Polylinear part
of advanced linear algerba MAT 315

Oleg Viro

1. Dual space

1.1. Vectors and covectors

Let V be a vector space over a field $\mathbb{F}$. Linear maps $V \to \mathbb{F}$ has many names. In Axler's textbook they are called *functionals*. They are called also *linear forms*, *linear functionals*, *dual vectors* and *covectors*. Below we call them *covectors*.

1.2. Dual vector space

Let V be a vector space over a field $\mathbb{F}$.

The set $\mathcal{L}(V, \mathbb{F})$ of all covectors is said to be *dual* to V and denoted by $V^{\checkmark}$.

As we know, this is a vector space over $\mathbb{F}$. Recall that the linear operations in $V^{\checkmark}$ are defined by formulas

$$(\varphi + \psi)(u) = \varphi(u) + \psi(u) \quad \text{for } \varphi, \psi \in V^{\checkmark} \text{ and } u \in V,$$

$$(a\varphi)(u) = a(\varphi(u)) \quad \text{for } \phi \in V^{\checkmark}, a \in \mathbb{F} \text{ and } u \in V.$$

1.3. Dual linear map

Let $T : V \to W$ be a linear map.

Define a map $T^{\checkmark} : W^{\checkmark} \to V^{\checkmark}$ by formula $T^{\checkmark}(\varphi) = \varphi \circ T$.
The map $T^{\checkmark}$ is said to be *dual* to T.

1.A. Theorem. *For a linear map T, the dual map $T^{\checkmark}$ is linear.*

Proof. $T^{\checkmark}(\varphi + \psi)(u) = (\varphi + \psi)(T(u))$
$= \varphi(T(u)) + \psi(T(u))$
$= T^{\checkmark}\varphi(u) + T^{\checkmark}\psi(u)$
and $T^{\checkmark}(a\varphi)(u) = (a\varphi)(T(u))$
$= a(\varphi(T(u)))$
$= a(T^{\checkmark}\varphi)(u)$. □

1.B. Theorem. $\mathrm{id}^{\checkmark} = \mathrm{id}$ *and* $(T \circ S)^{\checkmark} = S^{\checkmark} \circ T^{\checkmark}$. □

1.C Corollary. *If T is an isomorphism then $T^{\checkmark}$ is an isomorphism, and $(T^{\checkmark})^{-1} = (T^{-1})^{\checkmark}$.* □

1.4. Duality between monomorphisms and epimorphisms

1.D Theorem. Duality between surjective and injective
Let V and W be finite dimensional vector spaces, and let $T : V \to W$ be a linear map. Then
- *If T is injective, then $T^{\checkmark} : W^{\checkmark} \to V^{\checkmark}$ is surjective.*
- *If T is surjective, then $T^{\checkmark} : W^{\checkmark} \to V^{\checkmark}$ is injective.*

1.E Lemma. Injective $\iff$ left invertible
Under assumptions of 1.D,
T is injective $\iff$ $\exists$ a linear map $S : W \to V$ such that $S \circ T = \mathrm{id}$.

Proof of Lemma 1.E.
$\Longrightarrow$ Assume that T is injective. Choose a basis $u = (u_1, \dots, u_p)$ of V. Then $(Tu_1, \dots, Tu_p)$ are linearly independent and can be extended to a basis of W. Define $S : W \to V$ on this basis by mapping Tu_i back to u_i for $i = 1, \dots, p$ and mapping the rest of the basis to 0. Then $ST : u_i \mapsto u_i$. Hence $S \circ T = \mathrm{id}$. □
$\Longleftarrow$ Let $u \in \operatorname{null} T$. The $ST(u) = S0 = 0$. On the other hand, $ST(u) = \mathrm{id}(u) = u$. Hence $u = 0$. Thus $\operatorname{null} T = 0$ and T is injective. □

1.F Lemma. Surjective $\iff$ right invertible
Under assumptions of 1.D,
T is surjective $\iff$ $\exists$ a linear map $S : W \to V$ such that $T \circ S = \mathrm{id}$.

Proof of Lemma 1.F.
$\implies$ Assume that T is surjective. Choose a basis $w = (w_1, \dots, w_p)$ of W. Since T is surjective, $T^{-1}(w_i) \neq \varnothing$ for each i. Choose $v_i \in T^{-1}(w_i)$. There exists a unique linear map $S : W \to V$ such that $S(w_i) = v_i$ for each i. Then $TS(w_i) = T(v_i) = w_i$. Hence $T \circ S = \mathrm{id}$. $\square$
$\impliedby$ Let $u \in W$. Since $T \circ S = \mathrm{id}$, $u = TS(u) = T(S(u))$. Hence $u \in \operatorname{range} T$. Hence $W = \operatorname{range} T$ and T is surjective. $\square$

Proof of Theorem 1.D.
Assume T is injective. Then by Lemma 1.E there exists a linear map $S : W \to V$ with $S \circ T = \mathrm{id}$. By Theorem 1.B, $T^{\checkmark} \circ S^{\checkmark} = (S \circ T)^{\checkmark} = \mathrm{id}^{\checkmark} = \mathrm{id}$. Hence, by Lemma 1.F, $T^{\checkmark}$ is surjective.

Assume T is surjective. Then by Lemma 1.F there exists a linear map $S : W \to V$ with $T \circ S = \mathrm{id}$. By Theorem 1.B, $S^{\checkmark} \circ T^{\checkmark} = (T \circ S)^{\checkmark} = \mathrm{id}^{\checkmark} = \mathrm{id}$. Hence, by Lemma 1.E, $T^{\checkmark}$ is injective. $\square$

Remark. Right and left invertibility of a morphism can be defined for morphisms of any category:
a morphism f is *left invertible*
if there exists a morphism g such that $g \circ f = \mathrm{id}$;
a morphism f is *right invertible*
if there exists a morphism g such that $f \circ g = \mathrm{id}$.
On the other hand, surjectivity and injectivity of a linear map are defined in terms of elements. These are notions from the set theory. Lemmas 1.F and 1.E relate them to right and left invertibilities for the category of vector spaces. This allows to prove that surjactivity and injectivity are dual to each other, because right and left invertibilities are dual to each other.

1.5. Space dual to the coordinate vector space

The dual to $\mathbb{F}^0$ is $\mathbb{F}^0$. Indeed, any linear map maps 0, the only element of $\mathbb{F}^0$, to 0. Thus there is only one linear map $\mathbb{F}^0 \to \mathbb{F}$, and this is zero.

$\mathbb{F}^{\checkmark} = \mathbb{F}$. Indeed, an element of $\mathbb{F}^{\checkmark} = \mathcal{L}(\mathbb{F}, \mathbb{F})$, that is a linear map $\mathbb{F} \to \mathbb{F}$. It is defined, due to its linearity, by the image of 1, and any element of $\mathbb{F}$ may be the image of 1.

Recall the following theorem.

1.G. Parametrization of $\mathcal{L}(\mathbb{F}^n, V)$ by lists of n vectors
There is a natural bijection $V^n \to \mathcal{L}(\mathbb{F}^n, V)$. It maps a list $u = (u_1, \dots, u_n) \in V^n$ to a linear map

$$T_u : \mathbb{F}^n \to V : (x_1, \dots, x_n) \mapsto \sum_{i=1}^{n} x_i u_1.$$

The inverse map maps $T : \mathbb{F}^n \to V$ to the list $(T(\mathfrak{e}_1), \dots T(\mathfrak{e}_n))$, where $\mathfrak{e}_1, \dots, \mathfrak{e}_n \in \mathbb{F}^n$ are the standard basis vectors. □

1.H. Self-duality to the coordinate space: $(\mathbb{F}^n)^\checkmark = \mathbb{F}^n$.

Indeed, according to 1.G, we have a bijection $(\mathbb{F}^n)^\checkmark = \mathcal{L}(\mathbb{F}^n, \mathbb{F}) \to \mathbb{F}^n$. A covector $\varphi : \mathbb{F}^n \to \mathbb{F}$ corresponds to the list $(\varphi(\mathfrak{e}_1), \dots, \varphi(\mathfrak{e}_n)) \in \mathbb{F}^n$, which can be an arbitrary element of $\mathbb{F}^n$. Verify that this bijection is linear. □

The values $\varphi_1 = \varphi(\mathfrak{e}_1), \dots, \varphi_1 = \varphi(\mathfrak{e}_n)$ of a functional φ on the standard basis vectors $\mathfrak{e}_1, \dots, \mathfrak{e}_n$ can be considered as coordinates of φ in $(\mathbb{F}^n)^\checkmark$.

The basis $\mathfrak{e}^1, \dots, \mathfrak{e}^n$ of $(\mathbb{F}^n)^\checkmark$ corresponding to these coordinates is defined by formulas $\mathfrak{e}^j(x_1, \dots, x_n) = x_j$.

