Introduction to Geometry

it is a draft of lecture notes of H.M. Khudaverdian.
Manchester, 22 March 2017

Contents

1 Euclidean space

We recall important notions from linear algebra.

1.1 Vector space.

To study geometry we need the conception of *vector space* and associated with it *affine space*. Now we will recall the definition of vector space. Later we will consider affine space.

Vector space V on real numbers is a set of vectors with operations " $+$ "—addition of vector and " $\cdot$ "—multiplication of vector Lon real number (sometimes called coefficients, scalars). These operations obey the following axioms

- $\forall \mathbf{a}, \mathbf{b} \in V, \mathbf{a} + \mathbf{b} \in V$,
- $\forall \lambda \in \mathbf{R}, \forall \mathbf{a} \in V, \lambda \mathbf{a} \in V$.
- $\forall \mathbf{a}, \mathbf{b} \mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ (commutativity)
- $\forall \mathbf{a}, \mathbf{b}, \mathbf{c},\ \mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$ (associativity)
- $\exists\ \mathbf{0}$ such that $\forall \mathbf{a}$, $\mathbf{a} + \mathbf{0} = \mathbf{a}$
- $\forall \mathbf{a}$ there exists a vector $-\mathbf{a}$ such that $\mathbf{a} + (-\mathbf{a}) = 0$.
- $\forall \lambda \in \mathbf{R}, \lambda(\mathbf{a} + \mathbf{b}) = \lambda \mathbf{a} + \lambda \mathbf{b}$
- $\forall \lambda, \mu \in \mathbf{R} (\lambda + \mu) \mathbf{a} = \lambda \mathbf{a} + \mu \mathbf{a}$
- $(\lambda \mu) \mathbf{a} = \lambda (\mu \mathbf{a})$
- $1 \mathbf{a} = \mathbf{a}$

It follows from these axioms that in particularly $\mathbf{0}$ is unique and $-\mathbf{a}$ is uniquely defined by $\mathbf{a}$. (Prove it.)

Remark We denote by 0 real number 0 and *vector* $\mathbf{0}$. Sometimes we have to be careful to distinguish between zero vector $\mathbf{0}$ and number zero.

Examples of vector spaces... Consider now just one non-trivial example: a space of polynomials of order ≤ 2:

$$V = \{ax^2 + bx + c, a, b, c \in \mathbf{R}\} .$$

It is easy to see that polynomials are 'vectors' with respect to operation of addition and multiplication on numbers.

Consider **conterexample**: a space of polynomials of order 2 such that leading coefficient is equal to 1:

$$V = \{x^2 + bx + c, a, b, c \in \mathbf{R}\} .$$

This is not vcto space: why? since the for any two polynomials f, g from thsi space the polynomials $f - g$, $f + g$ *does not belong to this space.*

1.2 Basic example of (n-dimensional) vector space—$\mathbf{R}^n$

A basic example of vector space (over real numbers) is a space of ordered n-tuples of real numbers.
$\mathbf{R}^2$ is a space of pairs of real numbers. $\mathbf{R}^2 = \{(x, y),\ x, y \in \mathbf{R}\}$
$\mathbf{R}^3$ is a space of triples of real numbers. $\mathbf{R}^3 = \{(x, y, z),\ x, y, z \in \mathbf{R}\}$
$\mathbf{R}^4$ is a space of quadruples of real numbers. $\mathbf{R}^4 = \{(x, y, z, t),\ x, y, z, t, \in \mathbf{R}\}$
and so on...

$\mathbf{R}^n$—is a space of n-typles of real numbers:

$$\mathbf{R}^n = \{(x^1, x^2, \dots, x^n),\ x^1, \dots, , x^n \in \mathbf{R}\} \tag{1.1}$$

If $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$ are two vectors, $\mathbf{x} = (x^1, \dots, x^n)$, $\mathbf{y} = (y^1, \dots, y^n)$ then

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, \dots, x_n + y_n) .$$

and multiplication on scalars is defined as

$$\lambda\mathbf{x} = \lambda \cdot (x^1, \dots, x^n) = (\lambda x^1, \dots, \lambda x^n) , \quad (\lambda \in \mathbf{R}) .$$

($\lambda \in \mathbf{R}$).

Remark Why $\mathbf{R}^n$ is n-dimensional vector space? We see it later in the subsection 1.4

1.3 Linear dependence of vectors

We often consider linear combinations in vector space:

$$\sum_i \lambda_i \mathbf{x}_i = \lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \dots + \lambda_m \mathbf{x}_m , \tag{1.2}$$

where $\lambda_1, \lambda_2, \dots, \lambda_m$ are coefficients (real numbers), $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ are vectors from vector space V. We say that linear combination (1.2) is *trivial* if all coefficients $\lambda_1, \lambda_2, \dots, \lambda_m$ are equal to zero.

$$\lambda_1 = \lambda_2 = \dots = \lambda_m = 0\,.$$

We say that linear combination (1.2) is *not trivial* if at least one of coefficients $\lambda_1, \lambda_2, \dots, \lambda_m$ is not equal to zero:

$$\lambda_1 \neq 0, \text{or} \lambda_2 \neq 0, \text{or} \dots \text{or} \lambda_m \neq 0\,.$$

Recall definition of linearly dependent and linearly independent vectors:

Definition The vectors $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ in vector space V are *linearly dependent* if there exists a non-trivial linear combination of these vectors such that it is equal to zero.

In other words we say that the vectors $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ in vector space V are *linearly dependent* if there exist coefficients $\mu_1, \mu_2, \dots, \mu_m$ such that at least one of these coefficients is not equal to zero and

$$\mu_1\mathbf{x}_1 + \mu_2\mathbf{x}_2 + \dots + \mu_m\mathbf{x}_m = 0\,. \tag{1.3}$$

Respectively vectors $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ are *linearly independent* if they are not linearly dependent. This means that an arbitrary linear combination of these vectors which is equal zero is trivial.

In other words vectors $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_m\}$ are *linearly independent* if the condition

$$\mu_1\mathbf{x}_1 + \mu_2\mathbf{x}_2 + \dots + \mu_m\mathbf{x}_m = 0$$

implies that $\mu_1 = \mu_2 = \dots = \mu_m = 0$.

Very useful and workable

Proposition *Vectors $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ in vector space V are linearly dependent if and only if at least one of these vectors is expressed via linear combination of other vectors:*

$$\mathbf{x}_i = \sum_{j\neq i} \lambda_j \mathbf{x}_j\,.$$

Proof. If the condition (1.3) is obeyed then $x_i - \sum_{j\neq i}\lambda_j\mathbf{x}_j = 0$. This non-trivial linear combination is equal to zero. Hence vectors $\{x_1, \dots, \mathbf{x}_m\}$ are linearly dependent.

Now suppose that vectors $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ are linearly dependent. This means that there exist coefficients $\mu_1, \mu_2, \dots, \mu_m$ such that at least one of these coefficients is not equal to zero and the sum (1.3) equals to zero. WLOG suppose that $\mu_1 \neq 0$. We see that to

$$\mathbf{x}_1 = -\frac{\mu_2}{\mu_1}\mathbf{x}_2 - \frac{\mu_3}{\mu_1}\mathbf{x}_3 - \cdots - \frac{\mu_m}{\mu_1}\mathbf{x}_m \,,$$

i.e. vector $\mathbf{x}_1$ is expressed as linear combination of vectors $\{\mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_m\}$ $\blacksquare$.

1.4 Dimension of vector space. Basis in vector space.

Definition Vector space V has a dimension n if there exist n linearly independent vectors in this vector space, and any $n+1$ vectors in V are linearly dependent.

In the case if in the vector space V for an arbitrary N there exist N linearly independent vectors then the space V is *infinite-dimensional.* An example of infinite-dimensional vector space is a space V of all polynomials of an arbitrary order. One can see that for an arbitrary N polynomials

$$\{1, x, x^2, x^3, \dots, x^N\}$$

are linearly idependent. (Try to prove it!). This implies V is infinite-dimensional vector space.

Basis

Definition Let V be n-dimensional vector space. The ordered set $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ of n linearly independent vectors in V is called a basis of the vector space V.

Remark We say 'a basis', not 'the basis' since there are many bases in the vector space (see also Homeworks 1.2).

Remark Focus your attention: basis is *an ordered* set of vectors, not just a set of vectors[1].

Proposition *Let $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be an arbitrary basis in n-dimensional vector space V. Then any vector $\mathbf{x} \in V$ can be expressed as a linear combination of vectors $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ in a unique way, i.e. for every vector $\mathbf{x} \in V$ there exists an ordered set of coefficients $\{x^1, \dots, a^n\}$ such that*

$$\mathbf{x} = x^1\mathbf{e}_1 + \cdots + x^n\mathbf{e}_n \tag{1.4}$$

and if

[1] See later on orientation of vector spaces, where the ordering of vectors of basis will be highly important.

$$\mathbf{x} = a^1\mathbf{e}_1 + \cdots + a^n\mathbf{e}_n = b^1\mathbf{e}_1 + \cdots + b^n\mathbf{e}_n\,, \tag{1.5}$$

then $a^1 = b^1, a^2 = b^2, \ldots, a^n = b^n$. *In other words for any vector* $\mathbf{x} \in V$ *there exists an ordered* n*-tuple* $(x^1, \ldots, x^n)$ *of coefficients such that* $\mathbf{x} = \sum_{i=1}^n x^i\mathbf{e}_i$ *and this* n*-tuple is unique.*

Proof Let $\mathbf{x}$ be an arbitrary vector in vector space V. The dimension of vector space V equals to n. Hence $n+1$ vectors $(\mathbf{e}_1, \ldots, \mathbf{e}_n, \mathbf{x})$ are linearly dependent: $\lambda_1\mathbf{e}_1 + \cdots + \lambda_n\mathbf{e}_n + \lambda_{n+1}\mathbf{x} = 0$ and this combination is non-trivial. If $\lambda_{n+1} = 0$ then $\lambda_1\mathbf{e}_1 + \cdots + \lambda_n\mathbf{e}_n = 0$ and this combination is non-trivial, i.e. vectors $(\mathbf{e}_1, \ldots, \mathbf{e}_n$ are linearly dependent. Contradiction. Hence $\lambda_{n+1} \neq 0$, i.e. vector $\mathbf{x}$ can be expressed via vectors $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$: $\mathbf{x} = x^1\mathbf{e}_1 + \ldots x^n\mathbf{e}_n$ where $x^i = -\frac{\lambda_i}{\lambda_{n+1}}$. We proved that any vector can be expressed via vectors of basis. Prove now the uniqueness of this expansion. Namely, if (1.5) holds then $(a^1 - b^1)\mathbf{e}_1 + (a^2 - b^2)\mathbf{e}_2 + \cdots + (a^n - b^n)\mathbf{e}_n = 0$. Due to linear independence of basis vectors this means that $(a^1 - b^1) = (a^2 - b^2) = \cdots = (a^n - b^n) = 0$, i.e. $a^1 = b^1, a^2 = b^2, \ldots, a^n = b^n$ ∎

In other words:

Basis is a set of linearly independent vectors in vector space V which span (generate) vector space V.

(Recall that we say that vector space V is *spanned* by vectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ (or vectors vectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ *span* vector space V) if any vector $\mathbf{a} \in V$ can be expresses as a linear combination of vectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$.

Definition Coefficients $\{a^1, \ldots, a^n\}$ are called *components of the vector* $\mathbf{x}$ *in the basis* $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ or just shortly *components of the vector* $\mathbf{x}$.

Remark Basis is a maximal set of linearly independent vectors in a linear space V.

This leads to definition of a basis in infinite-dimensional space. We have to note that in infinite-dimensional space more useful becomes the conception of *topological basis* when infinite sums are considered.

Canonical basis in $\mathbf{R}^n$

We considered above the basic example of n-dimensional vector space—a space of ordered n-tuples of real numbers: $\mathbf{R}^n = \{(x^1, x^2, \ldots, x^n), x^i \in \mathbf{R}\}$ (see the subsection 1.2). What is the meaning of letter 'n' in the definition of $\mathbf{R}^n$?

Consider vectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n \in \mathbf{R}^n$:

$$\begin{array}{rl} \mathbf{e}_1 = & (1,0,0\dots,0,0) \\ \mathbf{e}_2 = & (0,1,0\dots,0,0) \\ \dots & \quad\quad \dots \\ \mathbf{e}_n = & (0,0,0\dots,0,1) \end{array} \tag{1.6}$$

Then for an arbitrary vector $\mathbf{R}^n \ni \mathbf{a} = (a^1, a^2, a^3, \dots, a^n)$

$$\mathbf{a} = (a^1, a^2, a^3, \dots, a^n) =$$

$$a^1(1,0,0\dots,0,0)+a^2(0,1,0\dots,0,0)+a^3(0,0,1,0\dots,0,0)+\dots+a^n(0,1,0\dots,0,1) =$$

$$= \sum_{i=1}^{m} a^i \mathbf{e}_i = a^i \mathbf{e}_i \qquad \text{(we will use sometimes condensed notations } \mathbf{x} = x^i \mathbf{e}_i)$$

Thus we see that for every vector $\mathbf{a} \in \mathbf{R}^n$ we have unique expansion via the vectors (1.6).

The basis (1.6) is the distinguished basis. Sometimes it is called *canonical basis in* $\mathbf{R}^n$. One can find another basis in $\mathbf{R}^n$–just take an arbitrary ordered set of n linearly independent vectors. (See exercises in Homework 0).

1.5 Scalar product. Euclidean space

In vector space one have additional structure: *scalar product of vectors.*

Definition Scalar product in a vector space V is a function $B(\mathbf{x}, \mathbf{y})$ on a pair of vectors which takes real values and satisfies the the following conditions:

$$\begin{array}{c} B(\mathbf{x}, \mathbf{y}) = B(\mathbf{y}, \mathbf{x}) \quad \text{(symmetricity condition)} \\ B(\lambda\mathbf{x} + \mu\mathbf{x}', \mathbf{y}) = \lambda B(\mathbf{x}, \mathbf{y}) + \mu B(\mathbf{x}', \mathbf{y}) \quad \text{(linearity condition)} \\ B(\mathbf{x}, \mathbf{x}) \geq 0 \ , B(\mathbf{x}, \mathbf{x}) = 0 \Leftrightarrow \mathbf{x} = 0 \ \text{(positive-definiteness condition)} \end{array} \tag{1.7}$$

Definition Euclidean space is a vector space equipped with a scalar product.

One can easy to see that the function $B(\mathbf{x}, \mathbf{y})$ is bilinear function, i.e. it is linear function with respect to the second argument also[2]. This follows from previous axioms:

$$B(\mathbf{x}, \lambda\mathbf{y}+\mu\mathbf{y}') \underbrace{=}_{\text{symm.}} B(\lambda\mathbf{y}+\mu\mathbf{y}', \mathbf{x}) \underbrace{=}_{\text{linear.}} \lambda B(\mathbf{y}, \mathbf{x})+\mu B(\mathbf{y}', \mathbf{x}) \underbrace{=}_{\text{symm.}} \lambda B(\mathbf{x}, \mathbf{y})+\mu B(\mathbf{x}, \mathbf{y}') \,.$$

[2]Here and later we will denote scalar product $B(\mathbf{x}, \mathbf{y})$ just by $(\mathbf{x}, \mathbf{y})$. Scalar product sometimes is called inner product. Sometimes it is called dot product.

A bilinear function $B(\mathbf{x},\mathbf{y})$ on pair of vectors is called sometimes *bilinear form* on vector space. Bilinear form $B(\mathbf{x},\mathbf{y})$ which satisfies the symmetricity condition is called *symmetric bilinear form.* Scalar product is nothing but symmetric bilinear form on vectors which is positive-definite: $B(\mathbf{x},\mathbf{x}) \geq 0$) and is non-degenerate $((\mathbf{x},\mathbf{x}) = 0 \Rightarrow \mathbf{x} = 0$.

Example We considered the vector space $\mathbf{R}^n$, the space of n-tuples (see the subsection 1.2). One can consider the vector space $\mathbf{R}^n$ as Euclidean space provided by the scalar product

$$B(\mathbf{x},\mathbf{y}) = x^1y^1 + \cdots + x^ny^n \tag{1.8}$$

This scalar product sometimes is called *canonical scalar product.*

Exercise Check that it is indeed scalar product.

Example We consider in 2-dimensional vector space V with basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ and $B(\mathbf{X},\mathbf{Y})$ such that $B(\mathbf{e}_1,\mathbf{e}_1) = 3$, $B(\mathbf{e}_2,\mathbf{e}_2) = 5$ and $B(\mathbf{e}_1,\mathbf{e}_2) = 0$. Then for every two vectors $\mathbf{X} = x^1\mathbf{e}_1 + x^2\mathbf{e}_2$ and $\mathbf{Y} = y^1\mathbf{e}_1 + y^2\mathbf{e}_2$ we have that

$$B(\mathbf{X},\mathbf{Y}) = (\mathbf{X},\mathbf{Y}) = \left(x^1\mathbf{e}_1 + x^2\mathbf{e}_2, y^1\mathbf{e}_1 + y^2\mathbf{e}_2\right) =$$

$$x^1y^1(\mathbf{e}_1,\mathbf{e}_1) + x^1y^2(\mathbf{e}_1,\mathbf{e}_2) + x^2y^1(\mathbf{e}_2,\mathbf{e}_1) + x^2y^2(\mathbf{e}_2,\mathbf{e}_2) = 3x^1y^1 + 5x^2y^2\,.$$

One can see that all axioms are obeyed.

Notations!

Scalar product sometimes is called "inner" product or "dot" product. Later on we will use for scalar product $B(\mathbf{x},\mathbf{y})$ just shorter notation $(\mathbf{x},\mathbf{y})$ (or $\langle\mathbf{x},\mathbf{y}\rangle$). Sometimes it is used for scalar product a notation $\mathbf{x}\cdot\mathbf{y}$. Usually this notation is reserved only for the canonical case (1.8).

Counterexample Consider again 2-dimensional vector space V with basis $\{\mathbf{e}_1, \mathbf{e}_2\}$.

Show that operation such that $(\mathbf{e}_1,\mathbf{e}_1) = (\mathbf{e}_2,\mathbf{e}_2) = 0$ and $(\mathbf{e}_1,\mathbf{e}_2) = 1$ does not define scalar product. *Solution.* For every two vectors $\mathbf{X} = x^1\mathbf{e}_1 + x^2\mathbf{e}_2$ and $\mathbf{Y} = y^1\mathbf{e}_1 + y^2\mathbf{e}_2$ we have that

$$(\mathbf{X},\mathbf{Y}) = \left(x^1\mathbf{e}_1 + x^2\mathbf{e}_2, y^1\mathbf{e}_1 + y^2\mathbf{e}_2\right) = x^1y^2 + x^2y^1$$

hence for vector $\mathbf{X} = (1,-1)$ $(\mathbf{X},\mathbf{X}) = -2 < 0$. Positive-definiteness is not fulfilled.

Another **Counterexample** Show that operation $(\mathbf{X},\mathbf{Y}) = x^1y^1 - x^2y^2$ does not define scalar product. *Solution.* Take $\mathbf{X} = (0,-1)$. Then $(\mathbf{X},\mathbf{X}) = -1$. The condition of positive-definiteness is not fulfilled. (See also exercises in Homework 2.)

1.6 Orthonormal basis in Euclidean space

One can see that for scalar product (1.8) and for the basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ defined by the relation (1.6) the following relations hold:

$$(\mathbf{e}_i, \mathbf{e}_j) = \delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases} \tag{1.9}$$

Let $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ be an ordered set of n vectors in n-dimensional Euclidean space which obeys the conditions (1.9). One can see that this ordered set is a basis [3].

Definition-Proposition The ordered set of vectors $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ in n-dimensional Euclidean space which obey the conditions (1.9) is a basis. This basis is called *an orthonormal basis*.

One can prove that every (finite-dimensional) Euclidean space possesses orthonormal basis.

Later by default we consider only orthonormal bases in Euclidean spaces. Respectively scalar product will be defined by the formula (1.8). Indeed let $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ be an orthonormal basis in Euclidean space. Then for an arbitrary two vectors $\mathbf{x}, \mathbf{y}$, such that $\mathbf{x} = \sum x^i \mathbf{e}_i$, $\mathbf{y} = \sum y^j \mathbf{e}_j$ we have:

$$(\mathbf{x}, \mathbf{y}) = \left(\sum x^i \mathbf{e}_i, \sum y^j \mathbf{e}_j\right) = \sum_{i,j=1}^{n} x^i y^j (\mathbf{e}_i, \mathbf{e}_j) = \sum_{i,j=1}^{n} x^i y^j \delta_{ij} = \sum_{i=1}^{n} x^i y^i$$

We come to the canonical scalar product (1.8). Later on we usually will consider scalar product defined by the formula (1.8) i.e. scalar product in orthonormal basis.

Remark We consider here general definition of scalar product then came to conclusion that in a special basis, (*orthonormal basis*), this is nothing but usual 'dot' product (1.8).

1.7 Affine spaces and vector spaces

Let V be a vector space. A set A whose elements will be called 'points' is an *affine space* associated with the vector space V if the following rule is

[3] Indeed prove that conditions (1.9) imply that these n vectors are linear independent. Suppose that $\lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 + \cdots + \lambda_n \mathbf{e}_n = 0$. For an arbitrary i multiply the left and right hand sides of this relation on a vector $\mathbf{e}_i$. We come to condition $\lambda_i = 0$. Hence vectors $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ are linearly dependent.

defined: to every point $P \in A$ and an arbitrary vector $\mathbf{x} \in V$ a point Q is assigned: $(P, \mathbf{x}) \mapsto Q$. We denote $Q = P + \mathbf{x}$.

The following properties must be satisfied:

- For arbitrary two vectors $\mathbf{x}, \mathbf{y} \in V$ and arbitrary point $P \in A$, $P + (\mathbf{x} + \mathbf{y}) = (P + \mathbf{x}) + \mathbf{y}$.
- For an arbitrary point $P \in A$, $P + \mathbf{0} = P$.

  For arbitrary two points $P, Q \in A$ there exists unique vector $\mathbf{y} \in V$ such that $P + \mathbf{y} = Q$.

If $P + \mathbf{x} = Q$ we often denote the vector $\mathbf{x} = Q - P = \vec{PQ}$. We say that vector $\mathbf{x} = \vec{PQ}$ starts at the point P and it ends at the point Q.

One can see that if vector $\mathbf{x} = \vec{PQ}$, then $\vec{QP} = -x$; if P, Q, R are three arbitrary points then $\vec{PQ} + \vec{QR} = \vec{PR}$.

Examples of affine space.

Every vector space can be considered as an affine space in the following way. We define affine space A as a same set as vector space V, but we consider vectors of V as points of this affine space. If $A = \mathbf{a}$ is an arbitrary point of the affine space, and $\mathbf{b}$ is an arbitrary vector of vector space V, then $A + \mathbf{b}$ is equal to the vector $\mathbf{a} + \mathbf{b}$. We assign to two 'points' $A = \mathbf{a}, B = \mathbf{b}$ the vector $\mathbf{x} = \mathbf{b} - \mathbf{a}$.

