all 8 axioms of linear space hold. It is easy for axioms (1)–(2), (5)–(8).

Check axiom (3). Let's prove that $o \in W$. Let $\alpha = 0$, $a \in W$, then $o = 0a \in W$ by condition 2.

Check axiom (4). Let's show that $-a \in W$ for each $a \in W$. Let $\alpha = -1$, then $-a = (-1)a \in W$ by condition 2. □

Examples

$\{o\}$ is *trivial* (or *zero*) space. It is subspace of V.

The set of all radius vectors whose heads are on the same stright line which passes through the origin is a subspace of $\mathbf{V}_2$ and $\mathbf{V}_3$.

The set of all radius vectors whose heads are on the same plane which passes through the origin is a subspace of $\mathbf{V}_3$.

What about all radius vectors whose heads are on the same stright line (or plane) which does not pass through the origin?

Let W be a set of all real 3-tuples whose the first component is zero, i.e.

$$W = \left\{(0,\ x_2,\ x_3)^\top :\ x_2, x_3 \in \mathbb{R}\right\}$$

Let's prove that W is a subspace in $\mathbb{R}^3$.

0. W is not empty, because, for instance, zero vector $(0,0,0)^\top$ is in W.

1. W is closed under vector addition:

$$(0,\ x_2,\ x_3)^\top + (0,\ y_2,\ y_3)^\top = (0,\ x_2 + y_2,\ x_3 + y_3)^\top \in W$$

(the sum of two vectors whose the first components are zeros is a vector whose the first component is zero).

2. W is closed under scalar multiplication:

$$\alpha(0,\ x_2,\ x_3)^\top = (0,\ \alpha x_2,\ \alpha x_3)^\top \in W.$$

(the product of a vector whose the first component is zero by a scalar is a vector whose the first component is zero).

Let W be a set of all real 3-tuples such that the sum of the components of each such a vector is 0.

$$W = \left\{(x_1,\, x_2,\, x_3)^\top :\; x_1 + x_2 + x_3 = 0\right\}$$

Let's prove that W is a subspace in $\mathbb{R}^3$.

0. W is not empty, because, for instance, zero vector $(0, 0, 0)^\top$ is in W.

1. W is closed under vector addition:

$$(x_1,\, x_2,\, x_3)^\top + (y_1,\, y_2,\, y_3)^\top = (x_1 + y_1,\, x_2 + y_2,\, x_3 + y_3)^\top.$$

If $x_1 + x_2 + x_3 = 0$ and $y_1 + y_2 + y_3 = 0$ then $(x_1 + y_1) + (x_2 + y_2) + (x_3 + y_3) = 0$. Hence, $(x_1 + y_1,\, x_2 + y_2,\, x_3 + y_3)^\top \in W$.

2. W is closed under scalar multiplication:

$$\alpha \cdot (x_1,\, x_2,\, x_3)^\top = (\alpha x_1,\, \alpha x_2,\, \alpha x_3)^\top.$$

If $x_1 + x_2 + x_3 = 0$ then $\alpha x_1 + \alpha_2 x_2 + \alpha_3 x_3 = 0$. Hence, $(\alpha x_1,\, \alpha x_2,\, \alpha x_3)^\top \in W$.

The following proposition generalize the last two examples.

Proposition 4.9. *The set W of all solutions to the homogenius system of m equations with n unknowns is a linear subspace in F^n.*

Proof. 0. $W \neq \emptyset$ because $(0, 0, \ldots, 0)^\top \in W$.

1. Let $(x_1, x_2, \ldots, x_n)^\top$, $(y_1, y_2, \ldots, y_n)^\top$ are partial solutions to the system. Let's prove that $(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)^\top$ is a solution too. The vector $(x_1, x_2, \ldots, x_n)^\top$, $(y_1, y_2, \ldots, y_n)^\top$ satisfy the i-th equation:

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = 0, \quad a_{i1}y_1 + a_{i2}y_2 + \ldots + a_{in}y_n = 0 \quad (i = 1, 2, \ldots, m)$$

We have

$$a_{i1}(x_1 + y_1) + a_{i2}(x_2 + y_2) + \ldots + a_{in}(x_n + y_n) = 0 \quad (i = 1, 2, \ldots, m)$$

Thus $(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)^\top$ satisfies all equations $(i = 1, 2, \ldots, m)$. Hence it is a solution to the system.

2. Let $(x_1, x_2, \ldots, x_n)^\top$ be a partial solution to the system. Let's prove that $(\alpha x_1, \alpha x_2, \ldots, \alpha x_n)^\top$ is a solution too:

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = 0 \quad (i = 1, 2, \ldots, m)$$

We have

$$a_{i1}(\alpha x_1) + a_{i2}(\alpha x_2) + \ldots + a_{in}(\alpha x_n) = 0 \quad (i = 1, 2, \ldots, m)$$

Thus $(\alpha x_1, \alpha x_2, \ldots, \alpha x_n)^\top$ satisfies all equations $(i = 1, 2, \ldots, m)$. Hence it is a solution to the system. $\square$

What about non-homogenius system?

The set W of all solutions to a non-homogenius system is not a subspace.

Example: Let's consider system

$$x_1 + x_2 + x_3 = 1$$

Vectors $(1,0,0)^\top$, $(0,1,0)^\top$ are two partial solutions.

But their sum $(1,1,0)$ is not apartial solution to the system. So W is not closed under addition. W is not subspace.

4.4. Linear combinations and linear hulls

The expression $\alpha_1 a_1 + \ldots + \alpha_n a_n$ is called *a linear combination of vectors* $a_1, \ldots, a_n$ *with coefficients* $\alpha_1, \ldots, \alpha_n$.

The set of all linear combinations of fixed vectors $a_1, \ldots, a_n$ is called *linear hull* and denoted as $\mathrm{Lin}(a_1, \ldots, a_n)$:

$$\mathrm{Lin}(a_1, \ldots, a_n) = \{\alpha_1 a_1 + \ldots + \alpha_n a_n : \ \alpha_1, \ldots, \alpha_n \in F\}\,.$$

Exercize 4.10. Describe the linear hull of a non-zero vector in $\mathbf{V}_2$ (or $\mathbf{V}_3$).

Exercize 4.11. Describe the linear hull of two non-collinear vectors in $\mathbf{V}_2$ (or $\mathbf{V}_3$).

Exercize 4.12. Describe the linear hull of three non-coplanar vectors in $\mathbf{V}_3$.

Theorem 4.13. *Let $a_1, \ldots, a_n$ be vectors in V. Then $\operatorname{Lin}(a_1, \ldots, a_n)$ is a subspace in V and for any subspace W containing vectors $a_1, \ldots, a_n$, it holds that $\operatorname{Lin}(a_1, \ldots, a_n) \subseteq W$.*

Proof. I. First let's prove that $\operatorname{Lin}(a_1, \ldots, a_n)$ is a subspace in V.

It's clear that $\operatorname{Lin}(a_1, \ldots, a_n) \neq \emptyset$.

If $a = \alpha_1 a_1 + \ldots + \alpha_n a_n$ and $b = \beta_1 a_1 + \ldots + \beta_n a_n$, then
$a + b = (\alpha_1 + \beta_1)a_1 + \ldots + (\alpha_n + \beta_n)a_n$. Thus, $a + b \in \operatorname{Lin}(a_1, \ldots, a_n)$.

If $\alpha \in F$, then $\alpha a = (\alpha\alpha_1)a_1 + \ldots + (\alpha\alpha_n)a_n$. Thus, $\alpha a \in F$.

Sufficient conditions in Theorem 4.8 hold.

II. Now let W be arbitrary subspace containing vectors $a_1, \ldots, a_n$.

Then for any numbers $\alpha_1, \ldots, \alpha_n$ in F it holds that $\alpha_1 a_1 + \ldots + \alpha_n a_n \in W$ (closeness under vector addition) and $\alpha_j a_j \in W$ $(j = 1, \ldots, n)$ (closeness under scalar multiplication).

hence $b \in W$ for any $b \in \mathrm{Lin}(a_1, \ldots, a_n)$, so $\mathrm{Lin}(a_1, \ldots, a_n) \subseteq W$. $\square$

The subspace $\mathrm{Lin}(a_1, \ldots, a_n)$ is called the subspace *spanned* (or *generated*) by $a_1, \ldots, a_n$.

The following result is not hard to prove.

Proposition 4.14. *$L(b_1, \dots, b_m) \subseteq L(a_1, \dots, a_n)$ iff each of $b_1, \dots, b_m$ is a linear conbination of $a_1, \dots, a_n$.*

Vector systems $a_1, \dots, a_n$ and $b_1, \dots, b_m$ are said to be *equivalent* iff each of $a_1, \dots, a_n$ is a linear conbination of $b_1, \dots, b_m$, and each of $b_1, \dots, b_m$ is a linear conbination of $a_1, \dots, a_n$.

Corollary 4.15. *Vector systems $a_1, \dots, a_n$ and $b_1, \dots, b_m$ are equivalent if $\mathrm{Lin}(a_1, \dots, a_n) = \mathrm{Lin}(b_1, \dots, b_m)$.*

4.5. Linear dependence

The linear combination $\alpha_1 a_1 + \ldots + \alpha_n a_n$ of vectors $a_1, \ldots, a_n \in V$ is called *trivial*, iff $\alpha_1 = \ldots = \alpha_n = 0$.

It is obvious that trivial linear cobination equals to o.

Vectors $a_1, \ldots, a_n$ are said to be *linearly dependent* if there exists their *non-trivial* linear combination equal to o, i.e. there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$, *not all zero*, such that

$$\alpha_1 a_1 + \ldots + \alpha_n a_n = 0. \tag{4.1}$$

Vectors $a_1, \ldots, a_n$ are said to be *linearly independent* iff they are not linearly dependent, i.e. iff equality (4.1) is possible only in the case when $\alpha_1 = \ldots = \alpha_n = 0$.

Example 4.16. Determine if the following system is linearly dependent?

Proposition 4.17 (System consisting of one vector)**.** *A system consisting of one vector is linearly dependent iff the vector is o.*

Proof. For a vector system consisting of the only vector a, equality (4.1) has the form $\alpha a = 0$, hence $\alpha = 0$ (then the combination is trivial) or $a = o$ (what's claimed). □

Proposition 4.18 (System consisting of two vectors)**.** *System consisting of two vectors is linearly dependent iff the vectors are proportional.*

Proof. Equality (4.1) has the form $\alpha_1 a_1 + \alpha_2 a_2 = 0$, where $\alpha_1 \neq 0$ or/and $\alpha_2 \neq 0$.

In the first case we have $a_1 = (-\alpha_2/\alpha_1)a_2$,

In the second one we have $a_2 = (-\alpha_1/\alpha_2)a_1$. □

Corollary 4.19. *Two vectors in* $\mathbf{V}_2$ (*or* $\mathbf{V}_3$) *are linearly dependent iff they are collinear.*

Proposition 4.20. *If a subsystem of some vector system is linearly dependent then the whole of system is linearly dependent too.*

Proof. Let $\{a_1, \ldots, \alpha_n\}$ be a vector system and $\{a_{j_1}, \ldots, a_{j_m}\}$ is its linearly dependent subsystem $(1 \leq j_1 < j_2 < \ldots < j_m \leq n)$.

By definition of linearly dependence, there exist scalars $\alpha_{j_1}, \ldots, \alpha_{j_m}$, not all zeros, such that

$$\alpha_{j_1} a_{j_1} + \ldots + \alpha_{j_m} a_{j_m} = o.$$

In this equality let's add to the l.h.s. the trivial combination of remaining vectors:

$$0a_{i_1} + \ldots + 0a_{i_{n-m}},$$

where $\{i_1, \ldots, i_{n-m}\} = \{1, \ldots, n\} \setminus \{j_1, \ldots, j_m\}$. It's clear that the new combination $\alpha_1 a_1 + \ldots + \alpha_m a_m$ equals to o but it's not trivial. □

Corollary 4.21. *Any subsystem of linearly independent system is linearly independent.*

(proof by contradiction)

Exercize 4.22. Is the following statement correct:

If a subsystem of some vector system is linearly independent then the whole of system is linearly independent too?

Exercize 4.23. Is the following statement correct:

If a system is linearly dependent then any its subsystem is linearly dependent too?

Theorem 4.24 (Linearly dependence criterion)**.** *Vectors* $a_1, \ldots, a_n$ $(n \geq 2)$ *are linearly dependent iff there exists* j *such that*

$$a_j = \mathrm{Lin}(a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n)$$

(i.e. $\exists\, j$ *such that* a_j *can be represented as linear combination of the remaining vectors).*

Proof. Necessity. Let $a_1, \ldots, a_n$ be linearly dependent, then

$$\alpha_1 a_1 + \ldots + \alpha_n a_n = 0,$$

and also there exists j such that $\alpha_j \neq 0$, from which

$$a_j = \left(-\frac{\alpha_1}{\alpha_j}\right) a_1 + \ldots + \left(-\frac{\alpha_{j-1}}{\alpha_j}\right) a_{j-1} + \left(-\frac{\alpha_{j+1}}{\alpha_j}\right) a_{j+1} + \ldots + \left(-\frac{\alpha_n}{\alpha_j}\right) a_n.$$

Sufficiency. Let

$$a_j = \alpha_1 a_1 + \dots \alpha_{j-1} a_{j-1} + \alpha_{j+1} a_{j+1} + \dots + \alpha_n a_n.$$

Then

$$\alpha_1 a_1 + \dots \alpha_{j-1} a_{j-1} + \alpha_j a_j + \alpha_{j+1} a_{j+1} + \dots + \alpha_n a_n = 0,$$

where $\alpha_j = -1$. We've got the non-trivial (since $\alpha_j = -1 \neq 0$) linear combination that equals to 0. $\square$

Corollary 4.25. *Three vectors in* $\mathbf{V}_3$ *are linearly dependent iff they are co-planar.*

Lemma 4.26. *If vectors $a_1, \ldots, a_n$ are expressed in terms of $b_1, \ldots, b_m$ and $n > m$ then $a_1, \ldots, a_n$ are linearly dependent.*

Proof. For some values α_{ij} we have that

$$a_j = \beta_{1j} b_1 + \beta_{2j} b_2 + \ldots + \beta_{mj} b_m \quad (j = 1, 2, \ldots, n) \tag{4.2}$$

We have to prove that if $n \geq m$ then there exist scalars $\alpha_1, \ldots, \alpha_n$, not all zero, such that

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_n a_n = o. \tag{4.3}$$

Substituting (4.2) into (4.3) we obtain

$$\alpha_1(\beta_{11} b_1 + \beta_{21} b_2 + \ldots + \beta_{m1} b_m) + \ldots + \alpha_n(\beta_{1n} b_1 + \beta_{2n} b_2 + \ldots + \beta_{mn} b_m) = o,$$

or, equivalently,

$$(\alpha_1 \beta_{11} + \ldots + \alpha_n \beta_{1n}) b_1 + \ldots + (\alpha_1 \beta_{m1} + \ldots + \alpha_n \beta_{mn}) b_m = o.$$

This holds if

$$\begin{cases} \alpha_1 \beta_{11} & + \ldots + & \alpha_n \beta_{1n} & = & 0, \\ \multicolumn{5}{c}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} \\ \alpha_1 \beta_{m1} & + \ldots + & \alpha_n \beta_{mn} & = & 0 \end{cases}$$

This is the homogeneous system of linear equations (with unknowns $\alpha_1, \ldots, \alpha_n$). Since $m < n$ then the system has a non-trivial solution $\alpha_1, \ldots, \alpha_n$, hence the linear combination in (4.3) is non-trivial and vectors $a_1, \ldots, a_n$ are linearly dependent. □

Corollary 4.27. *Let* $a_1, \ldots, a_n$ *are linearly independent. If vectors* $a_1, \ldots, a_n$ *are expressed in terms of* $b_1, \ldots, b_m$ *then* $n \leq m$*.*

Proof. (By contradiction) □

Corollary 4.28. *Two equivalent linear independent systems contain the same number of vectors.*

Proof. Let $a_1, \ldots, a_n$, $b_1, \ldots, b_m$ are two equivalent linear independent systems.

Vectors $a_1, \ldots, a_n$ can be expressed in terms of $b_1, \ldots, b_m$. Hence $n \leq m$.

But $b_1, \ldots, b_m$ can be expressed in terms of $a_1, \ldots, a_n$. Hence $m \leq n$.

So, $m = n$

□

Any linear independent subsystem equivalent to whole of system is called *a base* (or *basis*) *of the system.*

Corollary 4.29. *Any two bases of a system* $a_1, \ldots, a_n$ *contain the same number of vectors.*

This number is called the *rank* of the system and is written $\operatorname{rank}\{a_1, \ldots, a_n\}$.

Example 4.30.

4.6. Basis of linear space

Vector system $a_1, \ldots, a_n$ in V is called *complete* in V iff every vector in V is a linear combination of $a_1, \ldots, a_n$.

