LA-3 Lecture Notes

Karl-Heinz Fieseler

Uppsala 2014

Contents

1 Dual space **2**

2 Direct sums and quotient spaces **4**

3 Bilinear forms **7**

4 Jordan normal form **8**

5 Minimal and characteristic polynomial **17**

6 Spectral Theorem **21**

7 Determinants and exterior algebra **28**

1 Dual space

Definition 1.1. A *linear form* or *linear functional* on a vector space V is a linear map $\alpha : V \longrightarrow K$. They form a vector space

$$V^* := \mathcal{L}(V, K)$$

called the *dual space* of V.

Remark 1.2. If we, following the tradition, identify $\mathcal{L}(K^n, K^m)$ with the space $K^{m,n}$ of $m \times n$-matrices, we may write linear forms in the dual space $(K^n)^*$ of K^n as row vectors

$$K^{1,n} \cong (K^n)^*, \mathbf{y} = (y_1, ..., y_n) \mapsto \alpha_{\mathbf{y}}$$

with

$$\alpha_{\mathbf{y}} : V \longrightarrow K, \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \mathbf{yx} = \sum_{\nu=1}^{n} y_\nu x_\nu.$$

Definition 1.3. If $B = (\mathbf{v}_1, ..., \mathbf{v}_n)$ is a basis of V, we define the linear forms

$$\mathbf{v}_j^* : V \longrightarrow K, j = 1, ..., n,$$

by

$$\mathbf{v}_j^*(\mathbf{v}_i) = \delta_{ij}, i = 1, ..., n.$$

They form a basis $B^* = (\mathbf{v}_1^*, ..., \mathbf{v}_n^*)$ of V^*, it is called the dual basis for the basis $\mathbf{v}_1, ..., \mathbf{v}_n$.

Remark 1.4. The notation $\mathbf{v}_j^*$ may be a bit confusing: The definition of $\mathbf{v}_j^*$ involves the entire basis $\mathbf{v}_1, ..., \mathbf{v}_n$ – not just $\mathbf{v}_j \in V$. Indeed, there is **no** "natural" linear map $V \longrightarrow V^*, \mathbf{v} \mapsto \mathbf{v}^*$ – though of course there is a linear isomorphism

$$\varphi_B : V \longrightarrow V^*, \mathbf{v}_i \mapsto \mathbf{v}_i^*, i = 1, ..., n,$$

but it depends on the choice of the basis:

Example 1.5. If $B = (\mathbf{v}_1, ..., \mathbf{v}_n)$ is a basis of K^n, then

$$\varphi_B : K^n \longrightarrow (K^n)^*, \mathbf{x} \mapsto \mathbf{x}^T C^T C,$$

with the matrix $C \in K^{n,n}$ satisfying $C\mathbf{v}_i = \mathbf{e}_i, i = 1, ..., n$, i.e.

$$C = (\mathbf{v}_1, ..., \mathbf{v}_n)^{-1}.$$

To see that we simply check that

$$(\mathbf{v}_j^T C^T C)\mathbf{v}_i = (C\mathbf{v}_j)^T C\mathbf{v}_i = \mathbf{e}_j^T \mathbf{e}_i = \delta_{ij}.$$

Remark 1.6. A linear map $F : V \longrightarrow W$ induces a linear map in the reverse direction, the pull back of linear functionals:

$$F^* : W^* \longrightarrow V^*, \alpha \mapsto \alpha \circ F : V \xrightarrow{F} W \xrightarrow{\alpha} K.$$

Choose bases B and C of V resp. W. If F has matrix A w.r.t. B and C, then F^* has matrix A^T w.r.t. C^* and B^*.

For the proof it suffices to consider $V = K^n, W = K^m$ and

$$F = T_A : K^n \longrightarrow K^m, \mathbf{x} \mapsto A\mathbf{x}.$$

Now

$$F^*(\alpha_{\mathbf{y}})(\mathbf{x}) = \mathbf{y}(A\mathbf{x}) = (\mathbf{y}A)(\mathbf{x}),$$

thus using the isomorphism $(K^n)^* \cong K^{1,n}, (K^m)^* \cong K^{1,m}$ we find

$$F^* : K^{1,m} \longrightarrow K^{1,n}, \mathbf{y} \mapsto \mathbf{y}A.$$

In order to obtain the matrix of F^* with respect to the respective dual standard bases we have to transpose: The map

$$K^\ell \longrightarrow (K^\ell)^*, \mathbf{x} \mapsto \mathbf{x}^T,$$

is nothing but the coordinate map

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto x_1\mathbf{e}_1^T + ... + \mathbf{e}_\ell^T$$

w.r.t. the dual basis. Hence we finally arrive at

$$K^m \longrightarrow K^n, \mathbf{v} \mapsto \mathbf{v}^T \mapsto \mathbf{v}^T A \mapsto (\mathbf{v}^T A)^T = A^T\mathbf{v}.$$

2 Direct sums and quotient spaces

Definition 2.1. Given vector spaces $V_1, ..., V_r$, we define their direct sum $V_1 \oplus .... \oplus V_r$ or $\bigoplus_{i=1}^r V_i$ as the cartesian product

$$V_1 \times ... \times V_r$$

endowed with the componentwise addition and scalar multiplication.

Remark 2.2. If $V_i \hookrightarrow V, i = 1, .., r$, are subspaces of a given vector space V, there is a natural linear map

$$\sigma : \bigoplus_{i=1}^r V_i \longrightarrow V, (\mathbf{v}_1, ..., \mathbf{v}_r) \mapsto \mathbf{v}_1 + ... + \mathbf{v}_r.$$

If it is injective, we identify $\bigoplus_{i=1}^r V_i$ with

$$\sigma(V) = \sum_{i=1}^r V_i \hookrightarrow V$$

and express that by writing

$$\sum_{i=1}^r V_i = \bigoplus_{i=1}^r V_i.$$

Example 2.3. For subspaces $U, W \subset V$ we have

$$\ker(\sigma) = \{(\mathbf{v}, -\mathbf{v}); \mathbf{v} \in U \cap W\}$$

and thus

$$U + W = U \oplus W \Longleftrightarrow U \cap W = \{0\}.$$

Definition 2.4. Let $U \subset V$ be a subspace. A subspace $W \subset V$ is called a complementary subspace for $U \subset V$ if

$$V = U + W = U \oplus W.$$

Any subspace U admits complementary subspaces, but there is in general no distinguished choice. Sometimes it can be helpful to replace the complementary subspace with an abstract construction not depending on any choices.

Definition 2.5. Let $U \subset V$ be a subspace. Then

$$\mathbf{v} \sim \mathbf{w} :\Longleftrightarrow \mathbf{v} - \mathbf{w} \in U$$

defines an equivalence relation on V. The equivalence classes are the sets

$$\mathbf{v} + U := \{\mathbf{v} + \mathbf{u}; \mathbf{u} \in U\},$$

also called cosets of U. We denote

$$V/U := V/\sim = \{\mathbf{v} + U; \mathbf{v} \in V\}$$

the set of all equivalence classes and call it the quotient space of V mod U. The set V/U is made a vector space with the addition

$$(\mathbf{v} + U) + (\mathbf{w} + U) = (\mathbf{v} + \mathbf{w}) + U$$

and the scalar multiplication

$$\lambda(\mathbf{v} + U) := \lambda\mathbf{v} + U,$$

and the quotient projection

$$\varrho : V \longrightarrow V/U$$

then becomes a linear map.