On the other hand if A is an affine space with associated vector space V, then choose an arbitrary point $O \in A$ and consider the vectors starting at the at the origin. We come to the vector space V.

One can say that vector space is an affine space with fixed origin.

For example vector space $\mathbf{R}^2$ of pairs of real numbers can be considered as a set of points. If we choose arbitrary two points $A = (a^1, a^2), B = (b^1, b^2)$, then the vector $\vec{AB} = B - A = (b^1 - a^1, b^2 - a^2)$.

Geometrical properties of scalar product: length of the vectors, angle between vectors

The scalar product of vector on itself defines the *length of the vector*:

$$\text{Length of the vector } \mathbf{x} = |\mathbf{x}| = \sqrt{(\mathbf{x}, \mathbf{x})} = \sqrt{(x^1)^2 + \cdots + (x^n)^2} \tag{1.10}$$

If we consider Euclidean space $\mathbf{E^n}$ as the set of points (affine space) then the distance between two points $\mathbf{x}, \mathbf{y}$ is the length of corresponding vector:

$$\text{distance between points } \mathbf{x}, \mathbf{y} = |\mathbf{x} - \mathbf{y}| = \sqrt{\left(y^1 - x^1\right)^2 + \cdots + \left(y^n - x^n\right)^2}$$

We recall very important formula how scalar (inner) product is related with the angle between vectors:

$$(\mathbf{x}, \mathbf{y}) = x^1 y^1 + x^2 y^2 = |\mathbf{x}||\mathbf{y}| \cos\varphi$$

where φ is an angle between vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbf{E}^2$.

This formula is valid also in the three-dimensional case and any n-dimensional case for $n \geq 1$. It gives as a tool to calculate angle between two vectors:

$$(\mathbf{x}, \mathbf{y}) = x^1 y^1 + x^2 y^2 + \cdots + x^n y^n = |\mathbf{x}||\mathbf{y}| \cos\varphi \tag{1.11}$$

In particulary it follows from this formula that

$$\begin{gathered}\textit{angle between vectors } \mathbf{x}, \mathbf{y} \textit{ is acute if scalar product } (\mathbf{x}, \mathbf{y}) \textit{ is positive}\\ \textit{angle between vectors } \mathbf{x}, \mathbf{y} \textit{ is obtuse if scalar product } (\mathbf{x}, \mathbf{y}) \textit{ is negative}\\ \textit{vectors } \mathbf{x}, \mathbf{y} \textit{ are perpendicular if scalar product } (\mathbf{x}, \mathbf{y}) \textit{ is equal to zero}\end{gathered} \tag{1.12}$$

Remark Geometrical intuition tells us that cosinus of the angle between two vectors has to be less or equal to one and it is equal to one if and only if vectors $\mathbf{x}, \mathbf{y}$ are collinear. Comparing with (1.11) we come to the inequality:

$$\begin{gathered}(\mathbf{x}, \mathbf{y})^2 = \left(x^1 y^1 + \cdots + x^n y^n\right)^2 \leq \left((x^1)^2 + \cdots + (x^n)^2\right)\left((y^1)^2 + (\cdots + (y^n)^2\right) = (\mathbf{x}, \mathbf{x})(\mathbf{y}, \mathbf{y})\\ \text{and} (\mathbf{x}, \mathbf{y})^2 = (\mathbf{x}, \mathbf{x})(\mathbf{y}, \mathbf{y}) \quad \text{if vectors are colienar, i.e. } x^i = \lambda y^i\end{gathered} \tag{1.13}$$

This is famous Cauchy–Buniakovsky–Schwarz inequality, one of most important inequalities in mathematics. (See for more details Homework 2)

1.8 Transition matrices. Orthogonal bases and orthogonal matrices

One can consider different bases in vector space.

Let A be $n \times n$ matrix with real entries, $A = ||a_{ij}||$, $i, j = 1, 2, \ldots, n$:

$$A = \begin{pmatrix} a_{11} & a_{12} \ldots & a_{1n} \\ a_{21} & a_{22} \ldots & a_{2n} \\ a_{31} & a_{32} \ldots & a_{3n} \\ \ldots & \ldots\ldots & \ldots \\ a_{(n-1)\,1} & a_{(n-1)2} \cdots & a_{(n-1)n} \\ a_{n\,1} & a_{n2} \cdots & a_{nn} \end{pmatrix}$$

Let $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ be an arbitrary basis in n-dimensional vector space V.

The basis $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ can be considered as row of vectors, or $1 \times n$ matrix with entries–vectors.

Multiplying $1 \times n$ matrix $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ on matrix A we come to new row of vectors $\{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n\}$ such that

$$\{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n\} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}A = \tag{1.14}$$

$$\{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n\} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\} \begin{pmatrix} a_{11} & a_{12} \dots & a_{1n} \\ a_{21} & a_{22} \dots & a_{2n} \\ a_{31} & a_{32} \dots & a_{3n} \\ \dots & \dots\dots & \dots \\ a_{(n-1)\,1} & a_{(n-1)2} \dots & a_{(n-1)n} \\ a_{n\,1} & a_{n2} \dots & a_{nn} \end{pmatrix} \tag{1.15}$$

,

$$\begin{cases} \mathbf{e}'_1 = a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2 + a_{31}\mathbf{e}_3 + \dots + a_{(n-1)\,1}\mathbf{e}_{n-1} + a_{n\,1}\mathbf{e}_n \\ \mathbf{e}'_1 = a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2 + a_{32}\mathbf{e}_3 + \dots + a_{(n-1)\,2}\mathbf{e}_{n-1} + a_{n\,2}\mathbf{e}_n \\ \mathbf{e}'_1 = a_{13}\mathbf{e}_1 + a_{23}\mathbf{e}_2 + a_{33}\mathbf{e}_3 + \dots + a_{(n-1)\,3}\mathbf{e}_{n-1} + a_{n\,1}\mathbf{e}_n \\ \dots = \dots \dots + \dots \dots + \dots \dots + \dots + \dots\dots\dots\dots \\ \mathbf{e}'_n = a_{1n}\mathbf{e}_1 + a_{2n}\mathbf{e}_2 + a_{3n}\mathbf{e}_3 + \dots + a_{(n-1)\,n}\mathbf{e}_{n-1} + a_{n\,n}\mathbf{e}_n \end{cases}$$

or shortly:

$$\mathbf{e}'_i = \sum_{k=1}^{n} \mathbf{e}_k a_{ki}\,. \tag{1.16}$$

Definition Matrix A which transforms a basis $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ to the row of vectors $\{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n\}$ (see equation (1.16)) is *transition matrix* from the basis $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ to the row $\{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n\}$.

What is the condition that the row $\{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n\}$ is a basis too? The row, ordered set of vectors, $\{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n\}$ is a basis if and only if vectors $(\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n)$ are linearly independent. Thus we come to

Proposition 1 *Let $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ be a basis in n-dimensional vector space V, and let A be an $n \times n$ matrix with real entries. Then*

$$\{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n\} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}A \tag{1.17}$$

is a basis if and only if the transition matrix A has rank n, i.e. it is non-degenerate (invertible) matrix.

Recall that $n\times$ matrix A is nondegenerate (invertible) $\Leftrightarrow \det A \neq 0$.

Remark Recall that the condition that $n \times n$ matrix A is non-degenerate (has rank n) is equivalent to the condition that it is invertible matrix, or to the condition that $\det A \neq 0$.

Now suppose that $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is orthonoromal basis in n-dimensional Euclidean vector space. What is the condition that the new basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n\} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}A$ is an orthonormal basis too?

Definition We say that $n \times n$ matrix is orthogonal matrix if its product on transposed matrix is equal to unity matrix:

$$A^T A = I \,. \tag{1.18}$$

Exercise. Prove that determinant of orthogonal matrix is equal to ± 1:

$$A^T A = I \Rightarrow \det A = \pm 1 \,. \tag{1.19}$$

Solution $A^T A = I$. Hence $\det(A^T A) = \det A^T \det A = (\det A)^2 = \det I = 1$. Hence $\det A = \pm 1$. We see that in particular orthogonal matrix is non-degenerate ($\det A \neq 0$). Hence it is a transition matrix from one basis to another. The following Proposition is valid:

Proposition 2 Let $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ be an orthonormal basis in n-dimensional Euclidean vector space. Then the new basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n\} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}A$ is orthonormal basis if and only if the transition matrix A is orthogonal matrix.

Proof The basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n\}$ is orthonormal means that $(\mathbf{e}'_i, \mathbf{e}'_j) = \delta_{ij}$. We have:

$$\delta_{ij} = (\mathbf{e}'_i, \mathbf{e}'_j) = \left(\sum_{m=1}^{n} \mathbf{e}_m A_{mi}, \mathbf{e}'_j = \sum_{n=1}^{n} \mathbf{e}_n A_{nj} \right) = \sum_{m,n=1}^{n} A_{mi} A_{nj} (\mathbf{e}_m, \mathbf{e}_n) =$$

$$\sum_{m,n=1}^{n} A_{mi} A_{nj} \delta_{mn} = \sum_{m=1}^{n} A_{mi} A_{mj} = \sum_{m=1}^{n} A^T_{im} A_{mj} = (A^T A)_{ij} \,, \tag{1.20}$$

Hence $(A^T A)_{ij} = \delta_{ij}$, i.e. $A^T A = I$.

We know that determinant of orthogonal matrix equals to ± 1. It is very useful to consider the following groups:

- The group $O(n)$—group of orthogonal $n \times n$ matrices:

$$O(n) = \{A\colon \quad A^T A = I\} \,. \tag{1.21}$$

- The group $SO(n)$ special orthogonal group of $n \times n$ matrices:

$$SO(n) = \{A\colon \quad A^T A = I, \det A = 1\} \,. \tag{1.22}$$

1.9 Linear operators.

1.9.1 Matrix of linear operator in a given basis

Recall here facts about linear operators in vector space

Let P be a linear operator in vector space V:

$$P\colon\ V \to V, \qquad P(\lambda \mathbf{x} + \mu \mathbf{y}) = \lambda P(\mathbf{x}) + \mu P(\mathbf{y}).$$

Let $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be an arbitrary basis in n-dimensional vector space V. Consider the action of operator P on basis vectors: $\mathbf{e}'_i = P(\mathbf{e}_i)$:

$$\begin{array}{c} \mathbf{e}'_1 = P(\mathbf{e}_1) = \mathbf{e}_1 p_{11} + \mathbf{e}_2 p_{21} + \mathbf{e}_3 p_{31} + \dots + \mathbf{e}_n p_{n1} \\ \mathbf{e}'_2 = P(\mathbf{e}_2) = \mathbf{e}_1 p_{12} + \mathbf{e}_2 p_{22} + \mathbf{e}_3 p_{32} + \dots + \mathbf{e}_n p_{n2} \\ \mathbf{e}'_3 = P(\mathbf{e}_3) = \mathbf{e}_1 p_{13} + \mathbf{e}_2 p_{23} + \mathbf{e}_3 p_{31} + \dots + \mathbf{e}_n p_{n3} \\ \dots \\ \mathbf{e}'_n = P(\mathbf{e}_n) = \mathbf{e}_1 p_{1n} + \mathbf{e}_2 p_{2n} + \mathbf{e}_3 p_{3n} + \dots + \mathbf{e}_n p_{nn} \end{array} \tag{1.23}$$

Definition Let $\{\mathbf{e}_i\}$ be a basis. Then the transition matrix $||p_{ik}||$ defined by relation (1.23) is called *matrix of operator P in the basis $\{\mathbf{e}_i\}$.*

$$\mathbf{e}'_i = P(\mathbf{e}_i) = \sum \mathbf{e}_k p_{ki}\,.$$

In the case if linear operator P is non-degenerate (invertible) then vectors $\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3, \dots, \mathbf{e}'_n$, form a basis. The matrix $P = ||p_{ik}||$ is the transition matrix from the basis $\{\mathbf{e}_i\}$ to the basis $\{\mathbf{e}'_i = P(\mathbf{e}_i)\}$.

Matrix of linear operator changes if we change the basis

Before studying this question consider just an example.

Let P be linear operator in 2-dimensional vector space V such that for a basis $\{\mathbf{e}_1, \mathbf{e}_2\}$,

$$P(\mathbf{e}_1) = p_{11}\mathbf{e}_1 + p_{21}\mathbf{e}_2\,, P(\mathbf{e}_2) = p_{12}\mathbf{e}_1 + p_{22}\mathbf{e}_2\,.$$

i.e. the matrix of operator P in the basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ is a matrix

$$\begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}.$$

Consider now instead the basis $\mathbf{e}_1, \mathbf{e}_2$ *another basis*

$$\{\mathbf{f}_1, \mathbf{f}_2\}\,, \text{such that } \mathbf{f}_1 = \frac{1}{2}\mathbf{e}_1\,, \mathbf{f}_2 = 3\mathbf{e}_2\,. \tag{1.24}$$

Then for new basis we have

$$P(\mathbf{f}_1) = P\left(\frac{\mathbf{e}_1}{2}\right) = \frac{1}{2}\left(p_{11}\mathbf{e}_1 + p_{21}\mathbf{e}_2\right) = \frac{1}{2}\left(p_{11}(2\mathbf{f}_1) + p_{21}\frac{\mathbf{f}_2}{3}\right) = p_{11}\mathbf{f}_1 + \frac{p_{21}}{6}\mathbf{f}_2\,,$$

and

$$P(\mathbf{f}_2) = P\left(3\mathbf{e}_2\right) = 3\left(p_{12}\mathbf{e}_1 + p_{22}\mathbf{e}_2\right) = 3\left(p_{12}(2\mathbf{f}_1) + p_{22}\frac{\mathbf{f}_2}{3}\right) = 6p_{12}\mathbf{f}_1 + p_{22}\mathbf{f}_2\,,$$

i.e. the matrix of operator P in the basis $\{\mathbf{f}_1, \mathbf{f}_2\}$ is a matrix

$$\begin{pmatrix} p_{11} & 6p_{12} \\ \frac{p_{21}}{6} & p_{22} \end{pmatrix}.$$

We see that changing of basis (1.24) implies the changing of matrix representing operator P.

Do it now in general case.

Consider a new basis $\{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ in the linear space V. Let A be transition matrix from the basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ to the new basis $\{\mathbf{f}_1, \dots, \mathbf{f}_n\}$:

$$\{\mathbf{f}_1, \dots, \mathbf{f}_n\} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}A, \text{ i.e.} f_i = \sum_k \mathbf{e}_k a_{ki}$$

(see equation (1.16)). Then the action of operator P in the new basis is given by the formula $\mathbf{f}'_i = P(\mathbf{f}_i)$. According to the formulae (1.9.1) and (1.23) we have

$$\mathbf{f}'_i = P(\mathbf{f}_i) = P\left(\sum_q \mathbf{e}_q a_{qi}\right) = \sum_q a_{qi}\left(\sum_r \mathbf{e}_r p_{rq}\right) = \sum_{q,r} \mathbf{e}_r p_{rq} a_{qi} = \sum_r \mathbf{e}_r (PA)_{ri} =$$

$$\sum_{r,k} \mathbf{f}_k (A^{-1})_{kr} (PA)_{ri} = \sum_k \mathbf{f}_k (A^{-1}PA)_{ki}\,.$$

We see that in the new basis $\{\mathbf{f}_i\}$ a matrix of linear operator is $A^{-1}PA$:

$$\text{If } \{\mathbf{e}'_1, \dots, \mathbf{e}'_n\} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}P, \text{ then } \{\mathbf{f}'_1, \dots, \mathbf{f}'_n\} = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}A^{-1}PA, \;, \tag{1.25}$$

where A is transition matrix from the basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ to the basis $\{\mathbf{f}_1, \dots, \mathbf{f}_n\}$,

Consider the following example.

Example Let P be a linear operator in 2-dimensional vector space V such that in a basis $\mathbf{e}_1, \mathbf{e}_2$ it is given by the following relation:

$$P(\mathbf{e}) = 2\mathbf{e}\,, \quad P(\mathbf{e}_2) = \mathbf{e}_2\,.$$

Then the matrix of operator P in this basis is obviously

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \tag{1.26}$$

Now consider another basis, $\{\mathbf{f}_1, \mathbf{f}_2\}$ in the space V:

$$\begin{cases} \mathbf{f}_1 = 7\mathbf{e}_1 + 5\mathbf{e}_2 \\ \mathbf{f}_2 = 4\mathbf{e}_1 + 3\mathbf{e}_2 \end{cases}, \qquad \text{respectively} \qquad \begin{cases} \mathbf{e}_1 = 3\mathbf{f}_1 - 5\mathbf{f}_2 \\ \mathbf{e}_2 = -4\mathbf{f}_1 + 7\mathbf{f}_2 \end{cases}. \tag{1.27}$$

Calculate matrix of the operator P on this new basis:

$$P(\mathbf{f}_1) = P(7\mathbf{e}_1 + 5\mathbf{e}_2) = 14\mathbf{e}_1 + 5\mathbf{e}_2 = 14(3\mathbf{f}_1 - 5\mathbf{f}_2) + 5(-4\mathbf{f}_1 + 7\mathbf{f}_2) = 22\mathbf{f}_1 - 35\mathbf{f}_2\,,$$

$$P(\mathbf{f}_2) = P(4\mathbf{e}_1 + 3\mathbf{e}_2) = 8\mathbf{e}_1 + 3\mathbf{e}_2 = 8(3\mathbf{f}_1 - 5\mathbf{f}_2) + 3(-4\mathbf{f}_1 + 7\mathbf{f}_2) = 12\mathbf{f}_1 - 19\mathbf{f}_2\,.$$

Hence the matrix of operator P in the basis $\{\mathbf{f}_1, \mathbf{f}_2\}$ is matrix

$$\begin{pmatrix} 22 & 12 \\ -35 & -19 \end{pmatrix}. \tag{1.28}$$

Matrices (1.26) and (1.28) are different matrices which are represented the same linear operator P in different bases. According to equation (1.27)

$$\begin{pmatrix} 22 & 12 \\ -35 & -19 \end{pmatrix} = \begin{pmatrix} 7 & 4 \\ 5 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 7 & 4 \\ 5 & 3 \end{pmatrix} = \begin{pmatrix} 3 & -4 \\ -5 & 7 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 7 & 4 \\ 5 & 3 \end{pmatrix}, \tag{1.29}$$

Note that if a matrix $P = ||p_{ij}||$ is the transition matrix from the basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ to the basis $\{\mathbf{e}'_1, \ldots, \mathbf{e}'_n\}$. For an arbitrary vector $\mathbf{x}$

$$\forall \mathbf{x} = \sum_{i=1}^{n} \mathbf{e}_i x^i = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n) \cdot \begin{pmatrix} x^1 \\ x^2 \\ \ldots \\ x^n \end{pmatrix}$$

$$
P\mathbf{x} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) \cdot P \cdot \begin{pmatrix} x^1 \\ x^2 \\ \dots \\ x^n \end{pmatrix} = \sum_{i=1}^{n} \mathbf{e}'_i x^i = \sum_{i,k=1}^{n} \mathbf{e}_k p_{ki} x^i \,.
$$

If x^i are components of vector $\mathbf{x}$ at the basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ and x'^i are components of the vector $\mathbf{x}$ at the new basis $\{\mathbf{e}'_i\}$ then $x'^i = \sum_i p_{ik} x^k$.

1.9.2 Determinant and Trace of linear operator

We recall the definition of determinant and explain what is the trace of linear operator,

Definition-Proposition Let P be a linear operator in vector space V and let $P_{ik} = ||p_{ik}||$ be transition matrix of this operator in an arbitrary basis in V (see construction (1.23).) Then determinant of linear operator P equals to determinant of transition matrix of this operator.

$$
\det P = \det (p_{ik})
$$

In the same way we define trace of operator via trace of matrix:

$$
\operatorname{Tr} P = \operatorname{Tr}\, (||p_{ik}||) = p_{11} + p_{22} + p_{33} + \dots + p_{nn} \,. \tag{1.30}
$$

Determinant and trace of operator are well-defined. since due to (1.25) determinant and trace of transition matrice do not change if we change the basis in spite of the fact that transition matrix changes: $P \mapsto A^{-1}PA$, but

$$
\det \left(A^{-1}PA\right) = \det A^{-1} \det P \det A = (\det A)^{-1} \det P \det A = \det P \,.
$$

In the example above (see equations (1.26) and (1.28)) we have different matrices which represent the same but one operator P in different bases. These matrices are related by equations (1.27) and (1.29) and

$$
\det P = \det \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = 2 \cdot 1 = \det \begin{pmatrix} 22 & 12 \\ -35 & -19 \end{pmatrix} = 22 \cdot (-19) - (-35) \cdot 12 = 2
$$

$$
\operatorname{Tr} P = \operatorname{Tr} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = 2 + 1 = \operatorname{Tr} \begin{pmatrix} 22 & 12 \\ -35 & -19 \end{pmatrix} = 22 - 19 = 3
$$

In the same way one can see that trace is invariant too:

$$
\operatorname{Tr} (A^{-1}PA) = \sum_i (A^{-1}PA)_{ii} = \sum_{i,k,p} \left(A^{-1}\right)_{ik} p_{kp} = \sum_{i,k,p} A_{pi} \left(A^{-1}\right)_{ik} p_{kp} =
$$

$$\sum_{p,k} \left(A \cdot A^{-1}\right)_{pk} p_{kp} = \sum_{p,k} \delta_{kp} p_{kp} = \sum_{k} p_{kk} = \mathrm{Tr}\, P\,.$$

Trace of linear operator is an infinitesimal version of its determinant:

$$\det(1 + tP) = 1 + t\mathrm{Tr}\, P + O(t^2)\,.$$

This is infinitesimal version for the followiong famous formula which relates trace and det of linear operator:

$$\det e^{tA} = e^{t\mathrm{Tr}\, A}\,. \tag{1.31}$$

where $e^{tA} = \sum \frac{t^n A^n}{n!}$. E.g. if $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, then $e^{tA} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$, $\det e^{tA} = 1$ and $e^{t\mathrm{Tr}\, A} = e^0 = 1$.

1.9.3 Orthogonal linear operators

Now we study geometrical meaning of orthogonal linear operators in Euclidean space.

Recall that linear operator P in Euclidean space $\mathbf{E}^n$ is called orthogonal operator if it preserves scalar product:

$$(P\mathbf{x}, P\mathbf{y}) = (\mathbf{x}, \mathbf{y}),\ \text{for arbitrary vectors } \mathbf{x}, \mathbf{y} \tag{1.32}$$

In particular if $\{\mathbf{e}_i\}$ is orthonormal basis in Euclidean space then due to (1.32) the new basis $\{\mathbf{e}'_i = P(\mathbf{e}_i)\}$ is orthonormal too. Thus we see that matrix of orthogonal operator P in a given orthogonal basis is orthogonal matrix:

$$P^T \cdot P = I \tag{1.33}$$

(see (1.18) in subsection 1.7). In particular we see that for orthogonal linear operator $\det P = \pm 1$ (compare with (1.19)).