Indeed, for any $\varphi \in (\mathbb{F}^n)^\checkmark$ and $x = (x_1, \dots, x_n) \in \mathbb{F}^n$ we have

$$\begin{aligned}
\varphi(x) &= \varphi(\sum_{i=1}^{n} x_i \mathfrak{e}_i) \\
&= \sum_{i=1}^{n} x_i \varphi(\mathfrak{e}_i) \\
&= \sum_{i=1}^{n} \mathfrak{e}^i(x_1, \dots, x_n)\varphi(\mathfrak{e}_i) \\
&= \sum_{i=1}^{n} \varphi(\mathfrak{e}_i)\mathfrak{e}^i(x) = \sum_{i=1}^{n} \varphi_i \mathfrak{e}^i.
\end{aligned}$$

Thus $\varphi = \sum_{i=1}^{n} \varphi_i \mathfrak{e}^i$. □

In particular,

$$\mathfrak{e}^j(\mathfrak{e}_i) = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}.$$

Here it is convenient to use the Kronecker delta symbol, which is defined by formula

$$\delta_i^j = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$$

With the Kronecker delta the relation between $\mathfrak{e}_i$ and e^j looks as follows: $\mathfrak{e}^j(e_i) = \delta_i^j$.

1.I Corollary. *Any finite dimensional vector space V is isomorphic to its dual vector space $V^{\checkmark}$.*

Proof. V is isomorphic to $\mathbb{F}^n$. Hence, $V^{\checkmark}$ is isomorphic to $(\mathbb{F}^n)^{\checkmark}$ by 1.C. As we have just seen, $(\mathbb{F}^n)^{\checkmark}$ is isomorphic to $\mathbb{F}^n$. $\square$

A basis $(v_1, \dots, v_n)$ of V defines an isomorphism $T : V \to \mathbb{F}^n$ which maps the basis $(v_1, \dots, v_n)$ of V to the standard basis $(\mathfrak{e}_1, \dots, \mathfrak{e}_n)$ of $\mathbb{F}^n$. The dual map $T^{\checkmark} : (\mathbb{F}^n)^{\checkmark} \to V^{\checkmark}$ maps the standard dual basis $(\mathfrak{e}^1, \dots, \mathfrak{e}^n)$ to some basis of $V^{\checkmark}$. Its elements $T^{\checkmark}\mathfrak{e}^1, \dots, T^{\checkmark}\mathfrak{e}^n$ are denoted by $v^1, \dots, v^n$. The basis $(v^1, \dots, v^n)$ $V^{\checkmark}$ is said to be ***dual*** to the basis $(v_1, \dots, v_n)$. The basis $(v^1, \dots, v^n)$ is related to $(v_1, \dots, v_n)$ by formula $v^j(v_i) = \delta_i^j$, which is similar to formula $\mathfrak{e}^j(\mathfrak{e}_i) = \delta_i^j$ established above, and, in fact, follows from it: $v^j(v_i) = T^{\checkmark}\mathfrak{e}^j(v_i) = \mathfrak{e}^j(Tv_i) = \mathfrak{e}^j(\mathfrak{e}_i) = \delta_i^j$.

1.6. The second dual

The proof of Corollary 1.I is indirect. To construct an isomorphism $V \to V^{\checkmark}$, we use an isomorphism between V and $\mathbb{F}^n$. The isomorphism $V \to V^{\checkmark}$ obtained in this way depends on the choice of isomorphism $V \to \mathbb{F}^n$. This dependence is not a defect of our presentation. There is no canonical isomorphism between V and $V^{\checkmark}$.

Contrary to this, the space $(V^{\checkmark})^{\checkmark}$ which is dual to $V^{\checkmark}$ is canonically isomorphic to V, as we will see in this section.

1.J. Theorem. Canonical map to the second dual
Let V be a vector space over a field $\mathbb{F}$. There is a canonical linear map $V \to (V^{\checkmark})^{\checkmark}$. It is defined by formula $u \mapsto (V^{\checkmark} \to \mathbb{F} : \varphi \mapsto \varphi(u))$.

Proof. Linearity of the map $V^{\checkmark} \to \mathbb{F} : \varphi \mapsto \varphi(u)$ (that is its belonging to $(V^{\checkmark})^{\checkmark}$) follows immediately from the definition of linear operations in $V^{\checkmark}$: $(a\varphi + b\psi)u = a\varphi(u) + b\psi(u)$. Hence this formula defines a map $V \to (V^{\checkmark})^{\checkmark}$. This map is linear. Indeed, $\varphi(au + bv) = a\varphi(u) + b\varphi(v)$ by linearity of φ. $\square$

The construction of the map $V \to (V^{\checkmark})^{\checkmark}$ above does not involve any choice, it is natural and universal.

1.K. Theorem. *If V is finite dimensional then the natural map $V \to (V^{\checkmark})^{\checkmark}$ is an isomorphism.*

1.L Lemma. *If V is finite dimensional, then the natural map $V \to (V^{\checkmark})^{\checkmark}$ is injective.*

Proof of Lemma 1.L. Let $u \in V$ be a non-zero vector. As a list of vectors which consists of a single non-zero vector, u is linear independent. Hence, it can be included into a basis of V. Let u, u_1, ..., u_n be such a basis. The first covector of the dual basis takes value 1 on u. Therefore the image of u under the canonical map $V \to (V^{\checkmark})^{\checkmark}$ takes value 1 on this covector. Hence u is not in the kernel of $V \to (V^{\checkmark})^{\checkmark}$. So, we proved that any non-zero vector does not belong to the kernel. Thus, the map is injective. □

Proof of Theorem 1.K. By 1.I, any finite dimensional vector space V is isomorphic to its dual $V^{\checkmark}$, which, in turn, is isomorphic to its dual $(V^{\checkmark})^{\checkmark}$. Thus, V and $(V^{\checkmark})^{\checkmark}$ have the same dimension. Hence our map $V \to (V^{\checkmark})^{\checkmark}$, being an injective linear map between vector spaces of the same finite dimension, is an isomorphism. □

Remark. In the proof of Lemma 1.L given above, we used the assumption that V is finite dimensional. However, Lemma 1.L holds true without this assumption. For any non-zero vector in any vector space one can find a linear functional which takes non-zero value on this vector. However, in general case a construction of such a functional requires tools that we do not need. Nonetheless, Theorem 1.K cannot be extended to an infinite dimensional situation: for any infinite dimensional space V the canonical map $V \to (V^{\checkmark})^{\checkmark}$ is not an isomorphism.

1.7. Bracket, bra and ket

We see that in the finite dimensional case a vector space and its dual have the same dimension and play symmetric rôles: the space dual to $V^{\checkmark}$ is identified with V. So, if we denote $V^{\checkmark}$ by U, then V becomes $U^{\checkmark}$. This suggests to make notations more symmetric. Let us denote the value $\varphi(u)$ taken by a linear functional $\varphi \in V^{\checkmark}$ on a vector $u \in V$ by $\langle\varphi|u\rangle$. This defines a map

$$V^{\checkmark} \times V \to \mathbb{F} : (\varphi, u) \mapsto \langle\varphi|u\rangle.$$

This map is *bilinear*, that is it is linear against each of the arguments: $\langle\varphi|au + bv\rangle = a\langle\varphi|u\rangle + b\langle\varphi|v\rangle$ and $\langle a\varphi + b\psi|u\rangle = a\langle\varphi|u\rangle + b\langle\psi|u\rangle$. The first of these equalities is linearity of φ, the second, definition of linear operations with covectors. Bilinearity means that if one of the arguments in the bracket is fixed, then the bracket turns to a linear map.

The definition of dual linear map gets a new look under the bracket notation. Recall that map $T^{\checkmark}$ dual to a linear map T is defined by identity $T^{\checkmark}(\varphi) = \varphi \circ T$, which on the level of vectors can be rewritten as follows: $T^{\checkmark}(\varphi)(u) = \varphi(Tu)$. In the bracket notation this equality looks as follows: $\langle T^{\checkmark}\varphi|u\rangle = \langle\varphi|Tu\rangle$.

There is a tendency, especially among physicists, to identify an object with its action to other objects, provided that this action can characterize the acting object completely. From this point of view, a covector is a functional on vectors (and we have accepted this in our definition of a covector) and a vector is a functional on vectors (although this is NOT how we have introduced vectors!).

Dirac went further. He suggested to rename vectors and covectors according these their rôles in the bracket. According to Dirac, covectors should always be enclosed in the left-hand half of the bracket, like this: $\langle\varphi|$, and called *bra vectors*, while vectors should be dressed in the right-hand half of the bracket, like that: $|u\rangle$, and called *ket vectors*.

Bra space (that is the vector space of bra vectors) comes for free together with any vector (ket) space. Let us formalize this structure.