This is equivalent to the statement that $V = \mathrm{Lin}(a_1, \ldots, a_n)$.

A space V is said to be *finite-dimensional* iff there exist vectors $a_1, \ldots, a_n$ such that $V = \mathrm{Lin}(a_1, \ldots, a_n)$.

Vectors $a_1, \ldots, a_n$ in V are said to form a *basis* (or *base*) of V iff

(1) $a_1, \ldots, a_n$ is complete system in V.

(2) $a_1, \ldots, a_n$ are linearly independent.

Proposition 4.31. *Every finite-dimensional space, except* $\{o\}$*, has a basis.*

Proof. V is finite-dimensional hence $V = \mathrm{Lin}(a_1, \dots, a_n)$ for certain vectors $a_1, \dots, a_n$. As a basis of the space V we can take a basis of the system $a_1, \dots, a_n$. □

Proposition 4.32. *Any two bases for a space* V *contain the same number of vectors.*

Proof. Let $a_1, \dots, a_n$ and $b_1, \dots, b_m$ be two bases for V. Systems $a_1, \dots, a_n$ and $b_1, \dots, b_m$ are equivalent and linearly independent hence $n = m$. □

This number is called the *dimension* of V. Notation: $\dim V$. Naturally we define $\dim \{o\} = 0$.

Proposition 4.33. *Linearly independent vector system in V forms a basis iff any vector system containing more vectors in V is linearly dependent.*

Proof. Sufficiency. Let $a_1, \ldots, a_n$ form a basis. Vectors $b_1, \ldots, b_m$ are expressed in terms of linearly independent vectors $a_1, \ldots, a_n$. If $m > n$ then by Lemma 4.26 vectors $b_1, \ldots, b_m$ are linearly dependent.

Necessity. Let b be a vector in V. Consider vectors $a_1, \ldots, a_n, b$. By condition they are linearly dependent. Hence $b \in \operatorname{Lin}(a_1, \ldots, a_n)$. Thus $V = \operatorname{Lin}\{a_1, \ldots, a_n\}$. □

Proposition 4.34. *Let $V = \operatorname{Lin}\{a_1, \ldots, a_n\}$, i.e. the system $a_1, \ldots, a_n$ is complete. Then $a_1, \ldots, a_n$ form a basis iff*

(a) $n = 1$ *and* $a_1 \neq o$,

or

(b) *any system containing fewer vectors is not complete.*

Proof. Necessity. Let $b_1, \ldots, b_m$ form complete system, then linear independent system $a_1, \ldots, a_n$ is expressed in terms of $b_1, \ldots, b_m$, whence, by Lemma 4.26, $m \geq n$.

Sufficiency. For each $j = \{1, \ldots, n\}$ vectors $a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n$ don't form the basis, hence $a_j \notin \operatorname{Lin}(a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n)$. So by criterion of linear dependence vectors $a_1, \ldots, a_n$ are linear independent. □

Proposition 4.35. *Let* $\dim V = n$ *and vectors* $a_1, \ldots, a_n$ *are linearly independent then system* $a_1, \ldots, a_n$ *is complete and, therefore, forms a basis.*

Proof. Since $\dim V = n$, then any system containing more vectors is linearly dependent. Hence the system $a_1, \ldots, a_n$ is complete by Proposition 4.33. □

Proposition 4.36. *Let* $\dim V = n$ *and system* $a_1, \ldots, a_n$ *is complete, then it is linearly independent and, therefore, forms a basis.*

Proof. Since $\dim V = n$, then any system containg fewer vectors is not complete. Hence the system $a_1, \ldots, a_n$ is linear independent by Proposition 4.34. □

Dimension and basis of arithmetical space. Let's consider the following system of vectors in F^n:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \tag{4.4}$$

where j-th component of e_j is 1 and all other components of e_j are 0 $(j = 1, 2, \ldots, n)$.

Let's show that $e_1, e_2, \ldots, e_n$ form a basis of F^n and, therefore, $\dim F^n = n$.

Really, system (4.4) is linearly independent since by equating all components in l.h.s. and r.h.s. of equality

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = o, \tag{4.5}$$

we get $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$, so the linear combination (4.5) is trivial.

The system (4.4) is complete, since for any vector

$$a = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \in F^n$$

we have $a = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n$.

The system (4.4) is called *standard basis of arithmetical n-dimensional space*.

Exercize 4.37. Prove that vectors

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

form a basis of F^n.

Dimension and basis of geometric spaces.

Proposition 4.38.

1. $\dim \mathbf{V}_1 = 1$. *A system containing one vector in* $\mathbf{V}_1$ *forms a basis iff the vector is not* o.

2. $\dim \mathbf{V}_2 = 2$. *A system containing two vectors in* $\mathbf{V}_2$ *forms a basis iff the vectors are not collinear.*

3. $\dim \mathbf{V}_3 = 3$. *A system containing three vectors in* $\mathbf{V}_3$ *forms a basis iff the vectors are not co-planar.*

NUll space (solution space)

Dimension of the space of all solutions to the homogenius sytem of linear equations.

Examples.

4.7. Vector coordinates

Let $e_1, \ldots, e_n$ form a basis of space V over field F. Then for any vector a in V there exist numbers $\alpha_1, \ldots, \alpha_n$ in F such that $a = \alpha_1 e_1 + \ldots + \alpha_n e_n$.

Coefficients $\alpha_1, \ldots, \alpha_n$ are called the *coordinates of* a *in basis* $e_1, \ldots, e_n$.

Proposition 4.39. *Given basis* $e_1, \ldots, e_n$ *and a vector* a *the coordinates of* a *in* $e_1, \ldots, e_n$ *are unique.*

Proof. Suppose that we have two representations:

$$a = \alpha_1 e_1 + \ldots + \alpha_n e_n = \beta_1 e_1 + \ldots + \beta_n e_n.$$

Then

$$(\alpha_1 - \beta_1)e_1 + \ldots + (\alpha_n - \beta_n)e_n = o.$$

But vectors $e_1, \ldots, e_n$ be linearly independent, so all coefficients in obtained linear conbination ar zeros, hence $\alpha_j = \beta_j$ $(j = 1, \ldots, n)$. □

Coordinates can be written in a column that's called the *coordinate column* and denoted as $[a]_e$.

Thus, if $a = \alpha_1 e_1 + \ldots + \alpha_n e_n$, then

$$[a]_e = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

A coordinate column can be considered as a vector in F^n.

Proposition 4.40. *For any vectors a, b in V and all α in F*

$$[a+b]_e = [a]_e + [b]_e,$$
$$[\alpha a]_e = \alpha [a]_e.$$

Proof. Let

$$a = \alpha_1 e_1 + \ldots + \alpha_n e_n,$$
$$b = \beta_1 e_1 + \ldots + \beta_n e_n.$$

Then

$$a + b = (\alpha_1 + \beta_1)e_1 + \ldots + (\alpha_n + \beta_n)e_n,$$

hence $[a+b]_e = [a]_e + [b]_e$. The proof of the second equality is similar. □

4.8. Rank of a matrix

Let $F^{m \times n}$ be the set of all matrix with m rows and n columns.

Let $A \in F^{m \times n}$.

Consider columns $a_1, \dots, a_n$ of $A = (a_1, \dots, a_n)$ as vectors of F^m.

Consider rows $\widetilde{a}_1, \dots, \widetilde{a}_m$ of A as vectors of F^n.

A basis of the system $a_1, \dots, a_n$ is called a *column basis* of A.

The rank of the system $a_1, \dots, a_n$ is called a *column rank* of A.

A basis of the system $\widetilde{a}_1, \dots, \widetilde{a}_n$ is called a *row basis* of A.

The rank of the system $\widetilde{a}_1, \dots, \widetilde{a}_n$ is called a *row rank* of A.

$$A = \begin{pmatrix} 1 & 0 & 3 & 0 & 2 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \begin{aligned} \widetilde{a}_1 &= (1, 0, 3, 0, 2), \\ \widetilde{a}_2 &= (0, 1, 1, 0, 1), \\ \widetilde{a}_3 &= (0, 0, 0, 1, 6), \\ \widetilde{a}_4 &= (0, 0, 0, 0, 0), \end{aligned}$$

$$a_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \ a_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \ a_3 = \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \ a_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \ a_5 = \begin{pmatrix} 2 \\ 1 \\ 6 \\ 0 \end{pmatrix},$$

a_1, a_2, a_4 form a column base. This sytem is l.i. and equivalent to the whole of system: $a_3 = 3a_1 + a_2, \quad a_5 = 2a_1 + a_2 + 6a_4$

$\widetilde{a}_1$, $\widetilde{a}_2$, $\widetilde{a}_3$ form a row base. This sytem is l.i. and equivalent to the whole of system.

Proposition 4.41. *The column rank of a matrix in r.r.e.f. is equal to its row rank.*

Proof.

$$A = \begin{pmatrix} 0 & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots & * & 0 & * & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots & * \\ \multicolumn{15}{c}{\dotfill} \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & * & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \multicolumn{15}{c}{\dotfill} \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Columns $j_1, \dots, j_r$ of a matrix in r.r.e.f. form a column base.

From the other hand, rows $1, 2, \dots, r$ form a row base.

So row rank and column rank equal to r. □

Proposition 4.42. *Let matrix* $A' = (a'_1, \ldots, a'_n)$ *is row equivalent to* $A = (a_1, \ldots, a_n)$*. If* $a_{j_1}, \ldots, a_{j_r}$ *form a column base of* A*, then* $a'_{j_1}, \ldots, a'_{j_r}$ *form a column base of* A'*.*

Proof. Let's prove that vectors $a'_{j_1}, \ldots, a'_{j_r}$ are l.i.

Since vectors $a_{j_1}, \ldots, a_{j_r}$ are l.i. then the system of linear equations (with unknowns $\alpha_1, \ldots, \alpha_r$)

$$\alpha_1 a_{j_1} + \ldots + \alpha_r a_{j_r} = 0 \tag{4.6}$$

has a unique (zero) solution.

Let's perform the same row operation with the matrix of this system as with A.

It's obvious that we get the system

$$\alpha'_1 a_{j_1} + \ldots + \alpha'_r a_{j_r} = 0.$$

This system is equivalent to (4.6), hence it also has the unique (zero) solution that means that vectors $a'_{j_1}, \ldots, a'_{j_r}$ are l.i.

Now prove that for every $j \in \{1, \ldots, n\}$ the vector a'_j is expressed in terms of $a'_{j_1}, \ldots, a'_{j_r}$.

The system

$$\alpha_1 a_{j_1} + \ldots + \alpha_r a_{j_r} = a_j$$

is consistent, so the system

$$\alpha_1 a'_{j_1} + \ldots + \alpha_r a'_{j_r} = a'_j,$$

is also consistent (because they are row equivalent). □

Corollary 4.43. *Elementary row operations don't change the column rank of a matrix.*

Corollary 4.44. *Elementary column operations don't change the row rank of a matrix.*

Example 4.45. Find a column base of the matrix A and express its column as a linear conbination of columns in the base:

$$A = \begin{pmatrix} 1 & 2 & 3 & -1 & 1 \\ -3 & -6 & -9 & 3 & -3 \\ 2 & 4 & 6 & 1 & 2 \\ 2 & 5 & 7 & -1 & -1 \end{pmatrix}.$$

Using elementary row operations transform the matrix to r.r.e.f.:

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

$$\to \begin{pmatrix} 1 & 0 & 1 & -3 & 7 \\ 0 & 1 & 1 & 1 & -3 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 & 0 & 7 \\ 0 & 1 & 1 & 0 & -3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

In the obtained matrix columns 1, 2, 4 form a base. Hence a_1, a_2, a_4 form a base in A.

$a_3 = a_1 + a_2$, $a_5 = 7a_1 - 3a_2$.

Proposition 4.46. *Let A' be row equivalent to A. If $\widetilde{a}_1, \ldots, \widetilde{a}_m$ are rows of A, $\widetilde{a}'_1, \ldots, \widetilde{a}'_m$ are rows of A', then $L(\widetilde{a}_1, \ldots, \widetilde{a}_m) = L(\widetilde{a}'_1, \ldots, \widetilde{a}'_m)$. So systems $\widetilde{a}_1, \ldots, \widetilde{a}_m$ and $\widetilde{a}'_1, \ldots, \widetilde{a}'_m$ are equivalent.*

Proof. Let's consider the third kind of row transfromations (add a multiple of a row to another row). It is not hard to see that systems

$$\widetilde{a}_1, \ldots, \widetilde{a}_i, \ldots, \widetilde{a}_j, \ldots, \widetilde{a}_m;$$
$$\widetilde{a}_1, \ldots, \widetilde{a}_i + \alpha\widetilde{a}_j, \ldots, \widetilde{a}_j, \ldots, \widetilde{a}_m$$

are equivalent. □

Corollary 4.47. *Elementary row operations don't change the row rank of a matrix.*

Corollary 4.48. *Elementary column operations don't change the column rank of a matrix.*

Theorem 4.49. *The row rank of a matrix equals to its column rank.*

Proof. A matrix can be reduced to r.r.e.f. by row operations. But row and column ranks of a matrix in r.r.e.f. coincide. Hence they coincide for the original matrix. □

The row rank (column rank) of a matrix is simply called a *rank of the matrix*. Notation: $\operatorname{rank} A$.

Theorem 4.50 (Kroneker, Capelli)**.** *The system of linear equation is consistent iff the rank of the coefficient matrix of the system equals to the rank of enlarged matrix of the system.*

Proof. Using elementary row operations transform the enlarged matrix (A, b) to r.r.e.f. (A', b').

But a system is consistent iff the r.r.e.f. of the enlarged matrix doesn't contain rows

$$(0, 0, \ldots, 0, 1).$$

So the system is consistent iff $\operatorname{rank}(A', b') = \operatorname{rank} A'$, but $\operatorname{rank}(A', b') = \operatorname{rank}(A, b)$, $\operatorname{rank} A' = \operatorname{rank} A$, that's equivalent to $\operatorname{rank}(A, b) = \operatorname{rank} A$. □

Corollary 4.51. *The number of dependent unknowns of a consistent system of linear equations equals to the rank of the matrix.*

4.9. Null space of a matrix

Consider a homogenius system of linear equations with coefficient matrix $A \in F^{m \times n}$.

Denote the set of all solutions to the system as $N(A)$.

We know that $N(A)$ is subspace in F^n.

It's call a *null space of a matrix*.

The dimension of $N(A)$ is called the *defect* (or *nullity*) of A. Notation: $\operatorname{def} A$

Exercize 4.52. Consider the system

$$\begin{cases} \ \ x_1 + 3x_2 + 2x_3 + 4x_4 + x_5 = 0, \\ 3x_1 + 2x_2 + 2x_3 + \ \ x_4 + x_5 = 0. \end{cases}$$

The coefficient matrix can be transformd to the following r.r.e.f.:

$$A = \begin{pmatrix} 1 & 3 & 2 & 4 & 1 \\ 3 & 2 & 2 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & {}^{2}/_{7} & {}^{-5}/_{7} & {}^{1}/_{7} \\ 0 & 1 & {}^{4}/_{7} & {}^{11}/_{7} & {}^{2}/_{7} \end{pmatrix}.$$

Thus the common solution to the system can written as

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -{}^{2}/_{7} \\ -{}^{4}/_{7} \\ 1 \\ 0 \\ 0 \end{pmatrix} \cdot t_1 + \begin{pmatrix} {}^{5}/_{7} \\ -{}^{11}/_{7} \\ 0 \\ 1 \\ 0 \end{pmatrix} \cdot t_2 + \begin{pmatrix} -{}^{1}/_{7} \\ -{}^{2}/_{7} \\ 0 \\ 0 \\ 1 \end{pmatrix} \cdot t_3$$

or (multiplying by 7)

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -2 \\ -4 \\ 7 \\ 0 \\ 0 \end{pmatrix} \cdot \tau_1 + \begin{pmatrix} 5 \\ -11 \\ 0 \\ 7 \\ 0 \end{pmatrix} \cdot \tau_2 + \begin{pmatrix} -1 \\ -2 \\ 0 \\ 0 \\ 7 \end{pmatrix} \cdot \tau_3$$

$\tau_1, \tau_2, \tau_3 \in F$.

$\operatorname{rank} A = 2$ (the number of non-zero rows in r.r.e.f. = the number of dependent unknowns),

$\operatorname{def} A = \dim N(A) = 3$ (the number of independent unknowns)

$$\operatorname{rank} A + \operatorname{def} A = n.$$

Theorem 4.53. *Let* $A \in F^{m\times n}$, rank $A = r$*. Then* def $A = n - r$*.*

A basis of $N(A)$ is called a *fundamental system of solutions* to the homogenius system of equations.

Corollary 4.54. *Any linearly independent system of* $n - r$ *partial solutions to a system of homogenius linear equations forms a fundamental system of solutions.*

4.10. Linear variety

Let V be a linear space over a field F.

Let $a_0 \in V$.

L is a subspace in V.