1.10 Orientation in vector space

You heard words "orientation...", ""

You heard expressions like: A basis $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ have the same orientation as the basis $\{\mathbf{a}', \mathbf{b}', \mathbf{c}'\}$ if they both obey right hand rule or if they both obey left hand rule. In the other case we say that these bases have opposite orientation...

Try to give the exact meaning to these words.

1.10.1 Orientation in vector space. Oriented vector space

Note that in three-dimensional Euclidean space except scalar (inner) product, one can consider another important operation: vector product. The conception of orientation is indispensable for defining this operation.

Consider the set of *all* bases in the given vector space V.

Let $(\mathbf{e}_1, \dots \mathbf{e}_n)$, $(\mathbf{e}'_1, \dots \mathbf{e}'_n)$ be two arbitrary bases in the vector space V and let T be transition matrix which transforms the basis $\{\mathbf{e}_i\}$ to the new basis $\{\mathbf{e}'_i\}$:

$$\{\mathbf{e}'_1, \dots \mathbf{e}'_n\} = \{\mathbf{e}_1, \dots \mathbf{e}_n\}T\,, \qquad (\mathbf{e}'_i = \sum_{k=1}^{n} \mathbf{e}_k t_{ki}) \tag{1.34}$$

(see also (1.15)).

Definition We say that two bases $\{\mathbf{e}_1, \dots \mathbf{e}_n\}$ and $\{\mathbf{e}'_1, \dots \mathbf{e}'_n\}$ in V have *the same orientation* if the determinant of transition matrix (1.34) from the first basis to the second one is positive: $\det T > 0$.

We say that the basis $\{\mathbf{e}_1, \dots \mathbf{e}_n\}$ has an orientation *opposite to the orientation* of the basis $\{\mathbf{e}'_1, \dots \mathbf{e}'_n\}$ (or in other words these two bases have opposite orientation) if the determinant of transition matrix from the first basis to the second one is negative: $\det T < 0$.

Remark Transition matrix from basis to basis is non-degenerate, hence its determinant cannot be equal to zero. It can be or positive or negative.

One can see that orientation establishes the equivalence relation in the set of all bases. Denote this relation by "$\sim$": $\{\mathbf{e}_1, \dots \mathbf{e}_n\} \sim \{\mathbf{e}'_1, \dots \mathbf{e}'_n\}$, if two bases $\{\mathbf{e}_1, \dots \mathbf{e}_n\}$ and $\{\mathbf{e}'_1, \dots \mathbf{e}'_n\}$ have the same orientation, i.e. $\det T > 0$ for transition matrix.

Show that "$\sim$" is an equivalence relation, i.e. this relation is reflexive, symmetric and transitive.

Check it:

- it is reflexive, i.e. for every basis $\{\mathbf{e}_1, \dots \mathbf{e}_n\}$

$$\{\mathbf{e}_1, \dots, \mathbf{e}_n\} \sim \{\mathbf{e}_1, \dots, \mathbf{e}_n\}\,, \tag{1.35}$$

  because in this case transition matrix $T = I$ and $detI = 1 > 0$.

* it is symmetric, i.e.

  If $\{\mathbf{e}_1, \dots, \mathbf{e}_n\} \sim \{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}$ then $\{\mathbf{e}'_1, \dots, \mathbf{e}'_n) \sim \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$,

  because if T is transition matrix from the first basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ to the second basis $\{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}$: $\{\mathbf{e}'_1, \dots, \mathbf{e}'_n\} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}T$,

  then the transition matrix from the second basis $\{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}$ to the first basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is the inverse matrix T^{-1}: $\{\mathbf{e}_1, \dots, \mathbf{e}_n\} = \{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}T^{-1}$. Hence $\det T^{-1} = \frac{1}{\det T} > 0$ if $\det T > 0$.

* Is transitive, i.e. if $\{\mathbf{e}_1, \dots, \mathbf{e}_n\} \sim \{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}$ and $\{\mathbf{e}'_1, \dots, \mathbf{e}'_n) \sim \{\tilde{\mathbf{e}}_1, \dots, \tilde{\mathbf{e}}_n\}$, then one can see that $\{\mathbf{e}_1, \dots, \mathbf{e}_n\} \sim \{\tilde{\mathbf{e}}_1, \dots, \tilde{\mathbf{e}}_n\}$.

  Do it in detail. For convenience call a basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ the 'I-st' basis, call a basis $\{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}$ the 'II-nd' basis and call a basis $\{\tilde{\mathbf{e}}_1, \dots, \tilde{\mathbf{e}}_n\}$ the 'III-rd' basis. Let T_{12} be a transition matrix from the I-st basis to the II-nd basis, T_{13} be a transition matrix from the I-st basis to the III-rd basis and T_{23} be a transition matrix from the II-nd basis to the III-rd basis:

  $$\begin{aligned} \{\mathbf{e}'_1, \dots, \mathbf{e}'_n\} &= \{\mathbf{e}_1, \dots, \mathbf{e}_n\}T_{12} \\ \{\tilde{\mathbf{e}}_1, \dots, \tilde{\mathbf{e}}_n\} &= \{\mathbf{e}_1, \dots, \mathbf{e}_n\}T_{13} \\ \{\tilde{\mathbf{e}}_1, \dots, \tilde{\mathbf{e}}_n\} &= \{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}T_{23}, \end{aligned} \tag{1.36}$$

  Hence $\{\tilde{\mathbf{e}}_1, \dots, \tilde{\mathbf{e}}_n\} = \{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}T_{23} =$

  $$(\{\mathbf{e}_1, \dots, \mathbf{e}_n\}T_{12})\,T_{23} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}T_{12} \circ T_{23} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}T_{13}.$$

  We see that $\underbrace{T_{13}}_{\text{I-st} \to \text{III-rd}} = \underbrace{T_{12}}_{\text{I-st} \to \text{II-nd}} \circ \underbrace{T_{23}}_{\text{II-nd} \to \text{II-rd}}$:

  $$T_{13} = T_{12} \circ T_{23} \Rightarrow \det T_{13} = \det(T_{12} \circ T_{23}) = \det T_{12} \cdot \det T_{23}\,. \tag{1.37}$$

  Transitivity immediately follows from this relation: if I-st $\sim$ II and II-nd $\sim$ III-rd, then determinants of matrices T_{12} and T_{23} are positive. Hence according to relation (1.37) $\det T_{13}$ is positive too, i.e. I-st $\sim$ III-rd.

Since it is equivalence relation the set of all bases is a union if disjoint equivalence classes. Two bases are in the same equivalence class if and only if they have the same orientation.

How many equivalence classes exist? One, two or more?

Show first that there sare at least two equivalence classes.

Example Let $\{\mathbf{e}_1, \mathbf{e}_2 \ldots, \mathbf{e}_n\}$ be an arbitrary basis in n-dimensional vector space V. Swap the vectors $\mathbf{e}_1, \mathbf{e}_2$. We come to a new basis: $\{\mathbf{e}'_1, \mathbf{e}'_2 \ldots, \mathbf{e}'_n\}$

$$\mathbf{e}'_1 = \mathbf{e}_2, \mathbf{e}'_2 = \mathbf{e}_1, \text{ all other vectors are the same: } \mathbf{e}_3 = \mathbf{e}'_3, \ldots, \mathbf{e}_n = \mathbf{e}'_n \tag{1.38}$$

We have:

$$\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3 \ldots, \mathbf{e}'_n\} = \{\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_3, \ldots, \mathbf{e}_n\} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n\} T_{\text{swap}}\,, \tag{1.39}$$

where one can easy see that the determinant for transition matrix T_{swap} is equal to -1, i.e. bases $\{\mathbf{e}_1, \mathbf{e}_2 \ldots, \mathbf{e}_n\}$ and $\{\mathbf{e}_2, \mathbf{e}_1 \ldots, \mathbf{e}_n\}$ have opposite orientation.

E.g. write down the transition matrix (1.43) in the case if dimension of vector space is equal to 5, $n = 5$. Then we have $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3, \mathbf{e}'_4, \mathbf{e}'_5\} = \{\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_4, \mathbf{e}_5\} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4, \mathbf{e}_5\} T$ where

$$T_{\text{swap}} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \qquad (\det T_{\text{swap}} = -1)\,. \tag{1.40}$$

We see that bases $\{\mathbf{e}_1, \mathbf{e}_2 \ldots, \mathbf{e}_n\}$ and $\{\mathbf{e}'_1, \mathbf{e}'_2 \ldots, \mathbf{e}'_n\}$ have opposite orientation.

We see that there are at least two equivalence classes.

One can see that there are exactly two equivalence classes.

Proposition *Let two bases $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and $\{\mathbf{e}'_1, \ldots, \mathbf{e}'_n\}$ in vector space V have opposite orientation. Let $\{\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n\}$ be an arbitrary basis in V. Then the basis $\{\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n\}$ and the first basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ have the same orientation or the basis $\{\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n\}$ and the second basis $\{\mathbf{e}'_1, \ldots, \mathbf{e}'_n\}$ have the same orientation. In other words if $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, $\{\mathbf{e}'_1, \ldots, \mathbf{e}'_n\}$ and $\{\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n\}$ are three bases in vector space V such that $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\} \not\sim \{\mathbf{e}'_1, \ldots, \mathbf{e}'_n\}$ then*

$$\{\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n\} \sim \{\mathbf{e}_1, \ldots, \mathbf{e}_n\} \text{ or } \{\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n\} \sim \{\mathbf{e}'_1, \ldots, \mathbf{e}'_n\}\,. \tag{1.41}$$

There are two equivalence classes of bases with respect to orientation. An arbitrary basis belongs to the equivalence class of the basis $\{\mathbf{e}_1, \mathbf{e}_2 \ldots, \mathbf{e}_n\}$, or *it belongs to the to the equivalence class of the basis* $\{\mathbf{e}'_1, \mathbf{e}_2 \dots, \mathbf{e}'_n\}$ *(in the case if bases* $\{\tilde{\mathbf{e}}'_1, \dots, \tilde{\mathbf{e}}'_n\}$, $\{\tilde{\mathbf{e}}_1, \dots, \tilde{\mathbf{e}}_n\}$ *have opposite orientation).*

Proof of the statement immediately follows from equations (1.36) and (1.37). In the same way like in these equations we call a basis $\{\mathbf{e}_1, \mathbf{e}_2 \dots, \mathbf{e}_n\}$ the "I-st basis", a basis $\{\mathbf{e}'_1, \mathbf{e}'_2 \dots, \mathbf{e}'_n\}$ the "II-nd basis" and a basis $\{\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2 \dots, \tilde{\mathbf{e}}_n\}$ the "III-rd basis". Determinant of transition matrix T_{12} is negative since I-st and II-nd bases have opposite orientation. Then it follows from relation (1.37) that determinants of transition matrices T_{13} and T_{23} have opposite signs. Hence $\det T_{13} > 0$, i.e. I-st and III-rd bases have the same orientation, or $\det T_{23} > 0$,i.e II-nd and III-rd bases have the same orientation. ∎

Example Let $\{\mathbf{e}_1, \mathbf{e}_2 \dots, \mathbf{e}_n\}$ be an arbitrary basis in n-dimensional vector space V. Swap the vectors $\mathbf{e}_1, \mathbf{e}_2$. We come to a new basis: $\{\mathbf{e}'_1, \mathbf{e}'_2 \dots, \mathbf{e}'_n\}$

$$\mathbf{e}'_1 = \mathbf{e}_2, \mathbf{e}'_2 = \mathbf{e}_1, \text{ all other vectors are the same: } \mathbf{e}_3 = \mathbf{e}'_3, \dots, \mathbf{e}_n = \mathbf{e}'_n \tag{1.42}$$

We have:

$$\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3 \dots, \mathbf{e}'_n\} = \{\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_3, \dots, \mathbf{e}_n\} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \dots, \mathbf{e}_n\} T_{\text{swap}} \,, \tag{1.43}$$

where one can easy see that the determinant for transition matrix T_{swap} is equal to -1, i.e. bases $\{\mathbf{e}_1, \mathbf{e}_2 \dots, \mathbf{e}_n\}$ and $\{\mathbf{e}_2, \mathbf{e}_1 \dots, \mathbf{e}_n\}$ have opposite orientation.

E.g. write down the transition matrix (1.43) in the case if dimension of vector space is equal to 5, $n = 5$. Then we have $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3, \mathbf{e}'_4, \mathbf{e}'_5\} = \{\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_4, \mathbf{e}_5\} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4, \mathbf{e}_5\} T$ where

$$T_{\text{swap}} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \qquad (\det T_{\text{swap}} = -1) \,. \tag{1.44}$$

We see that bases $\{\mathbf{e}_1, \mathbf{e}_2 \dots, \mathbf{e}_n\}$ and $\{\mathbf{e}'_1, \mathbf{e}'_2 \dots, \mathbf{e}'_n\}$ have opposite orientation.

E.g. In the example considered above (see swappingbasevectors) an arbitrary basis $\{\mathbf{e}'_1, \dots \mathbf{e}'_n\}$ have the same orientation as the basis $\{\mathbf{e}_1, \mathbf{e}_2 \dots, \mathbf{e}_n\}$, i.e. belongs to the equivalence class of basis $\{\mathbf{e}_1, \mathbf{e}_2 \dots, \mathbf{e}_n\}$, or it has the same orientation as the "swapped" basis $\{\mathbf{e}_2, \mathbf{e}_1 \dots, \mathbf{e}_n\}$, i.e. it belongs to the equivalence class of the "swappedd" basis $\{\mathbf{e}_2, \mathbf{e}_1 \dots, \mathbf{e}_n\}$.

The set of all bases is a union of two disjoint subsets.

Any two bases which belong to the same subset have the same orientation. Any two bases which belong to different subsets have opposite orientation.

Definition *An orientation of a vector space is an equivalence class of bases in this vector space.*

Note that fixing any basis we fix orientation, considering the subset of all bases which have the same orientation that the given basis.

There are two orientations. Every basis has the same orientation as a given basis or orientation opposite to the orientation of the given basis.

If we choose an arbitrary basis then all bases which belong to the equivalence class of this basis may be called "left" bases and all the bases which do not belong to the equivalence class of this basis may be called "right" bases

Definition *An oriented vector space is a vector space equipped with orientation.*

Consider examples.

Example (Orientation in two-dimensional space). Let $\{\mathbf{e}_x, \mathbf{e}_y\}$ be arbitrary two bases in $\mathbf{R}^2$ and let $\mathbf{a}, \mathbf{b}$ be arbitrary two vectors in $\mathbf{R}^2$. Consider an ordered pair $\{\mathbf{a}, \mathbf{b},\}$. The transition matrix from the basis $\{\mathbf{e}_x, \mathbf{e}_y\}$ to the ordered pair $\{\mathbf{a}, \mathbf{b}\}$ is $T = \begin{pmatrix} a_x & b_x \\ a_y & b_y \end{pmatrix}$:

$$\{\mathbf{a}, \mathbf{b}\} = \{\mathbf{e}_x, \mathbf{e}_y\}T = \{\mathbf{e}_x, \mathbf{e}_y\} \begin{pmatrix} a_x & b_x \\ a_y & b_y \end{pmatrix}, \qquad \begin{cases} \mathbf{a} = a_x\mathbf{e}_x + a_y\mathbf{e}_y \\ \mathbf{b} = b_x\mathbf{e}_x + b_y\mathbf{e}_y \end{cases}$$

One can see that the ordered pair $\{\mathbf{a}, \mathbf{b}\}$ also is a basis, (i.e. these two vectors are linearly independent in $\mathbf{R}^2$) if and only if transition matrix is not degenerate, i.e. $\det T \neq 0$. The basis $\{\mathbf{a}, \mathbf{b}\}$ has the same orientation as the basis $\{\mathbf{e}_x, \mathbf{e}_y\}$ if $\det T > 0$ and the basis $\{\mathbf{a}, \mathbf{b}\}$ has the orientation opposite to the orientation of the basis $\{\mathbf{e}_x, \mathbf{e}_y\}$ if $\det T < 0$.

Example Let $\{\mathbf{e}, \mathbf{f}\}$ be a basis in 2-dimensional vector space. Consider bases $\{\mathbf{e}, -\mathbf{f}\}$, $\{\mathbf{f}, -\mathbf{e}\}$ and $\{\mathbf{f}, \mathbf{e}\}$.

1) We come to basis $\{\mathbf{e}, -\mathbf{f}\}$ reflecting the second basis vector. Transition matrix from initial basis $\{\mathbf{e}, \mathbf{f}\}$ to the basis $\{\mathbf{e}, -\mathbf{f}\}$ is $T_{\{\mathbf{e},-\mathbf{f}\}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Its determinant is -1. Bases $\{\mathbf{e}, \mathbf{f}\}$ and $\{\mathbf{e}, -\mathbf{f}\}$ have opposite orientation.

2) Transition matrix from initial basis $\{\mathbf{e}, \mathbf{f}\}$ to the basis $\{\mathbf{f}, -\mathbf{e}\}$ is $T_{\{\mathbf{f},-\mathbf{e}\}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Its determinant is 1. Bases $\{\mathbf{e}, \mathbf{f}\}$ and $\{\mathbf{f}, -\mathbf{e}\}$ have same orientation. We come to basis $\{\mathbf{f}, -\mathbf{e}\}$ rotating the initial basis on the angle $\pi/2$.

3) Transition matrix from initial basis $\{\mathbf{e}, \mathbf{f}\}$ to the basis $\{\mathbf{f}, \mathbf{e}\}$ is $T_{\{\mathbf{f},\mathbf{e}\}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Its determinant is -1. Bases $\{\mathbf{e}, \mathbf{f}\}$ and $\{\mathbf{e}, -\mathbf{f}\}$ have opposite orientation.

We come to basis $\{\mathbf{f}, \mathbf{e}\}$ reflecting the initial basis.

We see that bases $\{\mathbf{e}, \mathbf{f}\}$ and $\{\mathbf{f}, -\mathbf{e}\}$ have the same orientation; i.e. they belong to the same equivalenceclass. Bases $\{\mathbf{e}, -\mathbf{f}\}$ and $\{\mathbf{f}, \mathbf{e}\}$ have the same orientation too, they belong to the another equivalence class. If we say that bases $\{\mathbf{e}, \mathbf{f}\}$ and $\{\mathbf{f}, -\mathbf{e}\}$ are *left* bases then bases $\{\mathbf{e}, -\mathbf{f}\}$ and $\{\mathbf{f}, \mathbf{e}\}$ are *right* bases.

(There are plenty exercises in the Homework 3.)

Example(Orientation in three-dimensional euclidean space.) Let $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ be any basis in $\mathbf{E}^3$ and $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are arbitrary three vectors in $\mathbf{E}^3$:

$$\mathbf{a} = a_x\mathbf{e}_x + a_y\mathbf{e}_y + a_z\mathbf{e}_z \ \mathbf{b} = b_x\mathbf{e}_x + b_y\mathbf{e}_y + b_z\mathbf{e}_z, \ \mathbf{c} = c_x\mathbf{e}_x + c_y\mathbf{e}_y + c_z\mathbf{e}_z \,.$$

Consider ordered triple $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$. The transition matrix from the basis $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ to the ordered triple $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ is $T = \begin{pmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_z \end{pmatrix}$:

$$\{\mathbf{a}, \mathbf{b}, \mathbf{c}\} = \{\mathbf{e_x}, \mathbf{e_y}, \mathbf{e_z}\}\mathbf{T} = \{\mathbf{e_x}, \mathbf{e_y}, \mathbf{e_z}\} \begin{pmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_z \end{pmatrix}$$

One can see that the ordered triple $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ also is a basis, (i.e. these three vectors are linearly independent) if and only if transition matrix is not degenerate $\det T \neq 0$. The basis $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ has the same orientation as the basis $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ if

$$\det T > 0 \,. \tag{1.45}$$

The basis $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ has the orientation opposite to the orientation of the basis $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ if

$$\det T < 0 \,. \tag{1.46}$$

Remark Note that in the example above we considered in $\mathbf{E}^3$ *arbitrary bases* not necessarily orthonormal bases.

Relations (1.45),(1.46) define equivalence relations in the set of bases. Orientation is equivalence class of bases. There are two orientations, every basis has the same orientation as a given basis or opposite orientation.

If two bases $\{\mathbf{e}_i\}$, $\{\mathbf{e}_{i'}\}$ have the same orientation then they can be transformed to each other by continuous transformation, i.e. there exists one-parametric family of bases $\{\mathbf{e}_i(t)\}$ such that $0 \leq t \leq 1$ and $\{\mathbf{e}_i(t)\}|_{t=0} = \{\mathbf{e}_i\}$, $\{\mathbf{e}_i(t)\}|_{t=1} = \{\mathbf{e}_{i'}\}$. (All functions $\mathbf{e}_i(t)$ are continuous) In the case of three-dimensional space the following statement is true : *Let* $\{\mathbf{e}_i\}, \{\mathbf{e}_{i'}\}$ $(i = 1, 2, 3)$ *be two orthonormal bases in* $\mathbf{E}^3$ *which have the same orientation. Then there exists an axis* $\mathbf{n}$ *such that basis* $\{\mathbf{e}_i\}$ *transforms to the basis* $\{\mathbf{e}_{i'}\}$ *under rotation around the axis.*(This is Euler Theorem (see it later).

Exercise Show that bases $\{\mathbf{e}, \mathbf{f}, \mathbf{g}\}$ and $\{\mathbf{f}, \mathbf{e}, \mathbf{g}\}$ have opposite orientation but bases $\{\mathbf{e}, \mathbf{f}, \mathbf{g}\}$ and $\{\mathbf{f}, \mathbf{e}, -\mathbf{g}\}$ have the same orientation.

Solution. Transformation from basis $\{\mathbf{e}, \mathbf{f}, \mathbf{g}\}$ to basis $\{\mathbf{f}, \mathbf{e}, \mathbf{g}\}$ is "swapping" of vectors $((\mathbf{e}, \mathbf{f}) \mapsto (\mathbf{f}, \mathbf{e})$. This is reflection and this transformation changes orientation. One can see it using transition matrix:

$$T\colon \{\mathbf{f}, \mathbf{e}, \mathbf{g}\} = \{\mathbf{e}, \mathbf{f}, \mathbf{g}\}T = \{\mathbf{e}, \mathbf{f}, \mathbf{g}\} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} . \det T = -1$$

Transformation from basis $\{\mathbf{e}, \mathbf{f}, \mathbf{g}\}$ to basis $\{\mathbf{f}, \mathbf{e}, -\mathbf{g}\}$ is composition of two transformations: "swapping" of vectors $((\mathbf{e}, \mathbf{f}) \mapsto (\mathbf{f}, \mathbf{e})$ and changing direction of vector $\mathbf{g}$ $(\mathbf{g} \mapsto -\mathbf{g})$. We have two reflections:

$$\{\mathbf{e}, \mathbf{f}, \mathbf{g}\} \quad \stackrel{\text{reflection}}{\longrightarrow} \quad \{\mathbf{f}, \mathbf{e}, \mathbf{g}\} \quad \stackrel{\text{reflection}}{\longrightarrow} \quad \{\mathbf{f}, \mathbf{e}, -\mathbf{g}\}$$

Any reflection changes orientation. Two reflections preserve orinetation. One may come to this result using transition matrix:

$$T\colon \{\mathbf{f}, \mathbf{e}, -\mathbf{g}\} = \{\mathbf{e}, \mathbf{f}, \mathbf{g}\}T = \{\mathbf{e}, \mathbf{f}, \mathbf{g}\} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} . \det T = 1. \quad \text{Orientation is not changed.} \tag{1.47}$$

(See also exercises in Homework 3)

1.10.2 Orientation of linear operator

. Let P be invertible linear operator, i.e. $\det P \neq 0$.