Let V and W be vector spaces over $\mathbb{F}$. A map $b : V \times W \to \mathbb{F}$ is called a *bilinear pairing*, if it is linear against each of the variables. To a bilinear pairing $b : V \times W \to \mathbb{F}$ one associates two linear maps,

$$V \to W^{\checkmark} : v \mapsto (w \mapsto b(v,w)) \quad \text{and} \quad W \to V^{\checkmark} : w \mapsto (v \mapsto b(v,w)).$$

In the case of the canonical pairing $V^{\checkmark} \times V \to \mathbb{F}$ the associated linear maps are the identities $V^{\checkmark} \to V^{\checkmark}$ and $V \to V$. A bilinear pairing $b : V \times W \to \mathbb{F}$ is said to be *non-singular*, if the associated maps are isomorphisms. A non-singular bilinear pairing $b : V \times W \to \mathbb{F}$ is essentially the canonical pairing $W^{\checkmark} \times W \to \mathbb{F}$: at least, replacement V by $W^{\checkmark}$ by the associated isomorphism $V \to W^{\checkmark}$ turns b to the canonical pairing.

1.8. Matrix of dual map

1.M. Theorem. *Let $(v_1, \dots, v_p)$ be a basis of a vector space V and $(w_1, \dots, w_q)$ be a basis of vector space W. Let $T : V \to W$ be a linear map with matrix A with respect to the bases $(v_1, \dots, v_p)$ and $(w_1, \dots, w_q)$. Then the matrix of $T^{\checkmark} : W^{\checkmark} \to V^{\checkmark}$ with respect to the dual bases $(v^1, \dots, v^p)$ and $(w^1, \dots, w^q)$ is obtained from A by transposition of rows and columns.*

Proof. Matrix A consists of scalars a_{ij} such that $Tv_i = \sum_{k=1}^{q} a_{ki}w_k$. Let $B = (b_{ij})$ be the matrix of the dual map $T^{\checkmark}$ with respect to the dual bases. This means that $T^{\checkmark} w^j = \sum_{k=1}^{p} b_{kj}w^k$.

Consider $\langle w^j | Tv_i\rangle$. On one hand,

$$\langle w^j | Tv_i\rangle = \left\langle w^j \middle| \sum_{i=1}^{q} a_{ki}w_k \right\rangle = \sum_{i=1}^{q} a_{ki}\langle w^j | w_k\rangle = \sum_{i=1}^{q} a_{ki}\delta_k^j = a_{ji}.$$

On the other hand, $\langle w^j | T v_i \rangle = \langle T^{\checkmark} w^j | v_i \rangle$ and

$$\langle T^{\checkmark} w^j | v_i \rangle = \left\langle \sum_{k=1}^{p} b_{kj} v^k \middle| v_i \right\rangle = \sum_{k=1}^{p} b_{kj} \langle v^k | v_i \rangle = \sum_{k=1}^{p} b_{kj} \delta_i^k = b_{ij},$$

where (b_{ik}) is the matrix of the dual map. We see that $a_{ji} = b_{ij}$, that is the ij entry of the matrix (b_{ij}) coincides with the symmetric ji entry of A. □

1.9. Rank of dual map

Recall that the *rank* of a linear map is the dimension of its range, the rank of T is denoted by $\operatorname{rk} T$.

1.N. Theorem. *The ranks of a linear maps dual to each other are equal.*

Proof. Any linear map $T : V \to W$ is represented as a composition $V \xrightarrow{S} \operatorname{range} T \xrightarrow{R} W$ of the surjection S defined by T (that is $Su = Tu$ for $u \in V$) and an inclusion $\operatorname{range} T \xrightarrow{R} W$.

The rank of a composition of a surjection followed by an injection is equal to the rank of each of them. Hence $\operatorname{rk} S = \operatorname{rk} R = \dim \operatorname{range} T = \operatorname{rk} T$.

The dual maps give rise to a decomposition $W^{\checkmark} \xrightarrow{R^{\checkmark}} (\operatorname{range} T)^{\checkmark} \xrightarrow{S^{\checkmark}} V^{\checkmark}$ of $T^{\checkmark}$. By 1.D, $S^{\checkmark}$ is injective and $R^{\checkmark}$ surjective. Hence $\operatorname{rk} S^{\checkmark} = \operatorname{rk} R^{\checkmark} = \operatorname{rk} T^{\checkmark} = \dim(\operatorname{range} T)^{\checkmark}$. Spaces $\operatorname{range} T$ and $(\operatorname{range} T)^{\checkmark}$ are finite dimensional dual. Hence their dimensions are equal. □

1.O Corollary. *For any matrix the maximal number of its linearly independent rows is equal to the maximal number of its linearly independent columns.*

Proof. The maximal number of linearly independent columns of a matrix is equal to the rank of the linear map defined by it. The maximal number of linearly independent rows of a matrix is equal to the rank of the dual linear map, by Theorem 1.M. By Theorem 1.N these two ranks are equal. □

1.10. Improving matrix notation

Traditionally, vectors of $\mathbb{F}^n$ in matrix notation are associated with matrices-columns. This is motivated by our addiction to functional notation: the image of a vector u under a linear map T is denoted by Tu, and if we associate to u a matrix-column X and to map T, its matrix A, then AX is the matrix-column associated to Tu.

If we dared to associate a matrix-row to a vector, it would be possible to multiply it from left only by a matrix-column. This cannot be corresponding to linear maps, because the dimension of the space of linear maps is greater than the dimension of the space of matrix-columns. Hence either we would need to use multiplication by a matrix from right, or to modify the very multiplication of matrices.

The choice of notation used above is commonly accepted, and we speculate on other possibilities not because we consider seriously change of a commonly accepted notation, but because we want to prepare the next twist of notation's development. Here are the new rules.

(1) *Covectors are associated to matrix-rows.*
(2) *Bracket pairing of covectors and vectors is identified with multiplication of matrices.*
(3) *Dual maps are represented in dual bases by the same matrix.*
(4) *Matrix representation of a linear map between dual spaces (whose elements are represented by matrices-rows) is multiplication of a matrix-row by a matrix from the right hand side.*

The first rule does not require a formal justification. It means that an element $\sum_{i=1}^n x_i \mathfrak{e}^i$ of $(\mathbb{F}^n)^\checkmark$ is represented by matrix

$$\begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix}.$$

This is similar to the traditional representation of a vector $\sum_{i=1}^n x_i \mathfrak{e}_i \in \mathbb{F}^n$ by matrix

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

The second rule requires proof. Here it is. For $\sum_{i=1}^n x_i \mathfrak{e}^i \in (\mathbb{F}^n)^\checkmark$ and $\sum_{i=1}^n y_i \mathfrak{e}_i \in \mathbb{F}^n$

$$\left\langle \sum_{i=1}^n x_i \mathfrak{e}^i \,\middle|\, \sum_{j=1}^n y_j \mathfrak{e}_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \langle \mathfrak{e}^i | \mathfrak{e}_j \rangle = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \delta^i_j = \sum_{i=1}^n x_i y_i = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Consider now the last two rules. Let $T : \mathbb{F}^p \to \mathbb{F}^q$ be a linear map and A be the matrix of T in the standard bases. The dual map $T^\checkmark$ is defined by $\langle T^\checkmark x|y\rangle = \langle x|Ty\rangle$ for all $x \in (\mathbb{F}^q)^\checkmark$ and $y \in \mathbb{F}^p$. Let $x = \sum_{i=1}^q x_i\mathfrak{e}^i$ and $y = \sum_{j=1}^p y_j\mathfrak{e}_j$. Denote matrix $\begin{pmatrix} x_1 & x_2 & \dots & x_q\end{pmatrix}$ by X and matrix

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix}$$

by Y. Then Ty is represented by matrix AY. Hence $\langle x|Ty\rangle = XAY$. Fix x. The covector $T^\checkmark x \in (\mathbb{F}^p)^\checkmark$ is defined by equality $\langle T^\checkmark x|y\rangle = \langle x|Ty\rangle$ for all $y \in \mathbb{F}^p$. Denote the matrix-row representing $T^\checkmark x$ by Z. The equality turns into $ZY = XAY$. It holds true for all Y. Taking Y with all entries 0 but one, we see that each entry of Z equals the corresponding entry of XA. It follows $Z = XA$. □

1.11. The Einstein notation

In formulas used in the mathematical literature objects are denoted by letters, but, due to lack of letters in the commonly used alphabets, and necessity to denote similar objects numerated by numbers, different things are denoted by the same letter equipped with indices.

Mathematicians place indices on the right hand side and below the main symbol. Usually they hesitate to put index on the left hand side of the main symbol (because of writing from left to right) or to the upper position on the right hand side of the main symbol (since this position is reserved for exponents). However, in situations when there are too many indices of different nature, these objections do not work.