Example 2.6. If $V = \mathbb{R}^2$ and $U = \mathbb{R}\mathbf{x}$ is a line, then all lines $W = \mathbb{R}\mathbf{y} \neq U$ are complementary subspaces. Cosets $\mathbf{v} + U = \mathbf{v} + \mathbb{R}\mathbf{x}$ are lines parallel to $\mathbb{R}\mathbf{x}$. Such a coset intersects any complementary subspace W in exactly one point.

Remark 2.7. If $V = U \oplus W$, then the composition

$$W \hookrightarrow V \longrightarrow V/U$$

of the inclusion of W into V and the quotient projection $\varrho : V \longrightarrow V/U$ is an isomorphism.

Proposition 2.8. *If $F : V \longrightarrow W$ is a linear map and $U \subset \ker(F)$, there is a unique map $\overline{F} : V/U \longrightarrow W$ with $F = \overline{F} \circ \varrho$, i.e. the diagram*

$$\begin{array}{cccc} & V & \xrightarrow{F} & W \\ \varrho & \downarrow & \nearrow & \\ & V/U & & \end{array}$$

is commutative. If, furthermore $U = \ker(F)$ and F is surjective, then $\overline{F} : V/U \xrightarrow{\cong} W$ is an isomorphism.

Proof. Indeed, uniqueness follows from the surjectivity of ϱ, while existence follows from the fact that $F|_{\mathbf{v}+U} \equiv F(\mathbf{v})$, i.e. we may define

$$\overline{F}(\mathbf{v} + U) := F(\mathbf{v}).$$

□

Example 2.9. For a subspace $U \subset V$ we describe U^* as a quotient of V^*: The restriction

$$R : V^* \longrightarrow U^*, \alpha \mapsto \alpha|_U,$$

has kernel

$$\ker(R) := U^\perp := \{\alpha \in V^*; \alpha|_U = 0\}$$

and is surjective, hence induces an isomorphism

$$\overline{R} : V^*/U^\perp \xrightarrow{\cong} U^*.$$

Recall that $R = j^*$ with the inclusion $j : U \hookrightarrow V$.

3 Bilinear forms

We have emphasized that there is no natural isomorphism $V \cong V^*$. Nevertheless there are often situations, where there is such an isomorphism, namely, when one considers vector spaces with an additional datum, a bilinear form $\beta : V \times V \longrightarrow K$.

Definition 3.1. A bilinear form on a vector space V is a map

$$\beta : V \times V \longrightarrow K$$

such that for any $\mathbf{u} \in V$ the maps

$$\beta(\mathbf{u}, ..) : V \longrightarrow K \text{ and } \beta(.., \mathbf{u}) : V \longrightarrow K$$

are linear.

Example 3.2. Take $V = K^n$ and a matrix $A \in K^{n,n}$. Then

$$\beta(\mathbf{u}, \mathbf{v}) := \mathbf{u}^T A \mathbf{v}$$

defines a linear form on V. In particular

1. the matrix $A = I_n$ defines the standard

Remark 3.3. One might wonder why we here write $V \times V$ instead of $V \oplus V$: The point is, that we want to avoid confusion. A bilinear form on V is not a linear form on $V \oplus V$:

$$\mathrm{Bi}(V) \ncong (V \oplus V)^* \cong V^* \oplus V^*,$$

indeed

$$\dim \mathrm{Bi}(V) = (\dim V)^2, \text{while } \dim(V \oplus V)^* = 2 \dim V.$$

Instead there are natural isomorphisms

$$\mathrm{Bi}(V) \cong \mathcal{L}(V, V^*), \beta \mapsto F_\beta : V \longrightarrow V^*,$$

where

$$V^* \ni F_\beta(\mathbf{v}) = \beta(\mathbf{v}, ..)$$

and

$$\mathrm{Bi}(V) \cong \mathcal{L}(V, V^*), \beta \mapsto G_\beta.V \longrightarrow V^*,$$

where

$$V^* \ni G_\beta(\mathbf{v}) = \beta(.., \mathbf{v}).$$

4 Jordan normal form

Given an endomorphism $T : V \longrightarrow V$ of a finite dimensional vector space V, we want to find a basis $B = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$, s.th. the matrix of T with respect to B is "simple".

We try to decompose V into smaller subspaces, which are T-invariant.

Definition 4.1. A subspace $U \subset V$ is called T-invariant, if $T(U) \subset U$.

Example 4.2. 1. The entire space V and the zero space $\{0\}$ are invariant subspaces for any operator $T \in \mathcal{L}(V)$.

2. Let $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in K^{2,2}$.

   (a) The linear map

   $$T := T_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \mathbf{x} \mapsto A\mathbf{x},$$

   is a counterclockwise rotation with an angle of 90 degrees, hence $\{0\}$ and $\mathbb{R}^2$ are its only invariant subspaces.

   (b) The linear map

   $$T := T_A : \mathbb{C}^2 \longrightarrow \mathbb{C}^2, \mathbf{z} \mapsto A\mathbf{z},$$

   has the invariant subspaces

   $$\{0\}, \mathbb{C}\begin{pmatrix} 1 \\ i \end{pmatrix}, \mathbb{C}\begin{pmatrix} 1 \\ -i \end{pmatrix}, \mathbb{C}^2,$$

   the two lines being the eigenspaces for the eigenvalues $-i$ resp. i.

3. Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in K^{2,2}$. The linear map

   $$T := T_A : K^2 \longrightarrow K^2$$

   has the invariant subspaces

   $$\{0\}, K\begin{pmatrix} 1 \\ 0 \end{pmatrix}, K^2.$$

4. A line $U = K\mathbf{v}, \mathbf{v} \neq 0$, is a T-invariant subspace iff $\mathbf{v}$ is an eigenvector of T.

5. The λ-eigenspaces of T, the subspaces

$$V_\lambda = \{\mathbf{v} \in V; T(\mathbf{v}) = \lambda\mathbf{v}\}, \lambda \in K,$$

are T-invariant.

Remark 4.3. 1. If $U \subset V$ is T-invariant, and $B = B_1 \cup B_2$ a basis, s.th. B_1 spans U, the matrix of T w.r.t. B looks as follows

$$\begin{pmatrix} A & C \\ 0 & D \end{pmatrix}.$$

2. If there is even a T-invariant complementary subspace W and B_2 a basis of W, the matrix becomes

$$\begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}.$$

3. Låt $\lambda \in K$. A subspace $U \subset V$ is T-invariant if it is $(T - \lambda \mathrm{id}_V)$-invariant.

Definition 4.4. A T-invariant subspace U is called

1. *irreducible* if it does not admit proper nontrivial invariant subspaces,

2. *indecomposable* if it can not be written $U = U_1 \oplus U_2$ with nontrivial T-invariant subspaces.

Remark 4.5. 1. Irreducible T-invariant subspaces are indecomposable, but not the other way around: Take $T = T_{\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}} : K^2 \longrightarrow K^2$. Then $U = K^2$ is not irreducible, but indecomposable.

2. Now, since $\dim V < \infty$, it is clear, that V is a direct sum of indecomposable subspaces, and it remains to study, what they look like.