If a linear operator P acting on the space V has positive determinant then under the action of this operator an arbitrary basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ transforms to the new basis $\{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}$ such that transition matrix from basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ to the new basis $\{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}$ has positive determinant, i.e. these bases have the same orientation. Respectively if a linear operator P acting on the space V has negative determinant then under the action of this operator an arbitrary basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ transforms to the new basis $\{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}$ such that transition matrix from basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ to the new basis $\{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}$ has negative determinant, i.e. these bases have opposite orientation. Thus we can define does the linear operator P acting in the vector space V changes an orientation or it does not change an orientation of this vector space.

Definition. Non-degenerate (invertible) linear operator P ($\det P \neq 0$) acting in vector space V preserves an orientation of the vector space V if $\det P > 0$. It changes the orientation if $\det P < 0$.

If $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is an arbitrary basis which transforms to the new basis $\{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}$ under the action of nvertible operator P: $\mathbf{e}'_i = P(\mathbf{e}_i)$ then these bases have the same orientation if and only if operator P preserves an orientation, i.e. $\det P > 0$, and these bases have opposite orientation if and only if the operator P changes an orientation, i.e. $\det P < 0$.

1.11 Rotations and orthogonal operators preserving orientation of $\mathbf{E}^n$ (n=2,3)

Recall the notion of orthogonal operator (see 1.9.3). We study here orthogonal operators in $\mathbf{E}^2$ and $\mathbf{E}^3$. In particular we will show that orthogonal oeprators preserving orientations define rotations.

1.11.1 Orthogonal operators in $\mathbf{E}^2$— Rotations and reflections

We show that an orthogonal operator in $\mathbf{E}^2$ 'rotates the space' or makes a 'reflection'.

Let A be an orthogonal operator acting in Euclidean space $\mathbf{E}^2$: $(A\mathbf{x}, A\mathbf{y}) = (\mathbf{x}, \mathbf{y})$. Let $\{\mathbf{e}, \mathbf{f}\}$ be an orthonormal basis in 2-dimensional Euclidean space $\mathbf{E}^2$: $(\mathbf{e}, \mathbf{e}) = (\mathbf{f}, \mathbf{f}) = 1$ (i.e. $|\mathbf{e}| = |\mathbf{f}| = 1$) and $(\mathbf{e}, \mathbf{f}) = 0$–vectors $\mathbf{e}, \mathbf{f}$ have unit length and are orthogonal to each other.

Consider a new basis $\{\mathbf{e}', \mathbf{f}'\}$, an image of basis $\mathbf{e}, \mathbf{f}$ under action of A: $\mathbf{e}' = A(\mathbf{e})$, $\mathbf{f}' = A(\mathbf{f})$. Let $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ be matrix of operator A in the basis $\mathbf{e}, \mathbf{f}$, (see equation (1.23) and defintion after this equation):

$$\{\mathbf{e}', \mathbf{f}'\} = \{\mathbf{e}, \mathbf{f}\}A = \{\mathbf{e}, \mathbf{f}\} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \text{ i.e. } \mathbf{e}' = \alpha\mathbf{e} + \gamma\mathbf{f},\ \mathbf{f}' = \beta\mathbf{e} + \delta\mathbf{f}$$

New basis is orthonormal basis also, $(\mathbf{e}', \mathbf{e}') = (\mathbf{f}', \mathbf{f}') = 1\,, \quad (\mathbf{e}', \mathbf{f}') = 0$.

Operator A is orthogonal operator, and its matrix is orthogonal matrix:

$$A^T A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}^t \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} \alpha^2 + \gamma^2 & \alpha\beta + \gamma\delta \\ \alpha\beta + \gamma\delta & \beta^2 + \delta^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{1.48}$$

Remark With some abuse of notation, (if it is not a reason of confusion) we sometimes use the same letter for linear operator and the matrix of this operator in orthonormal basis.

We have $\alpha^2 + \gamma^2 = 1$, $\alpha\beta + \gamma\delta = 0$ and $\beta^2 + \delta^2 = 1$.

Hence one can choose angles $\varphi, \psi\colon 0 \leq 2\pi$ such that $\alpha = \cos\varphi$, $\gamma = \sin\varphi$, $\beta = \sin\psi$, $\delta = \cos\psi$. The condition $\alpha\beta + \gamma\delta =$ means that

$$\cos\varphi \sin\psi + \sin\varphi \cos\psi = \sin(\varphi + \psi) = 0$$

Hence $\sin\varphi = -\sin\psi$, $\cos\varphi = \cos\psi$ $(\varphi + \psi = 0)$ or $\sin\varphi = \sin\psi$, $\cos\varphi = -\cos\psi$ $(\varphi + \psi = \pi)$

Note that condition (1.48) implies that $\det A = \pm 1$.

The first case: $\det A = 1$, operator A preserves orientation;

$\sin\varphi = -\sin\psi$, $\cos\varphi = \cos\psi$,

$$A_\varphi = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \ (\det A_\varphi = 1)\,. \tag{1.49}$$

The second case: $\det A = -1$, operator A changes orientation;

$\sin\varphi = \sin\psi$, $\cos\varphi = -\cos\psi$,

$$\tilde{A}_\varphi = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} \ (\det \tilde{A}_\varphi = -1) \tag{1.50}$$

In the first case matrix of operator A_φ is defined by the relation (1.49). In this case the new basis is:

$$(\mathbf{e}', \mathbf{f}') = (\mathbf{e}, \mathbf{f}) A_\varphi = (\mathbf{e}, \mathbf{f}) \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}, \quad \begin{array}{r} \mathbf{e}' = A_\varphi(\mathbf{e}) = \cos\varphi\,\mathbf{e} + \sin\varphi\,\mathbf{f} \\ \mathbf{f}' = A_\varphi(\mathbf{f}) - \sin\varphi\,\mathbf{e} + \cos\varphi\,\mathbf{f} \end{array} \tag{1.51}$$

For an arbitrary vector $\mathbf{x} = x\mathbf{e} + y\mathbf{f}$ $\mathbf{x} \to A_\varphi(\mathbf{x}) = A_\varphi(x\mathbf{e} + y\mathbf{f}) = x'\mathbf{e} + y'\mathbf{f}$,

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\varphi - y\sin\varphi \\ \sin\varphi + y\cos\varphi \end{pmatrix}. \tag{1.52}$$

Operator A_φ rotates basis vectors $\mathbf{e}, \mathbf{f}$ and arbitrary vector $\mathbf{x}$ on an angle φ

In the second case a matrix of operator $\tilde{A}_\varphi$ is defined by the relation (1.50). One can see that

$$\tilde{A}_\varphi = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = A_\varphi R \tag{1.53}$$

where we denote by $R = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ a transition matrix from the basis $\{\mathbf{e}, \mathbf{f}\}$ to the basis $\{\mathbf{e}, -\mathbf{f}\}$—"reflection"l.

We see that in the second case the orthogonal operator $\tilde{A}_\varphi$ is composition of rotation and reflection: $\{\mathbf{e}, \mathbf{f}\} \xrightarrow{\tilde{A}_\varphi = A_\varphi R} \{\tilde{\mathbf{e}}, \tilde{\mathbf{f}}\}$:

$$\{\mathbf{e}, \mathbf{f}\} \xrightarrow{A_\varphi} \{\mathbf{e}' = \cos\varphi\,\mathbf{e} + \sin\varphi\mathbf{l}, \mathbf{f}, \mathbf{f}' = -\sin\varphi\,\mathbf{e} + \cos\varphi\,\mathbf{f}\} \xrightarrow{R} \{\tilde{\mathbf{e}} = \mathbf{e}', \tilde{f} = -\mathbf{f}\} \tag{1.54}$$

We come to proposition

Proposition. *Let A be an arbitrary 2×2 orthogonal linear transformation, $A^T A = 1$, and in particularly* $\det A = \pm 1$. *(As usual we consider matrix of orthogonal operator in the orthonormal basis.)*

If $\det A = 1$ *then there exists an angle $\varphi \in [0, 2\pi)$ such that $A = A_\varphi$ is an operator which rotates basis vectors and any vector* (1.49) *on the angle φ.*

If $\det A = -1$ *then there exists an angle $\varphi \in [0, 2\pi)$ such that $A = \tilde{A}_\varphi$ is a composition of rotation and reflection (see* (1.54)*).*

Remark One can show that orthogonal operator $\tilde{A}_\varphi$ is a reflection with respect to the axis which have the angle $\frac{\varphi}{2}$ with x-axis.

Consider just examples:

$$a)\varphi = 0,\ \tilde{A}_\varphi = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} \mathbf{e} \\ \mathbf{f} \end{pmatrix} \mapsto \begin{pmatrix} \mathbf{e} \\ -\mathbf{f} \end{pmatrix}$$

(reflection with respect to x-axis)

$$b)\varphi = \pi,\ \tilde{A}_\varphi = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} \mathbf{e} \\ \mathbf{f} \end{pmatrix} \mapsto \begin{pmatrix} -\mathbf{e} \\ \mathbf{f} \end{pmatrix}$$

(reflection with respect to y-axis)

$$b)\varphi = \frac{\pi}{2},\ \tilde{A}_\varphi = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} \mathbf{e} \\ \mathbf{f} \end{pmatrix} \mapsto \begin{pmatrix} \mathbf{f} \\ \mathbf{e} \end{pmatrix}$$

(reflection with respect to axis $y = x$ ("swapping" of basis vectors))

Try to do it in general case.

1.11.2 Orthogonal operators in $\mathbf{E}^3$ and rotations

We see in the previous paragraph that orthogonal operator preserving orientation of $\mathbf{E}^2$ is rotation operator. The same is true in $\mathbf{E}^2$. The main result of this paragraph will be the Euler Theorem about rotation, that every orthogonal operator preserving orientation in $\mathbf{E}^3$ is rotation around some axis.

We will give an exact formulation of the Euler Theorem at the end of this paragraph. Now we will formualte just preliminary statement:

The Euler Theorem. (*Preliminary statement*) An orthogonal operator in $\mathbf{E}^3$ preserving orientation is rotation operator with respect to an axis l on the angle φ. The axis is directed along eigenvector $\mathbf{N}$ of the operator P, $P(\mathbf{N}) = \mathbf{N}$,and angle of rotation is defined by equation

$$\operatorname{Tr} P = 1 + 2\cos\varphi\,.$$

We will come to this statement gradually step by step, and then will formulate it completely.

Let $\mathbf{E}^n$ be oriented vector space. Recall that oriented vector space means that it is chosen the equivalence class of bases: all bases in this class have the same orientation. We call all bases in the equivalence class defining orientation "left" bases. All "left" bases have the same orientation. To define an orientation in vector space V one may consider an arbitrary basis $\{\mathbf{e}_i^{(0)}\}$ in V and claim that this basis is "left" basis. The basis $\{\mathbf{e}_i^{(0)}\}$ defines equivalence class of "left" bases: all bases $\{\mathbf{e}_i\}$ such that $\{\mathbf{e}_i\} \sim \{\mathbf{e}_i^{(0)}$ will be called "left" bases. We can say that basis $\{\mathbf{e}_i^{(0)}\}$ defines the orientation.

Later on considering oriented vector space we often call all bases defining the orientation (i.e. belonging to the equivalence class of bases defining orientation) "left" bases.

Now we define rotation in $\mathbf{E}^3$. Recall the definition of rotation in $\mathbf{E}^2$ (see 1.11.1):

Definition Let $\mathbf{E}^2$ be an oriented Euclidean space. We say that linear operator P rotates this space on an angle "φ" if for a given "left" orthonormal basis $\{\mathbf{e}, \mathbf{f}\}$

$$\begin{cases} \mathbf{e}' = P(\mathbf{e}) = \mathbf{e}\cos\varphi + \mathbf{f}\sin\varphi \\ \mathbf{f}' = P(\mathbf{f}) = -\mathbf{e}\sin\varphi + \mathbf{f}\cos\varphi \end{cases} \qquad \text{i.e.} \quad \{\mathbf{e}', \mathbf{f}'\} = \{\mathbf{e}, \mathbf{f}\}\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \tag{1.55}$$

i.e. transition matrix from basis $\{\mathbf{e}, \mathbf{f}\}$ to new basis $\{\mathbf{e}' = P(\mathbf{e}), \mathbf{f}' = P(\mathbf{f})\}$ is the rotation matrix (1.49) (see also (1.51)).

Remark One can show that the angle of rotation does not depend on the choice of "left" basis. If we will choose another left basis $\tilde{\mathbf{e}}, \tilde{\mathbf{f}}$ then the angle remains the same

Operator P rotates every vector rotates on the angle φ.

If we choose a basis with opposite orientation ("right" basis) then the angle will change: $\varphi \mapsto -\varphi$.

We already did it in 1.11.1 and we also see from formula (1.55) that the matrix of operator P is orthogonal matrix such that its determinant equals 1. In 2-dimensional case we came to simple Proposition (see Proposition in 1.11.1) which we will repeat again[4]:

Proposition Let P be an orthogonal operator in oriented 2-dimensional Euclidean space. If operator P preserves orientation ($\det P = 1$) then it is a rotation operator (1.55) on some angle φ.

The situation is little bit more tricky in 3-dimensional case.

Let $\mathbf{E}^3$ be an Euclidean vector space. (Problem of orientation we will discuss below.) Let $\mathbf{N} \neq 0$ be an arbitrary non-zero vector in $\mathbf{E}^3$. Consider the line $l_{\mathbf{N}}$, spanned by vector $\mathbf{N}$. This is *axis* directed along the vector $\mathbf{N}$. Choose a unit vector

$$\mathbf{n} = \pm\frac{\mathbf{N}}{|\mathbf{N}|} \tag{1.56}$$

Vector $\mathbf{n}$ fixes an orientation on $l_{\mathbf{N}}$. Changing $\mathbf{n} \mapsto -\mathbf{n}$ changes an orientation on opposite).

[4]Just here we denote the operator by letter 'P' instead letter 'A'

Choose an arbitrary orthonormal basis such that first vector of this basis is directed along the axis: a basis $\{\mathbf{n}, \mathbf{f}, \mathbf{g}\}$.

Definition We say that a linear operator P rotates the Euclidean space $\mathbf{E}^3$ on the angle φ with respect to an axis $l_{\mathbf{N}}$ directed along a vector $\mathbf{N}$ if the following conditions are satisfied:

- $$P(\mathbf{N}) = \mathbf{N}$$

  vector $\mathbf{N}$ (and all vectors proportional to this vector) are eigenvectors of operator P with eigenvalue 1, i.e. axis remain intact

- for an orthonormal basis $\{\mathbf{n}, \mathbf{f}, \mathbf{g}\}$ such that the first vector of this basis is equal to $\mathbf{n}$, ($\mathbf{n}$ is a unit vector, proportional to $\mathbf{N}$)

  $$\begin{cases} \mathbf{f}' = P(\mathbf{f}) = \mathbf{f}\cos\varphi + \mathbf{g}\sin\varphi \\ \mathbf{g}' = P(\mathbf{f}) = -\mathbf{f}\sin\varphi + \mathbf{g}\cos\varphi \end{cases} \quad \text{i.e.} \quad \{\mathbf{f}', \mathbf{g}'\} = \{\mathbf{f}, \mathbf{g}\} \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}. \tag{1.57}$$

  In other words plane (subspace) orthogonal to axis rotates on the angle φ: linear operator P rotates every vector orthogonal to axis on the angle φ in the plane (subspace) orthogonal to the axis.

Linear operator P transforms the basis $\{\mathbf{n}, , \mathbf{f}, \mathbf{g}\}$ to the new basis $\{\mathbf{n}, \mathbf{f}', \mathbf{g}'\}$ $= \{\mathbf{n}, \mathbf{f}\cos\varphi + \mathbf{g}\sin\varphi, -\mathbf{f}\sin\varphi + \mathbf{g}\cos\varphi\}$. The matrix of operator P, i.e. the transition matrix from the basis $\{\mathbf{n}, , \mathbf{f}, \mathbf{g}\}$ to the basis $\{\mathbf{n}, \mathbf{f}', \mathbf{g}'\}$ is defined by the relation:

$$\{\mathbf{n}, \mathbf{f}', \mathbf{g}'\} = \{\mathbf{n}, \mathbf{f}\cos\varphi + \mathbf{g}\sin\varphi, -\mathbf{f}\sin\varphi + \mathbf{g}\cos\varphi\} = \{\mathbf{n}, , \mathbf{f}, \mathbf{g}\} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix} \tag{1.58}$$

Recalling definition (1.30) of trace of linear operator we come to the following relation

$$TrP = 1 + 2\cos\varphi \tag{1.59}$$

where φ is angle of rotation. Note that Trace of the operator does not depend on the choice of the basis. This formula express cosine of the angle of rotation in terms of operator, irrelevant of the choice of the basis.

Remark This formula defines angle of rotation up to a sign.

If we change orientation then $\varphi \mapsto -\varphi$. For non-oriented Euclidean space rotation is defined up to a sign[5]

Careful reader maybe already noted that even fixing the orientation of $\mathbf{E}^3$ does not fix the "sign" of the angle: If we change the orientation of the axis (changing $\mathbf{n} \mapsto -\mathbf{n}$) then changing the corresponding "left" basis will imply that $\varphi \mapsto -\varphi$. In fact angle φ is the angle of rotation of oriented plane which is orthogonal to the axis of rotation. Orientation on the plane is defined by orientation in $\mathbf{E}^3$ and orientation of the axis which is orthogonal to this plane. In the case of 3-dimensional space sign of the angle depends not only on orientation of $\mathbf{E}^3$ but on orientation of axis. In what follows we will ignore this. This means that we define rotation on the angle $\pm\varphi$ up to a sign.... Rotation is defined for operators preserving orientation. The difference between angles of rotations φ and $-\varphi$ is depending not only on orientation of $\mathbf{E}^3$ but on orientation of axis too. But we ignore this difference. Note that $\cos\varphi$ in the formula is defined up to a sign

Rotation operator eviently is orthogonal operator preserving orientation. Is it true converse implication? We are ready to formulate the following remarkable result.

Theorem (the Euler Theorem) *Let P be an orthogonal operator preserving an orientation of Euclidean space $\mathbf{E}^3$, i.e. operator P preserves the scalar product and orientation. Then it is a rotation operator with respect to an axis l on the angle φ. Every vector $\mathbf{N}$ directed along the axis does not change, i.e. the axis is 1-dimensional space of eigenvectors with eigenvalue 1, $P(\mathbf{N}) = \mathbf{N}$. Every vector orthogonal to axis rotates on the angle φ in the plane orthogonal to the axis,*

$$\mathrm{Tr}\, P = 1 + 2\cos\varphi \,.$$

The angle φ is defined up to a sign. Changing orientation of the Euclidean space and of the axis change sign of φ.

This Theorem can be restated in the following way: every orthogonal operator P preserving orientation, ($\det P \neq 0$) has an eigenvector $\mathbf{N} \neq 0$ with eigenvalue 1. This eigenvector defines the axis of rotation. In an orthonormal basis $\{\mathbf{n}, \mathbf{f}, \mathbf{g}\}$ where $\mathbf{n}$ is a unit vector along the axis, the transition matrix of operator has an appearance (1.58). Angle of rotaion can be defined via Trace of operator by formula $\mathrm{Tr}\, P = 1 + 2\cos\varphi$.

Remark If P is an identity operator, $P = I$ then " there is no rotation", more precisely: any line can be considered as an axis of rotation (every vector is eigenvector of identity matrix with eigenvalue 1) and angle of rotation is equal to zero. If $P \neq I$ then axis of rotation is defiend uniquely.

Proof of the Euler Theorem. The proof of the Euler Theorem has two parts. First and central part is to prove the existence of the axis. The rest is trivial: we take an arbitrary

[5]Does it recall you expressions such as "clockwise", "anticlock-wise" rotation?

orthonormal basis $\mathbf{n}, \mathbf{f}, \mathbf{g}$ such that $\mathbf{n}$ is eigenvector and we come to relation (1.57). We expose here maybe the most beautiful proof which belongs to Coxeter.

Let P be linear orthogonal operator preserving orientation. Note that for any two not-zero distinct vectors $\mathbf{e}, \mathbf{f}$ one can consider orthogonal operator $R_{\mathbf{e},\mathbf{f}}$ which changes orientation and swaps the vectors $\mathbf{e}, \mathbf{f}$: it is reflection with respect to the plane spanned by the vectors $\mathbf{e} + \mathbf{f}$ and a vector $\mathbf{e} \times \mathbf{f}$.

Let $\{\mathbf{e}, \mathbf{f}, \mathbf{g}\}$ be an arbitrary orthonormal basis in $\mathbf{E}^3$ and let $\mathbf{e}', \mathbf{f}', \mathbf{g}'$ be image of this basis under operator P

$$P(\mathbf{e}) = \mathbf{e}', \; P(\mathbf{f}) = \mathbf{f}' \; P(\mathbf{g}) = \mathbf{g}' .$$

If $\mathbf{e} = \mathbf{e}'$ nothing to prove ($\mathbf{e}$ is eigenvector with eigenvalue 1). If this is not the case, apply reflection operator $R_{\mathbf{e},\mathbf{e}'}$ to the initial basis $\{\mathbf{e}, \mathbf{f}, \mathbf{g}\}$ we come to the orthonormal basis $\{\mathbf{e}', \tilde{f}, \tilde{g}\}$, Then applying reflection operator $R_{\tilde{f},\mathbf{f}'}$ to this basis we come to the basis $\mathbf{e}', \mathbf{f}', \tilde{\tilde{g}}$. The third vector has no choice it has to be equal to $\mathbf{g}'$ since in the case if it is equal to $-\mathbf{g}'$ orientation is opposite. Hence we see that operator P is the product of two reflections operators. Consider the line l, intersection of these planes, we come to eigenvectors with eigenvalue 1. ∎

There are many other proofs, for example:

Another proof: Any non-degenerate 3×3 matrix has at least one eigenvector $\mathbf{x}$: $P\mathbf{x} = \lambda\mathbf{x}$, since cubic equation $\det(P - \lambda I) = 0$ has at lest one real root. Since P is orthogonal operator, then $\lambda = \pm 1$. If $\lambda = 1$, then $\mathbf{x}$ defines the axis. If $\lambda = -1$, $P\mathbf{x} = -\mathbf{x}$, then eigenvector with eigenvalue 1 belongs to the plane orthogonal to $\mathbf{x}$. ∎

Example Consider linear operator P such that for orthonormal basis $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$

$$P(\mathbf{e}_x) = \mathbf{e}_y, P(\mathbf{e}_y) = \mathbf{e}_x, P(\mathbf{e}_z) = -\mathbf{e}_z \tag{1.60}$$

This is obviously orthogonal operator since it transforms orthogonal basis to orthogonal one. This operator swaps first two vectors and reflects the third one. It preserves orientation: matrix of operator in the basis $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$, i.e. the transition matrix from the basis $\{\mathbf{e}_x, , \mathbf{e}_y, \mathbf{e}_z\}$ to the basis $\{P(\mathbf{e}_x), P(\mathbf{e}_y), P(\mathbf{e}_z)\}$ is defined by the relation:

$$\{P(\mathbf{e}_x), P(\mathbf{e}_y), P(\mathbf{e}_z)\} = \{\mathbf{e}_y, \mathbf{e}_x, -\mathbf{e}_z\} = \{\mathbf{e}_x, , \mathbf{e}_y, \mathbf{e}_z\} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$\det P = 1$. This operator preserves orientation. Hence by Euler Theorem it is a rotation. Find first axis of rotation. It is easy to see from (1.60) that $\mathbf{N} = \lambda(\mathbf{e}_x + \mathbf{e}_y)$ is eigenvector with eigenvalue 1:

$$P(\mathbf{N}) = P(\mathbf{e}_x + \mathbf{e}_y) = \mathbf{e}_y + \mathbf{e}_x = \mathbf{N} .$$

Hence axis of rotation is directed along the vector $\mathbf{e}_x + \mathbf{e}_y$. $\mathrm{Tr}\, P = 1 + 2\cos\varphi = -1$. The angle of rotation $\varphi = \pi$.