This happens in the classical notation of polylinear algebra used extensively in the physics and geometry literature. Until this point we used mainly lower indices. The only exception appeared in notation for dual basis. In fact, this exception is the first manifestation of the whole system, according to which about half of all indices should be upper.

Basis vectors in a vector space are numerated by lower indices, as we did: $\mathfrak{e}_i$. The coordinates of vectors are to be equipped with upper indices, like this: $(x^1, \dots, x^n) \in \mathbb{F}^n$ and $v = \sum_{i=1}^n x^i\mathfrak{e}_i \in \mathbb{F}^n$. Vectors in the basis dual to a basis $v_1, \dots, v_n \in V$ are numerated with upper indices (as we did): $v^1, \dots, v^n \in V^\checkmark$. Coordinates of a covector with respect to a basis $v^1, \dots, v^n \in V^\checkmark$ are numerated with low indices: $x_1v^1+x_2v^2+\dots+x_nv^n \in V^\checkmark$.

In sums each index of summation appears twice, once as a lower index and once as an upper index. This is so usual that there is an agreement to skip summation sign in such a situation (i.e., when in a formula an index appears twice once as lower and once as upper index). For example, formula $x_i\mathfrak{e}^i$ should be understood as $\sum_i x_i\mathfrak{e}^i$. The range of summation is determined from the context.

We will use this skipping of a summation sign cautiously, repeating the same formulas with the summation sign in order to reduce the risk of confusion, until the reader will get comfortable with the shorthand notation and appreciate its flexibility and convenience.

Recall that entries of the matrix of a linear map are involved in the following formulas: the image of the basis vector $\mathfrak{e}_j$ under the linear map with matrix (a_{ij}) is $\sum_{i=1}^m a_{ij}\mathfrak{e}_i$ and the ith coordinate of the image of vector $(x_1, \dots, x_n)$ is $\sum_{j=1}^n a_{ij}x_j$. The first formula suggests to raise the first index of the entry a_{ij}. Then it would take the shape $\sum_{i=1}^m a^i_j\mathfrak{e}_i$ or even $a^i_j\mathfrak{e}_i$ (by skipping the summation sign). In the second formula we have to raise, first, the index at x_j, as it was stated above, and then raising the first index at the matrix entry would make it perfect: $\sum_{j=1}^n a^i_j x^j$. Again, we can skip the summation sign and shorthand $\sum_{j=1}^n a^i_j x^j$ till $a^i_j x^j$.

Thus, in matrices that we met so far, the index numerating lines should be raised to the upper position, while the index numerating rows should be left in the lower position. Then skipping sigmas makes notation very similar to matrix notation: we write $a^i_j x^j$ instead of AX, and, say, matrix expression XAY for $\langle x|Ty\rangle$ that we discussed in 1.10 turns to $x_i a^i_j y^j$ where double summation (both over i and j) is understood. However this similarity falls short when the number of indices increases. It could be preserved if one could use high dimensional matrices.

2. Tensors

2.1. Polylinear maps

Let $V_1, \dots V_n$, W be vector spaces over a field $\mathbb{F}$. A map

$$F : V_1 \times \cdots \times V_n \to W : (v_1, \dots, v_n) \mapsto F(v_1, \dots, v_n)$$

is said to be *polylinear* or *multilinear*, if it is linear as a function of each of its arguments, when the other arguments are fixed. In other words,

$$F(v_1, \dots, v_{i-1}, x + y, v_{i+1}, \dots, v_n) =$$
$$F(v_1, \dots, v_{i-1}, x, v_{i+1}, \dots, v_n) + F(v_1, \dots, v_{i-1}, y, v_{i+1}, \dots, v_n),$$
$$F(v_1, \dots, v_{i-1}, av_i, v_{i+1}, \dots, v_n) = aF(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_n)$$

for $i = 1, \dots, n$, $a \in \mathbb{F}$. If $W = \mathbb{F}$, a polylinear map is called also a *polylinear function*, or *polylinear functional*, or *polylinear form*.

The set of all polylinear maps $V_1 \times \cdots \times V_n \to W$ is denoted by $\mathcal{L}(V_1, \ldots, V_n; W)$. This is a subspace of the vector space of all maps $V_1 \times \cdots \times V_n \to W$.

2.2. Tensor algebra of a vector space

Let V be a finite dimensional vector space over $\mathbb{F}$. A polylinear functional

$$T : \underbrace{V \times \cdots \times V}_{p \text{ times}} \times \underbrace{V^{\checkmark} \times \cdots \times V^{\checkmark}}_{q \text{ times}} \to \mathbb{F}$$

is called a *tensor* on V of *type* (p, q) and *order* or *valency* $p+q$. It is also said to be a mixed tensor p times *covariant* and q times *contravariant*.

Denote by $\mathrm{Tens}_p^q(V)$ the set of all tensors on a vector space V of type (p, q). As a subspace of $\mathcal{L}(V, \ldots, V, V^{\checkmark}, \ldots, V^{\checkmark}; \mathbb{F})$, $\mathrm{Tens}_p^q(V)$ is a vector space over the same ground field $\mathbb{F}$ as V. If one of the numbers p and q is zero, it is not mentioned in the notation $\mathrm{Tens}_p^q(V)$. Then we write $\mathrm{Tens}_p(V)$ or $\mathrm{Tens}^q(V)$.

Special cases:

- A tensor $V \to \mathbb{F}$ of type $(1, 0)$ is a covector. Thus $\mathrm{Tens}_1(V) = V^{\checkmark}$.
- A tensor $V^{\checkmark} \to \mathbb{F}$ of type $(0, 1)$ is an element of the double dual space $(V^{\checkmark})^{\checkmark}$, and, via the canonical identification of $(V^{\checkmark})^{\checkmark}$ with V, this is a vector. Thus $\mathrm{Tens}^1(V) = V$.
- A tensor $V \times V \to \mathbb{F}$ of type $(2, 0)$ is a bilinear form on V.
- A tensor $V \times V^{\checkmark} \to \mathbb{F}$ of type $(1, 1)$ defines (and is defined by) a linear map $V \to (V^{\checkmark})^{\checkmark} = V$, thus it is identified with an operator $V \to V$. Therefore $\mathrm{Tens}_1^1(V) = \mathcal{L}(V)$.

2.3. Coordinates in the spaces of tensors

Let $\mathfrak{e}_1, \ldots, \mathfrak{e}_n$ be a basis in a vector space V and $\mathfrak{e}^1, \ldots \mathfrak{e}^n$ be the dual basis in $V^{\checkmark}$. Consider a tensor $T : V^p \times (V^{\checkmark})^q \to \mathbb{F}$. It is defined by its values on lists of base vectors

$$T(\mathfrak{e}_{i_1}, \ldots, \mathfrak{e}_{i_p}, \mathfrak{e}^{j_1}, \ldots, \mathfrak{e}^{j_q}) = T_{i_1,\ldots,i_p}^{j_1,\ldots,j_q}$$

These values are called *coordinates* of T. A tensor of type (p, q) on a vector space of dimension n has n^{p+q} coordinates. A tensor, as a polylinear function on vectors $v_1, \ldots, v_p$ and covectors $u^1, \ldots, u^q$ is determined by its coordinates as follows:

$$
\begin{aligned}
&T(v_1,\dots,v_p,u^1,\dots,u^q)\\
&\qquad= T(v_1^{i_1}\mathfrak{e}_{i_1},\dots,v_p^{i_p}\mathfrak{e}_{i_p},u^1_{j_1}\mathfrak{e}^{j_1},\dots,u^q_{j_q}\mathfrak{e}^{j_q})\\
&\qquad= v_1^{i_1}\dots v_p^{i_p}u^1_{j_1}\dots u^q_{j_q}T(\mathfrak{e}_{i_1},\dots,\mathfrak{e}_{i_p},\mathfrak{e}^{j_1},\dots,\mathfrak{e}^{j_q})\\
&\qquad= T^{j_1,\dots,j_q}_{i_1,\dots,i_p}v_1^{i_1}\dots v_p^{i_p}u^1_{j_1}\dots u^q_{j_q}
\end{aligned}
$$

In this formula, we use the Einstein notation. The vector v_k is presented as $\sum_i v_k^i\mathfrak{e}_i$ and we skip the summation symbol, so that $v_k = v_k^i\mathfrak{e}_i$. Similarly, for covectors $u^k = \sum_j u_j^k\mathfrak{e}^j = u_j^k\mathfrak{e}^j$.