We are now looking for invariant subspaces $U \subset V$ admitting an invariant complementary subspace W, i.e.

$$V = U \oplus W.$$

Assume T admits an eigenvalue $\lambda \in K$. Since we may replace T with $(T - \lambda \mathrm{id}_V)$, we may even assume $\ker(T) \neq \{0\}$. Obviously $U = \ker(T)$ and $W = T(V)$ are invariant subspaces, satisfying

$$\dim U + \dim W = \dim V,$$

since $T(V) \cong V/\ker(T)$. But in general $U \cap W = \{0\}$ does not hold: Take for example $T : K^2 \longrightarrow K^2, \mathbf{e}_1 \mapsto 0, \mathbf{e}_2 \mapsto \mathbf{e}_1$. Nevertheless we succeed under an additional assumption:

Proposition 4.6. *Assume $S \in \mathcal{L}(V), S^2(V) = S(V)$. Then we have*

$$V = \ker(S) \oplus S(V).$$

Proof. Take $\mathbf{v} \in V$ and write

$$S(\mathbf{v}) = S^2(\mathbf{w}), \mathbf{w} \in V.$$

Then we have

$$\mathbf{v} = (\mathbf{v} - S(\mathbf{w})) + S(\mathbf{w}) \in \ker(S) + S(V).$$

Hence

$$V = \ker(S) + S(V),$$

while $\ker(S) + S(V) = \ker(S) \oplus S(V)$ because of $\dim \ker(S) + \dim S(V) = \dim V$. □

Corollary 4.7. *Let $\lambda \in K$ be an eigenvalue of T. Then we have*

$$V = \ker((T - \lambda \cdot \mathrm{id}_V)^k) \oplus (T - \lambda \cdot \mathrm{id}_V)^k(V)$$

for $k \gg 1$.

Proof. Let $F := (T - \lambda \mathrm{id}_V)^k$. The sequence of subspaces $F^k(V), k \in \mathbb{N}$, is decreasing, hence becomes stationary after a while, i.e. there is $k \in \mathbb{N}$, such that $F^{k+\ell}(V) = F^k(V)$ for all $\ell \in \mathbb{N}$. Now take $S = F^k$ and apply Prop. 4.6. □

The above reasoning motivates the introduction of *generalized eigenspaces*

$$\widehat{V}_\lambda \supset V_\lambda = \ker(T - \lambda \mathrm{id}_V).$$

Definition 4.8. Let $T \in \mathcal{L}(V)$. The *generalized eigenspace* $\hat{V}_\lambda$ for the eigenvalue $\lambda \in K$ of T is defined as

$$\widehat{V}_\lambda := \bigcup_{\nu=1}^{\infty} \ker((T - \lambda \mathrm{id}_V)^\nu).$$

Remark 4.9. $\widehat{V}_\lambda = \ker((T - \lambda \mathrm{id}_V)^k)$ for $k \gg 0$.

Proposition 4.10. *Let $\lambda \in K$ be an eigenvalue of $T \in \mathcal{L}(V)$, Then*

$$V = \widehat{V}_\lambda \oplus U_\lambda$$

holds with the invariant subspace

$$U_\lambda := \bigcap_{\nu=1}^{\infty} (T - \lambda \mathrm{id}_V)^\nu (V).$$

Indeed $U_\lambda = (T - \lambda \mathrm{id}_V)^k (V)$. for $k \gg 0$.

In order that our inductive approach works we have to assure that our operator $T \in \mathcal{L}(V)$ has "sufficiently many" eigenvalues:

Definition 4.11. The endomorphism $T : V \longrightarrow V$ is called *split*, if its characteristic polynomial is the product of linear polynomials:

$$\chi_T(X) = \prod_{i=1}^{r} (X - \lambda_i)^{k_i}$$

with $\lambda_1, ..., \lambda_r \in K$ and exponents $k_1, ..., k_r > 0$.

Example 4.12. 1. Let $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in K^{2,2}$. The endomorphism

$$T := T_A : K^2 \longrightarrow K^2$$

has the characteristic polynomial

$$\chi_T = T^2 + 1,$$

it is not split for $K = \mathbb{R}$, but for $K = \mathbb{C}$, since

$$T^2 + 1 = (T - i)(T + i).$$

We have $\lambda_1 = i, \lambda_2 = -i$ and

$$\hat{V}_1 = V_1 = \mathbb{C}\begin{pmatrix} 1 \\ -i \end{pmatrix}, \hat{V}_2 = V_2 = \mathbb{C}\begin{pmatrix} 1 \\ i \end{pmatrix}.$$

2. Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in K^{2,2}$. The linear map

$$T := T_A : K^2 \longrightarrow K^2$$

has the characteristic polynomial

$$\chi_T = (T - 1)^2,$$

it is obviously split, there is one eigenvalue $\lambda_1 = 1$ and

$$\hat{V}_1 = K^2 \supset V_1 = K\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Remark 4.13. For $K = \mathbb{C}$ all linear operators are split.

Theorem 4.14. *Let $T \in \mathcal{L}(V)$ be a split operator with the pairwise different eigenvalues $\lambda_1, ..., \lambda_r \in K$. Then we have*

$$V = \bigoplus_{i=1}^{r} \widehat{V}_{\lambda_i}.$$

Proof. Induction on $\dim V$. Take an eigenvalue $\lambda \in K$ of the operator T, use the decomposition

$$V = \widehat{V}_\lambda \oplus U_\lambda$$

and the fact that

$$\chi_T = \chi_F \cdot \chi_G$$

with

$$F := T|_{\hat{V}_\lambda}, G := T|_{U_\lambda}.$$

Indeed

$$\chi_F = (X - \lambda)^k,$$

while $\chi_G(\lambda) \neq 0$, since G does not have λ as one of its eigenvalues: $V_\lambda \cap U_\lambda = \{0\}$ and F has no eigenvalue $\neq \lambda$, since $F - \lambda \mathrm{id}_{\hat{V}_\lambda}$ is nilpotent. The statement holds for $G \in \mathcal{L}(U_\lambda)$ by induction hypothesis. □

Remark 4.15. If we choose a basis $B = B_1 \cup ... \cup B_r$, such that B_i is a basis of $\hat{V}_{\lambda_i}$, the matrix A of T with respect to B looks as follows:

$$A = \begin{pmatrix} A_1 & 0 & .. & .. & 0 \\ 0 & A_2 & 0 & .. & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & .. & A_{r-1} & 0 \\ 0 & 0 & .. & 0 & A_r \end{pmatrix}.$$

The next step is to investigate the generalized eigenspaces separately. So let us assume $V = \hat{V}_\lambda$. Then we have

$$T = \lambda \mathrm{id}_V + N$$

with the nilpotent operator

$$N := T - \lambda \mathrm{id}_V.$$

Proposition 4.16. *For a nilpotent operator $N \in \mathcal{L}(V)$ we have*

$$N^{\dim V} = 0.$$

Proof. By induction on $\dim V$. For $\dim V = 1$ we have $N = 0$. Since a nilpotent operator is not an isomorphism, we have $N(V) \subsetneqq V$ and find

$$(N|_{N(V)})^k = 0, k = \dim N(V).$$

Hence

$$N^s = 0, \ s \geq k + 1,$$

in particular for $s = \dim V$. □

Definition 4.17. Let $N \in \mathcal{L}(V)$ be a nilpotent operator. A subspace $U \subset V$ is called N-cyclic, if there is an "*N-cyclic*" vector $\mathbf{u} \in U$, i.e. s.th.

$$U = \mathrm{span}(T^i(\mathbf{u}), i \geq 0).$$

Lemma 4.18. *If U is N-cyclic with cyclic vector $\mathbf{u} \neq 0$, then*

$$U = K\mathbf{u} \oplus KT(\mathbf{u}) \oplus ... \oplus KT^s(\mathbf{u}),$$

if $T^s(\mathbf{u}) \neq 0 = T^{s+1}(\mathbf{u})$.