One can calculate explicitly angle of rotation: Consider orthonormal basis $\{\mathbf{n},\, f, \mathbf{g}\}$ adjusted to the axis ($\mathbf{n}||\mathbf{N}$). We have that $\mathbf{n} = \frac{\mathbf{e}_x+\mathbf{e}_y}{\sqrt{2}}$ since $\mathbf{n}$ is proportional to $\mathbf{N}$ and it is unit vector. Choose $\mathbf{f} = \frac{-\mathbf{e}_x+\mathbf{e}_y}{\sqrt{2}}$ and $\mathbf{g} = \mathbf{e}_z$. Then it is easy to see that

$$\{\mathbf{n}, \mathbf{f}, \mathbf{g}\} = \left\{\frac{\mathbf{e}_x + \mathbf{e}_y}{\sqrt{2}}, \frac{-\mathbf{e}_x + \mathbf{e}_y}{\sqrt{2}}, \mathbf{g}\right\}$$

is orthonormal basis.Using (1.60)one can see that

$$P(\mathbf{n}) = P\left(\frac{\mathbf{e}_x + \mathbf{e}_y}{\sqrt{2}}\right) = \frac{\mathbf{e}_y + \mathbf{e}_x}{\sqrt{2}} = \mathbf{n}\,,$$

$$P(\mathbf{f}) = P\left(\frac{-\mathbf{e}_x + \mathbf{e}_y}{\sqrt{2}}\right) = \frac{-\mathbf{e}_y + \mathbf{e}_x}{\sqrt{2}} = -\mathbf{f}, \quad P(\mathbf{g}) = -\mathbf{g}$$

We see that

$$\{\mathbf{n}, \mathbf{f}, \mathbf{g}\} \xrightarrow{P} \{\mathbf{n}, -\mathbf{f}, -\mathbf{g}\}\,.$$

Comparing with (1.57) and (1.58) we see that the operator P is rotation of $\mathbf{E}^3$ on the angle π with respect to the axis directed along the vector $\mathbf{e}_x + \mathbf{e}_y$.

1.12 Vector product in oriented $\mathbf{E}^3$

Now we give a definition of vector product of vectors in 3-dimensional Euclidean space equipped with orientation.

Let $\mathbf{E}^3$ be three-dimensional oriented Euclidean space, i.e. Euclidean space equipped with an equivalence class of bases with the same orientation. To define the orientation it suffices to consider just one orthonormal basis $\{\mathbf{e}, \mathbf{f}, \mathbf{g}\}$ which is claimed to be left basis. Then the equivalence class of the left bases is a set of all bases which have the same orientation as the basis $\{\mathbf{e}, \mathbf{f}, \mathbf{g}\}$.

Definition Vector product $L(\mathbf{x}, \mathbf{y}) = \mathbf{x} \times \mathbf{y}$ is a function of two vectors which takes vector values such that the following axioms (conditions) hold

- The vector $L(\mathbf{x}, \mathbf{y}) = \mathbf{x} \times \mathbf{y}$ is orthogonal to vector $\mathbf{x}$ and vector $\mathbf{y}$:

  $$(\mathbf{x} \times \mathbf{y}) \perp \mathbf{x}\,, \quad (\mathbf{x} \times \mathbf{y}) \perp \mathbf{y} \qquad (1.61)$$

  In particular it is orthogonal to the the plane spanned by the vectors $\mathbf{x}, \mathbf{y}$ (in the case if vectors $\mathbf{x}, \mathbf{y}$ are linearly independent)

- $$\mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x}, \qquad \text{(anticommutativity condition)} \qquad (1.62)$$

- $$(\lambda\mathbf{x}+\mu\mathbf{y})\times\mathbf{z}=\lambda(\mathbf{x}\times\mathbf{z})+\mu(\mathbf{y}\times\mathbf{z})\,,\qquad \text{(linearity condition)} \quad (1.63)$$

- If vectors $\mathbf{x},\mathbf{y}$ are perpendicular each other then the magnitude of the vector $\mathbf{x}\times\mathbf{y}$ is equal to the area of the rectangle formed by the vectors $\mathbf{x}$ and $\mathbf{y}$:

$$|\mathbf{x}\times\mathbf{y}|=|\mathbf{x}|\cdot|\,\mathbf{y}|\,,\qquad \text{if } \mathbf{x}\perp\mathbf{y}\,, i.e.(\mathbf{x},\mathbf{y})=0\,. \quad (1.64)$$

- If the ordered triple of the vectors $\{\mathbf{x},\mathbf{y},\mathbf{z}\}$, where $\mathbf{z}=\mathbf{x}\times\mathbf{y}$ is a basis, then this basis and an orthonormal basis $\{\mathbf{e},\mathbf{f},\mathbf{g}\}$ defining orientation of $\mathbf{E}^3$ have the same orientation:

$$\{\mathbf{x},\mathbf{y},\mathbf{z}\}=\{\mathbf{e},\mathbf{f},\mathbf{g}\}T,\ \text{where for transition matrix } T,\ \det T>0. \quad (1.65)$$

Vector product depends on orientation in Euclidean space.

Comments on conditions (axioms) (1.61)—(1.65):

1. The condition (1.63) of linearity of vector product with respect to the first argument and the condition (1.62) of anticommutativity imply that vector product is an operation which is linear with respect to the second argument too. Show it:

$$\mathbf{z}\times(\lambda\mathbf{x}+\mu\mathbf{y})=-(\lambda\mathbf{x}+\mu\mathbf{y})\times\mathbf{z}=-\lambda(\mathbf{x}\times\mathbf{z})-\mu(\mathbf{y}\times\mathbf{z})=\lambda(\mathbf{z}\times\mathbf{x})+\mu(\mathbf{z}\times\mathbf{y})\,.$$

Hence vector product is bilinear operation. Comparing with scalar product we see that vector product is bilinear anticommutative (antisymmetric) operation which takes vector values, while scalar product is bilinear symmetric operation which takes real values.

2. The condition of anticommutativity immediately implies that vector product of two colinear (proportional) vectors $\mathbf{x},\mathbf{y}$ ($\mathbf{y}=\lambda\mathbf{x}$) is equal to zero. It follows from linearity and anticommuativity conditions. Show it: Indeed

$$\mathbf{x}\times\mathbf{y}=\mathbf{x}\times(\lambda\mathbf{x})=\lambda(\mathbf{x}\times\mathbf{x})=-\lambda(\mathbf{x}\times\mathbf{x})=-\mathbf{x}\times(\lambda\mathbf{x})=-\mathbf{x}\times\mathbf{y}. \quad (1.66)$$

Hence $\mathbf{x}\times\mathbf{y}=0$, if $\mathbf{y}=\lambda\mathbf{x}$ ■.

3. It is very important to emphasize again that vector product depends on orientation. According the condition (1.65) if $\mathbf{z}=\mathbf{x}\times\mathbf{y}$ and we change

the orientation of Euclidean space, then $\mathbf{z} \to -\mathbf{z}$ since the basis $\{\mathbf{x}, \mathbf{y}, -\mathbf{z}\}$ as an orientation opposite to the orientation of the basis $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$.

You may ask a question: Does this operation (taking the vector product) which obeys all the conditions (axioms) (1.61)—(1.65) exist? And if it exists is it unique? We will show that the vector product is well-defined by the axioms (1.61)—(1.65), i.e. there exists an operation $\mathbf{x} \times y$ which obeys the axioms (1.61)—(1.65) and these axioms define the operation uniquely.

We will assume first that there exists an operation $L(\mathbf{x}, \mathbf{y}) = \mathbf{x} \times \mathbf{y}$ which obeys all the axioms (1.61)—(1.65). Under this assumption we will construct explicitly this operation (if it exists!). We will see that the operation that we constructed indeed obeys all the axioms (1.61)—(1.65).

Let $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ be an *arbitrary* left orthonormal basis of oriented Euclidean space $\mathbf{E}^3$, i.e. a basis which belongs to the equivalence class of the basis $\{\mathbf{e}, \mathbf{f}, \mathbf{g}\}$ defining orientation of $\mathbf{E}^3$. Then it follows from the considerations above for vector product that

$$\begin{array}{ccc} \mathbf{e}_x \times \mathbf{e}_x = 0, & \mathbf{e}_x \times \mathbf{e}_y = \mathbf{e}_z, & \mathbf{e}_x \times \mathbf{e}_z = -\mathbf{e}_y \\ \mathbf{e}_y \times \mathbf{e}_x = -\mathbf{e}_z, & \mathbf{e}_y \times \mathbf{e}_y = 0, & \mathbf{e}_y \times \mathbf{e}_z = \mathbf{e}_x \\ \mathbf{e}_z \times \mathbf{e}_x = \mathbf{e}_y, & \mathbf{e}_z \times \mathbf{e}_y = -\mathbf{e}_x, & \mathbf{e}_z \times \mathbf{e}_z = 0 \end{array} \tag{1.67}$$

E.g. $\mathbf{e}_x \times \mathbf{e}_x = 0$, because of (1.62), $\mathbf{e}_x \times \mathbf{e}_y$ is equal to $\mathbf{e}_z$ or to $-\mathbf{e}_z$ according to (1.64), and according to orientation arguments (1.65) $\mathbf{e}_x \times \mathbf{e}_y = \mathbf{e}_z$.

Now it follows from linearity and (1.67) that for two arbitrary vectors $\mathbf{a} = a_x\mathbf{e}_x + a_y\mathbf{e}_y + a_z\mathbf{e}_z$, $\mathbf{b} = b_x\mathbf{e}_x + b_y\mathbf{e}_y + b_z\mathbf{e}_z$

$$\mathbf{a}\times\mathbf{b} = (a_x\mathbf{e}_x + a_y\mathbf{e}_y + a_z\mathbf{e}_z)\times(b_x\mathbf{e}_x + b_y\mathbf{e}_y + b_z\mathbf{e}_z) = a_xb_y\mathbf{e}_x\times\mathbf{e}_y + a_xb_z\mathbf{e}_x\times\mathbf{e}_z +$$

$$a_yb_x\mathbf{e}_y \times \mathbf{e}_x + a_yb_z\mathbf{e}_y \times \mathbf{e}_z + a_zb_x\mathbf{e}_z \times \mathbf{e}_x + a_zb_y\mathbf{e}_z \times \mathbf{e}_y =$$
$$(a_yb_z - a_zb_y)\mathbf{e}_x + (a_zb_x - a_xb_z)\mathbf{e}_y + (a_xb_y - a_yb_x)\mathbf{e}_z\,. \tag{1.68}$$

It is convenient to represent this formula in the following very familiar way:

$$L(\mathbf{a}, \mathbf{b}) = \mathbf{a} \times \mathbf{b} = \det \begin{pmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{pmatrix} \tag{1.69}$$

We see that the operation $L(\mathbf{x}, \mathbf{y}) = \mathbf{x} \times \mathbf{y}$ which obeys all the axioms (1.61)—(1.65), if it exists, has an appearance (1.69), where $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ is an arbitrary orthonormal basis (with rightly chosen orientation). On the other hand using the properties of determinant and the fact that vectors are orthogonal if and only if their scalar product equals to zero one can easy see that the vector product defined by this formula indeed obeys all the conditions (1.61)—(1.65).

Thus we proved that the vector product is well-defined by the axioms (1.61)—(1.65) and it is given by the formula (1.69) in an arbitrary orthonormal basis (with rightly chosen orientation).

Remark In the formula above we have chosen an arbitrary orthonormal basis which belongs to the equivalence class of bases defining the orientation. What will happen if we choose instead the basis $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ an arbitrary orthonormal basis $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$. We see that such that answer does not change if both bases $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ and $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ have the same orientation, Formulae (1.67) are valid for an arbitrary orthonormal basis which have the same orientation as the orthonormal basis $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$.— In oriented Euclidean space $\mathbf{E}^3$ we may take an arbitrary basis from the equivalence class of bases defining orientation. On the other hand if we will consider the basis with opposite orientation then according to the axiom (1.65) vector product will change the sign. (See also the question 6 in Homework 4)

1.12.1 Vector product—area of parallelogram

The following Proposition states that vector product can be considered as area of parallelogram:

Proposition 2 *The modulus of the vector* $\mathbf{z} = \mathbf{x} \times \mathbf{y}$ *is equal to the area of parallelogram formed by the vectors* $\mathbf{x}$ *and* $\mathbf{y}$.:

$$S(\mathbf{x}, \mathbf{y}) = S(\Pi(\mathbf{x}, \mathbf{y})) = |\mathbf{x} \times \mathbf{y}| , \tag{1.70}$$

where we denote by $S(\mathbf{x}, \mathbf{y})$ the area of parallelogram $\Pi(\mathbf{x}, \mathbf{y})$ formed by the vectors $\mathbf{x}, \mathbf{y}$.

Proof: Consider the expansion $\mathbf{y} = \mathbf{y}_{||} + \mathbf{y}_{\perp}$, where the vector $\mathbf{y}_{\perp}$ is orthogonal to the vector $\mathbf{x}$ and the vector $\mathbf{y}_{||}$ is parallel to to vector $\mathbf{x}$. The area of the parallelogram formed by vectors $\mathbf{x}$ and $\mathbf{y}$ is equal to the product of the length of of the vector $\mathbf{x}$ on the height. The height is equal to the length of the vector $\mathbf{y}_{\perp}$. We have $S(\mathbf{x}, \mathbf{y}) = |\mathbf{x}||\mathbf{y}_{\perp}|$. On the other $\mathbf{z} = \mathbf{x} \times \mathbf{y} = \mathbf{x} \times (\mathbf{y}_{||} + \mathbf{y}_{\perp}) = \mathbf{x} \times \mathbf{y}_{||} + \mathbf{x} \times \mathbf{y}_{\perp}$. But $\mathbf{x} \times \mathbf{y}_{||} = 0$, because these vectors are colinear. Hence $\mathbf{z} = \mathbf{x} \times \mathbf{y}_{\perp}$ and $|\mathbf{z}| = |\mathbf{x}||\mathbf{y}_{\perp}| = S(\mathbf{x}, \mathbf{y})$ because vectors $\mathbf{x}, \mathbf{y}_{\perp}$ are orthogonal to each other.

This Proposition is very important to understand the meaning of vector product. Shortly speaking *vector product of two vectors is a vector which is orthogonal to the plane spanned by these vectors, such that its magnitude is equal to the area of the parallelogram formed by these vectors. The direction is defined by orientation.*

Remark It is useful sometimes to consider area of parallelogram not as a positive number but as an real number positive or negative (see the next subsubsection.)

It is not worthless to recall the formula which we know from the school that area of parallelogram formed by vectors $\mathbf{x}, \mathbf{y}$ equals to the product of the base on the height. Hence

$$|\mathbf{x} \times \mathbf{y}| = |\mathbf{x}| \cdot |\mathbf{y}| |\sin \theta| \,, \tag{1.71}$$

where θ is an angle between vectors $\mathbf{x}, \mathbf{y}$.

Finally I would like again to stress:

Vector product of two vectors is equal to zero if these vectors are colinear (parallel). Scalar product of two vectors is equal to zero if these vector are orthogonal.

Exercise[†]Show that the vector product obeys to the following identity:

$$((\mathbf{a} \times \mathbf{b}) \times \mathbf{c}) + ((\mathbf{b} \times \mathbf{c}) \times \mathbf{a}) + ((\mathbf{c} \times \mathbf{a}) \times \mathbf{b}) = 0 \,. \quad \text{(Jacoby identity)} \tag{1.72}$$

This identity is related with the fact that heights of the triangle intersect in the one point.

Exercise[†] Show that $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a}, \mathbf{c}) - \mathbf{c}(\mathbf{a}, \mathbf{b})$.

1.12.2 Area of parallelogram in $\mathbf{E}^2$ and determinant of 2×2 matrices

.

Let $\mathbf{a}, \mathbf{b}$ be two vectors in 2-dimensional vector space $\mathbf{E}^2$.

One can consider $\mathbf{E}^2$ as a plane in 3-dimensional Euclidean space $\mathbf{E}^3$. Our aim is to calculate the area of the parallelogram $\Pi(\mathbf{a}, \mathbf{b})$ formed by vectors $\mathbf{a}, \mathbf{b}$. Let $\mathbf{n}$ be a unit vector in $\mathbf{E}^3$ which is orthogonal to $\mathbf{E}^2$. Then it is obvious that the vector product $\mathbf{a} \times \mathbf{b}$ is proportional to the normal vector $\mathbf{n}$ to the plane $\mathbf{E}^2$:

$$\mathbf{a} \times \mathbf{b} = A(\mathbf{a}, \mathbf{b}) \mathbf{n} \,, \tag{1.73}$$

and the area of the parallelogram $\Pi(\mathbf{a}, \mathbf{b})$ equals to the modulus of the coefficient $A(\mathbf{c}, \mathbf{b})$:

$$S\left(\Pi\left(\mathbf{a}, \mathbf{b}\right)\right) = |\mathbf{a} \times \mathbf{b}| = |A(\mathbf{a}, \mathbf{b})| \,. \tag{1.74}$$

The normal unit vector $\mathbf{n}$ and coefficient $A(\mathbf{a},\mathbf{b})$ are defined up to a sign: $\mathbf{n}\to-\mathbf{n}$, $A\to-A$. On the other hand the vector product $\mathbf{a}\times\mathbf{b}$ is defined up to a sign too: vector product depends on orientation. The answer for $\mathbf{a}\times\mathbf{b}$ is not changed if we perform calculations for vector product in an arbitrary basis $\{\mathbf{e}'_x,\mathbf{e}'_y,\mathbf{e}'_z\}$ which have the same orientation as the the basis $\{\mathbf{e},\mathbf{f},\mathbf{n}\}$ and $\mathbf{a}\times\mathbf{b}\mapsto\to-\mathbf{a}\times\mathbf{b}$. If we consider an arbitrary basis $\{\mathbf{e}'_x,\mathbf{e}'_y,\mathbf{e}'_z\}$ which have the orientation opposite to the orientation of the basis $\{\mathbf{e},\mathbf{f},\mathbf{n}\}$ (e.g. the basis $\{\mathbf{e},\mathbf{f},-\mathbf{n}\}$) then $A(\mathbf{a},\mathbf{b})\to-A(\mathbf{a},\mathbf{b})$. The magnitude $A(\mathbf{a},\mathbf{b})$ is so called algebraic area of parallelogram. It can positive and negative.

If (a_1,a_2), (b_1,b_2) are coordinates of the vectors $\mathbf{a},\mathbf{b}$ in the basis $\{\mathbf{e},\mathbf{f}\}$: $\mathbf{a}=a_1\mathbf{e}+a_2\mathbf{f}$, $\mathbf{b}=b_1\mathbf{e}+b_2\mathbf{f}$ and according to (1.69)

$$\mathbf{a}\times\mathbf{b}=\det\begin{pmatrix}\mathbf{e}&\mathbf{f}&\mathbf{n}\\ a_1&a_2&0\\ b_1&b_2&0\end{pmatrix}=\mathbf{n}\det\begin{pmatrix}a_x&a_y\\ b_x&b_y\end{pmatrix}\tag{1.75}$$

Thus $A(\mathbf{a},\mathbf{b})$ in equation (1.74) is equal to $\det\begin{pmatrix}a_x&a_y\\ b_x&b_y\end{pmatrix}$, and we come to the following formula for area of parallelogram

$$S(\Pi(\mathbf{a},\mathbf{b}))=|\mathbf{a}\times\mathbf{b}|=\left|\det\begin{pmatrix}a_x&a_y\\ b_x&b_y\end{pmatrix}\right|.\tag{1.76}$$

This is an important formula for relation between determinant of 2×2 matrix, area of parallelogram and vector product.

One can deduce this relation in other way:

Let $\mathbf{E}^2$ be a 2-dimensional Euclidean space. The function $A(\mathbf{a},\mathbf{b})$ defined by the relation (1.76) obeys the following conditions:

- It is anticommutative:
$$A(\mathbf{a},\mathbf{b})=-A(\mathbf{a},\mathbf{b})\tag{1.77}$$
- It is bilinear
$$A(\lambda\mathbf{a}+\mu\mathbf{b},\mathbf{c})=\lambda A(\mathbf{a},\mathbf{c})+\mu A(\mathbf{b},\mathbf{c});\ A(\mathbf{c},\lambda\mathbf{a}+\mu\mathbf{b})=\lambda A(\mathbf{c},\mathbf{a})+\mu A(\mathbf{c},\mathbf{b}).\tag{1.78}$$
- and it obeys normalisation condition:
$$A(\mathbf{e},\mathbf{f})=\pm1\tag{1.79}$$
for an arbitrary orthonormal basis.
(Compare with conditions (1.61)—(1.65).)

One can see that these conditions define uniquely $A(\mathbf{a},\mathbf{b})$ and these are the conditions which define the determinant of the 2×2 matrix.

1.12.3 Volumes of parallelograms and determinnants of linear operators in $\mathbf{E}^2$

Let A be an arbitrary linear operator in $\mathbf{E}^2$. One can see that the following formula holds.