The coordinates $T^{j_1,\dots,j_q}_{i_1,\dots,i_p}$ of a tensor T are its coordinates with respect to the basis $\mathfrak{e}^{i_1,\dots,i_p}_{j_1,\dots,j_q}$ in $\mathrm{Tens}^q_p(V)$ in the sense that any tensor $T\in\mathrm{Tens}^q_p(V)$ can be presented as a linear combination of the base tensors $\mathfrak{e}^{i_1,\dots,i_p}_{j_1,\dots,j_q}$ with coefficients $T^{j_1,\dots,j_q}_{i_1,\dots,i_p}$, that is $T = T^{j_1,\dots,j_q}_{i_1,\dots,i_p}\mathfrak{e}^{i_1,\dots,i_p}_{j_1,\dots,j_q}$. The tensor $\mathfrak{e}^{i_1,\dots,i_p}_{j_1,\dots,j_q}$ is defined by formula $(\mathfrak{e}^{i_1,\dots,i_p}_{j_1,\dots,j_q})(v_1,\dots,v_q,u^1,\dots,u^q) = v_1^{i_1}\dots v_p^{i_p}u^1_{j_1}\dots u^q_{j_q}$, where $v_i = v_i^k\mathfrak{e}_k$ and $u^j = u^j_k\mathfrak{e}^k$, as above.

2.4. Change of basis

Under a change of basis in V, the new basis is expressed in terms of the old one according to formula

$$\tilde{\mathfrak{e}}_\alpha = C^i_\alpha\mathfrak{e}_i = \sum_i C^i_\alpha\mathfrak{e}_i$$

and the old basis is expressed in terms of the new by formula

$$\mathfrak{e}_\alpha = \tilde{C}^i_\alpha\tilde{\mathfrak{e}}_i,$$

where $\tilde{C}^i_\alpha$ are terms of the matrix inverse to the transition matrix C^α_i, so that $\tilde{C}^i_\alpha C^\alpha_j = \delta^i_j$ and $\tilde{C}^i_\alpha C^\beta_i = \delta^\beta_\alpha$. Then the coordinates of tensor changes by the following formula:

$$\tilde{T}^{\beta_1\dots,\beta_q}_{\alpha_1,\dots,\alpha_p} = T^{j_1,\dots,j_q}_{i_1,\dots,i_p}C^{i_1}_{\alpha_1}\dots C^{i_p}_{\alpha_p}\tilde{C}^{\beta_1}_{j_1}\dots\tilde{C}^{\beta_q}_{j_q}.$$

Indeed,

$$
\begin{aligned}
\tilde{T}^{\beta_1\dots,\beta_q}_{\alpha_1,\dots,\alpha_p} &= T(\tilde{\mathfrak{e}}_{\alpha_1},\dots,\tilde{\mathfrak{e}}_{\alpha_p},\tilde{\mathfrak{e}}^{\beta_1},\dots,\tilde{\mathfrak{e}}^{\beta_q})\\
&= T(C^{i_1}_{\alpha_1}\mathfrak{e}_{i_1},\dots,C^{i_p}_{\alpha_p}\mathfrak{e}_{i_p},\tilde{C}^{\beta_1}_{j_1}\mathfrak{e}^{j_1},\dots,\tilde{C}^{\beta_q}_{j_q}\mathfrak{e}^{j_q})\\
&= C^{i_1}_{\alpha_1}\dots C^{i_p}_{\alpha_p}\tilde{C}^{\beta_q}_{j_q}\dots\tilde{C}^{\beta_q}_{j_q}\ T(\mathfrak{e}_{i_1},\dots,\mathfrak{e}_{i_p},\mathfrak{e}^{j_1},\dots,\mathfrak{e}^{j_q})\\
&= C^{i_1}_{\alpha_1}\dots C^{i_p}_{\alpha_p}\tilde{C}^{\beta_q}_{j_q}\dots\tilde{C}^{\beta_q}_{j_q}\ T^{j_1,\dots,j_q}_{i_1,\dots,i_p}
\end{aligned}
$$

2.5. Maps induced by a linear map

A linear map $F : V\to W$ defines linear maps only for Tens_k and Tens^k, for Tens^q_p with both $p\neq 0$ and $q\neq 0$, it does not induce any map.

The map $\mathrm{Tens}_k(F) : \mathrm{Tens}_k(W) \to \mathrm{Tens}_k(V)$, which is induced by a linear map $F : V \to W$, maps $T : W^k \to \mathbb{F}$ to the composition $V^k \xrightarrow{F\times\cdots\times F} W^k \xrightarrow{T} \mathbb{F}$. If there is no danger of confusion, we will use a shorthand notation F^* for $\mathrm{Tens}_k(F)$. The star here indicates that the map is induced by F, its upper index position means that it acts in the direction that is opposite to the direction of F.

The map $\mathrm{Tens}^k(F) : \mathrm{Tens}^k(V) \to \mathrm{Tens}^k(W)$, which is induced by a linear map $F : V \to W$, maps $T : (V^\checkmark)^k \to \mathbb{F}$ to the composition $(W^\checkmark)^k \xrightarrow{F^\checkmark\times\cdots\times F^\checkmark} (V^\checkmark)^k \xrightarrow{T} \mathbb{F}$. If there is no danger of confusion, we will use a shorthand notation F_* for $\mathrm{Tens}^k(F)$. The star here indicates that the map is induced by F, its low index position means that it acts in the same direction as F.

Both constructions respect compositions: for linear maps $F : U \to V$ and $G : V \to W$, the maps induced by their composition $U \xrightarrow{F} V \xrightarrow{G} W$ are the appropriate compositions of the maps induced by F and G. More specifically: $(G \circ F)^* = F^* \circ G^*$ and $(G \circ F)_* = G_* \circ F_*$. The proofs of these statements are straightforward.

2.6. Multiplication of tensors

The product of tensors (i.e., polylinear functionals)

$$T : V^p \times (V^\checkmark)^q \to \mathbb{F} \text{ and } S : V^r \times (V^\checkmark)^s \to \mathbb{F},$$

is a tensor $T \otimes S : V^p \times (V^\checkmark)^q \times V^r \times (V^\checkmark)^s \to \mathbb{F}$ defined by formula

$$\begin{aligned} T \otimes S(v_1, \dots, v_p, w^1, \dots, w^q, u_1, \dots, u_r, z^1, \dots, z^s) \\ = T(v_1, \dots, v_p, w^1, \dots, w^q)S(u_1, \dots, u_r, z^1, \dots, z^s). \end{aligned}$$

This multiplication is distributive with respect to addition of tensors. In other words, the multiplication of tensors defines a bilinear pairing

$$\mathrm{Tens}^q_p(V) \times \mathrm{Tens}^s_r(V) \to \mathrm{Tens}^{q+s}_{p+r}(V).$$

Example. Recall that the simplest tensors are vectors (i.e., elements of $\mathrm{Tens}^1(V) = V$) and covectors (elements of $\mathrm{Tens}_1(V) = V^\checkmark$). Let us fix a basis $\mathfrak{e}_1, \dots, \mathfrak{e}_n$ in V. Then $\mathfrak{e}^1, \dots, \mathfrak{e}^n$ is the dual basis (this is a basis of $V^\checkmark = \mathrm{Tens}_1(V)$). Consider $\mathfrak{e}^i \otimes \mathfrak{e}^j \in \mathrm{Tens}^0_2(V)$. This is a bilinear form $V \times V \to \mathbb{F}$. Let us calculate its value on vectors $v = v^k\mathfrak{e}_k$ and $w = w^m\mathfrak{e}_m$

$$\begin{aligned} \mathfrak{e}^i \otimes \mathfrak{e}^j : (v, w) \mapsto \langle \mathfrak{e}^i | v\rangle\langle \mathfrak{e}^j | w\rangle = \langle \mathfrak{e}^i | v^k\mathfrak{e}_k\rangle\langle \mathfrak{e}^j | w^m\mathfrak{e}_m\rangle \\ = v^k\langle \mathfrak{e}^i | e_k\rangle w^m\langle \mathfrak{e}^j | e_m\rangle = v^k\delta^i_k w^m\delta^j_m = v^i w^j \end{aligned}$$

In words: the tensor $\mathfrak{e}^i \otimes \mathfrak{e}^j$ evaluated on pair of vectors v and w gives the product of the ith coordinate of v and the jth coordinate of w. The coordinates of tensor $\mathfrak{e}^i \otimes \mathfrak{e}^j$ in the coordinate system defined by the basis $\mathfrak{e}_1, \dots, \mathfrak{e}_n$ are $(\mathfrak{e}^i \otimes \mathfrak{e}^j)_{p,q} = \delta^i_p \delta^j_q$. So, the (i, j) coordinate equals 1 and all others equal 0. This is one of the base vectors in $\mathrm{Tens}^2_0(V)$.

In general, tensors $\mathfrak{e}^{i_1} \otimes \dots \otimes \mathfrak{e}^{i_p} \otimes \mathfrak{e}_{j_1} \otimes \dots \otimes \mathfrak{e}_{j_q}$ form a basis in $\mathrm{Tens}^q_p(V)$.