Notation $a_0 + L$ means the set of all possible vectors $a_0 + x$, where x is arbitrary vectors in L.

$a_0 + L$ is called a *linear variety* (or *linear manifold*) generated form the vector a_0 and the subspace V.

The *dimension* of $a_0 + L$ is $\dim(a_0 + L) = \dim L$.

Proposition 4.55. *The set of all solutions to the consistent system of linear equtions is a linear variety.*

Example 4.56. Consider the system

$$\begin{cases} x_1 + 2x_2 + x_3 = 2, \\ 2x_1 + 3x_2 + 4x_3 = 1, \end{cases}$$

Transform its enlarged matrix to r.r.e.f.:

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & 5 & -4 \\ 0 & 1 & -2 & 3 \end{array}\right)$$

The common solution is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ -3 \\ 0 \end{pmatrix} + \begin{pmatrix} -5 \\ 2 \\ 1 \end{pmatrix} \cdot t, \qquad t \in F$$

The set of all solutions is linear variety $a_0 + L$, where

$$a_0 = \begin{pmatrix} 4 \\ -3 \\ 0 \end{pmatrix}, \qquad L = \operatorname{Lin}\{a_1\}, \qquad a_1 = \begin{pmatrix} -5 \\ 2 \\ 1 \end{pmatrix}.$$

Theorem 4.57. *Let* $A \in F^{m\times n}$, $b \in F^m$, $\operatorname{rank} A = r$. *The set of all solutions to the consistent system of linear equtions with enlarged matrix* (A, b) *is a linear variety* $a_0 + L$, *where* a_0 *is any partial solution of the system, and* L *is the common solution to the corresponding homogenius system (i.e. null space of* A*). Moreover,* $\dim(a_0 + L) = n - r$.

4.11. Representing linear subspaces and varieties as a set of all solutions to a system of linear equations

Theorem 4.58. *Any linear subspace in F^n can be represented as the set of all solutions to a system of linear hoogenius equtions.*

Example 4.59. Let $L = L(b_1, b_2) \subseteq \mathbb{R}^4$, where

$$b_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \qquad b_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Find a system of linear equations whoes the set of solutions is L.

Consider the matrix B, compounded of columns b_1, b_2, and the matrix $\widetilde{B}$, compounded of b_1, b_2, x. We have that $x \in L(b_1, b_2)$ iff $\operatorname{rank} B = \operatorname{rank} \widetilde{B}$. Using elementary operations with rows let's reduce $\widetilde{B}$ to row echelon form:

$$\widetilde{B} = \left(\begin{array}{cc|c} 1 & 1 & x_1 \\ 1 & 1 & x_2 \\ 0 & 1 & x_3 \\ 0 & 1 & x_4 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & 1 & x_1 \\ 0 & 0 & x_2 - x_1 \\ 0 & 1 & x_3 \\ 0 & 1 & x_4 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & 1 & x_1 \\ 0 & 1 & x_3 \\ 0 & 0 & x_2 - x_1 \\ 0 & 0 & x_4 - x_3 \end{array}\right).$$

The rank of the matrix without the last column is 2. The rank of whole of matrix is also 2 iff $x_2 - x_1 = 0$ and $x_4 - x_3$. These equations are compounded the requiered system:

$$\begin{cases} x_1 - x_2 \phantom{{}+x_3 - x_4} = 0, \\ \phantom{x_1 - x_2 +{}} x_3 - x_4 = 0 \end{cases} \tag{4.7}$$

Theorem 4.60. *Any linear variety $b_0 + L$ in F^n can be represented as the set of solution to a system of linear equtions.*

Example 4.61. Let L is the same as in Example 4.59, and

$$b_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Find the system of linear equtions whose the set of all solutions is $b_0 + L$.

Substituring the components of b_0 into the l.h.s. of (4.7) we get values 1, -1 correspondingly, hence, the set of all solutions to the system

$$\begin{cases} x_1 - x_2 \qquad\qquad = \;\; 1, \\ \qquad\qquad x_3 - x_4 = -1 \end{cases}$$

coincides with $b_0 + L$.

Example 4.62. Find the intersection of linear varieties $a_0 + L(a_1, a_2)$ and $b_0 + L(b_1, b_2)$, where

$$a_0 = \begin{pmatrix} 3 \\ 2 \\ 0 \\ 2 \end{pmatrix}, \quad a_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

$$b_0 = \begin{pmatrix} -4 \\ -3 \\ 1 \\ -2 \end{pmatrix}, \quad b_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 2 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

First, find the systems of linear equations describing each of the varieties:

$$\begin{cases} x_1 - x_2 - x_3 = 1, \\ 2x_1 - 2x_2 - x_4 = 0, \end{cases} \qquad \begin{cases} x_1 - 2x_2 = 2, \\ -2x_2 + x_4 = 4 \end{cases}$$

The intersection of the linear varieties is described by the system

$$\begin{cases} x_1 - x_2 - x_3 = 1, \\ 2x_1 - 2x_2 - x_4 = 0, \\ x_1 - 2x_2 = 2, \\ - 2x_2 + x_4 = 4. \end{cases}$$

the set of all its solutions is the requiered linear varieties $c_0 + L(c_1)$, where

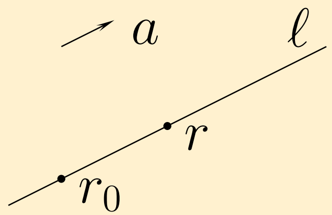
$$c_0 = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 4 \end{pmatrix}, \quad c_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 2 \end{pmatrix}.$$

Chapter 5

Lines and Planes

5.1. Lines in the plane

Consider a (straight) line ℓ in the plane passing throw a point r_0 and parallel to a vector $a \neq o$.



A point r is in the line $\ell \Leftrightarrow r - r_0$ is collinear to a

$\Leftrightarrow$ there exists a real scalar t such that $r - r_0 = t \cdot a$

$\Leftrightarrow r = r_0 + t \cdot a$

The last equation is the *vectorial equation of the line.*

a is called the *direction vector* of the line, t is parameter

Note 5.1. Each non zero multiple of a direction vector is a direction vector

A line is 1-dimensional linear variety!

Take a coordinate system. All vectors have unique coordinates. Say

$$[r] = (x, y), \qquad [r_0] = (x_0, y_0), \qquad [a] = (\alpha, \beta)$$

From $r = r_0 + ta$ we have

$$\begin{cases} x = x_0 + t\alpha, \\ y = y_0 + t\beta, \end{cases} \qquad t \in \mathbb{R}.$$

These are the *parametric* (or *explicit*) *equations* of the line ℓ

We can express t in each of the equations:

$$t = \frac{x - x_0}{\alpha}, \qquad t = \frac{y - y_0}{\beta}.$$

Eliminate t out of the parametric equations:

$$\frac{x - x_0}{\alpha} = \frac{y - y_0}{\beta}.$$

This is the *canonical equation* of the line. If one of the values α, β is zero, the corresponding numerator is zero.

Transforming the canonical equation we obtain

$$\beta x - \alpha y - \beta x_0 + \alpha y_0 = 0.$$

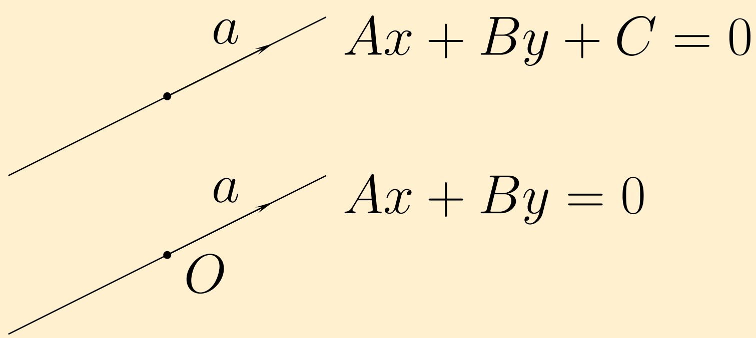
Denote $A = \beta$, $B = -\alpha$, $C = -\beta x_0 + \alpha y_0$. We get so called *general* (or *implicit*) *equation of the line in the plane*:

$$Ax + By + C = 0.$$

Proposition 5.2. *Let ℓ have equation $Ax + By + C = 0$.*

The coordinates of a direction vector of ℓ satisfy the equation $Ax + By = 0$.

So the line with equation $Ax + By = 0$ is parallel to the line ℓ and goes through the origin.



Exercize 5.3. $[A] = (1, 2)$, $[B] = (3, 1)$. Lets find a direction vector, write parametric equation, canonical equation, and general equation.

$a = \overrightarrow{AB}$, $[a] = (2, -1)$

$r_0 = A$, $[r_0] = (1, 2)$

Parametric (explicit) equation:

$$\begin{cases} x = 1 + 2t, \\ y = 2 - t, \end{cases} \qquad t \in \mathbb{R}.$$

Canonical equation:

$$\frac{x-1}{2} = \frac{y-2}{-1}.$$

General (implicit) equation:

$$x + 2y - 5 = 0.$$

Exercize 5.4. Draw lines

$2x + 3y = 6$

$2x - 3y = 6$

$-2x + 3y = 6$

$2x + 3y = -6$

$x = 2$

$y = -6$

$x = 0$

$y = 0$

$$\frac{x-5}{2} = \frac{y+2}{-1}.$$

$$\frac{x}{1} = \frac{y}{-1}.$$

$$\begin{cases} x = -1 - 3t, \\ y = 2 - 2t, \end{cases}$$

5.2. Normal vector of a line in the plane

Proposition 5.5. *Let the system of coordinates be rectangular cartesian. A line* ℓ *has equation* $Ax + By + C = 0$*. Let* $[n] = (A, B)$*.*

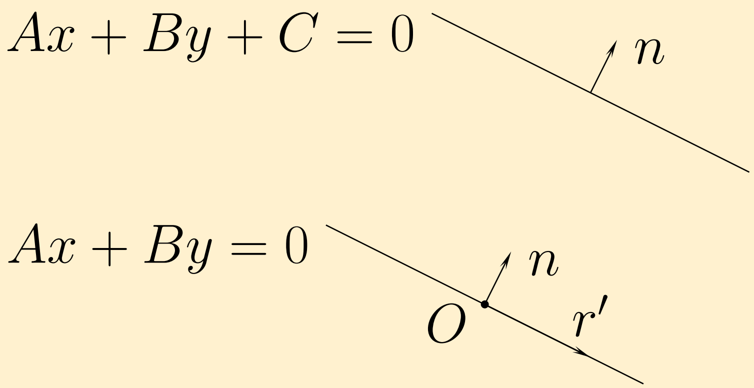
Then n *is orthogonal to all the vectors in line* ℓ *and the direction of* n *is ortogonal to* ℓ *and to all parallel lines.*

n *is called a* normal vector *to the line* ℓ*.*

Proof. The line with equation $Ax + By = 0$ is parallel to ℓ and goes through the origin.

For each point $r'(x', y')$ we have

r' is in $\ell \Leftrightarrow Ax' + By' = 0 \Leftrightarrow r'$ and n are orthogonal □



Note 5.6. Normal vector is orthogonal to a direction vector of the line

Note 5.7. Each non zero multiple of a normal vector is a normal vector

Normal equation of the line

$$(r - r_0, n) = 0.$$

Exercize 5.8. What are the direction vectors and normal vector of the following lines?

$2x + 3y = 6$

$2x - 3y = 6$

$-2x + 3y = 6$

$2x + 3y = -6$

$x = 2$

$y = -6$

$x = 0$

$y = 0$

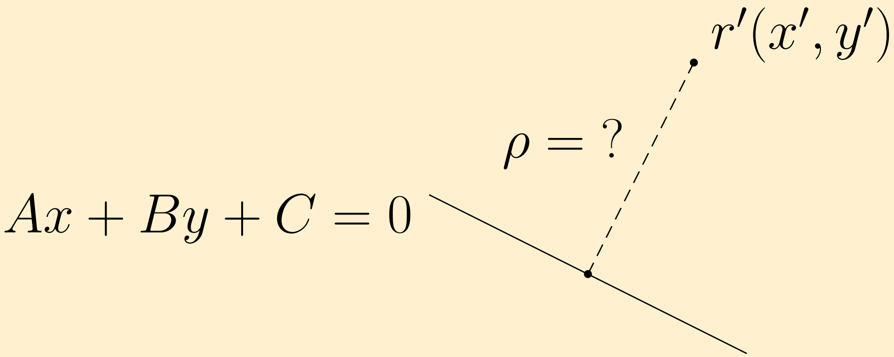
$$\frac{x-5}{2} = \frac{y+2}{-1}.$$

$$\frac{x}{1} = \frac{y}{-1}.$$

$$\begin{cases} x = -1 - 3t, \\ y = 2 - 2t, \end{cases}$$

5.2.1. Distance from a point to a line in the plane

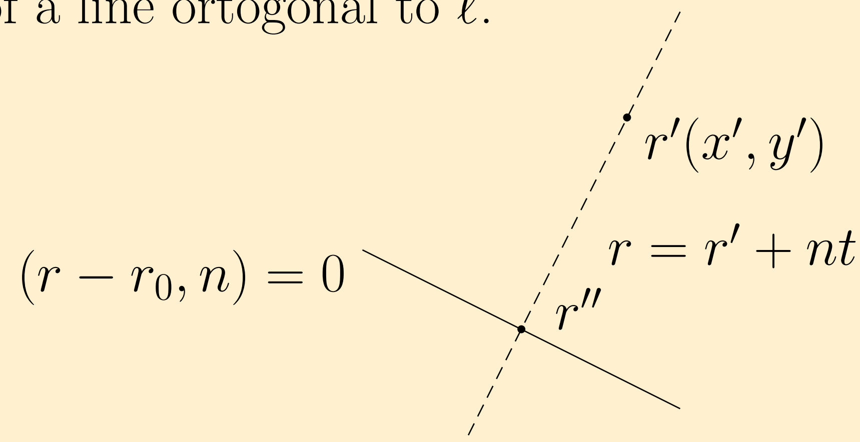
We want to find the distance ρ from a point $r'(x', y')$ to a line $Ax + By + C = 0$. Coordinate system is rectangular cartesian.



$$\rho = \frac{Ax' + By' + C}{\sqrt{A^2 + B^2}}$$

Proof. Let the line ℓ has a normal equation $(r - r_0, n) = 0$, where $n(A, B)$ is a normal vector to the line ℓ and r_0 is a point in the line.

So n is a direction vector of a line ortogonal to ℓ.



And the line passing through r' and orthogonal to ℓ has equation $r = r' + nt$.

Let the dot where ℓ and $r = r' + nt$ intersect be r''. Then the distance to be found is the distance between r' and r''.

Substitute the equation $r = r' + nt$ into the equation of ℓ:

$$(r' + nt - r_0, n) = 0$$

Whence

$$t = \frac{(r' - r_0, n)}{(n, n)} = \frac{(r' - r_0, n)}{|n|^2}$$

So

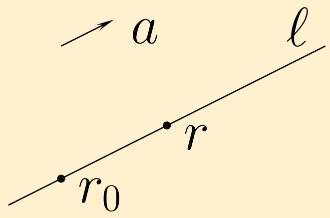
$$r'' = r' + nt = r' + n \cdot \frac{(r' - r_0, n)}{|n|^2}$$

$$\rho = |r'' - r'| = \left| n \cdot \frac{(r' - r_0, n)}{|n|^2} \right| = \frac{|n| \cdot (r' - r_0, n)}{|n|^2} = \frac{(r' - r_0, n)}{|n|}.$$

But $(r' - r_0, n) = Ax' + By' - Ax_0 - By_0 = Ax' + By' + C$. $\square$

5.3. Lines in space

Consider a (straight) line ℓ in space passing throw a point r_0 and parallel to a vector $a \neq o$.



A point r is in the line $\ell \Leftrightarrow r - r_0$ is collinear to a

$\Leftrightarrow$ there exists a real scalar t such that $r - r_0 = t \cdot a$

$\Leftrightarrow r = r_0 + t \cdot a$

The last equation is the *vectorial equation of the line.*

a is called the *direction vector* of the line, t is parameter

Note 5.9. Each non zero multiple of a direction vector is a direction vector

A line is 1-dimensional linear variety!

Take a coordinate system. All vectors have unique coordinates. Say

$$[r] = (x, y, z), \qquad [r_0] = (x_0, y_0, z_0), \qquad [a] = (\alpha, \beta, \gamma)$$

From $r = r_0 + ta$ we have

$$\begin{cases} x = x_0 + t\alpha, \\ y = y_0 + t\beta, \\ z = z_0 + t\gamma, \end{cases} \qquad t \in \mathbb{R}.$$

These are the *parametric* (or *explicit*) *equations* of the line ℓ

We can express t in each of the equations:

$$t = \frac{x - x_0}{\alpha}, \qquad t = \frac{y - y_0}{\beta}, \qquad t = \frac{z - z_0}{\beta},$$

Eliminate t out of the parametric equations:

$$\frac{x - x_0}{\alpha} = \frac{y - y_0}{\beta} = \frac{z - z_0}{\gamma}.$$

This is the *canonical equation* of the line in space. If one of the values α, β, γ is zero, the corresponding numerator is zero.