Proof. Obviously

$$U = K\mathbf{u} + KT(\mathbf{u}) + ... + KT^s(\mathbf{u}),$$

so we have to prove that the given vectors are linearly independent. Use induction on $\dim U$. If

$$\lambda_0\mathbf{u} + \lambda_1 T(\mathbf{u}) + ... + \lambda_s T^s(\mathbf{u}) = 0$$

we have

$$\lambda_0 T(\mathbf{u}) + \lambda_1 T^2(\mathbf{u}) + ... + \lambda_{s-1} T^s(\mathbf{u}) = 0$$

and the induction hypothesis for the T-cyclic subspace $T(U) \ni T(\mathbf{u})$ yields

$$\lambda_0 = ... = \lambda_{s-1} = 0.$$

Since $T^s(\mathbf{u}) \neq 0$ that implies $\lambda_s = 0$ as well. □

Theorem 4.19. *Let $N \in \mathcal{L}(V)$ be a nilpotent linear operator. Then*

$$V = U_1 \oplus ... \oplus U_t$$

with N-cyclic nonzero subspaces U_i. The number t is uniquely determined by N as well as, up to order, the dimensions $\dim U_i, i = 1, ..., t$.

Proof. Induction on $\dim V$. We have $N(V) \subsetneqq V$ - otherwise N would be an isomorphism. Hence by induction hypothesis for $N|_{N(V)} \in \mathcal{L}(N(V))$ we nay write

$$N(V) = W_1 \oplus ... \oplus W_q$$

with N-cyclic nonzero subspaces $W_i, i = 1, ..., q$. Choose $\mathbf{u}_1, ..., \mathbf{u}_q \in V$, s.th. $N(\mathbf{u}_i)$ is a cyclic vector for W_i, take

$$U_i = K\mathbf{u}_i + W_i, i = 1, ..., q.$$

and note that $\ker(N|_{U_i}) \subset W_i$. We have

$$U_1 + .. + U_q = U_1 \oplus .. \oplus U_q.$$

Assume that

$$\mathbf{v}_1 + ... + \mathbf{v}_q = 0$$

holds for the vectors $\mathbf{v}_i \in U_i, i = 1, ..., q$. Then

$$N(\mathbf{v}_1) + ... + N(\mathbf{v}_q) = 0.$$

By induction hypothesis we obtain $N(\mathbf{v}_1) = ... = N(\mathbf{v}_q) = 0$, in particular $\mathbf{v}_i \in W_i$, hence $\mathbf{v}_1 = ... = \mathbf{v}_q = 0$, once again by induction hypothesis. Finally, since

$$N : U := U_1 \oplus ... \oplus U_q \longrightarrow N(V)$$

is onto, we may write

$$V = U \oplus V_0$$

with $V_0 \subset \ker(N)$. Take any subspace $V_1 \subset V$ with $V = U \oplus V_1$ and choose a basis $\mathbf{v}_{q+1}, ..., \mathbf{v}_t$ of V_1. Choose $\mathbf{u}_{q+1}, ..., \mathbf{u}_t \in U$ with $N(\mathbf{u}_i) = N(\mathbf{w}_i), i = q+1, ..., t$. Then the vectors $\mathbf{w}_i := \mathbf{v}_i - \mathbf{u}_i, i = q+1, ..., t$, span a complementary subspace $V_0 \subset \ker(N)$. Now choose

$$U_i := K\mathbf{w}_i, i = q+1, ..., t.$$

Finally, we have $t = \dim\ker(N)$ and for $k > 0$ the numbers

$$\dim\ker(N^{k+1}) - \dim\ker(N^k) = |\{i; 1 \le i \le t, \dim U_i > k\}|$$

determine the dimensions $\dim U_i$. □

Theorem 4.20. *Let $T \in \mathcal{L}(V)$ be split. Then we may write*

$$V = \bigoplus_{i=1}^{r} V_i$$

with T-invariant subspaces V_i, such that

$$T|_{V_i} = \lambda_i \mathrm{id}_{V_i} + N_i$$

with an eigenvalue $\lambda_i \in K$ of T and V_i is N_i-cyclic.

Remark 4.21. It is well known that given a polynomial $f \in K[X]$ there is a field $L \supset K$, such that $f \in L[X]$ is split. E.g. for $K = \mathbb{R}$ and any polynomial $f \in K[X]$ the choice $L = \mathbb{C}$ is possible.

Now, if $V = K^n$ and

$$T = T_A : K^n \longrightarrow K^n, \mathbf{x} \mapsto A\mathbf{x},$$

we have a natural extension

$$\widetilde{T} : L^n \longrightarrow L^n, \mathbf{z} \mapsto A\mathbf{z}.$$

Hence we may write

$$L^n = \bigoplus_{i=1}^{r} W_i$$

as in Th.4.20, but that decomposition does not descend to $K^n \subset L^n$. Indeed, the subspaces $V_i := W_i \cap K^n$ will be in general too small.

Finally we indicate how to find a real Jordan form for

$$T := T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^n.$$

We consider its extension

$$\widetilde{T} : \mathbb{C}^n \longrightarrow \mathbb{C}^n.$$

Consider an eigenvalue $\lambda \in \mathbb{C}$ of T. There are two cases:

1. If $\lambda \in \mathbb{R}$, then λ is an eigenvalue of T and we may decompose $\widehat{V}_\lambda \cap \mathbb{R}$ as in the split case as a direct sum of $(T - \lambda \mathrm{id}_V)$-cyclic subspaces.
2. If $\lambda \in \mathbb{C}, \Im(\lambda) > 0$, then $\overline{\lambda}$ is an eigenvalue of $\widetilde{T}$ as well. We choose a decomposition of $\widehat{V}_\lambda$ and take for $\widehat{V}_{\overline{\lambda}}$ the complex conjugates of the subspaces in the decomposition of $\widehat{V}_\lambda$.

Hence we have to deal with the following situation: We have a T-invariant subspace

$$V \subset \mathbb{C}^n,$$

such that

$$T|_V = \lambda \mathrm{id}_V + N,\ \lambda = \alpha + i\beta$$

and V is N-cyclic. Now consider a basis $N^s(\mathbf{z}), N^{s-1}(\mathbf{z}), .., N(\mathbf{z}), \mathbf{z}$ of V. Write $\mathbf{z} = \mathbf{x} + i\mathbf{y}$ - note that $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are linearly independent, since $N^s(\mathbf{x}), N^s(\mathbf{y})$ are as real and imaginary part of an eigenvector belonging to an eigenvalue $\lambda \notin \mathbb{R}$. Then we obtain a basis for

$$(V \oplus \overline{V}) \cap \mathbb{R}^n$$

as follows:

$$N^s(\mathbf{y}), N^s(\mathbf{x}), ..., N(\mathbf{y}), N(\mathbf{x}), \mathbf{y}, \mathbf{x}.$$