Although the basis vectors of $\mathrm{Tens}^{q+s}_{p+r}(V)$ belong to the range of pairing $\mathrm{Tens}^q_p(V) \times \mathrm{Tens}^s_r(V) \to \mathrm{Tens}^{q+s}_{p+r}(V)$ the pairing is not surjective. It is not a linear, but a bilinear map. A linear map, whose range contains a basis, would be surjective. For a bilinear map, this is not true. Indeed, $\dim \mathrm{Tens}^q_p(V) \times \mathrm{Tens}^s_r(V) = n^{p+q} + n^{r+s}$ while $\dim \mathrm{Tens}^{q+s}_{p+r}(V) = n^{p+q+r+s}$, so usually the dimension of the target space is less than the dimension of the source.

For example, not any bilinear form $V \times V \to \mathbb{F}$ (i.e., a tensor of type $(2, 0)$) can be represented as a product of two covectors.

Those tensors, which can be presented as a product of tensors are called *decomposable*. Any tensor can be presented as a *sum* of products of tensors from $V^{\checkmark} = \mathrm{Tens}_1(V)$ and $V = \mathrm{Tens}^1(V)$.

There is a construction which for any two vector spaces V and W over $\mathbb{F}$ gives rise to a vector space $V \otimes W$. The dimension of $V \otimes W$ is the product of dimensions of V and W. The space $V \otimes V$ can be identified with $\mathrm{Tens}^2(V)$. It is generated by vectors which can be identified with $v \otimes u$, where $v, u \in V$.

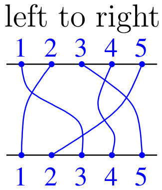
3. Symmetric and skew-symmetric

3.1. Digression on permutations and their parity

A permutation of a set X is a bijection $\sigma : X \to X$ of this set onto itself. Denote the set $\{1, 2, 3, \dots, n\}$ of the first n positive integers by $\mathbb{N}_n$. The set of all permutations of the set $\mathbb{N}_n$ is denoted by S_n and called the *symmetric group* of degree n.

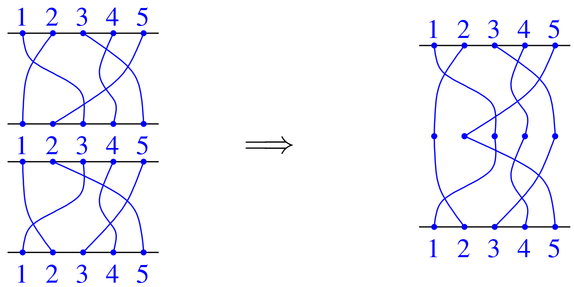
A permutation, mapping two elements of the set S to each other and every other element to itself, is called the *transposition* of these two elements.

Permutations belonging to S_n can be presented by pictures of the following type. Put n dots on two horizontal lines, numerate them from left to right by numbers from 0 to n and connect the dot k on the upper



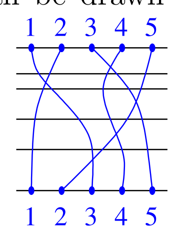
line to the dot $\sigma(k)$ on the lower line by a simple descending arc. One should draw the arcs clearly, avoiding intersection of several arcs in one point and points of tangency.

A picture for the composition $\sigma_1 \circ \sigma_1$ of permutations σ_0 and σ_2 can be obtained from the pictures for σ_0 and σ_2 by drawing them one over the other as follows.



3.A. Theorem. *Any permutation can be presented as a composition of transpositions.*

Proof. On a picture of arbitrary permutation, arcs can be drawn in such a way that no two intersection points of the arcs were on the same horizontal line. Then they can be separated from each other by horizontal lines. This gives a desired decomposition of the permutation into a composition of permutations each of which is presented by a picture with one intersection point. Those permutations are transpositions.



□

The arcs, which start at points i and j with $i < j$ and finish at $\sigma(i)$ and $\sigma(j)$, must intersect if $\sigma(i) > \sigma(j)$. They may intersect in several points, but the parity of the number of points depends on the mutual position of $\sigma(i)$ and $\sigma(j)$. Namely, if $\sigma(i) > \sigma(j)$, then the number of intersection points is odd, if $\sigma(i) < \sigma(j)$, then it is even.

A permutation which is a composition of odd number of transpositions is said to be ***odd***, otherwise it is said to be ***even***. The ***sign*** $\operatorname{sign} \sigma$ of a permutation σ is defined to be -1 if σ is odd and $+1$ if σ is even.

3.2. Symmetric tensors

A polylinear form $T : V^k \to \mathbb{F}$ is called ***symmetric*** if its values are not affected by any permutations of the arguments. In other words, T is symmetric, if, for any permutation $\sigma : \{1, \dots, k\} \to \{1, \dots, k\}$ and any $v_1, \dots, v_k \in V$,

$$T(v_1, \dots, v_k) = T(v_{\sigma(1)}, \dots, v_{\sigma(k)}).$$

For each k, symmetric polylinear functionals $V^k \to \mathbb{F}$ form a subspace of the vector space $\mathrm{Tens}_k(V)$ of all polylinear functionals $V^k \to \mathbb{F}$. This subspace is denoted by $\mathrm{Sym}_k(V)$.

Similarly, symmetric polylinear functionals $(V^{\checkmark})^k \to \mathbb{F}$ form a subspace of the space $\mathrm{Tens}^k(V)$. This subspace is denoted by $\mathrm{Sym}^k(V)$.

3.3. Digression on field's characteristic and arithmetic mean.

Consider the multiples of 1 in a field $\mathbb{F}$: $1, 1+1, 1+1+1, \dots$. Let us denote its term obtained as the sum of k units by $k \cdot 1$.

This sequence may be periodic, like in the field $\mathbb{F}_2 = \{0, 1\}$ of two elements, where $2 \cdot 1 = 1 + 1 = 0$.

The least k such that $k \cdot 1 = 0$ is called the *characteristic* of $\mathbb{F}$. If $k \cdot 1 \neq 0$ for any integer $k > 0$, then F is said to be a *field of characteristic zero*.

The fields $\mathbb{Q}$ of rational numbers, $\mathbb{R}$ of real numbers, $\mathbb{C}$ of complex numbers are all of characteristic zero.

In a field of characteristic zero, it is possible to divide any element of the field by any positive integer. In particular, in a field of characteristic zero one can define an *arithmetic mean* of any collection $a_1, \dots, a_k \in \mathbb{F}$ as $\frac{a_1+\dots+a_k}{k}$. This gives rise to a linear map $\mathbb{F}^k \to \mathbb{F}$ and it has a remarkable property that $\frac{a+\dots+a}{k} = a$. If the characteristic of the field is not zero and divides k, then an arithmetic mean of k elements of the field cannot be defined.

3.4. Symmetrization

Assume that our ground field $\mathbb{F}$ is of characteristic zero. Then any polylinear form can be symmetrized. Namely, there is a map which assigns to a polylinear form $T \in \mathrm{Tens}_k(V)$ a polylinear form defined by the formula

$$\operatorname{sym} T(v_1, \dots, v_k) = \sum_{\sigma \in S_k} \frac{1}{k!} T(v_{\sigma(1)}, \dots, v_{\sigma(k)})$$

Clearly, $\operatorname{sym} T$ is a symmetric polylinear form. It coincides with the original form T if T was already symmetric. Thus, sym is a projection of the space of all polylinear forms $\mathrm{Tens}_k(V)$ onto the subspace $\mathrm{Sym}_k(V)$ of symmetric forms.

3.5. Anti-symmetric polylinear maps

Let V be a vector space over a field $\mathbb{F}$. A bilinear form $T : V \times V \to F$ is said to be *anti-symmetric* or *skew-symmetric* if $T(v, w) = -T(w, v)$ for any $v, w \in V$.

More generally:

A polylinear map $T : V^k \to \mathbb{F}$ is said to be *anti-symmetric*, or *skew-symmetric*, or *alternating*, or *exterior k-form* on V if transposition of any two arguments implies multiplication of the value by -1.

In formula:

$$T(A, v, B, w, C) = -T(A, w, B, v, C),$$

where $v, w \in V$ and A, B, C are lists of vectors (some of which may be empty). Say, $A = a_1, \dots, a_i$, $B = b_1, \dots, b_j$, $C = c_1, \dots, c_l$.

3.B Reformulations. *Let $T : V^k \to \mathbb{F}$ be a polylinear map. Then the following statements are equivalent:*

(1) *T is anti-symmetric.*
(2) *T takes value zero on any list of vectors $v_1, v_2, \dots, v_k$ in which two of the vectors are equal (say, $v_i = v_j$ for some $i \neq j$).*
(3) *Adding to one of the arguments other arguments multiplied by an element of $\mathbb{F}$ does not change the value of T. In formula:*

$$T(A, v, B, w, C) = T(A, v, B, w + av, C),$$

where $v, w \in V$, $a \in \mathbb{F}$ and A, B, C are some lists of vectors.
(4) *T takes value zero on any linearly dependent list of vectors.*

Proof. (1) $\Longrightarrow$ (2): If T is anti-symmetric, then a transposition of the equal arguments multiplies the value of T by -1. On the other hand, the transposition does not change the list of arguments and hence does not change the value of T. Hence, the value is zero.