Transforming the canonical equation

$$\frac{x-x_0}{\alpha}=\frac{y-y_0}{\beta}=\frac{z-z_0}{\gamma}.$$

we obtain

$$\begin{cases} \beta x-\alpha y-\beta x_0+\alpha y_0=0, \\ \gamma y-\beta z-\gamma y_0+\beta z_0=0. \end{cases}$$

or

$$\begin{cases} A_1x+B_1y+C_1z+D_1=0, \\ A_2x+B_2y+C_2z+D_2=0. \end{cases}$$

These are so called *general* (or *implicit*) *equations of the line in space.*

Note 5.10. A line is 1-dimensional linear variety and in space it can be represented by implicit equation with rank equal to 2.

Example 5.11. Find the parametric, canonical and general equations of the line passing throw the point $(1, 3, 2)$ parallel to the vector $(1, -1, -2)$.

$$\begin{cases} x = 1 + t, \\ y = 3 - t, \\ z = 2 - 2t, \end{cases} \qquad t \in \mathbb{R}.$$

$$\frac{x-1}{1} = \frac{y-3}{-1} = \frac{z-2}{-2}$$

$$\begin{cases} -x + 1 = y - 3, \\ -2y + 6 = -z + 2 \end{cases} \qquad \Rightarrow \qquad \begin{cases} x + y - 4 = 0, \\ 2y - z - 4 = 0 \end{cases}$$

Example 5.12. Find the parametric, canonical and general equations of the line passing throw points $A(1,3,2)$ and $B(2,3,1)$.

Vector $\overrightarrow{AB}$ can be taken as a direction vector of the line: $[\overrightarrow{AB}] = (1, 0, -1)$.

$$\begin{cases} x = 1 \ + t, \\ y = 3, \\ z = 2 \ - t, \end{cases} \qquad t \in \mathbb{R}.$$

$$\frac{x-1}{1} = \frac{y-3}{0} = \frac{z-2}{-1} \qquad \Leftrightarrow \qquad \frac{x-1}{1} = \frac{z-2}{-1}, \quad y = 3$$

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

Example 5.13. Find parametric and canonical equations of the line

$$\begin{cases} x - 2y + 4z + 1 = 0, \\ x + 3y - 2z - 4 = 0 \end{cases}$$

Find the general solution of the system:

$$\begin{cases} x = 1 - 8t, \\ y = 1 + 6t, \\ z = \phantom{1 + {}} 5t. \end{cases}$$

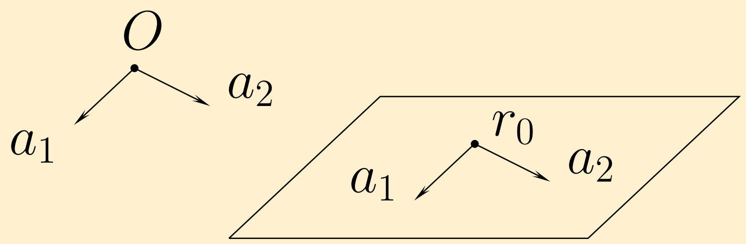
These are parametric equations of the line.

Canonical equations:

$$\frac{x-1}{-8} = \frac{y-1}{6} = \frac{z}{5}.$$

5.4. Planes in space

Take a plane passing through a point r_0 and parallel to non-collinear vectors a_1 and a_2.



r is in the plane $\Leftrightarrow$ $r - r_0$, a_1, a_2 are coplanar

$\Leftrightarrow$ there exist scalars t_1, t_2 such that $r - r_0 = t_1 a_1 + t_2 a_2$, or $r = r_0 + t_1 a_1 + t_2 a_2$. This is *vector parametric equations of the plane.*

It says that a plane in space is 2-dimensional linear variety!

Take a coordinate system.

$[r_0] = (x_0, y_0, z_0)$, $[a_1] = (\alpha_1, \beta_1, \gamma_1)$, $[a_2] = (\alpha_2, \beta_2, \gamma_2)$, $[r] = (x, y, z)$.

From $r = r_0 + t_1 a_1 + t_2 a_2$ we get parametric equations of the plane:

$$\begin{cases} x = x_0 + \alpha_1 t_1 + \alpha_2 t_2, \\ y = y_0 + \beta_1 t_1 + \beta_2 t_2, \\ z = z_0 + \gamma_1 t_1 + \gamma_2 t_2, \end{cases} \qquad t_1 \in \mathbb{R},\ t_2 \in \mathbb{R}.$$

We know that each 2-dimensional linear variety in 3-dimensional space can be represented as a set of all solution to a linear equation

$$Ax + By + Cx + D = 0.$$

This is the *general* (*implicit*) *equation* of the plane.

Given parametric equations of a plane we can find its general equation using the general technique (for representing linear varieties as a set of solutions to linear systems), but in this particular case we can use a special method.

Vectors $r - r_0$, a_1 and a_2 are coplanar $\Leftrightarrow$ their mixed product is 0:

$$(r - r_0, a_1, a_2) = 0 \qquad \Leftrightarrow \qquad \begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \end{vmatrix} = 0.$$

This expression gets us the general equation of the plane.

Example 5.14. Find parametric and general equations of the plane passing throuh the point $(1, 2, 3)$ parallel to vectors $(-2, 1, -1)$, $(1, -1, 2)$.

$$\begin{cases} x = 1 - 2t_1 + \ \ t_2, \\ y = 2 + \ \ t_1 - \ \ t_2, \\ z = 3 - \ \ t_1 + 2t_2, \end{cases} \qquad t_1 \in \mathbb{R},\ t_2 \in \mathbb{R}.$$

Let's find a general equation:

$$\begin{vmatrix} x-1 & y-2 & z-3 \\ -2 & 1 & -1 \\ 1 & -1 & 2 \end{vmatrix} = x + 3y + z - 10 = 0.$$

The general equation is $x + 3y + z - 10 = 0$.

Example 5.15. Find parametric and general equations of the plane passing throuh points $A(1,2,3)$, $B(2,1,1)$, $C(3,1,2)$.

We can take $\overrightarrow{AB}(1,-1,-2)$, $\overrightarrow{AC}(2,-1,-1)$ as direction vectors.

$$\begin{cases} x = 1 + t_1 + 2t_2, \\ y = 2 - t_1 - t_2, \\ z = 3 - 2t_1 - t_2, \end{cases} \qquad t_1 \in \mathbb{R},\ t_2 \in \mathbb{R}.$$

General equation:

$$\begin{vmatrix} x-1 & y-2 & z-3 \\ 1 & -1 & -2 \\ 2 & -1 & -1 \end{vmatrix} = -x - 3y + z + 4 = 0.$$

Example 5.16. Find parametric equations of the plane $2x - 3y + z - 2 = 0$.

Solving the equation we obtain:

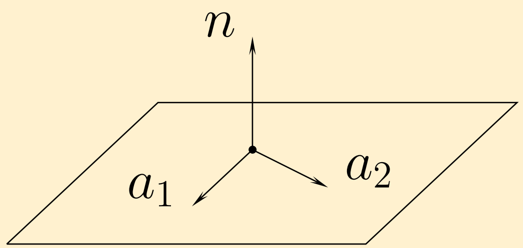
$$\begin{cases} x = t_1, \\ y = t_2, \\ z = 2 - 2t_1 + 3t_2, \end{cases} \qquad t_1 \in \mathbb{R},\ t_2 \in \mathbb{R}.$$

5.5. Normal equation of a plane in space

Proposition 5.17. *Let the system of coordinates be rectangular cartesian. A plane has equation* $Ax + By + Cz + D = 0$*. Let* $[n] = (A, B, C)$*.*

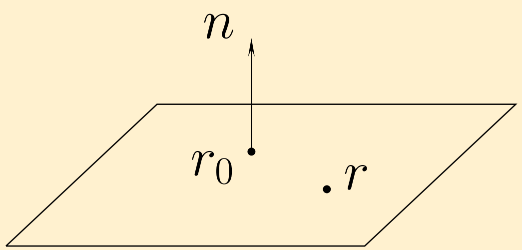
Then n *is orthogonal to all the vectors in the plane and the direction of* n *is ortogonal to the plane and to all parallel planes.*

n *is called a* normal vector *to the plane.*



a_1, a_2 are direction vectors
n is a normal vector

r is in the plane $\Leftrightarrow$ $r - r_0$ is perpendicular to n $\Leftrightarrow$ $(r - r_0, n) = 0$



$(r - r_0, n) = 0$ is *normal equation* of the plane

Let a coordinate system be rectangular. $[r] = (x, y, z)$, $[r_0] = (x_0, y_0, z_0)$, $[n] = (A, B, C)$

Then $A(x - x_0) + B(y - y_0) + C(z - z_0) = 0$ or $Ax + By + Cz - Ax_0 - By_0 - Cz_0 = 0$.
Denote $D = -Ax_0 - By_0 - Cz_0$ then the normal equation has the form
$Ax + By + Cz + D = 0$

Example 5.18. Find the normal vector of the plane $5x + z - 1 = 0$ (coordinate system is rectangular).

$[n] = (5, 0, 1)$

Example 5.19. Find the normal vector of the plane

$$\begin{cases} x = 1 - 2t_1 + 3t_2, \\ y = 1 + \ \ t_1 - 4t_2, \\ z = 1 \qquad\quad + 2t_2, \end{cases}$$

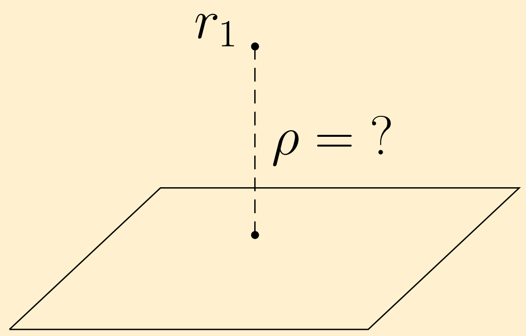
Direction vectors are $[a_1] = (-2, 1, 0)$, $[a_2] = (3, -4, 2)$

Vector $n = [a_1, a_2]$ (cross product of a_1 and a_2) is orthogonal to a_1 and a_2, hence it orthogonal to each vector parallel to the plane.

$$n = [a_1, a_2] = \begin{vmatrix} e_1 & e_2 & e_3 \\ -2 & 1 & 0 \\ 3 & -4 & 2 \end{vmatrix} = 2e_1 + 4e_2 + 5e_3$$

$[n] = (2, 4, 5)$

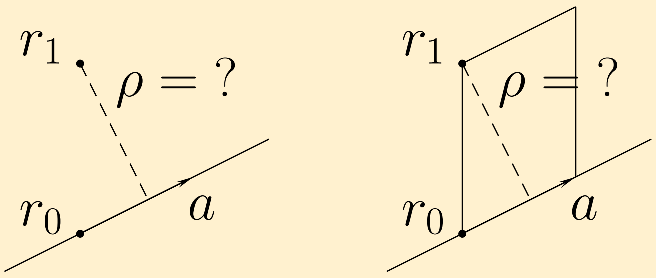
5.5.1. Distance from a point to a plane



$(r - r_0, n) = Ax + By + Cz + D = 0$, $[r_1] = (x_1, y_1, z_1)$

$$\rho = \frac{(r_1 - r_0, n)}{|n|} = \frac{|Ax_1 + By_1 + Cz_1 + D|}{\sqrt{A^2 + B^2 + C^2}}$$

5.5.2. Distance from a point to a line in space



ρ is a height of the parallelogramm generated from vectors a and $r_1 - r_0$

$$\rho = \frac{S}{|a|} = \frac{\big|[r_1 - r_0, a]\big|}{|a|}$$

Example 5.20. $[r_1] = (1,2,3)$,

$$\begin{cases} x = -1 + \ \ t, \\ y = \ \ \ 1 + 2t, \\ z = \ \ \ 2 - 3t, \end{cases}$$

$[r_0] = (-1,1,2)$, $[a] = (1,2,-3)$

$$[r_1 - r_0, a] = \begin{vmatrix} e_1 & e_2 & e_3 \\ 2 & 1 & 1 \\ 1 & 2 & -3 \end{vmatrix} = -5e_1 + 7e_2 + 3e_3$$

$$\rho = \frac{\big|[r_1 - r_0, a]\big|}{|a|} = \frac{\sqrt{(-5)^2 + 7^2 + 3^2}}{\sqrt{1^2 + 2^2 + (-3)^2}} = \sqrt{\frac{83}{14}}.$$

5.5.3. Distance between two lines in space

$$r = r_1 + a_1 t, \qquad r = r_2 + a_2 t$$

$$\rho = \frac{\left|(r_1 - r_2, a_1, a_2)\right|}{\left|[a_1, a_2]\right|}$$

Example 5.21.

$$\begin{cases} x = 1 + \ \ t, \\ y = 2 + 2t, \\ z = 3 - 3t, \end{cases} \qquad \begin{cases} x = \qquad\ \ t, \\ y = 1 + \ \ t, \\ z = 2 - 3t, \end{cases}$$

$[r_1] = (1, 2, 3)$, $[a_1] = (1, 2, -3)$, $[r_2] = (0, 1, 2)$, $[a_2] = (1, 1, -3)$

$$(r_1 - r_2, a_1, a_2) = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & -3 \\ 1 & 1 & -3 \end{vmatrix} = -4 \qquad [a_1, a_2] = \begin{vmatrix} e_1 & e_2 & e_3 \\ 1 & 2 & -3 \\ 1 & 1 & -3 \end{vmatrix} = -3e_1 - e_3$$

$$\rho = \frac{\left|(r_1 - r_2, a_1, a_2)\right|}{\left|[a_1, a_2]\right|} = \frac{4}{\sqrt{(-3)^2 + (-1)^2}} = \frac{4}{\sqrt{10}}$$

Chapter 6

Matrices

Let F be a field.

Let's recall that the rectangular table of the form

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \multicolumn{4}{c}{\dotfill} \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

where $a_{ij} \in F$, is called $m \times n$ *matrix* (or matrix with m rows and n columns).

Scalars a_{ij} are called *elements* or *entries* of the matrix.

Expression $A = (a_{ij})$ means that the element a_{ij} in the i-th row and j-th column ((i, j) element) equals to a_{ij}.

The set of all $m \times n$ matrices with elements in F is denoted as $F^{m \times n}$.

Two matrices A and B are called *equal* if they have the same sizes and corresponding elements are equal, i. e. $A = (a_{ij}) \in F^{m \times n}$, $B = (b_{ij}) \in F^{m \times n}$ and

$$a_{ij} = b_{ij} \qquad (i = 1, 2, \dots, m;\ j = 1, 2, \dots, n).$$

If $m = n$ then the matrix is called *square matrix* of *order* n.

Elements $a_{11}, a_{22}, \ldots, a_{nn}$ of a square matrix form its *diagonal*.

Elements $a_{1n}, a_{2,n-1}, \ldots, a_{n1}$ form its *secondary diagonal*.

In particular, $m \times 1$ matrix is called *column*.

$1 \times n$ matrix is called *row*.

6.1. Scalar multiple of a matrix and addition of matrices

Let $A = (a_{ij}) =\in F^{m\times n}$, $B = (b_{ij}) =\in F^{m\times n}$, $\alpha \in F$.

The *scalar multiple* of A with scalar α is a matrix denoted as αA and obtained by multiplying each element of A by α, i. e. the product αA is a matrix $C = (c_{ij}) \in F^{m\times n}$ such that

$$c_{ij} = \alpha a_{ij} \qquad (i = 1, 2, \dots, m;\ j = 1, 2, \dots, n).$$

The *sum* $A + B$ of the matrices A, B is a matrix denoted as $A + B$ and obtained by adding corresponding elements of A and B, i. e. the sum $A + B$ is a matrix $C = (c_{ij}) \in F^{m\times n}$ such that

$$c_{ij} = a_{ij} + b_{ij} \qquad (i = 1, 2, \dots, m;\ j = 1, 2, \dots, n).$$

It is easy to prove the following properties of the above operations.

Proposition 6.1. *Let A, B, C be matrices in $F^{m\times n}$. Let α, β be scalars in F. Then*

1) $A+(B+C)=(A+B)+C$;

2) $A+B=B+C$;

3) $A+O=A$, *where O is* zero $m\times n$ *matrix, i. e. the matrix in which all elements are zeros;*

4) *for any matrix A there exists a matrix B called* opposite *or* additive inverse *and denoted as $-A$, such that $A+B=O$; the elements of $-A$ equal to $b_{ij}=-a_{ij}$.*

5) $(\alpha\beta)A=\alpha(\beta A)$;

6) $1\cdot A=A$;

7) $(\alpha+\beta)A=\alpha A+\beta A$;

8) $\alpha(A+B)=\alpha A+\alpha B$.

Corollary 6.2. *With respect to multiplication with a scalar and sum of matrices the set* $F^{m\times n}$ *is a vector space over the filed* F. *The dimension of* $F^{m\times n}$ *is* mn.