Since $\mathbf{x} = \frac{1}{2}(\mathbf{z} + \overline{\mathbf{z}}), \mathbf{y} = \frac{1}{2i}(\mathbf{z} + \overline{\mathbf{z}})$, we obtain

$$T(\mathbf{x}) = \alpha\mathbf{x} - \beta\mathbf{y} + N(\mathbf{x}),\ T(\mathbf{y}) = \beta\mathbf{x} + \alpha\mathbf{y} + N(\mathbf{y})$$

as well as the analogous statements for $N^j(\mathbf{x}), N^j(\mathbf{y})$. Hence with respect to that basis T has the matrix

$$\begin{pmatrix} A & I_2 & 0 & .. & 0 \\ 0 & A & I_2 & .. & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & .. & A & I_2 \\ 0 & 0 & .. & 0 & A \end{pmatrix}$$

with

$$A = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}.$$

5 Minimal and characteristic polynomial

Polynomials $f \in K[X]$ may be evaluated not only at elements of the base field K, but even at linear operators $T \in \mathcal{L}(V)$ as follows: Given

$$f = a_r X^r + ... + a_1 X + a_0 \in K[X]$$

we define

$$f(T) := a_r T^r + ... + a_1 T + a_0 \mathrm{id}_V \in \mathcal{L}(V).$$

Proposition 5.1. *Given $T \in \mathcal{L}(V)$ there is a unique monic polynomial*

$$\mu_T \in K[X]$$

of minimal degree satisfying

$$\mu_T(T) = 0.$$

We have then for a polynomial $f \in K[T]$ the following equivalence:

$$f(T) = 0 \Longleftrightarrow \mu_T | f.$$

We call μ_T the minimal polynomial *of the operator T and obtain*

$$K[T] = K\mathrm{id}_V \oplus KT \oplus ... \oplus KT^{r-1},\ r = \deg(\mu_T).$$

Proof. Since $\dim \mathcal{L}(V) = (\dim V)^2 < \infty$, we may choose $r \in \mathbb{N}$ minimal such that $\mathrm{id}_V, T, ..., T^r \in \mathcal{L}(V)$ are linearly independent. Then there is a nontrivial relation

$$a_r T^r + ... + a_1 T + a_0 \mathrm{id}_V = 0$$

with $a_r \neq 0$, since $T^{r-1}, ..., T, \mathrm{id}_V$ are linearly independent. We may even assume $a_r = 1$. Take

$$\mu_T := X^r + a_{r-1}X^{r-1} + ... + a_1 X + a_0.$$

Obviously

$$K\mathrm{id}_V + KT + ... + KT^{r-1} = K\mathrm{id}_V \oplus KT \oplus ... \oplus KT^{r-1},$$

in order to see that the sum is $K[T]$, take some $f(T) \in K[T]$ and write

$$f = g\mu_T + h, \deg h < \deg \mu_T = r.$$

Then it follows

$$f(T) = h(T) \in K\mathrm{id}_V + KT + ... + KT^{r-1}.$$

□

Here is an explicit description of the minimal polynomial of a split operator:

Theorem 5.2. *Let $T \in \mathcal{L}(V)$ be a split operator and*

$$V = \bigoplus_{i=1}^{r} \widehat{V}_{\lambda_i}$$

the decomposition of V as the direct sum of the generalized eigenspaces of T. Then we have

$$\mu_T = \prod_{i=1}^{r} (X - \lambda_i)^{m_i},$$

where m_i is the maximal dimension of a $(T - \lambda_i \mathrm{id}_V)$-cyclic subspace of $\widehat{V}_{\lambda_i}$ in some direct sum decomposition.

For the proof we need:

Remark 5.3. 1. Since $T^k T^\ell = T^\ell T^k$ we see that the subspace

$$K[T] := \{f(T); f \in K[X]\}$$

forms a commutative subring

$$K[T] \subset \mathcal{L}(V)$$

of the endomorphism ring $\mathcal{L}(V)$ of the vector space V.

2. If $T = \lambda \mathrm{id}_V + N$ with a nilpotent operator we have

$$K[T] = K\mathrm{id}_V \oplus NK[N],$$

where $NK[N]$ consists of the nilpotent operators in $K[T]$ and the complement

$$K[T] \setminus NK[N]$$

consists of isomorphims only. For the proof we have to invert $\lambda \mathrm{id}_V + \tilde{N}$ with $\lambda \neq 0$ and $\tilde{N} \in NK[N]$. We may even assume $\lambda = 1$ and check that

$$(\mathrm{id}_V + \tilde{N})^{-1} = \mathrm{id}_V - \tilde{N} + \tilde{N}^2 - ... + (-1)^{n-1}\tilde{N}^{n-1}$$

with $n := \dim V$.

3. For $f \in K[X]$ we have:

$$f(T) = f(\lambda)\mathrm{id}_V + \tilde{N}$$

with some $\tilde{N} \in NK[N]$.

In particular:

1. If $f(\lambda) = 0$, then $f(T) \in NK[N]$ is nilpotent.

2. If $f(\lambda) \neq 0$, then $f(T) : V \longrightarrow V$ is an isomorphism.

Proof of 5.2. Let

$$\nu_T = \prod_{i=1}^{r} (X - \lambda_i)^{m_i}.$$

We show that

$$f(T) = 0 \Longleftrightarrow \nu_T | f$$

holds for $f \in K[X]$. First of all we have

$$f(T) = 0 \Longleftrightarrow f(T_i) = 0, i = 1, ..., r$$

with the restrictions

$$T_i := T|_{\widehat{V}_{\lambda_i}}.$$

Let us now show that $f(T) = 0$ implies $\nu_T | f$. According to Rem.5.3 we then have $f(\lambda_i) = 0, i = 1, ..., r$, and may write

$$f = hg,\ h = \prod_{i=1}^{r} (X - \lambda_i)^{k_i}$$

where $k_1, ..., k_r > 0$ and $g(\lambda_i) \neq 0$ for $i = 1, ..., r$. Then we see once again with Rem.5.3, that all $g(T_i)$ are isomorphisms, hence so is $g(T)$ and thus $h(T) = 0$. Furthermore writing

$$h = h_i p_i, h_i = (X - \lambda_i)^{k_i},$$

we see with the same argument that $p_i(T_i)$ is an isomorphism, hence $h_i(T_i) = 0$ and then necessarily $k_i \geq m_i$ resp. $\nu_T | f$. For the reverse implication it suffices to show $\nu_T(T) = 0$. But that follows immediately from $\nu_T(T_i) = 0$ for $i = 1, ..., r$. □

Theorem 5.4 (Cayley-Hamilton)**.** *For $T \in \mathcal{L}(V)$ we have*

$$\chi_T(T) = 0 \in \mathcal{L}(V).$$

Proof. We may assume $V = K^n$, and furthermore, by Rem.4.21, even that T is split - the operators T and $\widetilde{T}$ have the same matrices and thus the same characteristic polynomials. Now we look at the decomposition

$$V = \bigoplus_{i=1}^{r} \widehat{V}_{\lambda_i}$$

of V as the direct sum of the generalized eigenspaces of T. For the characteristic polynomial χ_T we have

$$\chi_T(X) = \prod_{i=1}^{r} (X - \lambda_i)^{n_i}$$

with $n_i := \dim \widehat{V}_{\lambda_i} \geq m_i$, where $m_i > 0$ is as in Th.5.2. □

Definition 5.5. A split operator $T \in \mathcal{L}(V)$ is called *diagonalizable* or *semisimple*, iff V admits a basis consisting of eigenvectors of T.

Proposition 5.6. *For a split operator $T \in \mathcal{L}(V)$ the following statements are equivalent:*

1. *$T \in \mathcal{L}(V)$ is diagonalizable.*

2. *The minimal polynomial μ_T has only simple roots.*

3. *$\widehat{V}_\lambda = V_\lambda$ for all eigenvalues λ of T.*

6 Spectral Theorem

An essential point in the proof of the Jordan normal form was to find, given an operator $T : V \longrightarrow V$, a decomposition of V as a direct sum of invariant subspaces: $V = U \oplus W$. For $K = \mathbb{R}, \mathbb{C}$ vector spaces come often with some additional structure, which allows to measure lengths and angles and to define canonical complementary subspaces $U^\perp$ for a given subspace U.