In formulas: $T(A, v, B, w, C) = -T(A, w, B, v, C)$. On the other hand, if $w = v$, then $T(A, v, B, w, C) = T(A, w, B, v, C)$.
Hence $T(A, v, B, v, C) = 0$. □

(2) $\Longrightarrow$ (1):

$$\begin{aligned} 0 &= T(A, v + w, B, v + w, C) \\ &= T(A, v, B, v, C) + T(A, v, B, w, C) + T(A, w, B, v, C) + T(A, w, B, w, C) \\ &= T(A, v, B, w, C) + T(A, w, B, v, C). \end{aligned}$$ □

(2) $\Longrightarrow$ (3):

$$\begin{aligned} T(A, v, B, w + av, C) &= T(A, v, B, w, C) + aT(A, v, B, v, C) \\ &= T(A, v, B, w, C) + a0 = T(A, v, B, w, C) \end{aligned}$$ □

(3) $\Longrightarrow$ (4): Assume v_j is a linear combination of all other vectors of the list $v_1, \dots, v_k$: $v_j = \sum_{i \neq j} a_i v_i$. Then by (3),

$$\begin{aligned} T(v_1, \dots, v_k) &= T\left(v_1, \dots, v_{j-1},\, v_j - \sum_{i \neq j} a_i v_i,\, v_{j+1}, \dots, v_k\right) \\ &= T(v_1, \dots, v_{j-1}, 0, v_{j+1}, \dots, v_k) = 0 \quad \square \end{aligned}$$

(4) $\Longrightarrow$ (2): A list of vectors in which two elements are equal is linearly dependent. $\square$

Since the value of a skew-symmetric form is multiplied by -1 under each transposition of arguments, under an arbitrary permutation of arguments the value is multiplied by the sign of permutation:

$$T(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(n)}) = \operatorname{sign}\sigma \quad T(v_1, v_2, \dots, v_n)$$

for any anti-symmetric n-linear form T.

Denote the space of anti-symmetric polylinear forms $V^k \to \mathbb{F}$ by symbol $\Lambda^k V$.

If $k = 1$, then all conditions of Theorem 3.B hold true tautologically. So, $\Lambda^1 V = V^{\checkmark}$.

If $\dim V = 1$, then $\Lambda^2 V = 0$. Indeed, in 1-dimensional space any two vectors are linearly dependent, therefore by statement (4) of Theorem 3.B the value of antisymmetric form must be zero.

More generally, for the same reason $\Lambda^k V = 0$ if $\dim V < k$.

3.6. Anti-symmetrization

Assume that the ground field $\mathbb{F}$ has characteristic zero. Any polylinear form can be *anti-symmetrized*. There is a linear map $alt : \operatorname{Tens}_k(V) \to \Lambda^k(V)$ defined by the formula

$$\operatorname{alt} T(v_1, \dots, v_k) = \sum_{\sigma \in S_k} \frac{\operatorname{sign}\sigma}{k!} T(v_{\sigma(1)}, \dots, v_{\sigma(k)})$$

If $T \in \operatorname{Tens}_k(V)$ is anti-symmetric, then $\operatorname{alt} T = T$. Thus, alt is a projection of the space of all polylinear forms $\operatorname{Tens}_k(V)$ onto the subspace $\Lambda^k(V)$ which consists of anti-symmetric forms.

3.7. Exterior k-forms on a k-dimensional space

3.C. Theorem. *For any integer $k > 0$, $\dim \Lambda^k(\mathbb{F}^k) = 1$.*

Proof. Notice first that $\Lambda^k(\mathbb{F}^k) \neq 0$, because it contains $\operatorname{alt}(\mathfrak{e}^1 \otimes \mathfrak{e}^2 \otimes \dots \otimes \mathfrak{e}^k)$ which is not zero. Hence $\dim \Lambda^k(\mathbb{F}^k) \geq 1$.

Consider a skew-symmetric tensor $T : \underbrace{\mathbb{F}^k \times \dots \times \mathbb{F}^k}_{k \text{ times}} \to \mathbb{F}$. It is defined by its coordinates in the space $\operatorname{Tens}_k(\mathbb{F}^k)$. Recall (see section 2.3) that the coordinates of a polylinear map are its values on sequences $(\mathfrak{e}_{i_1}, \mathfrak{e}_{i_2}, \dots, \mathfrak{e}_{i_k})$ of base vectors. If $i_p = i_q$ for some $p \neq q$, then the value of a skew-symmetric form is zero. If $i_p \neq i_q$ for any $p \neq q$, then $(i_1, i_2, \dots, i_k) = (\sigma(1), \sigma(2), \dots, \sigma(k))$ for some $\sigma \in S_k$. Hence $T(\mathfrak{e}_{\sigma(1)}, \mathfrak{e}_{\sigma(2)}, \dots, \mathfrak{e}_{\sigma(k)}) = \operatorname{sign}\sigma \quad T(\mathfrak{e}_1, \mathfrak{e}_2, \dots, \mathfrak{e}_k))$. Therefore, T can be recovered from $T(\mathfrak{e}_1, \mathfrak{e}_2, \dots, \mathfrak{e}_k) \in \mathbb{F}$. Hence $\dim \Lambda^k(\mathbb{F}^k) \leq 1$. □

An isomorphism $\Lambda^k(\mathbb{F}^k) \to \mathbb{F}$ is defied by $T \mapsto T(\mathfrak{e}_1, \mathfrak{e}_2 \dots, \mathfrak{e}_k)$.

3.D Corollary. *If* $\dim V = k > 0$, *then* $\dim \Lambda^k(V) = 1$.

4. Determinant

4.1. Determinant of an operator

Let V be a vector space of dimension n over a field $\mathbb{F}$, and $T : V \to V$ be a linear map. The map $T^* : \Lambda^n(V) \to \Lambda^n(V)$ induced by T, as any linear map of a 1-dimensional space to itself, is a multiplication by some element of $\mathbb{F}$. This element is called the *determinant* of T and is denoted by $\det T$.

4.2. Properties of determinants

4.A. *The determinant of a composition is the product of the determinants of the factors. In formula:* $\det(T \circ S) = \det T \det S$.

Indeed, $(T \circ S)^* = S^* \circ T^*$, composition of multiplications by $\det T$ and multiplication by $\det S$ is the multiplication by $\det T \det S$. □

4.B. $\det \mathrm{id} = 1$.

Indeed, $\mathrm{id}^* = \mathrm{id}$ and $\mathrm{id} : \mathbb{F} \to \mathbb{F}$ is multiplication by 1 □

4.C. *If T is invertible, then* $\det T^{-1} = \dfrac{1}{\det T}$. *In particular,* $\det T \neq 0$.

Indeed, the equality $\mathrm{id} = T \circ T^{-1}$ implies $1 = \det \mathrm{id} = \det T \det T^{-1}$. □

4.D. *If T is not invertible, then* $\det T = 0$.

Proof. Represent $T : V \to V$ as a composition of the surjective map $S : V \to \operatorname{range} T$ defined by T and inclusion $i : \operatorname{range} T \to V$. Since T is not invertible, it is not surjective and $\dim \operatorname{range} T < \dim V$. Therefore $\Lambda^{\dim V}(\operatorname{range} T) = 0$. Hence, $T^* : \Lambda^n(V) \to \Lambda^n(V)$ is factored through the zero space. So, it is zero. $\square$

The last two properties imply the following convenient criterion for non-invertibility of a linear map $T : V \to V$:

4.E. *A linear map $T : V \to V$ is not invertible* $\iff$ $\det T = 0$. $\square$

4.3. Invariance of determinant

4.F. *If operators $T : V \to V$ and $S : W \to W$ are isomorphic in the category of operators (that is there exists an linear isomorphism $L : V \to W$ such that $S = L \circ T \circ L^{-1}$), then* $\det S = \det T$.

Proof. Assume, first, that $W = V$. Then $\det S = \det(L \circ T \circ L^{-1}) = \det L \det T \det(L^{-1}) = \det T \det L (\det L)^{-1} = \det T$.

If $W \neq V$, these arguments are not applicable, because a determinant is defined only for operators, it is not defined for a linear map between different vector spaces. In order to make the arguments legitimate, we have to identify V and W somehow. Here is a straightforward proof. Notice that commutative diagrams in that proof are not necessary. Their purpose is to clarify the exposition, rather than to obscure.