All results obtained for general vector spaces are valid for $F^{m\times n}$. In particular in $F^{m\times n}$ the operation of matrix subtraction is introduced etc.

6.2. Product of matrices

Let $A = (a_{ij}) \in F^{m\times n}$, $B = (b_{jl}) \in F^{n\times k}$. The *product* of A by B is a matrix $C = (c_{il}) \in F^{m\times k}$, such that

$$c_{il} = \sum_{j=1}^{n} a_{ij}b_{jl} = a_{i1}b_{1l} + a_{i2}b_{2l} + \ldots + a_{in}b_{nl}. \qquad (6.1)$$

Remark that two matrices can be multiplied iff the number of columns of the first matrix equals to the number of the second one. (In this case they say that “sizes of operands agree”).

The notations for product of matrices are $A \cdot B$, $A \times B$ or simply AB.

Sometimes they say that matrices are multiplied “row by column” meaning that formula (6.1) saus that in order to find the element of i-th row and l-th column of the product AB one should multiply each elements in i-th row of A by the corresponding element in l-th column of B and then add all these products.

Example 6.3.

$$(a_1, a_2, \ldots, a_n) \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \sum_{j=1}^{n} a_j b_j = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n;$$

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \cdot (b_1, b_2, \ldots, b_n) = \begin{pmatrix} a_1 b_1 & a_1 b_2 & \ldots & a_1 b_n \\ a_2 b_1 & a_2 b_2 & \ldots & a_2 b_n \\ \ldots & \ldots & \ldots & \ldots \\ a_m b_1 & a_m b_2 & \ldots & a_m b_n \end{pmatrix};$$

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}.$$

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = b_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + b_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \dots + b_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix};$$

$$(a_1,\, a_2,\, \dots,\, a_m) \cdot \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix} =$$

$$= a_1 \cdot (b_{11},\, b_{12},\, \dots,\, b_{1n}) + a_2 \cdot (b_{21},\, b_{22},\, \dots,\, b_{2n}) + \dots + a_m \cdot (b_{m1},\, b_{m2},\, \dots,\, b_{mn}).$$

Example 6.4. Let $A = (a_{ij}) \in F^{m\times n}$, $b = (b_i) \in F^m$, $x = (x_j) \in F^n$.

Now we can write the system of linear equations

$$\begin{cases} a_{11}x_1 + \ a_{12}x_2 + \ldots + \ a_{1n}x_n = b_1; \\ a_{21}x_1 + \ a_{22}x_2 + \ldots + \ a_{2n}x_n = b_2; \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases}$$

in the following matrix notation: $Ax = b$.

The square matrix of order n in which all off-diagonal elements are 0's and diagonal elements are 1's is called an *identity matrices* and denoted as I_n or (if there is no ambiguity) simply I.

Identity matrix:

$$I = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Proposition 6.5.

1) *If* $A \in F^{m\times n}$, $B \in F^{n\times p}$, $C \in F^{p\times q}$, *then* $(AB)C = A(BC)$ *(associativity);*

2) *If* $A \in F^{m\times n}$, $B \in F^{n\times p}$, $C \in F^{n\times p}$, *then* $A(B + C) = AB + AC$ *(distributivity I);*

3) *If* $A \in F^{m\times n}$, $B \in F^{m\times n}$, $C \in F^{n\times p}$, *then* $(A + B)C = AC + BC$ *(distributivity II);*

4) *If* $A \in F^{m\times n}$, $B \in F^{n\times p}$, *then* $\alpha(AB) = (\alpha A)B$.

5) *If* $A \in F^{m\times n}$, *then* $AI_n = A$ *and* $I_m A = A$, *where* I_n, I_m *are* identity matrices *of order* n *and* m *correspondingly.*

Proof. All properties can be verified directly.

For example let's prove the first one.

First, note that all operations in LHS and RHS of

$$(AB)C = A(BC)$$

can be performed (sizes agree).

Second, sizes of the matrix in LHS coincide with sizes of the matrix in RHS and equal to $m \times q$.

Finally, let's prove that corresponding elements if matrices in LHS and RHS equal.

Let

$$A = (a_{ij}), \qquad B = (b_{jk}), \qquad C = (c_{kl}).$$

Denote

$$D = AB = (d_{ik}), \quad G = BC = (g_{jl}), \quad F = (AB)C = (f_{il}), \quad H = A(BC) = (h_{il}).$$

By definition of products of matrices, $d_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$.

Substituting this expression into the formula for f_{il}, we will have

$$f_{il} = \sum_{k=1}^{p} d_{ik} c_{kl} = \sum_{k=1}^{p} \left(\sum_{j=1}^{n} a_{ij} b_{jk} \right) c_{kl} = \sum_{k=1}^{p} \sum_{j=1}^{n} a_{ij} b_{jk} c_{kl}.$$

By definition of products of matrices, $g_{jl} = \sum_{k=1}^{p} b_{jk} c_{kl}$.

Substituting this expression into the formula for h_{il}, we will have

$$h_{il} = \sum_{j=1}^{n} a_{ij} g_{jl} = \sum_{j=1}^{n} a_{ij} \left(\sum_{k=1}^{p} b_{jk} c_{kl} \right) = \sum_{k=1}^{p} \sum_{j=1}^{n} a_{ij} b_{jk} c_{kl}.$$

Now we have $f_{il} = h_{il}$ $(i = 1, 2, \ldots, m,\ l = 1, 2, \ldots, q)$, hence $F = H$. $\square$

Proposition 6.6. *The j-th column of the product AB is the linear combination of columns of A with coefficients taken from the j-th column of B.*

The i-th column of the product AB is the linear combination of rows of A with coefficients taken from the i-th row of A.

Corollary 6.7. *If $A \in F^{m\times n}$ and $B \in F^{n\times p}$, then*

$$\operatorname{rank}(AB) \leq \operatorname{rank} A, \qquad \operatorname{rank}(AB) \leq \operatorname{rank} B.$$

It is obvious that both products AB and BA are defined if and only if A and B are square with the same order.

It is easy to find an example when $AB \neq BA$.

For instance,

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}.$$

Hence in general $AB \neq BA$.

Thus, the set $F^{n\times n}$ of all square matrices is closed under operation of multiplication but the multiplication does not possess the commutativity.

6.3. Elementary operations as matrix multiplication

Let E_{ij} be a matrix obtained from identity matrix by interchanging its i-th and j-th rows:

$$E_{ij} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}.$$

Let $E_i(\alpha)$ be a matrix obtained from the identity matrix by multiplying its i-th row with α:

$$E_i(\alpha) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \alpha & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}.$$

Let $E_{ij}(\alpha)$ matrix obtained from the identity matrix by adding the j-th row multiplied with α to the i-th row:

$$E_{ij}(\alpha) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \dots & \alpha & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}.$$

Matrices E_{ij}, $E_i(\alpha)$, $E_{ij}(\alpha)$ are called the *matrices of elementary operations.*

Proposition 6.8. *Elementary operations* with rows *of a matrix* $A \in F^{m\times n}$ *can be performed by multiplying* A from the left *with matrices of elementary operations. More precisely,*

- *multiplying with* E_{ij} *performs the interchanging the rows* i *and* j,
- *multiplying with* $E_i(\alpha)$ *performs the multiplication the* i*-th row by* α,
- *multiplying with* $E_{ij}(\alpha)$ *performs the addition the* j*-th row multiplied by* α *to the* i*-th row.*

Elementary operations with columns *of a matrix* $A \in F^{m\times n}$ *can be performed by multiplying* A from the right *with matrices of elementary operations. More precisely,*

- *multiplying with* E_{ij} *performs the interchanging the columns* i *and* j,
- *multiplying with* $E_i(\alpha)$ *performs the multiplication the* i*-th column by* α,
- *multiplying with* $E_{ij}(\alpha)$ *performs the addition the* i*-th column multiplied by* α *to the* j*-th column.*

Proof. The proposition follows from proposition 6.6. □

Thus, an elementary operation with rows of A is equivalent to multiplying A form the left with a matrix that can be obtained from the identity matrix with the same elementary operation.

Similarly for column operations but one should multiply from the right.

As illustration let's consider the 3×3 case, the third type of elementary operations with rows:

$$\begin{pmatrix} 1 & \alpha & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} + \alpha a_{21} & a_{12} + \alpha a_{22} & a_{13} + \alpha a_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Exercize 6.9. Consider all other elementary operations with columns and rows of 3×3 matrix.

6.4. Matrix transpose

Let $A = (a_{ij}) \in F^{m\times n}$.

The matrix $B = (b_{ji}) \in F^{n\times m}$ with $b_{ji} = a_{ij}$, is called *transposed* to A and denoted as $A^\top$.

Example 6.10.

$$\begin{pmatrix} 1 & 2 & -3 \\ 0 & -3 & 7 \\ -5 & -1 & 4 \end{pmatrix}^\top = \begin{pmatrix} 1 & 0 & -5 \\ 2 & -3 & -1 \\ -3 & 7 & 4 \end{pmatrix}.$$

Proposition 6.11.

1) If $A \in F^{m\times n}$, $B \in F^{m\times n}$, then $(A+B)^\top = A^\top + B^\top$.

2) If $A \in F^{m\times n}$, $B \in F^{n\times k}$, then $(AB)^\top = B^\top A^\top$.

Proof. The first property is obvious. Let's prove the second property.

First, note that sizes of matrices in LHS and RHS of the equality to prove coincide.

Second, let $A = (a_{ij})$, $B = (b_{jk})$, $C = AB = (c_{ik})$, $D = (AB)^\top = (d_{ki})$, then

$$d_{ki} = c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}.$$

$F = A^\top = (f_{ji})$, $G = B^\top = (g_{kj})$, $H = B^\top A^\top = (h_{ki})$, then

$$f_{ji} = a_{ij}, \quad g_{kj} = b_{jk}, \quad h_{ki} = \sum_{j=1}^{n} g_{kj} f_{ji} = \sum_{j=1}^{n} b_{jk} a_{ij}.$$

Now we have that $d_{ki} = h_{ki}$ $(k = 1, 2, \ldots, k;\ i = 1, 2, \ldots, m)$. $\square$

Chapter 7

Determinants

7.1. Definitions

Recall that the determinant of the 2-d order is

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

The determinant of the 3-d order can be expressed through the determinants of the 2-d order as follows:

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

Similarly we can define the determinant of the 4-th order:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} =$$

$$= a_{11} \cdot \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{12} \cdot \begin{vmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix} + a_{13} \cdot \begin{vmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{vmatrix} - a_{14} \cdot \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{vmatrix}.$$

And so on...

Let's consider this more formally.

We'll give a definition of the determinant *by induction.*

Base of the induction

Define the determinant of the 1×1 matrix $A = (a_{11}) \in F^{1\times 1}$ to be a_{11} (determinant of the 1-st order).

Step of the induction

Assume that we know what is the determinant of $(n-1) \times (n-1)$ matrix (determinant of the $(n-1)$-th order). Let $A \in F^{n\times n}$.

Let $\mathsf{M}_{ij}(A)$ (or simply M_{ij} if it's clear what matrix A meant) denote the determinant of the matrix obtained by deleting the i-th row and the j-th column of A.

$\mathsf{M}_{ij}(A)$ is called *minor* of A at location (i, j).

Then the determinant of A (determinant of the n-th order) is

$$\det A = a_{11}\mathsf{M}_{11} - a_{12}\mathsf{M}_{12} + \ldots + (-1)^{n+1}a_{1n}\mathsf{M}_{1n} = \sum_{j=1}^{n}(-1)^{1+j}a_{1j}\mathsf{M}_{1j}.$$

This formula is also called *the first-row expansion of the determinant.*

The determinant $\det A$ of the matrix A is also denoted as

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \multicolumn{4}{c}{\dotfill} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

7.2. Determinant of a triangular matrix

A matrix $A = (a_{ij}) \in F^{n\times n}$ is called *lower-triangular*, if $a_{ij} = 0$ for all i, j, such that $i < j$, i. e.

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \multicolumn{4}{c}{\dotfill} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

A matrix $A = (a_{ij})$ is called *upper-triangular*, if $a_{ij} = 0$ for all i, j, such that $i > j$, i. e.

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \multicolumn{4}{c}{\dotfill} \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}.$$

Theorem 7.1. *The determinant of a lower-triangular matrix is equal to the product of* *diagonal elements:*

$$\det A = a_{11}a_{22}\dots a_{nn}.$$

The same is for upper-triangular matrices.

Proof. Expanding by the first row:

$$\begin{vmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 \\ a_{31} & a_{32} & a_{33} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix} = a_{11} \cdot \begin{vmatrix} a_{22} & 0 & \dots & 0 \\ a_{32} & a_{33} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix} = a_{11} \cdot a_{22} \cdot \begin{vmatrix} a_{33} & \dots & 0 \\ \dots & \dots & \dots \\ a_{n3} & \dots & a_{nn} \end{vmatrix} =$$

$$= \dots = a_{11}a_{22}\dots a_{nn}.$$

□

7.3. Expanded form of the determinant

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

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

These are *expanded forms of the determinants*.

We can also try to obtain the expanded form of the 4-th order determinant, expanding each 3-th order determinant in the first-row expansion of the 4-th order determinant:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} =$$

$$= a_{11} \cdot \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{12} \cdot \begin{vmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix} + a_{13} \cdot \begin{vmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{vmatrix} - a_{14} \cdot \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{vmatrix} =$$

$$\begin{aligned} &= a_{11} \cdot (a_{22}a_{33}a_{44} + a_{23}a_{34}a_{42} + a_{24}a_{32}a_{43} - a_{24}a_{33}a_{42} - a_{23}a_{32}a_{44} - a_{22}a_{34}a_{43}) \\ &- a_{12} \cdot (a_{21}a_{33}a_{44} + a_{23}a_{34}a_{41} + a_{24}a_{31}a_{43} - a_{24}a_{33}a_{41} - a_{23}a_{31}a_{44} - a_{21}a_{34}a_{43}) \\ &+ a_{13} \cdot (a_{21}a_{32}a_{44} + a_{22}a_{34}a_{41} + a_{24}a_{31}a_{42} - a_{24}a_{32}a_{41} - a_{22}a_{31}a_{44} - a_{21}a_{34}a_{42}) \\ &- a_{14} \cdot (a_{21}a_{32}a_{43} + a_{22}a_{33}a_{41} + a_{23}a_{31}a_{42} - a_{23}a_{32}a_{41} - a_{22}a_{31}a_{43} - a_{21}a_{33}a_{42}) = \end{aligned}$$

$$\begin{aligned} &= a_{11}a_{22}a_{33}a_{44} + a_{11}a_{23}a_{34}a_{42} + a_{11}a_{24}a_{32}a_{43} - a_{11}a_{24}a_{33}a_{42} - a_{11}a_{23}a_{32}a_{44} - a_{11}a_{22}a_{34}a_{43} \\ &-a_{12}a_{21}a_{33}a_{44} - a_{12}a_{23}a_{34}a_{41} - a_{12}a_{24}a_{31}a_{43} + a_{12}a_{24}a_{33}a_{41} + a_{12}a_{23}a_{31}a_{44} + a_{12}a_{21}a_{34}a_{43} \\ &+a_{13}a_{21}a_{32}a_{44} + a_{13}a_{22}a_{34}a_{41} + a_{13}a_{24}a_{31}a_{42} - a_{13}a_{24}a_{32}a_{41} - a_{13}a_{22}a_{31}a_{44} - a_{13}a_{21}a_{34}a_{42} \\ &-a_{14}a_{21}a_{32}a_{43} - a_{14}a_{22}a_{33}a_{41} - a_{14}a_{23}a_{31}a_{42} + a_{14}a_{23}a_{32}a_{41} + a_{14}a_{22}a_{31}a_{43} + a_{14}a_{21}a_{33}a_{42}. \end{aligned}$$

In the general case we have the following theorem.

Theorem 7.2 (Expanded form of the determinant)**.** *The n-th order determinant is equal to the sum of $n!$ products $\pm a_{1j_1}a_{2j_2}\dots a_{nj_n}$, where $(j_1,\, j_2,\, \dots,\, j_n)$ is a permutation of $(1, 2, \dots, n)$ and the sign is $+$ if the number $\sigma(j_1,\, j_2,\, \dots,\, j_n)$ of inversions in $(j_1,\, j_2,\, \dots,\, j_n)$ is* even *and the sign is $-$ if the number of inversions is* odd *(the* inversion *occurs iff $j_i > j_{i'}$ but $i < i'$).*

This can be written as

$$\det A = \sum_{(j_1,\, j_2,\, \dots,j_n)} (-1)^{\sigma(j_1,\, j_2,\, \dots,j_n)} a_{1j_1}a_{2j_2}\dots a_{nj_n},$$

where the sum is over all permutations $(j_1,\, j_2,\, \dots,\, j_n)$ of $(1, 2, \dots, n)$ and $\sigma(j_1,\, j_2,\, \dots,\, j_n)$ denotes the number of inversions in the permutation.