Definition 6.1. An inner product on a real or complex vector space V is a real bilinear map

$$\sigma : V \times V \longrightarrow K$$

such that

1. $$\sigma(.., \mathbf{w}) \in V^*, \ \forall \mathbf{w} \in V,$$

   i.e. $\sigma(.., \mathbf{w})$ is $\mathbb{C}$-linear for all $\mathbf{w} \in V$,

2. $$\sigma(\mathbf{w}, \mathbf{v}) = \overline{\sigma(\mathbf{v}, \mathbf{w})}$$

3. and
   $$\sigma(\mathbf{v}, \mathbf{v}) \in \mathbb{R}_{>0}, \ \forall \mathbf{v} \in V \setminus \{0\}.$$

An inner product space is a pair (V, σ) with a real or complex vector space V and an inner product $\sigma : V \times V \longrightarrow K$.

Example 6.2. 1. An inner product $\sigma : K^n \times K^n \longrightarrow K$ can be written

$$\sigma(\mathbf{v}, \mathbf{w}) : \mathbf{v}^T S \overline{\mathbf{w}}$$

with a matrix $S \in K^{n,n}$, such that $\overline{S}^T = S$ and

$$\mathbf{v}^T S \overline{\mathbf{v}} > 0$$

holds for all $\mathbf{v} \in V \setminus \{0\}$.

2. $V = K[X]$ with

$$\sigma(f, g) = \int_a^b f(x)\overline{g(x)}dx,$$

where $a < b$.

Remark 6.3. 1. An $\mathbb{R}$-linear homomorphism $T : V \longrightarrow W$ between complex vector spaces is called antilinear iff $T(\lambda \mathbf{v}) = \overline{\lambda} T(\mathbf{v})$.

2. We have an (antilinear) isomorphism

$$\Phi_\sigma : V \longrightarrow V^*, \mathbf{w} \mapsto \sigma(.., \mathbf{w}).$$

Definition 6.4. The adjoint of an operator $T : V \longrightarrow V$ is the is the linear map

$$\Phi_\sigma \circ T^* \circ (\Phi_\sigma)^{-1} : V \longrightarrow V^* \longrightarrow V^* \longrightarrow V,$$

indeed, we shall identify V^* with V using the isomorphism Φ_σ and denote it T^* as well.

Remark 6.5. 1. The adjoint T^* of T is the unique linear map T^* satisfying

$$\sigma(T(\mathbf{v}), \mathbf{w}) = \sigma(\mathbf{v}, T^*(\mathbf{w})).$$

2. Conjugating that equality we obtain

$$\sigma(T^*(\mathbf{v}), \mathbf{w}) = \sigma(\mathbf{v}, T(\mathbf{w})),$$

3. so in particular

$$T^{**} = T.$$

4. $\mathrm{id}^* = \mathrm{id}, (R + T)^* = R^* + T^*, (RT)^* = T^* R^*, (\lambda T)^* = \overline{\lambda} T^*.$

5. If $\sigma(\mathbf{v},\mathbf{w}) = \mathbf{v}^T S\overline{\mathbf{w}}$ and $T = T_A \in \mathcal{L}(K^n)$, we have $T^* = T_{A^*}$, where $A^T S = S\overline{A^*}$, i.e.
$$A^* = \overline{S}^{-1}\overline{A}^T S,$$
in particular, for $S = I_n$ we find
$$A^* = \overline{A}^T.$$

Definition 6.6. A linear operator $T \in \mathcal{L}(V)$ on an inner product space is called

1. normal if it commutes with its adjoint
$$T^*T = TT^*,$$

2. self adjoint (symmetric for $K = \mathbb{R}$) if $T^* = T$.

Example 6.7. 1. $T = \lambda \mathrm{id}_V$ is self adjoint iff $\lambda \in \mathbb{R}$.

2. If $T^* = \pm T$ or $T^* = T^{-1}$, the operator T is normal.

3. With T is also λT normal, but the sum of normal operators need not be normal: Indeed, any operator is the sum of a self adjoint and a skew adjoint ($T^* = -T$) operator.

4. The self adjoint operators form a real subspace of $\mathcal{L}(V)$.

5. The composition of two self adjoint operators is again self adjoint, if they commute, otherwise in general not.

Proposition 6.8. *Let $T \in \mathcal{L}(V)$ be a normal operator. If $\lambda \in K$ is an eigenvalue of T, then $\overline{\lambda}$ is an eigenvalue of T^*, and the corresponding eigenspaces coincide.*

Proof. Since T and T^* commute, the eigenspace V_λ of T is T^*-invariant, and then
$$T^*|_{V_\lambda} = (T|_{V_\lambda})^* = (\lambda\ \mathrm{id}_{V_\lambda})^* = \overline{\lambda}\ \mathrm{id}_{V_\lambda}\ .$$
□

Corollary 6.9. *1. The eigenvalues of a self adjoint operator are real.*

2. *A symmetric operator on a real inner product space has a real eigenvalue.*

Proof. The first part is obvious with the preceding proposition, and for the second part we may assume $V = \mathbb{R}^n$. Both T and σ extend to a self adjoint operator resp. an inner product on $\mathbb{C}^n$. The extended operator has an eigenvalue, which is a real number - but the extended operator and T have the same characteristic polynomial, hence it is an eigenvalue of T as well. □

Proposition 6.10. *For a normal operator $T \in \mathcal{L}(V)$ we have*

$$V_\lambda \perp V_\mu$$

for eigenvalues $\mu \neq \lambda$.

Proof. For $\mathbf{u} \in V_\lambda, \mathbf{v} \in V_\mu$ we have

$$(\lambda - \mu)\sigma(\mathbf{u}, \mathbf{v}) = \sigma(\lambda\mathbf{u}, \mathbf{v}) - \sigma(\mathbf{u}, \overline{\mu}\mathbf{v})$$
$$(\sigma(T(\mathbf{u}), \mathbf{v}) - \sigma(\mathbf{u}, T^*(\mathbf{v}))) = 0,$$

hence $\lambda \neq \mu$ implies $\sigma(\mathbf{u}, \mathbf{v}) = 0$. □

Indeed, in the complex case V is the orthogonal sum of the eigenspaces $V_\lambda, \lambda \in \mathbb{C}$.

Theorem 6.11 (Spectral Theorem)**.** 1. *A normal operator $T \in \mathcal{L}(V)$ on a complex inner product space is diagonalizable. Indeed, there is an ON-basis of V consisting of eigenvectors of T.*

2. *A symmetric operator $T \in \mathcal{L}(V)$ on a real inner product space is diagonalizable. Indeed, there is an ON-basis of V consisting of eigenvectors of T.*

Proof. Induction on $\dim V$. In both cases there is an eigenvalue $\lambda \in K$ of T. We show that

$$U := V_\lambda^\perp$$

is both T- and T^*-invariant and then may apply the induction hypothesis to $T|_U$. Indeed for $\mathbf{v} \in V_\lambda, \mathbf{u} \in U$ we have:

$$\sigma(T(\mathbf{u}), \mathbf{v}) = \sigma(\mathbf{u}, T^*(\mathbf{v})) = \sigma(\mathbf{u}, \overline{\lambda}\mathbf{v}) = \lambda\sigma(\mathbf{u}, \mathbf{v}) = 0$$

as well as

$$\sigma(T^*(\mathbf{u}), \mathbf{v}) = \sigma(\mathbf{u}, T(\mathbf{v})) = \sigma(\mathbf{u}, \lambda\mathbf{v}) = \overline{\lambda}\sigma(\mathbf{u}, \mathbf{v}) = 0.$$