Let $\mathbf{a}, \mathbf{b}$ be two arbitrary vectors in $\mathbf{E}^2$. Let $\mathbf{a}', \mathbf{b}'$ be two vectors such that

$$\mathbf{a}' = A(\mathbf{a})\,, \quad \mathbf{b} = A(\mathbf{b}')\,.$$

Consider two parallelograms: Parallelogram $\Pi(\mathbf{a}, \mathbf{b})$ formed by vectors $\mathbf{a}, \mathbf{b}$, and the second parallelogram $\Pi(\mathbf{a}', \mathbf{b}')$ formed by vectors $\alpha'.b'$. Then one can deduce from equation (1.76) that

$$\text{Area of } \Pi(\mathbf{a}', \mathbf{b}') \ = |\det A| \cdot \text{Area of } \Pi(\mathbf{a}, \mathbf{b}) \ . \tag{1.80}$$

This formula relates volumes of parallelograms $\Pi(\mathbf{a}, \mathbf{b})$, $\Pi(\mathbf{a}', \mathbf{b}')$ with determinant of linear operator which transforms the first parallelogram to the second one. (See also exercise 9 in Homework 4).

1.12.4 Volume of parallelepiped

The vector product of two vectors is related with area of parallelogram. What about a volume of parallelepiped formed by three vectors $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$?

Consider parallelepiped $\Pi(\mathbf{a}, \mathbf{b}, \mathbf{c})$ formed by vectors $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$. The parallelogram $\Pi(\mathbf{a}, \mathbf{b})$ formed by vectors $\mathbf{b}, \mathbf{c}$ can be considered as a base of this parallelepiped.

Let θ be an angle between height and vector $\mathbf{a}$. It is just the angle between the vector $\mathbf{b} \times \mathbf{c}$ and the vector $\mathbf{a}$. Then the volume is equal to the length of the height multiplied on the area of the parallelogram, $V = Sh = S|\mathbf{a}|\cos\theta$, i.e. volume is equal to scalar product of the vectors $\mathbf{a}$ on the vector product of vectors $\mathbf{b}$ and $\mathbf{c}$:

$$V(\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}) = |(\mathbf{a}, \mathbf{b} \times \mathbf{c})| = \left| \left(a_x\mathbf{e}_x + a_y\mathbf{e}_y + a_z\mathbf{e}_z, \det \begin{pmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{pmatrix} \right) \right|$$

$$= |(a_x\mathbf{e}_x + a_y\mathbf{e}_y + a_z\mathbf{e}_z, (b_yc_z - b_zc_y)\mathbf{e}_x + (b_zc_x - b_xc_z)\mathbf{e}_y + (b_xc_y - b_yc_x)\mathbf{e}_z)| =$$

$$|a_x(b_yc_z - b_zc_y) + a_y(b_zc_x - b_xc_z) + a_z(b_xc_y - b_yc_x)| = \left| \det \begin{pmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{pmatrix} \right| .$$

We come to beautiful and useful formula:

$$\text{volume of } \Pi(\mathbf{a},\mathbf{b},\mathbf{c}) = |(\mathbf{a},[\mathbf{b}\times\mathbf{c}])| = \left|\det\begin{pmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{pmatrix}\right| . \tag{1.81}$$

Compare this formula for the formula (1.76) for the area of parallelogram.

Remark In these formulae we consider the volume of the parallelepiped as a positive number. It is why we put the sign of 'modulus' in all the formulae above. On the other hand often it is very useful to consider the volume as a real number (it could be positive and negative).

Exercise Consider the function $F(\mathbf{a},\mathbf{b},\mathbf{c}) = (\mathbf{a},\mathbf{b}\times\mathbf{c})$.

1. Show that $F(\mathbf{a},\mathbf{b},\mathbf{c}) = 0$ if and only if vectors $\mathbf{a},\mathbf{b},\mathbf{c}$ are linear dependent.

2. Show that for an arbitrary vector $\mathbf{a}$, $F(\mathbf{a},\mathbf{a},\mathbf{c}) = 0$.

3. Show that for arbitrary vectors $\mathbf{a},\mathbf{b}$, $F(\mathbf{a},\mathbf{b},\mathbf{c}) = -F(\mathbf{a},\mathbf{b},\mathbf{c})$. Can you deduce 3) from the 2)?

1.12.5 Volumes of parallelepipeds and determinnants of linear operators in $\mathbf{E}^3$

Write down an equation for the volumes of parallelepipeds analogous to equation (1.80) for the the areas of parallelograms. Now instead parallelogram we consider parallelepiped, and instead linear operator A in $\mathbf{E}^2$ we consider linear operator A in $\mathbf{E}^3$.

Let A be an arbitrary linear operator in $\mathbf{E}^3$. In the same way as in formula (1.80) the following formula holds:

Let $\mathbf{a},\mathbf{b},\mathbf{c}$ be three arbitrary vectors in $\mathbf{E}^3$. Linear operator A transforms these three vectors to three vectors $\mathbf{a}',\mathbf{b}',\mathbf{c}'$ where

$$\mathbf{a}' = A(\mathbf{a}) , \quad \mathbf{b} = A(\mathbf{b}') , \mathbf{c}' = P(\mathbf{c}') .$$

Consider two parallelepipeds: Parallelepiped $\Pi(\mathbf{a},\mathbf{b}\,c)$ formed by vectors $\mathbf{a},\mathbf{b},\mathbf{c}$ and the second parallelepiped $\Pi(\mathbf{a}',\mathbf{b}'\,c')$ formed by vectors $\alpha'.b',\mathbf{c}'$. Then it follows from (1.81) the following formula and determinant of operator A:

$$\text{Volume of } \Pi(\mathbf{a}',\mathbf{b}',\mathbf{c}') \ = |\det A| \cdot \text{Volume of } \Pi(\mathbf{a},\mathbf{b},\mathbf{c}) \ . \tag{1.82}$$

This formula relates volumes of parallelepipeds $\Pi(\mathbf{a},\mathbf{b},\mathbf{c})$, $\Pi(\mathbf{a}',\mathbf{b}',\mathbf{c}')$ with determinant of linear operator which transforms the first parallelepiped to the second one. (See also exercise 9 in Homework 4).

2 Differential forms

2.1 Tangent vectors, curves, velocity vectors on the curve

Tangent vector is a vector $\mathbf{v}$ applied at the given point $\mathbf{p} \in \mathbf{E}^n$.

The set of all tangent vectors at the given point $\mathbf{p}$ is a vector space. It is called tangent space of $\mathbf{E}^3$ at the point $\mathbf{p}$ and it is denoted $T_{\mathbf{p}}(\mathbf{E}^n)$.

One can consider *vector field* on $\mathbf{E}^n$, i.e.a function which assigns to every point $\mathbf{p}$ vector $\mathbf{v}(\mathbf{p}) \in T_{\mathbf{p}}(\mathbf{E}^n)$.

It is instructive to study the conception of tangent vectors and vector fields on the curves and surfaces embedded in $\mathbf{E}^n$. In this course we mainly consider tangent vectors to curves.

A curve in $\mathbf{E}^n$ with parameter $t \in (a, b)$ is a continuous map

$$C\colon \qquad (a,b) \to \mathbf{E}^n \qquad \mathbf{r}(t) = (x^1(t), \dots, x^n(t)), \quad a < t < b \tag{2.1}$$

For example consider in $\mathbf{E}^2$ the curve

$$C\colon \qquad (0, 2\pi) \to \mathbf{E}^2 \qquad \mathbf{r}(t) = (R\cos t, R\sin t),\ \ 0 \le t < 2\pi\,.$$

The image of this curve is the circle of the radius R. It can be defined by the equation:

$$x^2 + y^2 = R^2\,.$$

To distinguish between curve and its image we say that curve C in (2.1) is *parameterised* curve or *path*. We will call the image of the curve *unparameterised curve* (see for details the next subsection). It is very useful to think about parameter t as a "time" and consider parameterised curve like *point moving along a curve.* Unparameterised curve is the trajectory of the moving point. It is *locus of the points.* The using of word "curve" without adjective "parameterised" or "nonparameterised" sometimes is ambiguous.

Vectors tangent to curve—velocity vector

Let $\mathbf{r}(t)$ $\quad \mathbf{r} = \mathbf{r}(t)$ be a curve in $\mathbf{E}^n$.

Velocity $\mathbf{v}(t)$ it is the vector

$$\mathbf{v}(t) = \frac{d\mathbf{r}}{dt} = \big(\dot{x}^1(t), \dots, \dots \dot{x}^n(t)\big) = \big(v^1(t), \dots, v^n(t)\big)$$

in $\mathbf{E}^n$. Velocity vector is *tangent vector to the curve.*

Let $C\colon \mathbf{r} = \mathbf{r}(t)$ be a curve and $\mathbf{r}_0 = \mathbf{r}(t_0)$ any given point on it. Then the set of all vectors tangent to the curve at the point $\mathbf{r}_0 = \mathbf{r}(t_0)$ is one-dimensional vector space $T_{\mathbf{r}_0}C$. It is linear subspace in vector space $T_{\mathbf{r}_0}C$. The points of the tangent space $T_{\mathbf{r}_0}C$ are the points of tangent line.

Remark We consider by default only *smooth, regular* curves. Curve $\mathbf{r}(t) = (x^1(t), \dots, x^n(t))$ is called smooth if all functions $x^i(t)$, $(i = 1, 2, \dots, n)$ are smooth functions (Function is called smooth if it has derivatives of arbitrary order.) Curve $\mathbf{r}(t)$ is called regular if velocity vector $\mathbf{v}(t) = \frac{d\mathbf{r}(t)}{dt}$ is not equal to zero at all t.

2.2 Reparameterisation

One can move along trajectory with different velocities, i.e. one can consider different parameterisation. E.g. consider

$$C_1\colon \quad \begin{cases} x(t) = t \\ y(t) = t^2 \end{cases} \quad 0 < t < 1\,, \qquad C_2\colon \quad \begin{cases} x(t) = \sin t \\ y(t) = \sin^2 t \end{cases} \quad 0 < t < \frac{\pi}{2}$$

Images of these two parameterised curves are the same. In both cases point moves along a piece of the same parabola but with different velocities.

Definition

Two smooth curves $C_1\colon \quad \mathbf{r}_1(t)\colon (a_1, b_1) \to \mathbf{E}^n$ and $C_2\colon \quad \mathbf{r}_2(\tau)\colon (a_2, b_2) \to \mathbf{E}^n$ are called equivalent if there exists reparameterisation map:

$$t(\tau)\colon (a_2, b_2) \to (a_1, b_1),$$

such that

$$r_2(\tau) = r_1(t(\tau)) \tag{2.2}$$

Reparameterisation $t(\tau)$ is diffeomorphism, i.e. function $t(\tau)$ has derivatives of all orders and first derivative $t'(\tau)$ is not equal to zero.

E.g. curves in (2.2) are equivalent because a map $\varphi(t) = \sin t$ transforms first curve to the second.

Equivalence class of equivalent parameterised curves is called non-parameterised curve.

Equivalent curves have the same image.

It is useful sometimes to distinguish curves in the same equivalence class which differ by orientation.

Definition Let C_1, C_2 be two equivalent curves. We say that they have same orientation (parameterisations $\mathbf{r}_1(t$ and $\mathbf{r}_(\tau)$ have the same orientation) if reparameterisation $t = t(\tau)$ has positive derivative, $t'(\tau) > 0$. We say that they have opposite orientation (parameterisations $\mathbf{r}_1(t$ and $\mathbf{r}_(\tau)$ have the opposite orientation) if reparameterisation $t = t(\tau)$ has negative derivative, $t'(\tau) < 0$.

Changing orientation means changing the direction of "walking" around the curve.

Equivalence class of equivalent curves splits on two subclasses with respect to orientation.

Non-formally: Two curves are equivalent curves (belong to the same equivalence class) if these parameterised curves (paths) have the same images. Two equivalent curves have the same image. They define the same set of points in $\mathbf{E}^n$. Different parameters correspond to moving along curve with different velocity. Two equivalent curves have opposite orientation If two parameterisations correspond to moving along the curve in different directions then these parameterisations define opposite orientation.

What happens with velocity vector if we change parameterisation? It changes its value, but it can change its direction only on opposite (If these parameterisations have opposite orientation of the curve):

$$\mathbf{v}(\tau) = \frac{d\mathbf{r}_2(\tau)}{d\tau} = \frac{d\mathbf{r}(t(\tau))}{d\tau} = \frac{dt(\tau)}{d\tau} \cdot \frac{d\mathbf{r}(t)}{dt}\Big|_{t=t(\tau)} \tag{2.3}$$

Or shortly: $\mathbf{v}(\tau)\big|_\tau = t_\tau(\tau)\mathbf{v}(t)\big|_{t=t(\tau)}$

We see that velocity vector is multiplied on the coefficient (depending on the point of the curve), i.e. velocity vectors for different parameterisations are collinear vectors.

(We call two vectors $\mathbf{a}, \mathbf{b}$ collinear, if they are proportional each other, i,e, if $\mathbf{a} = \lambda\mathbf{b}$.)

Example Consider following curves in $\mathbf{E}^2$:

$$C_1\colon \quad \begin{cases} x = \cos\theta \\ y = \sin\theta \end{cases} ,0 < \theta < \pi, \qquad C_2\colon \quad \begin{cases} x = u \\ y = \sqrt{1-u^2} \end{cases} ,-1 < u < 1,$$

$$\begin{cases} x = \tan t \\ y = \frac{\sqrt{\cos 2t}}{\cos t} \end{cases} ,-\frac{\pi}{4} < t < \frac{\pi}{4} \tag{2.4}$$

These three parameterised curves,(paths) define the same non-parameterised curve: the upper piece of the circle: $x^2 + y^2 = 1, y > 0$. The reparameterisation $u(\theta) = \cos\theta$ transforms the second curve to the first one.

The reparameterisation $u(\theta) = \cos\theta$ transforms the second curve to the first one.

The reparameterisation $u(\theta) = \tan t$ transforms the second curve to the third one one: $\frac{\sqrt{\cos 2t}}{\cos t} = \frac{\sqrt{\cos^2 t - \sin^2 t}}{\cos t} = \sqrt{1 - \tan^2 t}$.

Curves C_1, C_2 have opposite orientation because $u'(\theta) < 0$. Curves C_2, C_3 have the same orientation, because $u'(t) > 0$. Curves C_1 and C_2 have opposite orientations too (Why?).

In the first case point moves with constant pace $|\mathbf{v}(\theta)| = 1$ anti clock-wise "from right to left" from the point $A = (1, 0)$ to the point $B = (-1, 0)$. In the second case pace is not constant, but $v_x = 1$ is constant. Point moves clockwise "from left to right", from the point $B = (-1, 0)$ to the point $A = (1, 0)$. In the third case point also moves clock-wise "from the left to right".

There are other examples in the Homeworks.

2.3 Differential 0-forms and 1-forms

2.3.1 Definition and examples of 0-forms and 1-formsf

Most of considerations of this and next subsections can be considered only for $\mathbf{E}^2$ or $\mathbf{E}^3$. All examples for differential forms is only for $\mathbf{E}^2$, $\mathbf{E}^3$.

0-form on $\mathbf{E}^n$ it is just function on $\mathbf{E}^n$ (all functions under consideration are differentiable)

Now we define 1-forms.

Definition Differential 1-form ω on $\mathbf{E}^n$ is a function on tangent vectors of $\mathbf{E}^n$, such that it is linear at each point:

$$\omega(\mathbf{r}, \lambda\mathbf{v}_1 + \mu\mathbf{v}_2) = \lambda\omega(\mathbf{r}, \mathbf{v}_1) + \mu\omega(\mathbf{r}, \mathbf{v}_2) \,. \tag{2.5}$$

Here $\mathbf{v}_1, \mathbf{v}_2$ are vectors tangent to $\mathbf{E}^n$ at the point $\mathbf{r}$, $(\mathbf{v}_1, \mathbf{v}_2 \in T_x\mathbf{E}^n)$ (We recall that vector tangent at the point $\mathbf{r}$ means vector attached at the point $\mathbf{r}$). We suppose that ω is smooth function on points $\mathbf{r}$.

If $\mathbf{X}(\mathbf{r})$ is vector field and ω-1-form then evaluating ω on $\mathbf{X}(\mathbf{r})$ we come to the function $w(\mathbf{r}, \mathbf{X}(\mathbf{r}))$ on $\mathbf{E}^3$.

Let $\mathbf{e}_1, \dots, \mathbf{e}_n$ be a basis in $\mathbf{E}^n$ and $(x^1, \dots, x^n)$ corresponding coordinates: an arbitrary point with coordinates $(x^1, \dots, x^n)$ is assigned to the vector $\mathbf{r} = x^1\mathbf{e}_1 + x^2\mathbf{e}_2 + \dots x^n\mathbf{e}_n$ starting at the origin.

Translating basis vectors $\mathbf{e}_i$ $(i = 1, \dots, n)$ from the origin to other points of $\mathbf{E}^n$ we come to vector field which we also denote $\mathbf{e}_i$ $(i = 1, \dots, n)$. The value of vector field $\mathbf{e}_i$ at the point $(x^1, \dots, x^n)$ is the vector $\mathbf{e}_i$ attached at this point (tangent to this point).

Let ω be an 1-form on $\mathbf{E}^n$. Consider an arbitrary vector field $\mathbf{A}(\mathbf{r}) = \mathbf{A}(x^1, \dots, x^n)$:

$$\mathbf{A}(\mathbf{r}) = A^1(\mathbf{r})\mathbf{e}_1 + \dots + A^n(\mathbf{r})\mathbf{e}_n = \sum_{i=1}^{n} A^i(\mathbf{r})\mathbf{e}_i$$

Then by linearity

$$\omega(\mathbf{r}, \mathbf{A}(\mathbf{r})) = \omega\left(\mathbf{r}, A^1(\mathbf{r})\mathbf{e}_1 + \dots + A^n(\mathbf{r})\mathbf{e}_n\right) = A^1\omega(\mathbf{r}, \mathbf{e}_1) + \dots + A^n\omega(\mathbf{r}, \mathbf{e}_n)\,.$$

Consider *basic* differential forms $dx^1, dx^2, \dots, dx^n$ such that

$$dx^i(\mathbf{e}_j) = \delta^i_j = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases} . \tag{2.6}$$

Then it is easy to see that

$$dx^1(\mathbf{A}) = \mathbf{A}^1, dx^2(\mathbf{A}) = \mathbf{A}^2, ..., \text{i.e.} dx^i(\mathbf{A}) = A^i$$

Hence

$$\omega(\mathbf{r}, \mathbf{A}(\mathbf{r})) = \left(\omega_1(\mathbf{r})dx^1 + \omega_2(\mathbf{r})dx^2 + \dots + \omega_n(\mathbf{r})dx^n\right)(\mathbf{A}(\mathbf{r}))$$

where components $\omega_i(\mathbf{r}) = \omega(\mathbf{r}, \mathbf{e}_i)$.

In the same way as an arbitrary vector field on $\mathbf{E}^n$ can be expanded over the basis $\{\mathbf{e}_i\}$ (see (2.3.1)), an arbitrary differential 1-form ω can be expanded over the basis forms(2.3.1)

$$\omega = \omega_1(x^1, \dots, x^n)dx^1 + \omega_2(x^1, \dots, x^n)dx^2 + \dots + \omega_n(x^1, \dots, x^n)dx^n\,.$$

Example Consider in $\mathbf{E}^3$ a basis $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$ and corresponding coordinates (x, y, z). Then

$$\begin{array}{c} dx(\mathbf{e}_x) = 1, dx(\mathbf{e}_y) = 0, dx(\mathbf{e}_z) = 0 \\ dy(\mathbf{e}_x) = 0, dy(\mathbf{e}_y) = 1, dy(\mathbf{e}_z) = 0 \\ dz(\mathbf{e}_x) = 0, dz(\mathbf{e}_y) = 0, dz(\mathbf{e}_z) = 1 \end{array} \tag{2.7}$$

The value of a differential 1-form $\omega = a(x,y,z)dx + b(x,y,z)dy + c(x,y,z)dz$ on vector field $\mathbf{X} = A(x,y,z)\mathbf{e}_x + B(x,y,z)\mathbf{e}_y + C(x,y,z)\mathbf{e}_z$ is equal to

$$\omega(\mathbf{r}, \mathbf{X}) = a(x,y,z)dx(\mathbf{X}) + b(x,y,z)dx(\mathbf{X}) + c(x,y,z)dx(\mathbf{X}) =$$

$$a(x,y,z)A(x,y,z) + b(x,y,z)B(x,y,z) + c(x,y,z)C(x,y,z)$$

It is very useful (see below) introduce for basic vectors new notations:

$$\mathbf{e}_i \mapsto \frac{\partial}{\partial x^i} \quad \text{for basic vectors } \mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z \text{ in } \mathbf{E}^3\, \mathbf{e}_x \mapsto \frac{\partial}{\partial x} \quad \mathbf{e}_y \mapsto \frac{\partial}{\partial y} \quad \mathbf{e}_z \mapsto \frac{\partial}{\partial z}\,. \tag{2.8}$$

In these new notations the formula (2.3.1) looks like

$$dx^i\left(\frac{\partial}{\partial x^j}\right) = \delta^i_j = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases} \quad .$$

and the formula (2.7) looks like

$$\begin{array}{l} dx\left(\frac{\partial}{\partial x}\right) = 1, dx\left(\frac{\partial}{\partial y}\right) = 0, dx\left(\frac{\partial}{\partial z}\right) = 0 \\ dy\left(\frac{\partial}{\partial x}\right) = 0, dy\left(\frac{\partial}{\partial y}\right) = 1, dy\left(\frac{\partial}{\partial z}\right) = 0 \\ dz\left(\frac{\partial}{\partial x}\right) = 0, dz\left(\frac{\partial}{\partial y}\right) = 0, dz\left(\frac{\partial}{\partial z}\right) = 1 \end{array}$$

It is very useful to introduce new notation for vectors $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$.

In the next subsection we will consider the directional derivative of function along vector fields. The directional derivative will justify our new notations (2.8).

2.3.2 Vectors—directional derivatives of functions

Let $\mathbf{R}$ be a vector in $\mathbf{E}^n$ tangent to the point $\mathbf{r} = \mathbf{r}_0$ (attached at a point $\mathbf{r} = \mathbf{r}_0$). Define the operation of derivative of an arbitrary (differentiable) function at the point $\mathbf{r}_0$ along the vector $\mathbf{R}$— directional derivative of function f along the vector $\mathbf{R}$

Definition

Let $\mathbf{r}(t)$ be a curve such that

- $\mathbf{r}(t)\big|_{t=0} = \mathbf{r}_0$

- Velocity vector of the curve at the point $\mathbf{r}_0$ is equal to $\mathbf{R}$: $\frac{d\mathbf{r}(t)}{dt}\Big|_{t=0} = \mathbf{R}$

Then directional derivative of function f with respect to the vector $\mathbf{R}$ at the point $\mathbf{r}_0$ $\partial_{\mathbf{R}} f\big|_{\mathbf{r}_0}$ is defined by the relation

$$\partial_{\mathbf{R}} f\big|_{\mathbf{r}_0} = \frac{d}{dt}\left(f\left(\mathbf{r}(t)\right)\right)\big|_{t=0}. \tag{2.9}$$

Using chain rule one come from this definition to the following important formula for the directional derivative:

$$\text{If } \mathbf{R} = \sum_{i=1}^{n} R^i \mathbf{e}_i \text{ then } \partial_{\mathbf{R}} f\big|_{\mathbf{r}_0} = \sum_{i=1}^{n} R^i \frac{\partial}{\partial x^i} f(x^1, \dots, x^n)\big|_{\mathbf{r}=\mathbf{r}_0} \tag{2.10}$$

It follows form this formula that

One can assign to every vector $\mathbf{R} = \sum_{i=1}^{n} R^i \mathbf{e}_i$ *the operation* $\partial_{\mathbf{R}} = R^1 \frac{\partial}{\partial x^1} + R^2 \frac{\partial}{\partial x^2} + \cdots + R^n \frac{\partial}{\partial x^n}$ *of taking directional derivative:*

$$\mathbf{R} = \sum_{i=1}^{n} R^i \mathbf{e}_i \mapsto \partial_{\mathbf{R}} = \sum_{i=1}^{n} R^i \frac{\partial}{\partial x^i} \tag{2.11}$$

Thus we come to notations (2.8). The symbols $\partial_x, \partial_y, \partial_z$ correspond to partial derivative with respect to coordinate x or y or z . Later we see that these new notations are very illuminating when we deal with arbitrary coordinates, such as polar coordinates or spherical coordinates, The conception of orthonormal basis is ill-defined in arbitrary coordinates, but one can still consider the corresponding partial derivatives. Vector fields $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$ (or in new notation $\partial_x, \partial_y, \partial_z$) can be considered as a basis[6] in the space of all vector fields on $\mathbf{E}^3$.