For example, in the permutation $(2, 3, 1, 4)$ we have 2 (even) inversions: $(2,\ 1)$ and $(3,\ 1)$.

In the permutation $(2, 4, 1, 3)$ we have 3 (odd) inversions: $(2,\ 1)$, $(4,\ 1)$, $(4,\ 3)$.

Proof. Omitted □

Exercize 7.3. Verify the theorem for $n = 2, 3, 4$.

7.4. Row and column expansions of the determinant

Theorem 7.4 (Row and column expansions of the determinant)**.** *Let* $A \in F^{n \times n}$*.*

For any row i

$$\det A = (-1)^{i+1} a_{i1} \mathsf{M}_{i1} + (-1)^{i+2} a_{i2} \mathsf{M}_{i2} + \ldots + (-1)^{i+n} a_{in} \mathsf{M}_{in} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \mathsf{M}_{ij}$$

(the i*-th row expansion).*

For any column j

$$\det A = (-1)^{1+j} a_{1j} \mathsf{M}_{1j} + (-1)^{2+j} a_{2j} \mathsf{M}_{2j} + \ldots + (-1)^{n+j} a_{nj} \mathsf{M}_{nj} = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \mathsf{M}_{ij}$$

(the j*-th column expansion).*

Proof. Omitted □

Example 7.5. 2-nd row expansion of the 3-d order determinant:

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

3-rd column expansion of the 3-rd order determinant:

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

$(-1)^{i+j}\mathsf{M}_{ij}(A)$ is called *cofactor* at location (i, j) and denoted as $\mathsf{C}_{ij}(A)$ (or C_{ij} if there is no ambiguity).

Now we can write formulas in Theorem 7.4 as follows

$$\det A = \sum_{j=1}^{n} a_{ij}\mathsf{C}_{ij}, \qquad \det A = \sum_{i=1}^{n} a_{ij}\mathsf{C}_{ij}.$$

Theorem 7.6. *A matrix and its transpose have equal determinants, i.e.*

$$\det A = \det A^\top.$$

Proof. Induction by n.

Base of the induction: $n = 1$. For a matrix of size 1, the transpose and the matrix itself are equal, so their determinants are equal too.

Step of the induction. Suppose that theorem is true for matrices of size $n - 1$. Hence

$$\det A = a_{11}\mathsf{C}_{11}(A)+a_{12}\mathsf{C}_{12}(A)+\ldots+a_{1n}\mathsf{C}_{1n}(A) = a_{11}\mathsf{C}_{11}(A^\top)+a_{12}\mathsf{C}_{21}(A^\top)+\ldots+a_{1n}\mathsf{C}_{n1}(A^\top).$$

The last expression is the first-column expansion of $\det(A^\top)$. □

Theorem 7.7. *If two rows or columns are interchanged, the determinant changes sign.*

Corollary 7.8. *The determinant of a matrix with two equal rows or columns is zero.*

Theorem 7.9. *Let $A = (a_{ij}) \in F^{n \times n}$. If $i \neq k$*

$$\sum_{j=1}^{n} a_{kj} \mathsf{C}_{ij} = 0. \tag{7.1}$$

If $j \neq k$

$$\sum_{i=1}^{n} a_{ik} \mathsf{C}_{ij} = 0.$$

Proof. Let B be the matrix obtained from A by replacing i-th row by k-th row.

$\det B = 0$ since B has two identical rows.

Expanding $\det B$ along i-th row we get (7.1).

Similarly for columns. □

Theorem 7.10. *Determinant is a linear function of each row, i. e.*

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \multicolumn{4}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\ a'_{i1} + a''_{i1} & a'_{i2} + a''_{i2} & \dots & a'_{in} + a''_{in} \\ \multicolumn{4}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \multicolumn{4}{c}{\dots\dots\dots\dots\dots} \\ a'_{i1} & a'_{i2} & \dots & a'_{in} \\ \multicolumn{4}{c}{\dots\dots\dots\dots\dots} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \multicolumn{4}{c}{\dots\dots\dots\dots\dots} \\ a''_{i1} & a''_{i2} & \dots & a''_{in} \\ \multicolumn{4}{c}{\dots\dots\dots\dots\dots} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix},$$

$$
\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \multicolumn{4}{c}{\dotfill} \\ \alpha a'_{i1} & \alpha a'_{i2} & \dots & \alpha a'_{in} \\ \multicolumn{4}{c}{\dotfill} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \alpha \cdot \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \multicolumn{4}{c}{\dotfill} \\ a'_{i1} & a'_{i2} & \dots & a'_{in} \\ \multicolumn{4}{c}{\dotfill} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.
$$

Similarly for columns.

Proof.

Expand the first determinant about the i-th row:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \multicolumn{4}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\ a'_{i1}+a''_{i1} & a'_{i2}+a''_{i2} & \dots & a'_{in}+a''_{in} \\ \multicolumn{4}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = (a'_{i1}+a''_{i1})\mathsf{C}_{i1} + (a'_{i2}+a''_{i2})\mathsf{C}_{i2} + \ldots + (a'_{in}+a''_{in})\mathsf{C}_{in} =$$

$$= (a'_{i1}\mathsf{C}_{i1} + \ldots + a'_{in}\mathsf{C}_{in}) + (a''_{i1}\mathsf{C}_{i1} + \ldots + a''_{in}\mathsf{C}_{in}) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \multicolumn{4}{c}{\dots\dots\dots\dots\dots\dots} \\ a'_{i1} & a'_{i2} & \dots & a'_{in} \\ \multicolumn{4}{c}{\dots\dots\dots\dots\dots\dots} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \multicolumn{4}{c}{\dots\dots\dots\dots\dots\dots} \\ a''_{i1} & a''_{i2} & \dots & a''_{in} \\ \multicolumn{4}{c}{\dots\dots\dots\dots\dots\dots} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

The proof of the second equality is similar. □

Corollary 7.11. *Adding a row multiplied by a scalar of a square matrix to another row of the same matrix does not change the determinant.*

Similarly for columns

Proof. Let's add the j-th row multiplied with α to the i-th row:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \multicolumn{4}{c}{\dotfill} \\ a_{j1}+\alpha a_{i1} & a_{j2}+\alpha a_{i2} & \dots & a_{jn}+\alpha a_{in} \\ \multicolumn{4}{c}{\dotfill} \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \multicolumn{4}{c}{\dotfill} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \multicolumn{4}{c}{\dotfill} \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \multicolumn{4}{c}{\dotfill} \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \multicolumn{4}{c}{\dotfill} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \alpha \cdot \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \multicolumn{4}{c}{\dotfill} \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \multicolumn{4}{c}{\dotfill} \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \multicolumn{4}{c}{\dotfill} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

The last determinant is zero because it has two equal rows. □

The convenient way to evaluate the determinant of a matrix is to reduce it to row-echelon form, recording any sign changes caused by row interchanges and recording all factors cased by multiplying rows with scalars.

Example 7.12. Evaluate the determinant

$$\Delta = \begin{vmatrix} 2 & -1 & 3 & 2 \\ 1 & 2 & 3 & 4 \\ 3 & 1 & -2 & 3 \\ 1 & 2 & -3 & -1 \end{vmatrix}.$$

Interchange the first and the second rows. The determinant changes sign. Subtract the first row from all other rows with suitable multipliers (2, 3, 1 correspondingly) in order to annul entries below the diagonal:

$$\Delta = -\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -5 & -3 & -6 \\ 0 & -5 & -11 & -9 \\ 0 & 0 & -6 & -5 \end{vmatrix}.$$

Subtract the second row from the third one:

$$\Delta = -\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -5 & -3 & -6 \\ 0 & 0 & -8 & -3 \\ 0 & 0 & -6 & -5 \end{vmatrix}.$$

Subtract the third row multiplied with $\dfrac{6}{8} = \dfrac{3}{4}$ from the forth one:

$$\Delta = -\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -5 & -3 & -6 \\ 0 & 0 & -8 & -3 \\ 0 & 0 & 0 & -\dfrac{11}{4} \end{vmatrix} = -1 \cdot (-5) \cdot (-8) \cdot \left(-\frac{11}{4}\right) = 110$$

7.5. Vandermonde determinant

Let $x_1, x_2, \ldots, x_n$ be some scalars in F. *Vandermonde determinant* is

$$W(x_1, x_2, \ldots, x_n) = \begin{vmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \multicolumn{5}{c}{\dotfill} \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{vmatrix}. \tag{7.2}$$

Theorem 7.13.

$$W(x_1, x_2, \ldots, x_n) = \prod_{1 \le i < j \le n} (x_j - x_i). \tag{7.3}$$

Proof. Induction by n.

If $n = 2$ then

$$W(x_1, x_2) = \begin{vmatrix} 1 & x_1 \\ 1 & x_2 \end{vmatrix} = x_2 - x_1.$$

Suppose that (7.3) holds for Vandermonde determinants of the $(n-1)$-th order. Let's prove it for determinants of the n-th order.

Subtract the $(n-1)$-th column multiplied with x_1 from the n-th column; then subtract the $(n-2)$-th column multiplied with x_1 from the $(n-2)$-th column and so on. We obtain

$$W(x_1, x_2, \dots, x_n) = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & x_2 - x_1 & x_2^2 - x_2 x_1 & \dots & x_2^{n-1} - x_2^{n-2} x_1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n - x_1 & x_n^2 - x_n x_1 & \dots & x_n^{n-1} - x_n^{n-2} x_1 \end{vmatrix}.$$

Expanding the determinant about the first row we obtain

$$W(x_1, x_2, \dots, x_n) = \begin{vmatrix} x_2 - x_1 & x_2^2 - x_2 x_1 & \dots & x_2^{n-1} - x_2^{n-2} x_1 \\ \dots & \dots & \dots & \dots \\ x_n - x_1 & x_n^2 - x_n x_1 & \dots & x_n^{n-1} - x_n^{n-2} x_1 \end{vmatrix}.$$

Taking out factors $(x_2 - x_1), (x_3 - x_1), \dots, (x_n - x_1)$ of all rows we obtain:

$$W(x_1, x_2, \dots, x_n) = (x_2 - x_1)(x_3 - x_1)\dots(x_n - x_1) \begin{vmatrix} 1 & x_2 & x_2^2 & \dots & x_2^{n-2} \\ 1 & x_3 & x_3^2 & \dots & x_3^{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-2} \end{vmatrix}$$

$$= (x_2 - x_1)(x_3 - x_1)\dots(x_n - x_1) \cdot W(x_2, x_3, \dots, x_n).$$

It remains to apply inductive hypothesis to $W(x_2, x_3, \dots, x_n)$:

$$W(x_1, x_2, \dots, x_n) = (x_2 - x_1)(x_3 - x_1)\dots(x_n - x_1) \prod_{2 \le i < j \le n} (x_j - x_i) = \prod_{1 \le i < j \le n} (x_j - x_i).$$

□

Corollary 7.14. *Vandermonde determinant $W(x_1, x_2, \dots, x_n)$ is zero if and only if there are at least two equal numbers among $x_1, x_2, \dots, x_n$.*

Example 7.15.

$$\begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix} = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2).$$

7.6. Non-singular matrices

Let $A \in F^{n\times n}$.

If $\operatorname{rank} A < n$ then A is called *singular*.

If $\operatorname{rank} A = n$ then A is called *non-singular*.

Lemma 7.16. *Let $A \in F^{n\times n}$. and A' is obtained from A by means of row elementary operations. Then* $\det A = 0$ *if and only if* $\det A' = 0$.

Proof. Interchanging the rows changes only sign of the determinant.

Multiplication of the row by a scalar $\alpha \neq 0$ multiplies the determinant by α. But if the determinant was 0 it remains to be 0. If it was not 0 it remains to be not 0.

Adding a row to another row does not change the determinant. □

Theorem 7.17. *Let $A \in F^{n\times n}$. Then* $\det A = 0$ *if and only if A is singular.*

Proof. Let A' be r.r.e.f. of A.

A' can be obtained from A by means of row elementary operations, hence by Lemma 7.16 $\det A = 0$ if and only if $\det A' = 0$.

It's clear that $\det A' = 0$ if and only if $\operatorname{rank} A < n$ because A' is r.r.e.f. □

7.7. Matrix inverse

Let $A \in F^{n\times n}$.

The transpose of the matrix composed from cofactors of A is called the *adjoint* of A. The adjoint of A is denoted by $\operatorname{adj} A$:

$$\operatorname{adj} A = \begin{pmatrix} \mathsf{C}_{11} & \mathsf{C}_{21} & \dots & \mathsf{C}_{n1} \\ \mathsf{C}_{12} & \mathsf{C}_{22} & \dots & \mathsf{C}_{n2} \\ \dots & \dots & \dots & \dots \\ \mathsf{C}_{1n} & \mathsf{C}_{2n} & \dots & \mathsf{C}_{nn} \end{pmatrix}.$$

Theorem 7.18. *If $A \in F^{n \times n}$ then*

$$A \cdot \operatorname{adj} A = \operatorname{adj} A \cdot A = \det A \cdot I.$$

Proof. Let's prove that $A \cdot \operatorname{adj} A = \det A \cdot I$.

Consider the (i, j) element of $A \cdot \operatorname{adj} A = (b_{ij})$.

$$b_{ij} = a_{i1}\mathsf{C}_{j1} + a_{i2}\mathsf{C}_{j2} + \ldots + a_{in}\mathsf{C}_{jn} = \begin{cases} \det A, & \text{if } i = j, \\ 0, & \text{if } i \neq j \end{cases}$$

by Theorems 7.4, 7.9.

Thus, $A \cdot \operatorname{adj} A = \det A \cdot I$.

Equality $\operatorname{adj} A \cdot A = \det A \cdot I$ is proved similarly. □

Let $A \in F^{n \times n}$. The matrix $B \in F^{n \times n}$ such that $AB = BA = I$ is called *inverse* of A.

The inverse of A is denoted by A^{-1}.

Theorem 7.19. *If* $\det A \neq 0$ *(A is non-singular) then the inverse* A^{-1} *of A is unique and*

$$A^{-1} = \frac{1}{\det A} \cdot \operatorname{adj} A. \tag{7.4}$$

If $\det A = 0$ *(A is singular) then the inverse does not exist.*

Proof. Let $\det A \neq 0$. Then by Theorem 7.18

$$A \cdot \frac{1}{\det A} \cdot \operatorname{adj} A = \frac{1}{\det A} \cdot \operatorname{adj} A \cdot A = I,$$

hence $(1/\det A) \cdot \operatorname{adj} A$ is the inverse.

Let's prove that if $\det A \neq 0$ then the inverse is unique. Assume the contrary: let both B and B' are inverses of A.

Then $B \cdot A \cdot B' = B$ since $A \cdot B' = I$.

On the other hand, $B \cdot A \cdot B' = B'$ since $B \cdot A = I$.

Now we have $B = B'$.

It remains to prove that if $\det A = 0$ then the inverse does not exist.

Really, since $\det A = 0$ then $\operatorname{rank} A < n$.

If $A \cdot B = I$ then by Corollary 6.7

$$n > \operatorname{rank} A \geq \operatorname{rank} I = n.$$

Contradiction. $\square$

Example 7.20. Find the inverse of 2×2 matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Example 7.21. Evaluate inverse of

$$A = \begin{pmatrix} 2 & 4 & 1 \\ -1 & 2 & 3 \\ 3 & -5 & -3 \end{pmatrix}.$$

$$\det A = 41, \quad \mathsf{C}_{11} = \begin{vmatrix} 2 & 3 \\ -5 & -3 \end{vmatrix} = 9, \quad \mathsf{C}_{12} = -\begin{vmatrix} -1 & 3 \\ 3 & -3 \end{vmatrix} = 6, \quad \mathsf{C}_{13} = \begin{vmatrix} -1 & 2 \\ 3 & -5 \end{vmatrix} = -1,$$

$$\mathsf{C}_{21} = -\begin{vmatrix} 4 & 1 \\ -5 & -3 \end{vmatrix} = 7, \quad \mathsf{C}_{22} = \begin{vmatrix} 2 & 1 \\ 3 & -3 \end{vmatrix} = -9, \quad \mathsf{C}_{23} = -\begin{vmatrix} 2 & 4 \\ 3 & -5 \end{vmatrix} = 22,$$

$$\mathsf{C}_{31} = \begin{vmatrix} 4 & 1 \\ 2 & 3 \end{vmatrix} = 10, \quad \mathsf{C}_{32} = -\begin{vmatrix} 2 & 1 \\ -1 & 3 \end{vmatrix} = -7, \quad \mathsf{C}_{33} = \begin{vmatrix} 2 & 4 \\ -1 & 2 \end{vmatrix} = 8,$$

Hence

$$A^{-1} = \frac{1}{41} \cdot \begin{pmatrix} 9 & 7 & 10 \\ 6 & -9 & -7 \\ -1 & 22 & 8 \end{pmatrix}.$$

Proposition 7.22. *Let* $A \in F^{n\times n}$, $B \in F^{n\times n}$.