□

Corollary 6.12. *For a normal operator $T \in \mathcal{L}(V)$ we have:*

$$||T(\mathbf{v})|| = ||T^*(\mathbf{v})||.$$

Proof. We may assume $V = \mathbb{C}^n$ and apply the complex spectral theorem, i.e. the first part of Th.6.11. But there is a straight forward argument as well:

$$0 = \sigma((TT^* - T^*T)(\mathbf{v}), \mathbf{v}) = \sigma(TT^*(\mathbf{v}), \mathbf{v})) - \sigma(T^*T(\mathbf{v}), \mathbf{v})$$

$$\sigma(T^*(\mathbf{v}), T^*(\mathbf{v})) - \sigma(T(\mathbf{v}), T(\mathbf{v})) = ||T^*(\mathbf{v})||^2 - ||T(\mathbf{v})||^2.$$

□

Definition 6.13. A linear operator $T \in \mathcal{L}(V)$ is called an isometry iff

$$\sigma(T(\mathbf{u}), T(\mathbf{v})) = \sigma(\mathbf{u}, \mathbf{v})$$

holds for all $\mathbf{u}, \mathbf{v} \in V$. Such an isometry is also called

1. an *orthogonal transformation* if $K = \mathbb{R}$,
2. a *unitary transformation* if $K = \mathbb{C}$.

Remark 6.14.

1. T is an isometry iff $T^* = T^{-1}$.
2. A linear operator is an isometry iff it preserves lengths, i.e.

   $$||T(\mathbf{u})|| = ||\mathbf{u}||.$$

   The proof follows from the fact that

   $$4\Re(\sigma(\mathbf{u}, \mathbf{v})) = ||\mathbf{u} + \mathbf{v}||^2 - ||\mathbf{u} - \mathbf{v}||^2$$

   and

   $$4\Im(\sigma(\mathbf{u}, \mathbf{v})) = 2\Re(\sigma(\mathbf{u}, i\mathbf{v})) = ||\mathbf{u} + i\mathbf{v}||^2 - ||\mathbf{u} - i\mathbf{v}||^2.$$

3. An isometry is normal.
4. An isometry $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ with $\det(T) = 1$ is a rotation.

Definition 6.15. A rotation on a two dimensional real inner product space is an isometry $T \in \mathcal{L}(V)$ with $\det(T) = 1$.

Theorem 6.16. *Let $T \in \mathcal{L}(V)$ be an isometry of the inner product space V.*

1. *If $K = \mathbb{C}$ there is an ON-basis $\mathbf{v}_1, ..., \mathbf{v}_n$ of V, such that*

$$T(\mathbf{v}_i) = \lambda_i \mathbf{v}_i, i = 1, ..., n$$

with $\lambda_i \in \mathbb{C}, |\lambda_i| = 1$.

2. *If $K = \mathbb{R}$ there is an orthogonal decomposition*

$$V = \bigoplus_{i=1}^{r} V_i$$

with T-invariant subspaces of dimension 1 or 2, such that $T|_{V_i}$ is multiplication with ± 1 $(\dim V_i = 1)$ or a rotation with an angle between $\vartheta, 0 < |\vartheta| < \pi$ $(\dim V_i = 2)$.

Definition 6.17. A self adjoint linear operator is called positive if all its eigenvalues are non-negative real numbers.

Remark 6.18. A self adjoint linear operator is positive iff

$$\sigma(T(\mathbf{v}), \mathbf{v}) \geq 0$$

holds for all $\mathbf{v} \in V$. The condition is obviously necessary, but it is sufficient as well: If $\mathbf{v} = \sum_{i=1}^{n} \mathbf{v}_i$ with pairwise orthogonal eigenvectors $\mathbf{v}_i \in V$ of T, we have

$$\sigma(T(\mathbf{v}), \mathbf{v}) = \sum_{i=1}^{n} \lambda_i ||\mathbf{v}_i||^2.$$

A positive operator $T \in \mathcal{L}(V)$ can uniquely be written

$$T = F^2$$

with a positive operator $F \in \mathcal{L}(V)$. We write

$$\sqrt{T} := F.$$

Example 6.19. For an isomorphism $S \in \mathcal{L}(V)$ the operator $T = S^*S$ is positive. Indeed

$$T^* = (S^*S)^* = S^*S^{**} = S^*S = T$$

and

$$\sigma(T(\mathbf{u}), \mathbf{u}) = \sigma(S(\mathbf{u}), S(\mathbf{u})) > 0$$

for $\mathbf{u} \neq 0$.

Theorem 6.20 (Polar decomposition)**.** *Any $T \in \mathcal{L}(V)$ can be decomposed*

$$T = SF$$

with an isometry S and a positive operator F. If T is an isomorphism, the decomposition is unique.

Proof. Uniqueness: If T is an isomorphism, so is F. We obtain

$$T^*T = (SF)^*SF = F^*S^*SF = F^*F = F^2,$$

i.e. $F = \sqrt{T^*T}$. Sedan följer

$$S = TF^{-1}.$$

Existence: We set

$$F := \sqrt{T^*T}$$

and show

$$||T(\mathbf{v})|| = ||F(\mathbf{v})||,$$

in particular $U := \ker(T) = \ker(F)$. Indeed

$$||T(\mathbf{v})||^2 = \sigma(T(\mathbf{v}), T(\mathbf{v})) = \sigma(T^*T(\mathbf{v}), \mathbf{v})$$

$$= \sigma(F^2(\mathbf{v}), \mathbf{v}) = \sigma(F(\mathbf{v}), F(\mathbf{v})) = ||F(\mathbf{v})||^2.$$

We obtain a commutative triangle

$$\begin{array}{cccc} & F(V) & \stackrel{S_0}{\longrightarrow} & T(V) \\ \overline{F} & \uparrow & \nearrow & \\ & V/U & \overline{T} & \end{array}$$

with isomorphisms $\overline{F} : V/U \longrightarrow F(V), \overline{T} : V/U \longrightarrow T(V)$ and an isometry $S_0 : F(V) \longrightarrow T(V)$. Now take any isometry

$$S_1 : F(V)^{\perp} \longrightarrow T(V)^{\perp}$$

and set

$$S = S_0 \oplus S_1.$$

□

7 Determinants and exterior algebra

Definition 7.1. Let V be a vector space, $\dim V = n$.

1. A k-linear form (or simply k-form) on V is a map

$$\alpha : V^k := \underbrace{V \times ... \times V}_{k \text{ times}} \longrightarrow K,$$

such that

$$V \longrightarrow K, \mathbf{v} \mapsto \alpha(\mathbf{v}_1, ..., \mathbf{v}_{i-1}, \mathbf{v}, \mathbf{v}_{i+1}, ..., \mathbf{v}_n)$$

is a linear form for all $i = 1, ..., n$ and $\mathbf{v}_1, ..., \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, ..., \mathbf{v}_n \in V$. We denote $M_k(V)$ (multilinear) the vector space of all k-forms on V.

2. An alternating k-form on V is a k-form α, such that

$$\exists\, i \neq j : \mathbf{v}_i = \mathbf{v}_j \Longrightarrow \alpha(\mathbf{v}_1, ..., \mathbf{v}_k) = 0.$$

We denote $A_k(V) \subset M_k(V)$ the vector space of all alternating k-forms on V.

Remark 7.2. Let $(\mathbf{e}_1, ..., \mathbf{e}_n)$ be a basis of the vector space V.

1. A k-form $\alpha : V^k \longrightarrow K$ is determined by the values

$$\alpha(\mathbf{e}_{i_1}, ..., \mathbf{e}_{i_k}), 1 \leq i_\nu \leq n,$$

and they can be prescribed arbitrarily. In particular $\dim M_k(V) = n^k$.