An arbitrary vector field (2.3.1) can be rewritten in the following way:

$$\mathbf{A}(\mathbf{r}) = A^1(\mathbf{r})\mathbf{e}_1 + \cdots + A^n(\mathbf{r})\mathbf{e}_n = A^1(\mathbf{r})\frac{\partial}{\partial x^1} + A^2(\mathbf{r})\frac{\partial}{\partial x^2} + \cdots + A^n(\mathbf{r})\frac{\partial}{\partial x^n} \tag{2.12}$$

[6] Coefficients of expansion are functions, elements of algebra of functions, not numbers ,elements of field. To be more careful, these vector fields are basis of the *module* of vector fields on $\mathbf{E}^3$

2.3.3 Differential acting 0-forms $\to$ 1-forms

Now we introduce very important operation: Differential d which acts on 0-forms and transforms them to 1 forms.

$$\boxed{\begin{array}{l}\text{Differential}\\ \text{0-forms}\end{array}} \xrightarrow{d} \boxed{\begin{array}{l}\text{Differential}\\ \text{1-forms}\end{array}}$$

Later we will learn how differential acts on 1-forms transforming them to 2-forms.

Definition Let $f = f(x)$-be 0-form, i.e. function on $\mathbf{E}^n$. Then

$$df = \sum_{i=1}^{n} \frac{\partial f(x^1, \dots, x^n)}{\partial x^i} dx^i \,. \tag{2.13}$$

The value of 1-form df on an arbitrary vector field (2.12) is equal to

$$df(\mathbf{A}) = \sum_{i=1}^{n} \frac{\partial f(x^1, \dots, x^n)}{\partial x^i} dx^i(\mathbf{A}) = \sum_{i=1}^{n} \frac{\partial f(x^1, \dots, x^n)}{\partial x^i} A^i = \partial_{\mathbf{A}} f \tag{2.14}$$

We see that *value of differential of* 0*-form* f *on an arbitrary vector field* $\mathbf{A}$ *is equal to directional derivative of function* f *with respect to the vector* $\mathbf{A}$.

The formula (2.14) defines df in invariant way without using coordinate expansions. Later we check straightforwardly the coordinate-invariance of the definition (2.13).

Exercise Check that

$$dx^i(\mathbf{A}) = \partial_{\mathbf{A}} x^i \tag{2.15}$$

Example If $f = f(x, y)$ is a function $(0 - form)$ on $\mathbf{E}^2$ then

$$df = \frac{\partial f(x, y)}{\partial x} dx + \frac{\partial f(x, y)}{\partial y} dy$$

and for an arbitrary vector field $\mathbf{A} = \mathbf{A} = A_x \mathbf{e}_x + \mathbf{A}_y \mathbf{e}_y = A_x(x, y)\partial_x + A_y(x, y)\partial_y$

$$df(\mathbf{A}) = \frac{\partial f(x, y)}{\partial x} dx(\mathbf{A}) + A_y(x, y) \frac{\partial f(x, y)}{\partial y} dy(\mathbf{A}) =$$

$$\mathbf{A}_x(x, y) \frac{\partial f(x, y)}{\partial x} + A_y(x, y) \frac{\partial f(x, y)}{\partial y} = \partial_{\mathbf{A}} f \,.$$

Example Find the value of 1-form $\omega = df$ on the vector field $\mathbf{A} = x\partial_x + y\partial_y$ if $f = \sin(x^2 + y^2)$.

$\omega(\mathbf{A}) = df(\mathbf{A})$. One can calculate it using formula (2.13) or using formula (2.14).

Solution (using (2.13)*)*:

$$\omega = df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = 2x\cos(x^2 + y^2)dx + 2y\cos(x^2 + y^2)dy\,.$$

$$\omega(\mathbf{A}) = 2x\cos(x^2 + y^2)dx(\mathbf{A}) + 2y\cos(x^2 + y^2)dy(\mathbf{A}) =$$

$$2x\cos(x^2 + y^2)A_x + 2y\cos(x^2 + y^2)dA_y = 2(x^2 + y^2)\cos(x^2 + y^2)\,,$$

Another solution (using (2.14)*)*

$$df(\mathbf{A}) = \partial_{\mathbf{A}} f = A_x \frac{\partial f}{\partial x} + A_x \frac{\partial f}{\partial y} = 2(x^2 + y^2)\cos(x^2 + y^2)\,.$$

See other examples in Homeworks.

2.4 Differential 1-form in arbitrary coordinates

Why differential forms? Why so strange notations for vector fields.

If we use the technique of differential forms we in fact do not care about what coordinates we work in: calculations are the same in arbitrary coordinates.

Consider first some examples

Example (Polar coordinates) Consider polar coordinates in $\mathbf{E}^2$:

$$\begin{cases} x(r,\varphi) = r\cos\varphi \\ y(r,\varphi) = r\sin\varphi \end{cases} \qquad (0 \le \varphi < 2\pi, 0 \le r < \infty),$$

We have that for basic 1-forms

$$dr = r_x dx + r_y dy = \frac{x}{\sqrt{x^2 + y^2}} dx + \frac{y}{\sqrt{x^2 + y^2}} dy = \frac{xdx + ydy}{r} \tag{2.16}$$

Respectively

$$dx = x_r dr + x_\varphi d\varphi = \cos\varphi dr - r\sin\varphi d\varphi$$

and

$$dy = y_r dr + y_\varphi d\varphi = \sin\varphi dr + r\cos\varphi d\varphi \tag{2.17}$$

For basic vector fields

$$\partial_r = \frac{\partial x}{\partial r}\partial_x + \frac{\partial y}{\partial r}\partial_y = \cos\varphi\partial_x + \sin\varphi\partial_y = \frac{x\partial_x + y\partial_y}{r},$$

$$\partial_\varphi = \frac{\partial x}{\partial \varphi}\partial_x + \frac{\partial y}{\partial \varphi}\partial_y = -r\sin\varphi\partial_x + r\cos\varphi\partial_y = x\partial_y - y\partial_x, \tag{2.18}$$

Example Calculate the value of forms $\omega = xdx + ydy$ and $\sigma = xdy - ydx$ on vector fields $\mathbf{A} = x\partial_x + y\partial_y$, $\mathbf{B} = x\partial_y - y\partial_x$. Perform calculations in Cartesian and in polar coordinates.

In Cartesian coordinates:

$$\omega(\mathbf{A}) = xdx(x\partial_x + y\partial_y) + ydy(x\partial_x + y\partial_y) = x^2 + y^2,\ \omega(\mathbf{B}) = xdx(\mathbf{B}) + ydy(\mathbf{B}) = 0,$$

$$\sigma(\mathbf{A}) = xdy(\mathbf{A}) - ydx(\mathbf{A}) = 0,\ \sigma(\mathbf{B}) = xdy(\mathbf{B}) - ydx(\mathbf{B}) = x^2 + y^2\,.$$

Now perform calculations in polar coordinates. According to relation (2.16)

$$\omega = xdx + ydy = rdr,\ \sigma = xdy - ydx = r^2 d\varphi$$

and according to relations (2.18) and (**??**)

$$\mathbf{A} = x\partial_x + y\partial_y = r\partial_r,\ \mathbf{B} = x\partial_y - y\partial_x = \partial_\varphi$$

Hence $\omega(\mathbf{A}) = rdr(\mathbf{A}) = r^2 = x^2 + y^2, \omega(\mathbf{B} = rdr(\partial_\varphi) = 0,$

$$\sigma(\mathbf{A}) = r^2 d\varphi(r\partial_r) = 0, \quad \sigma(\mathbf{B}) = r^2 d\varphi(\partial_\varphi) = r^2 = x^2 + y^2\,.$$

Answers coincide.

2.4.1 Calculations in arbitrary coordinates *

Consider an arbitrary (local) coordinates $u^1, \dots, u^n$ on $\mathbf{E}^n$: $u^i = u^i(x^1, \dots, x^n)$, $i = 1, \dots, n$. Show first that

$$du^i = \sum_{k=1}^n \frac{\partial u^i(x^1, \dots, x^n)}{\partial x^k} dx^k\,. \tag{2.19}$$

It is enough to check it on basic fields:

$$du^i\left(\frac{\partial}{\partial x^m}\right) = \partial_{\left(\frac{\partial}{\partial x^m}\right)} u^i = \frac{\partial u^i(x^1, \dots, x^n)}{x^m} = \sum_{k=1}^n \frac{\partial u^i(x^1, \dots, x^n)}{\partial x^k} dx^k\left(\left(\frac{\partial}{\partial x^m}\right)\right)\,.$$

because (see (2.3.1)):

$$dx^i\left(\frac{\partial}{\partial x^j}\right)=\delta^i_j=\begin{cases}1 \text{ if } i=j\\ 0 \text{ if } i\neq j\end{cases}\quad. \tag{2.20}$$

(We rewrite the formula (2.3.1) using new notations ∂_i instead $\mathbf{e}_i$). In the previous formula (2.3.1) we considered *cartesian* coordinates.

Show that the formula above is valid in an *arbitrary coordinates.*

One can see using chain rule that

$$\frac{\partial}{\partial u^i}=\frac{\partial x^1}{\partial u^i}\frac{\partial}{\partial x^1}+\frac{\partial x^2}{\partial u^i}\frac{\partial}{\partial x^2}+\cdots+\frac{\partial x^n}{\partial u^i}\frac{\partial}{\partial x^n}=\sum_{k=1}^n\frac{\partial x^k}{\partial u^i}\frac{\partial}{\partial x^k} \tag{2.21}$$

Calculate the value of differential form du^i on vector field $\frac{\partial}{\partial u^j}$ using (2.19) and (2.21):

$$du^i\left(\frac{\partial}{\partial u^j}\right)=\sum_{k=1}^n\frac{\partial u^i(x^1,\dots,x^n)}{\partial x^k}dx^k\left(\sum_{r=1}^n\frac{\partial x^r}{\partial u^j}\frac{\partial}{\partial x^r}\right)= \tag{2.22}$$

$$\sum_{k,r=1}^n\frac{\partial u^i(x^1,\dots,x^n)}{\partial x^k}\frac{\partial x^r(u^1,\dots,u^n)}{\partial u^j}dx^k\left(\frac{\partial}{\partial x^r}\right)=$$

$$\sum_{k,r=1}^n\frac{\partial u^i(x^1,\dots,x^n)}{\partial x^k}\frac{\partial x^r(u^1,\dots,u^n)}{\partial u^j}\delta^k_r=\sum_{k=1}^n\frac{\partial x^k}{\partial u^j}\frac{\partial u^i}{\partial x^k}=\delta^j_i$$

We come to

$$du^i\left(\frac{\partial}{\partial u^j}\right)=\delta^i_j=\begin{cases}1 \text{ if } i=j\\ 0 \text{ if } i\neq j\end{cases}\quad. \tag{2.23}$$

We see that formula (2.20) has the same appearance in arbitrary coordinates. In other words it is invariant with respect to an arbitrary transformation of coordinates.

Exercise Check straightforwardly the invariance of the definition (2.13). In coordinates $(u^1,\dots,u^n)$

Solution We have to show that the formula (2.13) does not changed under changing of coordinates $u^i=u^i(x^1,\dots,x^n)$.

$$df=\sum_{i=1}^n\frac{\partial f(x^1,\dots,x^n)}{\partial x^i}dx^i=\sum_{i=1,k}^n\frac{\partial f(x^1,\dots,x^n)}{\partial x^i}\frac{\partial x^i}{\partial u^k}du^k==\sum_{i=1}^n\frac{\partial f}{\partial u^k}du^k\,,$$

because $\sum_{i=1}^n\frac{\partial f(x^1,\dots,x^n)}{\partial x^i}\frac{\partial x^i}{\partial u^k}=\frac{\partial f}{\partial u^k}$

Example

Consider more in detail $\mathbf{E}^2$. (For $\mathbf{E}^3$ considerations are the same, just calculations little bit more complicated) Let u,v be an arbitrary coordinates in $\mathbf{E}^2$, $u=u(x,y), v=v(x,y)$.

$$du=\frac{\partial u(x,y)}{\partial x}dx+\frac{\partial u(x,y)}{\partial y}dy,\qquad dv=\frac{\partial v(x,y)}{\partial x}dx+\frac{\partial v(x,y)}{\partial y}dy \tag{2.24}$$

and

$$\partial_u = \frac{\partial x(u,v)}{\partial u}\partial_x + \frac{\partial y(u,v)}{\partial u}\partial_y, \ \ \partial_v = \frac{\partial x(u,v)}{\partial v}\partial_x + \frac{\partial y(u,v)}{\partial v}\partial_y \tag{2.25}$$

(As always sometimes we use notation ∂_u instead $\frac{\partial}{\partial u}$, ∂_x instead $\frac{\partial}{\partial x}$ e.t.c.) Then

$$\begin{aligned} du(\partial_u) = 1, du(\partial_v) = 0 \\ dv(\partial_w) = 0, dv(\partial_v) = 1 \end{aligned} \tag{2.26}$$

This follows from the general formula but it is good exercise to repeat the previous calculations for this case:

$$du(\partial_u) = \left(\frac{\partial u(x,y)}{\partial x}dx + \frac{\partial u(x,y)}{\partial y}dy\right)\left(\frac{\partial x(u,v)}{\partial u}\partial_x + \frac{\partial y(u,v)}{\partial u}\partial_y\right) =$$

$$\frac{\partial u(x,y)}{\partial x}\frac{\partial x(u,v)}{\partial u} + \frac{\partial u(x,y)}{\partial y}\frac{\partial y(u,v)}{\partial u} = \frac{\partial x(u,v)}{\partial u}\frac{\partial u(x,y)}{\partial x} + \frac{\partial y(u,v)}{\partial u}\frac{\partial u(x,y)}{\partial y} = 1$$

We just apply chain rule to the function $u = u(x, y) = u(x(u, v), y(u, v))$:

Analogously

$$du(\partial_v) = \left(\frac{\partial u(x,y)}{\partial x}dx + \frac{\partial u(x,y)}{\partial y}dy\right)\left(\frac{\partial x(u,v)}{\partial v}\partial_x + \frac{\partial y(u,v)}{\partial v}\partial_y\right)$$

$$\frac{\partial u(x,y)}{\partial x}\frac{\partial x(u,v)}{\partial v} + \frac{\partial u(x,y)}{\partial y}\frac{\partial y(u,v)}{\partial v} = \frac{\partial x(u,v)}{\partial v}\frac{\partial u(x,y)}{\partial x} + \frac{\partial y(u,v)}{\partial v}\frac{\partial u(x,y)}{\partial y} = 0$$

The same calculations for dv.

2.4.2 Calculations in polar coordinates *

Example. Let $f = x^4 - y^4$ and vector field $\mathbf{A} = r\partial_r$. Calculate 1-form $\omega = df$ and $\omega(\mathbf{A})$. We have $\omega = df = 4x^3dx - 4y^3dy$. One has transforms form from Cartesian coordinates to polar or vector field from polar coordinates to Cartesian.
In Cartesian coordinates: $\mathbf{A} = r\frac{\partial}{\partial r} = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}$. Hence $\omega(\mathbf{A}) = df(\mathbf{A}) =$

$$(4x^3dx - 4y^3dy)\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right) = 4x^3dx\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right) - 4y^3dy\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right) = 4x^4 - 4y^4\,.$$

$$\text{Or using (2.14)}\,, \omega(A) = df(\mathbf{A}) = \partial_{\mathbf{A}} f = \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)(x^4 - y^4) = 4x^4 - 4y^4$$

In polar coordinates $f = x^4 - y^4 = (x^2 - y^2)(x^2 + y^2) = r^2(r^2 \cos^\varphi - r^2 \sin^2\varphi) = r^4 \cos 2\varphi$, $\omega = df = 4r^3 \cos 2\varphi dr - 2r^4 \sin 2\varphi d\varphi$, and $\omega(\mathbf{A}) = \omega(r\partial_r) = 4r^4 \cos 2\varphi$ since $dr(\partial_r) = 1, d\varphi(\partial_r) = 0$. Or using (2.14)

$$\omega(\mathbf{A}) = df(\mathbf{A}) = \partial_A f = r\frac{\partial}{\partial r}\left(r^4 \cos\varphi\right) = 4r^4 \cos 2\varphi\,.$$

Example Calculate the value of form $\omega = \frac{xdy - ydx}{x^2+y^2}$ on the vector field $\mathbf{A} = \partial_\varphi$. $\partial_{\mathbf{A}} F = r\frac{\partial}{\partial r}(r^4 \cos 2\varphi) = 4r^4 \cos 2\varphi = 4(x^4 - y^4)$. Or using 1-forms: We have to transform form from Cartesian coordinates to polar or vector field from polar to Cartesian.

$$\frac{xdy - ydx}{x^2 + y^2} = d\varphi, \quad \omega(\mathbf{A}) = d\varphi(\partial_\varphi) = 1$$

or

$$\partial_\varphi = x\partial_y - y\partial_x,\ \omega(\mathbf{A}) = \frac{xdy(x\partial_y - y\partial_x) - ydx(x\partial_y - y\partial_x)}{x^2 + y^2} = 1\,.$$

2.5 Integration of differential 1-forms over curves

Let $\omega = \omega_1(x^1, \ldots, x^n)dx^1 + \cdots + \omega_1(x^1, \ldots, x^n)dx^n = \sum_{i=1}^n \omega_i dx^i$ be an arbitrary 1-form in $\mathbf{E}^n$

and $C\colon \mathbf{r} = \mathbf{r}(t), t_1 \leq t \leq t_2$ be an arbitrary smooth curve in $\mathbf{E}^n$.

One can consider the value of one form ω on the velocity vector field $\mathbf{v}(t) = \frac{d\mathbf{r}(t)}{dt}$ of the curve:

$$\omega(\mathbf{v}(t)) = \sum_{i=1}^n \omega_i \left(x^1(t), \ldots, x^n(t)\right) dx^i(\mathbf{v}(t)) = \sum_{i=1}^n \omega_i \left(x^1(t), \ldots, x^n(t)\right) \frac{dx^i(t)}{dt}$$

We define now integral of 1-form ω over the curve C.

Definition The integral of the form $\omega = \omega_1(x^1, \ldots, x^n)dx^1 + \cdots + \omega_n(x^1, \ldots, x^n)dx^n$ over the curve $C\colon \quad \mathbf{r} = \mathbf{r}(t) \quad t_1 \leq t \leq t_2$ is equal to the integral of the function $\omega(\mathbf{v}(t))$ over the interval $t_1 \leq t \leq t_2$:

$$\int_C \omega = \int_{t_1}^{t_2} \omega(\mathbf{v}(t))dt = \int_{t_1}^{t_2} \left(\sum_{i=1}^n \omega_i \left(x^1(t), \ldots, x^n(t)\right) \frac{dx^i(t)}{dt} \right) dt\,. \quad (2.27)$$

Proposition *The integral $\int_C \omega$ does not depend on the choice of coordinates on $\mathbf{E}^n$. It does not depend (up to a sign) on parameterisation of the curve: if $C\colon \quad \mathbf{r} = \mathbf{r}(t) \quad t_1 \leq t \leq t_2$ is a curve and $t = t(\tau)$ is an arbitrary reparameterisation, i.e. new curve $C'\colon \quad \mathbf{r}'(\tau) = \mathbf{r}(t(\tau)) \quad \tau_1 \leq \tau \leq \tau_2$, then $\int_C \omega = \pm \int_C' \omega$:*

$$\int_C \omega = \int_{C'} \omega, \quad \textit{if orientaion is not changed, i.e. if } t'(\tau) > 0$$

and
$$\int_C \omega = -\int_{C'} \omega, \quad \textit{if orientaion is changed, i.e. if } t'(\tau) < 0$$

If reparameterisation changes the orientation then starting point of the curve becomes the ending point and vice versa.

Proof of the Proposition Show that integral does not depend (up to a sign) on the parameterisation of the curve. Let $t(\tau)$ ($\tau_1 \leq t \leq \tau_2$) be reparameterisation. We come to the new curve C': $\mathbf{r}'(\tau) = \mathbf{r}(t(\tau))$. Note that the new velocity vector $\mathbf{v}'(\tau) = \frac{dr(t(\tau))}{d\tau} = t'(\tau)\mathbf{v}(t(\tau))$. Hence $\omega(\mathbf{v}'(\tau)) = w(\mathbf{v}(t(\tau)))t'(\tau)$. For the new curve C'

$$\int_{C'} \omega = \int_{\tau_1}^{\tau_2} \omega(\mathbf{v}'(\tau))d\tau = \int_{\tau_1}^{\tau_2} \omega(\mathbf{v}(t(\tau))\frac{dt(\tau)}{d\tau}d\tau = \int_{t(\tau_1)}^{t(\tau_2)} \omega(\mathbf{v}(t))dt$$

$t(\tau_1) = t_1$, $t(\tau_2) = t_2$ if reparameterisation does not change orientation and $t(\tau_1) = t_2$, $t(\tau_2) = t_1$ if reparameterisation changes orientation.

Hence $\int_{C'} w = \int_{t_1}^{t_2)} \omega(\mathbf{v}(t))dt = \int_C \omega$ if orientation is not changed and $\int_{C'} w = \int_{t_2}^{t_1)} \omega(\mathbf{v}(t))dt = -\int_{t_1}^{t_2)} \omega(\mathbf{v}(t))dt = -\int_C \omega$ is orientation is changed.

Example

Let
$$\omega = a(x,y)dx + b(x,y)dy$$

be 1-form in $\mathbf{E}^2$ (x, y–are usual Cartesian coordinates). Let C: $\mathbf{r} = \mathbf{r}(t)$ $\begin{cases} x = x(t) \\ y = y(t) \end{cases}$, $t_1 \leq t \leq t_2$ be a curve in $\mathbf{E}^2$.