If $AB = I$ *then* $BA = I$*, i. e.* $B = A^{-1}$ *and* $A = B^{-1}$*.*

Proof. Since $AB = E$ and $\operatorname{rank} E = n$, then by corollary 6.7 $\operatorname{rank} A = \operatorname{rank} B = n$, hence $\det A \neq 0$ and $\det B \neq 0$. Thus, there exist matrices A^{-1} and B^{-1}.

Multiplying both parts of $AB = E$ from the left by A^{-1} we get $A^{-1}AB = A^{-1}$, hence $B = A^{-1}$.

Multiplying both parts of $AB = E$ from the right by B^{-1} we get $A = B^{-1}$. □

Proposition 7.23. *Let* $A \in F^{n\times n}$, $B \in F^{n\times n}$, $\det A \neq 0$, $\det B \neq 0$, *then*

$$(A^{-1})^{-1} = A, \quad (AB)^{-1} = B^{-1}A^{-1}, \quad (A^\top)^{-1} = (A^{-1})^\top.$$

Proof. The first property is proved in Proposition 7.22.

Let's prove that $(A^\top)^{-1} = (A^{-1})^\top$.

This is equivalent to proposition that $(A^{-1})^\top$ is the inverse of $A^\top$. Verify this:

$$(A^{-1})^\top A^\top = (AA^{-1})^\top = I.$$

The equality $(AB)^{-1} = B^{-1}A^{-1}$ can be proved similarly. □

Note 7.24. The formula (7.4) is convenient for finding the inverse only for small n or in special cases.

As a rule in order to find the matrix inverse one use the method of elementary operations.

Let $\det A \neq 0$ then A can be reduced to identity matrix I using elementary row operations. It turns out that if one perform the same sequence of elementary operations with the identity matrix, then one gets A^{-1}.

Really, by Proposition 6.8 each elementary operation with rows of A is equivalent to multiplying A from the left by a matrix of special form. Hence the sequence of elementary operations is equivalent to multiplying by such matrices $S_1, S_2, \ldots, S_t$.

If this sequence of elementary operations reduce the matrix to the identity matrix then

$$S_t \cdot (S_{t-1} \cdot \ldots \cdot (S_2 \cdot (S_1 \cdot A)) \ldots) = I. \tag{7.5}$$

Let $B = S_t S_{t-1} \ldots S_2 S_1$, then from (7.5) we get $B = A^{-1}$.

Hence performing the same operations with the identity matrix we obtain

$$S_t \cdot (S_{t-1} \cdot \ldots \cdot (S_2 \cdot (S_1 \cdot I)) \ldots) = B = A^{-1}.$$

As a rule for implementation of this algorithm one write matrix (A, I). By means of elementary operations with rows of this matrix one can obtain the identity matrix on the place of A, then on the place of I there will be A^{-1}.

Example 7.25. Find the inverse of the matrix in Example 7.21.

$$(A, E) = \left(\begin{array}{rrr|rrr} 2 & 4 & 1 & 1 & 0 & 0 \\ -1 & 2 & 3 & 0 & 1 & 0 \\ 3 & -5 & -3 & 0 & 0 & 1 \end{array}\right) \to \begin{array}{l} R_1 \leftrightarrow R_2 \\ R_1 \leftarrow -R_1 \\ R_2 \leftarrow R_2 - 2R_1 \\ R_3 \leftarrow R_2 - 3R_1 \end{array}$$

$$\to \left(\begin{array}{rrr|rrr} 1 & -2 & -3 & 0 & -1 & 0 \\ 0 & 8 & 7 & 1 & 2 & 0 \\ 0 & 1 & 6 & 0 & 3 & 1 \end{array}\right) \to \begin{array}{l} R_2 \leftrightarrow R_3 \\ R_1 \leftarrow R_1 + 2R_2 \\ R_3 \leftarrow R_3 - 8R_2 \end{array}$$

$$\to \left(\begin{array}{rrr|rrr} 1 & 0 & 9 & 0 & 5 & 2 \\ 0 & 1 & 6 & 0 & 3 & 1 \\ 0 & 0 & -41 & 1 & -22 & -8 \end{array}\right) \to \begin{array}{l} R_3 \leftarrow -R_3/41 \\ R_1 \leftarrow R_1 - 9R_3 \\ R_2 \leftarrow R_2 - 6R_3 \end{array} \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{9}{41} & \frac{7}{41} & \frac{10}{41} \\ 0 & 1 & 0 & \frac{6}{41} & -\frac{9}{41} & -\frac{7}{41} \\ 0 & 0 & 1 & -\frac{1}{41} & \frac{22}{41} & \frac{8}{41} \end{array}\right).$$

To the right from the vertical line we have A^{-1}.

Theorem 7.26. *Let A, B be matrices in $F^{n\times n}$. Then*

$$\det(AB) = \det A \cdot \det B.$$

Proof. Omitted □

Corollary 7.27. *If $\det A \neq 0$, then*

$$\det(A^{-1}) = \frac{1}{\det A}.$$

7.8. Cramer's rule

Let's consider the system of linear equations $Ax = b$ with square non-singular matrix $A \in F^{n\times n}$.

Theorem 7.28. *If $A \in F^{n\times n}$ and $\det A \neq 0$ (A is non-singular), then the unique solution $x \in F^n$ of the system $Ax = b$ can be found as*

$$x = A^{-1}b.$$

Proof. Since $\det A \neq 0$ then there exists A^{-1} and the unique solution $x \in F^n$ of $Ax = b$ can be found by multiplying both part of this equality by A^{-1} from the left. We obtain

$$A^{-1}Ax = A^{-1}b.$$

Taking into account that $A^{-1}A = I$ we get $x = A^{-1}b$. □

Theorem 7.29 (Cramer's rule)**.** *If $A \in F^{n\times n}$ and $\det A \neq 0$ (A is non-singular), then the components of the unique solution $x \in F^n$ of the system $Ax = b$ can be found as*

$$x_j = \frac{\Delta_j}{\Delta},$$

where $\Delta = \det A$ and Δ_j is the determinant of the matrix obtained from A replacing the j-th column by b.

Proof. We have $x = A^{-1}b$.

The j-th component of x equals to the product of the j-th column of A and b. Using formula (7.4) we get

$$x_j = \frac{1}{\det A} \cdot (\mathsf{C}_{1j}, \mathsf{C}_{2j}, \ldots, \mathsf{C}_{nj}) \cdot b = \frac{1}{\Delta} \cdot (\mathsf{C}_{1j} b_1 + \mathsf{C}_{2j} + b_2 + \ldots + \mathsf{C}_{nj} b_n) = \frac{\Delta_j}{\Delta}.$$

□

The Cramer's rule has theoretical significance. In practice it is used for small n (for example $n = 2, 3$) or for systems of special form.

Example 7.30. Let's apply the Cramer's rule to the 2×2 system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1, \\ a_{21}x_1 + a_{22}x_2 = b_2. \end{cases}$$

If $\Delta = a_{11}a_{22} - a12a_{21} \neq 0$ then the (unique) solution of the system has components

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} = \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{12}a_{21}}, \qquad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} = \frac{a_{11}b_2 - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}}.$$

7.9. Change of basis

Suppose that

$$\mathbf{e} = \langle e_1, e_2, \ldots, e_n \rangle, \qquad \mathbf{e}' = \langle e'_1, e'_2, \ldots, e'_n \rangle$$

are two bases of a vector space V. We can call $\mathbf{e}$ an *old* basis, and $\mathbf{e}'$ a *new* basis.

There are unique coefficients α_{ij} such that

$$e'_j = \alpha_{1j} e_1 + \alpha_{2j} e_n + \ldots + \alpha_{nj} e_n, \qquad \text{i.e.} \qquad [e'_j]_{\mathbf{e}} = \begin{pmatrix} \alpha_{1j} \\ \alpha_{2j} \\ \vdots \\ \alpha_{nj} \end{pmatrix} \qquad (j = 1, 2, \ldots, n). \quad (7.6)$$

The matrix $[\mathbf{e}']_{\mathbf{e}} = (\alpha_{ij})$ is called the *matrix of basis change.*

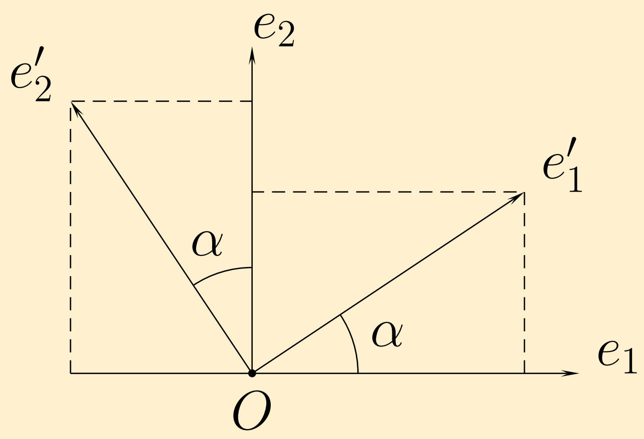
In other words, the matrix of basis change is the matrix composed from coordinate columns of vectors of the new basis:

$$[\mathbf{e}']_{\mathbf{e}} = ([e'_1]_{\mathbf{e}}, [e'_2]_{\mathbf{e}}, \ldots, [e'_n]_{\mathbf{e}}).$$

Example 7.31. Let in $\mathbf{V}_2$ there be two right orthonormal bases $\mathbf{e} = \langle e_1, e_2 \rangle$ and $\mathbf{e}' = \langle e'_1, e'_2 \rangle$.

Let vectors e'_1 and e'_2 obtained from e_1, e_2 correspondingly by rotating on the angle α.

Let's write the matrix of basis change.



$$e'_1 = \cos\alpha \cdot e_1 + \sin\alpha \cdot e_2, \qquad e'_2 = -\sin\alpha \cdot e_1 + \cos\alpha \cdot e_2.$$

$$[\mathbf{e}']_{\mathbf{e}} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

Find the connection between coordinates of a vector x in old and new bases.

Let

$$x = x_1 e_1 + \ldots + x_n e_n = x'_1 e'_1 + \ldots + x'_n e'_n, \qquad \text{i. e.} \qquad [x]_{\mathbf{e}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad [x]_{\mathbf{e}'} = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}.$$

We have

$$x = \sum_{j=1}^{n} x'_j e'_j = \sum_{j=1}^{n} x'_j \sum_{i=1}^{n} \alpha_{ij} e_i = \sum_{i=1}^{n} \underbrace{\left(\sum_{j=1}^{n} \alpha_{ij} x'_j \right)}_{x_i} e_i.$$

Since the coordinates of the vector are unique (if a basis is fixed) then

$$x_i = \sum_{j=1}^{n} \alpha_{ij} x'_j \qquad (i = 1, 2, \ldots, n). \tag{7.7}$$

Formulas (7.7) can be written in matrix notation as $[x]_{\mathbf{e}} = [\mathbf{e}']_{\mathbf{e}} [x]_{\mathbf{e}'}$.

Example 7.32. [Continuation]

$$\begin{cases} x_1 = x_1' \cos\alpha - x_2' \sin\alpha, \\ x_2 = x_1' \sin\alpha + x_2' \cos\alpha. \end{cases}$$

Since columns of $[\mathbf{e}']_{\mathbf{e}}$ are linear independent then rank$[\mathbf{e}']_{\mathbf{e}} = n$ and $[\mathbf{e}']_{\mathbf{e}}$ is non-singular.

Multiplying both parts of equation $[x]_{\mathbf{e}} = [\mathbf{e}']_{\mathbf{e}}[x]_{\mathbf{e}'}$ from the left by $[\mathbf{e}']_{\mathbf{e}}^{-1}$ we get

$$[x]_{\mathbf{e}'} = [\mathbf{e}']_{\mathbf{e}}^{-1}[x]_{\mathbf{e}}.$$

Hence

$$[\mathbf{e}]_{\mathbf{e}'} = [\mathbf{e}']_{\mathbf{e}}^{-1}.$$

7.10. Change of coordinate system

In this section sometimes vectors of vector space will called points. This change in names appeals to geometric interpretation: a radius-vector in $\mathbf{V}_3$ can be associated with the its endpoint.

Coordinate system of a linear space V is an aggregate of a point $\mathcal{O}$ (called *origin*) in V and basis $\mathbf{e} = \langle e_1, e_2, \ldots, e_n \rangle$ of V.

The *coordinates of a point* $\mathcal{A} \in V$ in this coordinate system are the coordinates of the vector $\mathcal{A} - \mathcal{O}$ in basis $e_1, e_2, \ldots, e_n$.

The column of coordinates of the point $\mathcal{A}$ is denoted by $[\mathcal{A}]_{\mathcal{O},\mathbf{e}}$, i. e. by definition

$$[\mathcal{A}]_{\mathcal{O},\mathbf{e}} = [\mathcal{A} - \mathcal{O}]_e.$$

Let's find the connection between coordinates of a point in different coordinate systems.

Let there be an old coordinates system $\mathcal{O}, e_1, \ldots, e_n$ and new coordinate system $\mathcal{O}', e'_1, \ldots, e'_n$, with formulas (7.6) holding and

$$\mathcal{O}' - \mathcal{O} = \sum_{j=1}^{n} \gamma_j e_j, \quad \text{i. e.} \quad [\mathcal{O}']_{\mathcal{O},\mathbf{e}} = [\mathcal{O}' - \mathcal{O}]_{\mathbf{e}} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{pmatrix}.$$

Let $\mathcal{A}$ be a point in V, with

$$\mathcal{A} - \mathcal{O}' = x_1 e_1 + \ldots + x_n e_n = x'_1 e'_1 + \ldots + x'_n e'_n,$$

i. e.

$$[\mathcal{A}]_{\mathcal{O},\mathbf{e}} = [\mathcal{A} - \mathcal{O}]_{\mathbf{e}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad [\mathcal{A}]_{\mathcal{O}',\mathbf{e}'} = [\mathcal{A} - \mathcal{O}']_{\mathbf{e}'} = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}.$$

Since $\mathcal{A} - \mathcal{O} = (\mathcal{O}' - \mathcal{O}) + (\mathcal{A} - \mathcal{O}')$, then

$$[\mathcal{A}]_{\mathcal{O},\mathbf{e}} = [\mathcal{A} - \mathcal{O}]_{\mathbf{e}} = [\mathcal{O}' - \mathcal{O}]_{\mathbf{e}} + [\mathcal{A} - \mathcal{O}']_{\mathbf{e}} = [\mathcal{O}']_{\mathcal{O},\mathbf{e}} + [\mathcal{A} - \mathcal{O}']_{\mathbf{e}} = [\mathcal{O}']_{\mathcal{O},\mathbf{e}} + [e']_{\mathbf{e}}[\mathcal{A} - \mathcal{O}']_{\mathbf{e}'}.$$

Thus,

$$[\mathcal{A}]_{\mathcal{O},\mathbf{e}} = [\mathcal{O}']_{\mathcal{O},\mathbf{e}} + [e']_{e}[\mathcal{A}]_{\mathcal{O}',\mathbf{e}'},$$

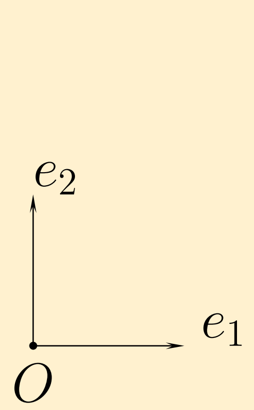
or

$$x_i = \gamma_i + \sum_{j=1}^{n} \alpha_{ij} x'_j \qquad (i = 1, 2, \ldots, n).$$

Example 7.33. [Continuation] Let O, e_1, e_2 and O', e'_1, e'_2 be two coordinate systems in $\mathbf{V}_2$.

The coordinates of O' in the old coordinate system are $(4, 1)$.

e'_1, e'_2 obtained from e_1, e_2 correspondingly by rotating on the angle $\pi/6$.



Let's find the connection between coordinates of a point in these coordinate systems.

$$\begin{cases} x_1 = \gamma_1 + x'_1 \cos\alpha - x'_2 \sin\alpha = 4 + \dfrac{\sqrt{3}}{2} x'_1 - \dfrac{1}{2} x'_2, \\ x_2 = \gamma_2 + x'_1 \sin\alpha + x'_2 \cos\alpha = 1 + \dfrac{1}{2} x'_1 + \dfrac{\sqrt{3}}{2} x'_2. \end{cases}$$

Chapter 8

Linear transformations

8.1. Definitions and examples

Let V and W be two linear spaces over the same field F. A function $\varphi : V \to W$ is called a *linear transformation* from V to W if for all vectors x, y in V and scalars α in F,

1) $\varphi(x + y) = \varphi x + \varphi y$;

2) $\varphi(\alpha x) = \alpha \cdot \varphi x$.

The set of all linear transformations mapping V to W is denoted by $\Phi(V, W)$.