Linear maps T, S, and L form a commutative diagram
$\begin{array}{ccc} V & \xrightarrow{L} & W \\ \downarrow{\scriptstyle T} & & \downarrow{\scriptstyle S} \\ V & \xrightarrow{L} & W \end{array}$. These maps induce maps T^*, S^*, and L^*, which form a commutative diagram

$$\begin{array}{ccc} \Lambda^n V & \xleftarrow[L^*]{} & \Lambda^n W \\ {\scriptstyle T^*}\uparrow & & {\scriptstyle S^*}\uparrow \\ \Lambda^n V & \xleftarrow[L^*]{} & \Lambda^n W \end{array} \tag{1}$$

where $n = \dim V = \dim W$. The maps T^* and S^* are multiplications by $\det T$ and $\det S$, respectively. Let us fix isomorphisms $\Lambda^n V \to \mathbb{F}$ and $\Lambda^n W \to \mathbb{F}$ and, in diagram (1), replace $\Lambda^n V$ and $\Lambda^n W$ with $\mathbb{F}$ via

these isomorphisms. We get a diagram
$$\begin{tikzcd} \mathbb{F} & \mathbb{F} \arrow[l, "L^*"'] \\ \mathbb{F} \arrow[u, "\det T"] & \mathbb{F} \arrow[l, "L^*"'] \arrow[u, "\det S"'] \end{tikzcd}$$
. Linear map $L^* : \mathbb{F} \to \mathbb{F}$ appears twice in this diagram. It is an isomorphism $\mathbb{F} \to \mathbb{F}$, thus this is a multiplication by some $C \in \mathbb{F}$, $C \neq 0$. The commutativity of the diagram (1) means that multiplication by $C \det(S)$ coincides with the multiplication by $\det(T)C$. Hence $\det S = \det T$. $\square$

4.4. Formula for determinant

4.G. *Let $T : \mathbb{F}^n \to \mathbb{F}^n$ be a linear map with $T\mathfrak{e}_j = T^i_j \mathfrak{e}_i$. Then*
$$\det T = \sum_{\sigma \in S_n} \operatorname{sign}\sigma \;\; T_1^{\sigma(1)} T_2^{\sigma(2)} \cdots T_n^{\sigma(n)}.$$

Proof. Let $D \in \Lambda^n(\mathbb{F}^n)$ be the base vector which is characterized by the property that $D(\mathfrak{e}_1, \dots, \mathfrak{e}_n) = 1$. Then

$$\begin{aligned}
(T^*D)(\mathfrak{e}_1, \dots, \mathfrak{e}_n) &= D(T\mathfrak{e}_1, \dots, T\mathfrak{e}_n) \\
&= D(T_1^{j_1}\mathfrak{e}_{j_1}, \dots, T_n^{j_n}\mathfrak{e}_{j_n}) \\
&= T_1^{j_1} \dots T_n^{j_n} D(\mathfrak{e}_{j_1}, \dots, \mathfrak{e}_{j_n}) \\
&= \sum_{\sigma \in S_n} T_1^{\sigma(1)} \dots T_n^{\sigma(n)} D(\mathfrak{e}_{\sigma(1)}, \dots, \mathfrak{e}_{\sigma(n)}) \\
&= \sum_{\sigma \in S_n} T_1^{\sigma(1)} \dots T_n^{\sigma(n)} \operatorname{sign}\sigma \;\; D(\mathfrak{e}_1, \dots, \mathfrak{e}_n) \\
&= \sum_{\sigma \in S_n} \operatorname{sign}\sigma \;\; T_1^{\sigma(1)} \dots T_n^{\sigma(n)} \qquad \square
\end{aligned}$$

4.5. Characteristic polynomial of an operator

Let V be a finite-dimensional vector space over a field $\mathbb{F}$ and $T : V \to V$ be an operator. For each $\lambda \in \mathbb{F}$ consider the determinant of operator $\lambda I - T : V \to V$. (Here we denote by I the identity operator id_V.)

By Theorem 4.G, the determinant is a sum of products of linear functions of λ. Each of the products contains $\dim V$ factors. Therefore this is a polynomial in λ of degree $\dim V$.

$\det(\lambda I - T)$ is called the *characteristic polynomial* of T.

4.H. *The characteristic polynomial of T is an invariant of the isomorphism class of T.*

Proof. Recall that any operator isomorphic to $T : V \to V$ can be presented as $L^{-1} \circ T \circ L$, where $L : W \to V$ is an invertible linear map. It follows from Theorem 4.F, that the values of the characteristic polynomials of T and S at each $\lambda \in \mathbb{F}$ are equal, because operators $\lambda I - T$ and $\lambda I - L^{-1} \circ T \circ L$ are isomorphic for any value of λ. Indeed, $\lambda I - L^{-1} \circ T \circ L = \lambda L^{-1} \circ L - L^{-1} \circ T \circ L = L^{-1} \circ (\lambda I - T) \circ L$.

These arguments suffice if the characteristic of the ground field $\mathbb{F}$ is zero, because two polynomials over such a field equals iff they have the same values at each $\lambda \in \mathbb{F}$. However, if $\mathbb{F}$ has a finite characteristic, then a polynomials are not defined by their values. For example, a polynomial x^2 and x have the same values if $\mathbb{F} = \mathbb{Z}/_2$, and, more generally, if the characteristic is 2.

Theorem 4.5 holds true for operators over any field $\mathbb{F}$. It can be proved similarly to Theorem 4.F. We leave the proof as an exercise. $\square$

Recall (see Theorem 5.6 from Axler's textbook) that λ is an eigenvalue of a linear operator T in a finite-dimensional space iff $T - \lambda\,\mathrm{id}$ is not invertible. Together with the criterion of non-invertibility from Section 4.2, it implies that

4.I. *λ is an eigenvalue of a linear operator in a finite dimensional vector space iff λ is a root of the characteristic polynomial of this operator.*

Exercise. Prove that $(-1)^{\dim V} \det f$ is the free term of the characteristic polynomial of $f : V \to V$.

4.6. Trace

Let V be a finite-dimensional vector space over a field $\mathbb{F}$ and $T : V \to V$ be an operator. Let in some basis $v_1, \dots, v_n$ the operator T has matrix T_i^j. Then $T_i^i (= \sum_i T_i^i = T_1^1 + \dots + T_n^n)$ is called the *trace* of T and denoted by $\operatorname{tr} T$.

This definition requires a proof, because it involves a choice of basis, while in the name no basis is mentioned. A proof comes as a reference to Theorem 4.I and the following statement.

4.J. *For any linear operator T in an n-dimensional vector space V represented by a matrix (T_i^j) for some basis of V, the coefficient at λ^{n-1} in the characteristic polynomial of T equals $-\sum_i T_i^i$.*

The Theorem 4.J together with Exercise above can be summarized in the following formula:

$$\det(\lambda I - T) = \lambda^n - \operatorname{tr} T \lambda^{n-1} + \dots + (-1)^n \det T$$

Proof of 4.J. Expand $\det(\lambda I - T)$ according to Theorem 4.G. The summands which contribute to the monomial of degree $n-1$ correspond to permutations σ which leave $n-1$ elements of $\{1, 2, \dots, n\}$ fixed. Only one permutation has this property: $\sigma = \mathrm{id}$. The summand corresponding to $\sigma = \mathrm{id}$ is $(\lambda - T_1^1)(\lambda - T_2^2)\dots(\lambda - T_n^n)$. Expand it:

$$\begin{aligned}(\lambda - T_1^1)(\lambda - T_2^2)\dots(\lambda - T_n^n) = \lambda^n - \lambda^{n-1}\sum_i T_i^i + \dots\\ = \lambda^n - \operatorname{tr} T \lambda^{n-1} + \dots\end{aligned}$$

□

Thus, the trace $\operatorname{tr} T$ and determinant $\det T$ are, up to sign, coefficients of the characteristic polynomial of T. Other coefficients also are numerical invariants of the operator. The trace and determinant occupy the extreme positions and they have special properties distinguishing them from other numerical invariants which come from the characteristic. For the determinant, this is its multiplicativity: $\det(S \circ T) = \det S \det T$. The next theorem is a distinctive property of trace.

4.K. *For any $S, T \in \mathcal{L}(V)$, $\operatorname{tr}(S \circ T) = \operatorname{tr}(T \circ S)$.*

Proof. Fix a basis in V. Let T_i^j and S_i^j be marices of T and S in this basis. Then in the Einstein notation the products TS and ST have matrices $(T_i^j S_k^i)$ and $(S_i^j T_k^i)$. Let us find the traces:

$$\operatorname{tr}(TS) = T_i^j S_j^i \text{ and } \operatorname{tr}(ST) = S_i^j T_j^i$$

These two numbers are equal, because the summation indices are damn, their renaming would not effect the sum. □

Remark. Observe that the trace is neither multiplicative, nor additive. Indeed, consider $S = T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Then

$$\operatorname{tr}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = 0, \quad \text{while } \operatorname{tr}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^2 = \operatorname{tr}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 2$$