2. For an alternating k-form and a permutation $\pi : \{1, ..., k\} \longrightarrow \{1, ..., k\}$ we have

$$\alpha(\mathbf{v}_{i_{\pi(1)}}, ..., \mathbf{v}_{i_{\pi(k)}}) = \varepsilon(\pi)\alpha(\mathbf{v}_{i_1}, ..., \mathbf{v}_{i_k}).$$

This follows from the fact that permutations can be factorized into transpositions and a transposition τ has sign $\varepsilon(\tau) = -1$.

3. A k-form is alternating iff the above condition is satisfied for basis vectors $\mathbf{v}_1, ..., \mathbf{v}_k \in \{\mathbf{e}_1, ..., \mathbf{e}_n\}$.

4. An alternating k-form α is is determined by the values

$$\alpha(\mathbf{e}_{i_1}, ..., \mathbf{e}_{i_k}), i_1 < ... < i_k$$

and they can be prescribed arbitrarily. In particular $\dim A_k(V) = \binom{n}{k}$.

5. An alternating n-form $\alpha \in A_n(V)$ is trivial iff $\alpha(\mathbf{e}_1, ..., \mathbf{e}_n) = 0$.

6. A linear operator $F : V \longrightarrow W$ induces pull back homomorphisms

$$T^* : M_k(W) \longrightarrow M_k(V), A_k(W) \longrightarrow A_k(V),$$

not to confused with the adjoint of T.

Since for an n-dimensional vector space V we have $\dim A_n(V) = 1$, we may define the determinant of a linear operator $T \in \mathcal{L}(V)$ as follows:

Definition 7.3. The determinant $\det(T) \in K$ of an operator $T \in \mathcal{L}(V)$ is defined by

$$T^* = \mu_{\det(K)},$$

where $T^* : A_n(V) \longrightarrow A_n(V)$ denotes the pull back of n-forms and $\mu_a : A_n(V) \longrightarrow A_n(V)$ scalar multiplication with $a \in K$.

There are some immediate remarks:

Proposition 7.4. *1.* $\det(ST) = \det(S)\det(T)$.

2. $\det(\lambda \mathrm{id}_V) = \lambda^n$

3. $\det(T) \neq 0$ *iff* T *is an isomorphism.*

Proof. Exercise! □

We want to define a product for alternating forms. First of all there is a bilinear map, the tensor product of multilinear forms:

$$M_k(V) \times M_\ell(V) \longrightarrow M_{k+\ell}, (\alpha, \beta) \mapsto \alpha \otimes \beta,$$

with

$$\alpha \otimes \beta(\mathbf{v}_1, ..., \mathbf{v}_{k+\ell}) := \alpha(\mathbf{v}_1, ..., \mathbf{v}_k)\beta(\mathbf{v}_{k+1}, ..., \mathbf{v}_{k+\ell}).$$

Next we define a projection operator $M_k(V) \longrightarrow A_k(V)$ as follows:

Definition 7.5. Let $\mathrm{char}(K) = 0$. We define

$$\mathrm{Alt}_k : M_k(V) \longrightarrow A_k(V)$$

by

$$\mathrm{Alt}_k(\varphi)(\mathbf{v}_1, ..., \mathbf{v}_k) := \frac{1}{k!} \sum_{\pi \in \mathbf{S}_k} \varepsilon(\pi)\varphi(\mathbf{v}_{\pi(1)}, ..., \mathbf{v}_{\pi(k)}),$$

where $\mathbf{S}_k$ denotes the set (group) of all permutations $\pi : \{1, ..., k\} \longrightarrow \{1, ..., k\}$. In order to see that the resulting form is alternating use the fact that $\varepsilon(\pi \circ \varphi) = \varepsilon(\pi)\varepsilon(\varphi)$.

Remark 7.6. 1. The map Alt_k is a projection:

$$\mathrm{Alt}_k|_{A_k(V)} = \mathrm{id}$$

Definition 7.7. 1. The exterior product of the alternating forms $\alpha \in A_k(V), \beta \in A_\ell(V)$ is

$$\alpha \wedge \beta := \frac{(k+\ell)!}{k! \cdot \ell!} \mathrm{Alt}_{k+\ell}(\alpha \otimes \beta) \in A_{k+\ell}(V)$$

2.

$$A_*(V) := \bigoplus_{k=0}^{n} A_k(V)$$

is an associative algebra with the linear extension of the above wedge product. Furthermore it satisfies the "graded commutativity rule"

$$\alpha \wedge \beta = (-1)^{k\ell} \beta \wedge \alpha$$

for $\alpha \in A_k(V), \beta \in A_\ell(V)$.

Proof. We comment on the associativity: For convenience of notation let us write

$$[\alpha] := \mathrm{Alt}_k(\alpha), \alpha \in A_k(V),$$

such that

$$\alpha \wedge \beta = \frac{(k+\ell)!}{k! \cdot \ell!} [\alpha \otimes \beta]$$

First of all one proves

$$[\alpha \otimes \beta] = [[\alpha] \otimes [\beta]].$$

Then one obtains for $\alpha \in A_k(V), \beta \in A_\ell(V), \gamma \in A_m(V)$ the following:

$$(\alpha \wedge \beta) \wedge \gamma = \frac{(k+\ell+m)!}{(k+\ell)! \cdot m!} [(\alpha \wedge \beta) \otimes \gamma]$$

$$= \frac{(k+\ell+m)!}{(k+\ell)! \cdot m!} \left[\frac{(k+\ell)!}{k! \cdot \ell!} [\alpha \otimes \beta] \otimes \gamma \right]$$

$$= \frac{(k+\ell+m)!}{k! \cdot \ell! \cdot m!} [\alpha \otimes \beta \otimes \gamma],$$

using $[\gamma] = \gamma$. The same reasoning works for $\alpha \wedge (\beta \wedge \gamma)$.

In order to show the above "graded commutativity" of the wedge product one checks immediately that $\beta \wedge \alpha = -\alpha \wedge \beta$ holds for 1-forms. Then we may assume $\alpha = \alpha_1 \wedge ... \wedge \alpha_k$ and $\beta = \beta_1 \wedge ... \wedge \beta_\ell$, see the below Prop.7.8. We move the factors α_i successively:

$$\alpha \wedge \beta = \alpha_1 \wedge ... \wedge \alpha_k \wedge \beta_1 \wedge ... \wedge \beta_\ell = (-1)^\ell \alpha_1 \wedge ... \wedge \alpha_{k-1} \wedge \beta_1 \wedge ... \wedge \beta_\ell \wedge \alpha_k$$

and obtain our formula after having done that k times. □

Proposition 7.8. *Let* $\mathbf{e}_1^*, ..., \mathbf{e}_n^* \in V^* = A_1(V)$ *be the dual basis of* $\mathbf{e}_1, ..., \mathbf{e}_n \in V$. *Assume* $1 \leq i_1 < .... < i_k \leq n$ *and* $1 \leq j_1 < ... < j_k \leq n$. *Then*

$$\mathbf{e}_{i_1}^* \wedge ... \wedge \mathbf{e}_{i_k}^*(\mathbf{e}_{j_1}, ..., \mathbf{e}_{j_k}) = \begin{cases} 1 & , \quad \textit{if } j_\nu = i_\nu, \nu = 1, ..., k \\ 0 & , \quad \textit{otherwise} \end{cases}$$

and

$$A_k(V) = \bigoplus_{1 \leq i_1 < ... < i_k \leq n} K \cdot \mathbf{e}_{i_1}^* \wedge ... \wedge \mathbf{e}_{i_k}^*.$$