Consider velocity vector field of this curve

$$\mathbf{v}(t) = \frac{d\mathbf{r}(t)}{dt} = \begin{pmatrix} v_x(t) \\ v_y(t) \end{pmatrix} = \begin{pmatrix} x_t(t) \\ y_t(t) \end{pmatrix} = x_t\partial_x + y_t\partial_y \tag{2.28}$$

($x_t = \frac{dx(t)}{dt}$, $y_t = \frac{dy(t)}{dt}$).

One can consider the value of one form ω on the velocity vector field $\mathbf{v}(t)$ of the curve: $\omega(\mathbf{v}) = a(x(t), y(t))dx(\mathbf{v}) + b(x(t), y(t))dy(\mathbf{v}) =$

$$a(x(t), y(t))x_t(t) + b(x(t), y(t))y_t(t)\,.$$

The integral of the form $\omega = a(x,y)dx + b(x,y)dy$ over the curve C: $\mathbf{r} = \mathbf{r}(t)$ $t_1 \leq t \leq t_2$ is equal to the integral of the function $\omega(\mathbf{v}(t))$ over the interval $t_1 \leq t \leq t_2$:

$$\int_C \omega = \int_{t_1}^{t_2} \omega(\mathbf{v}(t))dt = \int_{t_1}^{t_2} \left(a(x(t), y(t)) \frac{dx(t)}{dt} + b(x(t), y(t)) \frac{dy(t)}{dt} \right) dt \,. \tag{2.29}$$

Example Consider an integral of the form $\omega = 3dy + 3y^2dx$ over the curve C: $\mathbf{r}(t) \begin{cases} x = \cos t \\ y = \sin t \end{cases}, 0 \leq t \leq \pi/2$. ($C$ is the arc of the circle $x^2 + y^2 = 1$ defined by conditions $x, y \geq 0$).

Velocity vector $\mathbf{v}(t) = \frac{d\mathbf{r}(t)}{dt} = \begin{pmatrix} v_x(t) \\ v_y(t) \end{pmatrix} = \begin{pmatrix} x_t(t) \\ y_t(t) \end{pmatrix} = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}$. The value of the form on velocity vector is equal to

$$\omega(\mathbf{v}(t)) = 3y^2(t)v_x(t) + 3v_y(t) = 3\sin^2 t(-\sin t) + 3\cos t = 3\cos t - 3\sin^3 t$$

and

$$\int_C \omega = \int_0^{\frac{\pi}{2}} w(\mathbf{v}(t))dt = \int_0^{\frac{\pi}{2}} (3\cos t - 3\sin^3 t)dt = 3\left(\sin t + \cos t - \frac{\cos^3 t}{3}\right)\Big|_0^{\frac{\pi}{2}}$$

Example Now consider the integral of 1-form over the curve C which is the upper half of the circle $x^2 + y^2 = 1$: C: $\begin{cases} x^2 + y^2 = 1 \\ y \geq 0 \end{cases}$. Curve is given as an image. We have the image of the curve not the parameterised curve. We have to define a parameterisation ourself.

We consider three different parameterisations of this curve. Sure to calculate the integral it suffices to calculate $\int_C \omega$ in an arbitrary given parameterisation $\mathbf{r} = \mathbf{r}(t)$ of the curve C, then note that for an arbitrary reparameterisation $t = t(\tau)$, the integral will remain the same or it will change a sign depending on the reparameterisation $t = t(\tau)$ preserves orientation or not.

$$\mathbf{r}_1(t)\colon \begin{cases} x = R\cos t \\ y = R\sin t \end{cases}, 0 \leq t \leq \pi\,, \quad \mathbf{r}_2(t)\colon \begin{cases} x = R\cos \Omega t \\ y = R\sin \Omega t \end{cases}, 0 \leq t \leq \frac{\pi}{\Omega}\,, (\Omega > 0)$$

and

$$\mathbf{r}_3(t)\colon \begin{cases} x = t \\ y = \sqrt{R^2 - t^2} \end{cases}, -R \leq t \leq R, , \tag{2.30}$$

All these curves are the same image. If $\Omega = 1$ the second curve coincides with the first one. First and second curve have the same orientation (reparameterisation $t \mapsto \Omega t$) The third curve have orientation opposite to first and second (reparameterisation $t \mapsto \cos t$, the derivative $\frac{d\cos t}{dt} < 0$).

Calculate integrals $\int_{C_1}\omega$, $\int_{C_2}\omega$, $\int_{C_3}\omega$ in the case if $\omega = xdy - ydx$ and check straightforwardly that these integrals coincide if orientation is the same or they have different signs if orientation is opposite. For the form $\omega = xdy - ydx$. $\omega(\mathbf{v}) = xy_t - yx_t$. We have

$$\int_{C_1}\omega = \int_0^\pi (xy_t - yx_t)dt = \int_0^\pi (R^2\cos^2 t + R^2\sin^2 t)dt = \pi R^2$$

$$\int_{C_2}\omega = \int_0^{\frac{\pi}{\Omega}} (xy_t - yx_t)dt = \int_0^\pi (R^2\Omega\cos^2\Omega t + R^2\Omega\sin^2\Omega t)dt = \pi R^2\,.$$

These answers coincide: both parameterisation have the same orientation.

For the third parameterisation:

$$\int_{C_3}\omega = \int_0^R (xy_t - yx_t)dt = \int_0^R \left(t\left(\frac{-t}{\sqrt{R^2-t^2}}\right) - \sqrt{R^2-t^2}\right)dt =$$

$$-R^2\int_0^R \frac{dt}{\sqrt{R^2-t^2}} = -R^2\int_0^1 \frac{du}{\sqrt{1-u^2}} = -\pi R^2$$

We see that the sign is changed.

Finally consider the integral of the form $\omega = xdy - ydx$ over the semicircle in polar coordinates instead Cartesian coordinates, We have that in polar coordinates semicircle is $\begin{cases} r(t) = R \\ \varphi(t) = t \end{cases}$, $0 \leq t \leq \pi$. The form $\omega = xdy - ydx = r\cos\varphi d(r\sin\varphi) - r\sin\varphi d(r\sin\varphi) = r^2 d\varphi$ and $\mathbf{v}(t) = (r_t, \varphi_t) = (0,1)$, i.e. $\mathbf{v}(t) = \partial_\varphi$. We have that $\omega(\mathbf{v}(t)) = r(t)^2 d\varphi(\partial_\varphi) = R^2$. Hence $\int_C \omega = \int_0^\pi R^2 dt = \pi R^2$. Answer is the same: The value of integral does not change if we change coordinates in the plane.

For other examples see Homeworks.

2.6 Integral over curve of exact form

1-form ω is called exact if there exists a function f such that $\omega = df$.

Of course not any form is an exact form (see exercises in Homeworks.) E.g. 1-form $xdy - ydx$ is not an exact form. Indeed suppose that this is an exact form, i.e. $xdy - ydx = df = f_x dx + f_y dy$, then $f_y = x$ and $f_x = -y$. We see that on one hand $f_{xy} = (f_x)_y = -1$ and on the other hand $f_{yx} = (f_y)_x = 1$. Contradiction.

Theorem
Let ω be an exact 1-form in $\mathbf{E}^n$, $\omega = df$.
Then the integral of this form over an arbitrary curve C: $\quad \mathbf{r} = \mathbf{r}(t) \quad t_1 \le t \le t_2$ is equal to the difference of the values of the function f at starting and ending points of the curve C:

$$\int_C \omega = f\big|_{\partial C} = f(\mathbf{r}_2) - f(\mathbf{r}_1)\,, \quad \mathbf{r}_1 = \mathbf{r}(t_1), \mathbf{r}_2 = \mathbf{r}(t_2)\,. \tag{2.31}$$

Proof: $\int_C df = \int_{t_1}^{t_2} df(\mathbf{v}(t)) = \int_{t_1}^{t_2} \frac{d}{dt} f(\mathbf{r}(t))dt = f(\mathbf{r}(t))|_{t_1}^{t_2}$.

Example Calculate an integral of the form $\omega = 3x^2(1+y)dx + x^3dy$ over the arc of the semicircle $x^2 + y^2 = 1, y \ge 0$.

One can calculate the integral naively using just the formula (2.29): Choose a parameterisation of C,e.g., $x = \cos t, y = \sin t$, then $\mathbf{v}(t) = -\sin t\partial_x + \cos t\partial_x$ and $\omega(\mathbf{v}(t)) = (3x^2(1+y)dx + x^3dy)(-\sin t\partial_x + \cos t\partial_y) = -3\cos^2 t(1+\sin t)\sin t + \cos^3 t \cdot \cos t$ and

$$\int_C \omega = \int_0^\pi (-3\cos^2 t \sin t - 3\cos^2 t \sin^2 t + \cos^4 t)dt = ...$$

Calculations are little bit long.

But for the form $\omega = 3x^2(1+y)dx + x^3dy$ one can calculate the integral in a much more efficient way noting that it is an exact form:

$$\omega = 3x^2(1+y)dx + x^3dy = d\left(x^3(1+y)\right) \tag{2.32}$$

Hence it follows from the Theorem that

$$\int_C \omega = f(\mathbf{r}(\pi)) - f(\mathbf{r}(0)) = x^3(1+y)\big|_{x=1,y=0}^{x=-1,y=0} = -2 \tag{2.33}$$

Remark If we change the orientation of curve then the starting point becomes the ending point and the ending point becomes the starting point.— The integral changes the sign in accordance with general statement, that integral of 1-form over parameterised curve is defined up to reparameterisation.

Corollary *The integral of an exact form over an arbitrary closed curve is equal to zero.*

Proof. According to the Theorem $\int_C \omega = \int_C df = f\big|_{\partial C} = 0$, because the starting and ending points of closed curve coincide.

Example. Calculate the integral of 1-form $\omega = x^5 dy + 5x^4 y dx$ over the ellipse $x^2 + \frac{y^2}{9} = 1$.

The form $\omega = x^5 dy + 5x^4 y dx$ is exact form because $\omega = x^5 dy + 5x^4 y dx = d(x^5 y)$. Hence the integral over ellipse is equal to zero, because it is a closed curve.

2.7 †Differential 2-forms (in $\mathbf{E}^2$)

2.7.1 † 2-form–area of parallelogram

We give first general ideas about what is it differential k-form ($k = 2, 3$)

1-form is a linear function on vectors:

$$\omega(\mathbf{A}):\ \omega(\lambda\mathbf{A} + \mu\mathbf{B}) = \lambda\omega(\mathbf{A}) + \mu\omega(\mathbf{B}),$$

2-form is a bilinear function on two vectors:

$$\omega(\mathbf{A}, \mathbf{K}):\ \omega(\lambda\mathbf{A}+\mu\mathbf{B}, \mathbf{K}) = \lambda\omega(\mathbf{A}, K)+\mu\omega(\mathbf{B}, K)\,, \omega(\mathbf{K}, \lambda\mathbf{A}+\mu\mathbf{B}) = \lambda\omega(\mathbf{K}, \mathbf{A})+\mu\omega(\mathbf{K}, \mathbf{B})$$

which obey to the following condition

$$\omega(\mathbf{A}, \mathbf{B}) = -\omega(\mathbf{B}.\mathbf{A}) \tag{2.34}$$

This condition implies that the value of of 2-form on vectors $\mathbf{A}, \mathbf{B}$ is proportional to the area of parallelogram $\Pi_{\mathbf{A},\mathbf{B}}$ formed vy these vectors. Explain it on a simple example.

Consider differential 2-form $dx \wedge dy$ in $\mathbf{E}^2$:

$$dx \wedge dy \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) = 1$$

(In the same way as 1-forms dx, dy are basic forms for 1-form.)

Linearity conditions and condition (2.34) imply that for an arbitrary 2-form ω in $\mathbf{E}^2$ $\omega = a(x, y) dx \wedge dy$.

Take two vector fields $\mathbf{A}, \mathbf{B}$, $\mathbf{A} = A_x \frac{\partial}{\partial x} + A_y \frac{\partial}{\partial y}$, Then due to conditions (2.34) above we have

$$\omega(\mathbf{A}, \mathbf{B}) = \omega\left(A_x \frac{\partial}{\partial x} + A_y \frac{\partial}{\partial y}, B_x \frac{\partial}{\partial x} + B_y \frac{\partial}{\partial y}\right) =$$

$$A_xB_x\omega(\partial_x,\partial_x)A_xB_y\omega(\partial_x,\partial_y)A_yB_x\omega(\partial_y,\partial_x)A_yB_y\omega(\partial_y,\partial_y)=$$

$$a\left(A_xB_x\underbrace{dx\wedge dy(\partial_x,\partial_x)}_{=0}A_xB_y\underbrace{dx\wedge dy(\partial_x,\partial_y)}_{=1}A_yB_x\underbrace{dx\wedge dy(\partial_y,\partial_x)}_{=-1}A_yB_y\underbrace{dx\wedge dy(\partial_y,\partial_y)}_{=0}\right)=$$

$$a(A_xB_y-A_yB_x)=a\cdot\text{area of parallelogram }\Pi_{\mathbf{A},\mathbf{B}}=a\det\begin{pmatrix}A_x & A_y\\ B_x & B_y\end{pmatrix}$$

In a analogous way 3-forms are related with volume of parallelipiped, ... k-form with volume of k-parallelipiped...

2.7.2 † Wedge product

We considered detailed definition of 1-forms. Now we give some formal approach to describe 2-forms. Differential forms on $\mathbf{E}^2$ is an expression obtained by adding and multiplying functions and differentials dx, dy. These operations obey usual associativity and distributivity laws but multiplications is not moreover of one-forms on each other is *anti-commutative*:

$$\omega\wedge\omega'=-\omega'\wedge\omega\quad\text{if }\omega,\omega'\text{ are 1-forms}\tag{2.35}$$

In particular

$$dx\wedge dy=-dy\wedge dx, dx\wedge dx=0, dy\wedge dy=0\tag{2.36}$$

Example If $\omega=xdy+zdx$ and $\rho=dz+ydx$ then

$$\omega\wedge\rho=(xdy+zdx)\wedge(dz+ydx)=xdy\wedge dz+zdx\wedge dz+xydy\wedge dx$$

and

$$\rho\wedge\omega=(dz+ydx)\wedge(xdy+zdx)=xdz\wedge dy+zdz\wedge dx+xydx\wedge dy=-\omega\wedge\rho$$

Changing of coordinates. If $\omega=a(x,y)dx\wedge dy$ be two form and $x=x(u,v), y=y(u,v)$ new coordinates then $dx=x_udu+x_vdv$, $dy=y_udu+y_vdv$ ($x_u=\frac{\partial x(u,v)}{\partial u}$, $x_v=\frac{\partial x(u,v)}{\partial v}$, $y_u=\frac{\partial y(u,v)}{\partial u}$, $y_v=\frac{\partial y(u,v)}{\partial v}$). and

$$a(x,y)dx\wedge dy=a\left(x(u,v),y(u,v)\right)(x_udu+x_vdv)\wedge(y_udu+y_vdv)=\tag{2.37}$$

$$a\left(x(u,v),y(u,v)\right)(x_udu+x_vdv)\left(x_uy_vdu\wedge dv+x_vy_udv\wedge du\right)=$$

$$a\left(x(u,v),y(u,v)\right)(x_uy_v-x_vy_u)du\wedge dv$$

Example Let $\omega=dx\wedge dy$ then in polar coordinates $x=r\cos\varphi, y=r\sin\varphi$

$$dx\wedge dy=(\cos\varphi dr-r\sin\varphi d\varphi)\wedge(\sin\varphi dr+r\cos\varphi d\varphi)=rdr\wedge d\varphi\tag{2.38}$$

2.7.3 † 0-forms (functions) $\xrightarrow{d}$ 1-forms $\xrightarrow{d}$ 2-forms

We introduced differential d of functions (0-forms) which transform them to 1-form. It obeys the following condition:

- d: is linear operator: $d(\lambda f + \mu g) = \lambda df + \mu dg$
- $d(fg) = df \cdot g + f \cdot dg$

Now we introduce differential on 1-forms such that

- d: is linear operator on 1-forms also
- $d(f\omega) = df \wedge \omega + f d\omega$
- $ddf = 0$

Remark Sometimes differential d is called *exterior differential*.

Perform calculations using this definition and (2.35):

$$d\omega = d(\omega_1 dx + \omega_2 dy) = dw_1 \wedge dx + dw_2 \wedge dy = \left(\frac{\partial \omega_1(x,y)}{\partial x} dx + \frac{\partial \omega_1(x,y)}{\partial y} dy\right) \wedge dx +$$

$$\left(\frac{\partial \omega_2(x,y)}{\partial x} dx + \frac{\partial \omega_2(x,y)}{\partial y} dy\right) \wedge dy = \left(\frac{\partial \omega_2(x,y)}{\partial x} - \frac{\partial \omega_1(x,y)}{\partial y}\right) dx \wedge dy$$

Example Consider 1-form $\omega = xdy$. Then $d\omega = d(xdy) = dx \wedge dy$.

2.7.4 †Exact and closed forms

We know that it is very easy to integrate exact 1-forms over curves (see the subsection "Integral over curve of exact form")

How to know is the 1-form exact or no?

Definition We say that one form ω is *closed* if two form $d\omega$ is equal to zero.

Example 1-form $xdy + ydx$ is closed because $d(xdy + ydx) = 0$.

It is evident that exact 1-form is closed:

$$\omega = d\rho \Rightarrow d\omega = d(d\rho) = d \circ d\rho = 0 \tag{2.39}$$

We see that the condition that form is closed is necessary condition that form is exact.

So if $d\omega \neq 0$, i.e. the form is not closed, then it is not exact.

Is this condition sufficient? Is it true that a closed form is exact?

In general the answer is: *No.*

E.g. we considered differential 2-form

$$\omega = \frac{xdy - ydx}{x^2 + y^2} \tag{2.40}$$

defined in $\mathbf{E}^2\backslash 0$. It is closed, but it is not exact (See non-compulsory exercises 11,12,13 in the Homework 6).

How to recognize for 1-form ω is it exact or no?
Inverse statement (Poincaré lemma) is true if 1-form is well-defined in $\mathbf{E}^2$:
A closed 1-form ω in $\mathbf{E}^n$ is exact if it is well-defined at all points of $\mathbf{E}^n$, i.e. if it is differentiable function at all points of $\mathbf{E}^n$.

Sketch a proof for 1-form in $\mathbf{E}^2$: if ω is defined in whole $\mathbf{E}^2$ then consider the function

$$F(\mathbf{r}) = \int_{C_{\mathbf{r}}} \omega \tag{2.41}$$

where we denote by $C_{\mathbf{r}}$ an arbitrary curve which starts at origin and ends at the point $\mathbf{r}$. It is easy to see that the integral is well-defined and one can prove that $\omega = df$.

The explicit formula for the function (2.41) is the following: If $\omega = a(x,y)dx + b(x,y)dy$ then $F(x,y) = \int_0^1 \left(a(tx,ty)x + b(tx,ty)y\right) dt$.

Exercise Check by straightforward calculation that $\omega = dF$ (See exercise 14 in Homework 6).

2.7.5 † Integration of two-forms. Area of the domain

We know that 1-form is a linear function on tangent vectors. If $\mathbf{A}, \mathbf{B}$ are two vectors attached at the point $\mathbf{r}_0$, i.e. tangent to this point and ω, ρ are two 1-forms then one defines the value of $\omega \wedge \rho$ on $\mathbf{A}, B$ by the formula

$$\omega \wedge \rho(\mathbf{A}, \mathbf{B}) = \omega(\mathbf{A})\rho(B) - \omega(B)\rho(A) \tag{2.42}$$

We come to bilinear anisymmetric function on tangent vectors. If $\sigma = a(x,y)dx \wedge dy$ is an arbitrary two form then this form defines bilinear form on pair of tangent vectors: $\sigma(\mathbf{A}, \mathbf{B}) =$

$$a(x,y)dx \wedge dy(A,B) = a(x,y)\left(dx(\mathbf{A})dy(\mathbf{B}) - dx(\mathbf{B})dy(\mathbf{A})\right) = a(x,y)(A_xB_y - A_yB_y) \tag{2.43}$$

One can see that in the case if $a = 1$ then right hand side of this formula is nothing but the area of parallelogram spanned by the vectors $\mathbf{A}, \mathbf{B}$.

This leads to the conception of integral of form over domain.

Let $\omega = a(x)dx \wedge dy$ be a two form and D be a domain in $\mathbf{E}^2$. Then by definition

$$\int_D \omega = \int_D a(x,y)dxdy \tag{2.44}$$

If $\omega = dx \wedge dy$ then

$$\int_D w = \int_D (x,y)dxdy = \text{Area of the domain } D \tag{2.45}$$

The advantage of these formulae is that we do not care about coordinates[7]

Example Let D be a domain defined by the conditions

$$\begin{cases} x^2 + y^2 \leq 1 \\ y \geq 0 \end{cases} \tag{2.47}$$

Calculate $\int_D dx \wedge dy$.

$\int_D dx \wedge dy = \int_D dxdy =$ area of the $D = \frac{\pi}{2}$.

If we consider polar coordinates then according (2.38)

$$dx \wedge dy = rdr \wedge d\varphi$$

Hence $\int_D dx \wedge dy = \int_D rdr \wedge d\varphi = \int_D rdrd\varphi = \int_0^1 \left(\int_0^\pi d\varphi\right) rdr = \pi \int_0^1 rdr = \pi/2$.

Another example

Example Let D be a domain in $\mathbf{E}^2$ defined by the conditions

$$\begin{cases} \frac{(x-c)^2}{a^2} + \frac{y^2}{b^2} \leq 1 \\ y \geq 0 \end{cases} \tag{2.48}$$

D is domain restricted by upper half of the ellipse and x-axis. Ellipse has the centre at the point $(c, 0)$. Its area is equal to $S = \int_D dx \wedge dy$. Consider new variables x', y': $x = c + ax', y = by'$. In new variables domain D becomes the domain from the previous example:

$$\frac{(x-c)^2}{a^2} + \frac{y^2}{b^2} = {x'}^2 + {y'}^2$$

and $dx \wedge dy = abdx' \wedge dy'$. Hence

$$S = \int_{\frac{(x-c)^2}{a^2} + \frac{y^2}{b^2} \leq 1, y \geq 0} dx \wedge dy = ab \int_{x'^2 + y'^2 \leq 1, y' \geq 0} dx' \wedge dy' = \frac{\pi ab}{2} \tag{2.49}$$

Theorem 2 (Green formula) Let ω be 2-form such that $\omega = d\omega'$ and D be a domain– interior of the closed curve C. Then

$$\int_D \omega = \int_C \omega' \tag{2.50}$$

[7] If we consider changing of coordinates then jacobian appears: If u, v are new coordinates, $x = x(u, v)$, $y = y(u, v)$ are new coordinates then

$$\int a(x, y)dxdy = \int a(x(u, v), y(u, v)) \det \begin{pmatrix} x_u & x_v \\ x_u & x_v \end{pmatrix} dudv \tag{2.46}$$

In formula(2.44) it appears under as a part of coefficient of differential form.