By induction we get

$$\varphi\left(\sum_{i=1}^{s} \alpha_i a_i\right) = \sum_{i=1}^{s} \alpha_j \varphi a_i, \tag{8.1}$$

where $a_i \in V$, $\alpha_i \in F$ $(i = 1, 2, \ldots, s)$.

In property 2) putting $\alpha = 0$, we obtain $\varphi o = o$.

Note that in LHS and RHS of this equation o denotes zero vectors in different (if $V \neq W$) spaces.

Consider a few examples of linear transformations.

Define a map $\theta : V \to W$ by equality $\theta x = o$ for all $x \in V$. It is obvious that θ is a linear transformation. It is called a *zero* transformation. Note, that a map ψ, such that $\psi x = a$, where a is a fixed non-zero vector in W, is not a linear transformation.

Define a map $\varepsilon : V \to V$ by equality $\varepsilon x = x$ for all $x \in V$. The map ε is obviously linear transformation. It is called the *identity* transformation.

Let the transformation $\varphi : \mathbf{V}_2 \to \mathbf{V}_2$ maps each vector x to a vector φx, obtained from x by rotating on the angle α about origin. This transformation is called *rotation transformation* (*on the angle* α). It is not hard to see that it is linear transformation. Also, rotation the space $\mathbf{V}_3$ about a line passing throw origin is the linear transformation.

If $\varphi : V \to V$ maps each vector x to the vector $\varphi x = a + x$, where a is fixed, then φ is called *translation*. If $a \neq 0$ the translation is not a linear transformation.

Let $A = (a_{ij}) \in F^{m\times n}$. Consider the transformation $\varphi : F^n \to F^m$, mapping a vector $x = (x_1, x_2, \dots, x_n)^\top \in F^n$ to $\varphi x = (y_1, y_2, \dots, y_m)^\top \in F^m$, such that

$$
\begin{aligned}
y_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n, \\
y_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n, \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
y_m &= a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n.
\end{aligned}
$$

Using matrix notation this can be written shorter:

$$\varphi x = Ax. \tag{8.2}$$

It is not hard to see that φ is linear transformation. Further we will see that the formula (8.2) defines the general form of linear transformation from F^n to F^m.

The *differentiation transformation*, D, from $F[x]$ to $F[x]$ maps each polynomial $f(x)$ to its derivative:

$$Df(x) = f'(x).$$

The differentiation transformation is often denoted by $\dfrac{d}{dx}$.

The *integration transformation* I from $F[x]$ to $F[x]$ maps each polynomial to its primitive:

$$If(x) = \int_0^x f(t)dt.$$

Both these transformations are linear.

8.2. The matrix of linear transformation

Let $\mathbf{e} = \{e_1, e_2, \ldots, e_n\}$ be a basis of a vector space V.

Let

$$x = \sum_{j=1}^{n} \alpha_j e_j \in V$$

i. e. $[x]_{\mathbf{e}} = (\alpha_1, \alpha_2, \ldots, \alpha_n)^\top$. From (8.1) we have

$$\varphi x = \sum_{j=1}^{n} \alpha_j \varphi e_j. \tag{8.3}$$

We have proved the following

Proposition 8.1. *A linear transformation is restored uniquely from the images of basis vectors.*

Let $\mathbf{f} = \{f_1, f_2, \ldots, f_m\}$ be a basis of a vector space W. Since $\varphi e_j \in W$ $(j = 1, 2, \ldots, n)$, then φe_j can be represented as

$$\varphi e_j = \sum_{i=1}^{m} \alpha_{ij} f_i,$$

i. e. $[\varphi e_j]_{\mathbf{f}} = (\alpha_{1j}, \alpha_{2j}, \ldots, \alpha_{mj})^\top$.

The matrix in which the j-th column is the column vector $[\varphi e_j]_{\mathbf{f}}$ $(j = 1, 2, \ldots, n)$ is called the *matrix of transformation* φ relative to bases$\mathbf{e}$ and $\mathbf{f}$.

Notation: $[\varphi]_{\mathbf{f},\mathbf{e}}$ or simply $[\varphi]$ (if there is no ambiguity).

From the definition we have

$$[\varphi]_{\mathbf{f},\mathbf{e}} = (\alpha_{ij}) \in F^{m \times n}.$$

Proposition 8.2. *For any vector x in V*

$$[\varphi x]_{\mathbf{f}} = [\varphi]_{\mathbf{f},\mathbf{e}}[x]_{\mathbf{e}}. \tag{8.4}$$

Proof. Coming to coordinates in (8.3) we get

$$[\varphi x]_{\mathbf{f}} = \sum_{j=1}^{n} \alpha_j [\varphi e_j]_{\mathbf{f}}.$$

The last equation can be written in matrix notation as (8.4). □

Consider equation

$$[\varphi x]_{\mathbf{f}} = A[x]_{\mathbf{e}}, \tag{8.5}$$

we A is a matrix in $F^{m\times n}$. This formula to each vector $x \in V$ puts in correspondence a vector $\varphi x \in W$ and hence define a mapping $\varphi : V \to W$.

Proposition 8.3. *The mapping* $\varphi : V \to W$*, defined by (8.5), is the linear transformation, with* A *being its matrix:* $[\varphi]_{\mathbf{f},\mathbf{e}} = A$*.*

Proof. For all $x, y \in V$ we have

$$[\varphi(x+y)]_{\mathbf{f}} = A[x+y]_{\mathbf{e}} = A([x]_{\mathbf{e}} + [y]_{\mathbf{e}}) = A[x]_{\mathbf{e}} + A[y]_{\mathbf{e}} = [\varphi x]_{\mathbf{f}} + [\varphi y]_{\mathbf{f}},$$

hence $\varphi(x+y) = \varphi x + \varphi y$. For any $x \in V$ and any $\alpha \in F$ we have

$$[\varphi(\alpha x)]_{\mathbf{f}} = A[\alpha x]_{\mathbf{e}} = A(\alpha[x]_{\mathbf{e}}) = \alpha A[x]_{\mathbf{e}} = \alpha[\varphi x]_{\mathbf{f}},$$

hence, $\varphi(\alpha x) = \alpha\varphi x$. Thus, the mapping φ is linear transformation.

Verify that $[\varphi]_{\mathbf{f},\mathbf{e}} = A$. Substitute e_j $(j = 1, 2, \ldots, n)$ in (8.5): $[\varphi e_j]_{\mathbf{f}} = A[e_j]_{\mathbf{e}}$. Since j-th entry of the column $[e_j]_{\mathbf{e}}$ is 1, and all other entries are 0, then $A[e_j]_{\mathbf{e}}$ is the j-th column of A, hence A by definition is the matrix of transformation φ. □

Thus, formula (8.5) define the general form of linear transformation.

Note 8.4. The matrix of linear transformation from F^n to F^m defined by (8.2) in standard bases of these spaces is A. Hence formula (8.2) defines the general case of linear transformation from F^n to F^m. This allows us to identify linear transformations with matrices and transfer the terminology from transformations to matrices.

8.3. Changing the matrix of linear transformation if bases are changed

Let $\varphi \in \Phi(V, W)$,

$$\mathbf{e} = \{e_1, e_2, \ldots, e_n\}\,, \; \mathbf{e}' = \{e'_1, e'_2, \ldots, e'_n\} \text{ are two bases of } V,$$
$$\mathbf{f} = \{f_1, f_2, \ldots, f_n\}\,, \; \mathbf{f}' = \{f'_1, f'_2, \ldots, f'_n\} \text{ are two bases of } W.$$

Consider what happens with matrix of transformation φ if bases $\mathbf{e}$ and $\mathbf{f}$ will be changed on $\mathbf{e}'$ and $\mathbf{f}'$ correspondingly.

Proposition 8.5.

$$[\varphi]_{f',e'} = [\mathbf{f}']_{\mathbf{f}}^{-1}[\varphi]_{f,e}[\mathbf{e}']_{\mathbf{e}}. \tag{8.6}$$

Proof. From

$$[x]_{\mathbf{e}} = [\mathbf{e}']_{\mathbf{e}}[x]_{e'}, \qquad [\varphi x]_{\mathbf{f}} = [\mathbf{f}']_{\mathbf{f}}[\varphi x]_{f'}, \qquad [\varphi x]_{\mathbf{f}} = [\varphi]_{f,e}[x]_{e}$$

we can easily obtain the following

$$[\varphi x]_{\mathbf{f}'} = [\mathbf{f}']_{\mathbf{f}}^{-1}[\varphi]_{\mathbf{f},\mathbf{e}}[\mathbf{e}']_{\mathbf{e}}[x]_{\mathbf{e}'},$$

for all $x \in V$.

Denote $A = [\mathbf{f}']_{\mathbf{f}}^{-1}[\varphi]_{\mathbf{f},\mathbf{e}}[\mathbf{e}']_{\mathbf{e}}$, by Proposition 8.3 we obtain what is required. □

8.4. The image and the kernel of a linear transformation

The *image* of a transformation $\varphi \in \Phi(V, W)$ is the set

$$\operatorname{Im}\varphi = \varphi V = \{y \in W : \ \exists\, x \in V, \varphi x = y\}\,.$$

The *kernel* of $\varphi \in \Phi(V, W)$ is

$$\operatorname{Ker}\varphi = \{x \in V : \varphi x = o\}\,.$$

Exercize 8.6. Let $\varphi \in \Phi(V, W)$. Prove that $\operatorname{Ker}\varphi$ and $\operatorname{Im}\varphi$ are subspaces in W and V correspondingly.

The dimension of the image of a linear transformation is called its *rank* and denotes by $\operatorname{rank}\varphi$.

The dimension of the kernel of a linear transformation is called its *defect* or *nullity* and denoted by $\operatorname{def}\varphi$.

Theorem 8.7.

1. *Let* $\mathbf{e} = \{e_1, e_2, \ldots, e_n\}$ *and* $\mathbf{f} = \{f_1, f_2, \ldots, f_m\}$ *be bases of spaces* V, W *correspondingly. Then* $\operatorname{rank}\varphi = \operatorname{rank}[\varphi]_{\mathbf{f},\mathbf{e}}$.

2. *The following equality holds:* $\operatorname{def}\varphi + \operatorname{rank}\varphi = \dim V$.

Proof. 1) Consider the set $[\varphi V]$ of coordinate columns of all vectors in $\varphi V = \{y = \varphi x : x \in V\}$. We have $[\varphi V] = \{y = [\varphi]_{\mathbf{f},\mathbf{e}} x : x \in F^n\}$, hence $[\varphi V]$ is a linear hull of all columns of the matrix $[\varphi]_{\mathbf{f},\mathbf{e}}$ and $[\varphi V]$ has the dimension $\operatorname{rank}[\varphi]_{\mathbf{f},\mathbf{e}}$. Obviously, $\dim V = \dim[\varphi V]$.

2) Consider the set $[\operatorname{Ker} V]$ of all coordinate columns of all vectors in $\operatorname{Ker}\varphi$. We have $[\operatorname{Ker}\varphi] = \{x \in F^n : [\varphi]_{\mathbf{f},\mathbf{e}} x = 0\}$, hence $[\operatorname{Ker}\varphi]$ is a set of all solutions to the homogeneous system of linear equations, hence $\dim[\operatorname{Ker}\varphi] = n - \operatorname{rank}[\varphi]_{\mathbf{f},\mathbf{e}}$, or $\operatorname{def}\varphi = n - \operatorname{rank}\varphi$. $\square$

8.5. Linear transformations mapping the space to itself

Let's consider linear transformations mapping the vector space V to itself. We denote the set of all such transformations by $\Phi(V)$.

The *matrix of linear transformation* φ relative to (one!) basis $\mathbf{e} = \langle e_1, e_2, \ldots, e_n \rangle$ is the matrix composed from coordinate columns $[\varphi e_j]_{\mathbf{e}}$ $(j = 1, 2, \ldots, n)$.

Notation: $[\varphi]_{\mathbf{e}}$.

8.6. Eigenvalues and eigenvectors of a linear transformation

Let $\varphi \in \Phi(V)$.

Non-zero vector $x \in V$ is called the *eigenvector* of φ if there exists $\lambda \in F$ such that

$$\varphi x = \lambda x. \tag{8.7}$$

The scalar λ is called the *eigenvalue* of φ.

Consider the set of all eigenvectors corresponding to the fixed eigenvalue λ. Amplify this set by the zero vector and denote the resulting set by V_λ. We have

$$V_\lambda = \{x \in V : \varphi x = \lambda x\}.$$

The equality $\varphi x = \lambda x$ can be equivalently written as $(\varphi - \lambda\varepsilon)x = o$, hence

$$V_\lambda = \operatorname{Ker}(\varphi - \lambda\varepsilon).$$

Corollary 8.8. *The set V_λ is a subspace in V.*

Proposition 8.9. *If x is an eigenvector then for any non-zero scalar $\alpha \in F$ the vector αx is also an eigenvector corresponding to the same eigenvalue.*

Proof.

$$\varphi(\alpha x) = \alpha(\varphi x) = \alpha(\lambda x) = \lambda(\alpha x),$$

i. e. αx is an eigenvalue corresponding to the eigenvalue λ. □

Consider the problem of finding all eigenvalues of the transformation φ.

Let $\mathbf{e} = \{e_1, \ldots, e_n\}$ be a basis of V.

A vector x is an eigenvector if and only if $x \neq o$ and $(\varphi - \lambda\varepsilon)x = o$ for some $\lambda \in F$.

This is euivalent to

$$\Big([\varphi]_{\mathbf{e}} - \lambda E\Big)\,[x]_{\mathbf{e}} = o. \tag{8.8}$$

Thus x is an eigenvector if and only if its coordinate column $[x]_{\mathbf{e}}$ is a non-zero solution to the square system of homogenius linear equations (8.8).

This vector exists if and only if

$$\det\left([\varphi]_{\mathbf{e}} - \lambda E\right) = 0. \tag{8.9}$$

By analogy with (8.9) we can consider the equation (respective to the unknown λ):

$$\det(A - \lambda E) = 0,$$

where $A \in F^{n \times n}$

This equation is called *characteristic equation of the matrix* A.

The polynomial

$$\chi_A(\lambda) = \det(\lambda E - A)$$

is also called the *characteristic polynomial* of A.

We have proved the following

Theorem 8.10. *A scalar $\lambda \in F$ is an eigenvalue of a linear transformation $\varphi \in \Phi(V)$ if and only if λ is a root of characteristic polynomial of φ.*

Let $\varphi \in \Phi(V)$ and vectors $e_1, \ldots, e_n$ are eigenvectors and $e_1, \ldots, e_n$ form the basis of V.

$$\varphi e_j = \lambda_j e_j \qquad (j = 1, \ldots, n),$$

Then

$$[\varphi]_{\mathbf{e}} = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}.$$

Example 8.11. Over the field $\mathbb{R}$ find eigenvalues and eigenvectors of the matrix

$$\begin{pmatrix} -3 & 12 \\ -2 & 7 \end{pmatrix}$$

Find a nonsingular matrix Q such that $Q^{-1}AQ$ is a diagonal matrix.

$$A - \lambda I = \begin{vmatrix} -3-\lambda & 12 \\ -2 & 7-\lambda \end{vmatrix} = (-3-\lambda)(7-\lambda) - 12 \cdot (-2) = \lambda^2 - 4\lambda + 3 = (\lambda - 1)(\lambda - 3)$$

The matrix has two eigenvalues: $\lambda_1 = 1$, $\lambda_2 = 3$.

Let's find corresponding eigenvectors.

$$(A - \lambda_1 I)x = \begin{pmatrix} -4 & 12 \\ -2 & 6 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The solution is $(x_1, x_2)^\top = t \cdot (3, 1)^\top$, $t \in \mathbb{R}$

$$(A-\lambda_1 I)x = \begin{pmatrix} -6 & 9 \\ -2 & 4 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The solution is $(x_1, x_2)^\top = t \cdot (2,3)^\top$, $t \in \mathbb{R}$

Vectors $(3,1)^\top$, $(2,3)^\top$ form the basis.

Thus, if

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, \qquad Q = \begin{pmatrix} 3 & 2 \\ 1 & 3 \end{pmatrix},$$

then $D = Q^{-1}AQ$.

Example 8.12. Over the field $\mathbb{R}$ find eigenvalues and eigenvectors of the matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}$$

Does exist a nonsingular matrix Q such that $Q^{-1}AQ$ is diagonal matrix?

$$A - \lambda I = \begin{vmatrix} -\lambda & 1 \\ -1 & 2-\lambda \end{vmatrix} = (-\lambda)(2-\lambda) - 1 \cdot (-1) = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2.$$

The matrix has one eigenvalue $\lambda_1 = 1$.

Let's find corresponding eigenvectors.

$$(A - \lambda_1 I)x = \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The solution is $(x_1, x_2)^\top = t \cdot (1, 1)^\top$, $t \in \mathbb{R}$.

There is no basis consisting of eigenvalues, so there is no matrix Q such that $Q^{-1}AQ$ is a diagonal matrix.