As of 11th of May 2016, I am not attending LinAlg2 anymore due to difficulties with my upcoming student exchange year and work on my master thesis.
I therefore suppose this document will not be updated any more.

Linear Algebra 2 – Lecture Notes

Lukas Prokop

summer term 2016

Contents

This lecture took place on 29th of Feb 2016 (Prof. Franz Lehner).

Exam: written and orally

Tutorial session:

- Every Monday, 18:30-20:00, SR 11.34
- Contact: gernot.holler@edu.uni-graz.at

Konversatorium:

- Every Monday, 10:00–10:45, SR 11.33

Topics, wie already discussed:

- Vector spaces
- Linear maps and their equivalence with matrices
- We introduced equivalence of matrices ($PAQ = B$)
- We defined the following techniques:
  - Rank
  - Linear equation system
  - Inverse matrices
  - Basis transformation

In this semester, we will discuss:

- PAP^{-1}, which is related to eigenvalues and diagonalization, hence $\bigvee^?_P PAP^{-1} = D$.

1 Linear maps (cont.)

1.1 Addition to chapter 5.2.4

$\mathrm{Hom}(V, W)$ in special case $W = \mathbb{K}$. We define,

$$V^* := \mathrm{Hom}(V, \mathbb{K})$$

also denoted V' is called *dual space* of vector space V. The elements $v* \in V*$ are called *linear forms* or *linear functionals*.

We denote,

$$v^*(v) =: \langle v*, v \rangle$$

1.2 Example

$$V = \mathbb{K}^n$$

$v^* : V \to \mathbb{K}$ is uniquely defined with values $v^*(e_i) =: a_i$.

$$\langle v^*, v \rangle = \left\langle v^*, \sum_{i=1}^n v_i e_i \right\rangle = \sum_{i=1}^n v_i \langle v^*, e_i \rangle$$

$$v^* \left(\sum_{i=1}^n v_i e_i \right) = \sum_{i=1}^n v_i v^*(e_i) = \sum_{i=1}^n a_i v_i$$

1.3 More general

We know, $\dim \mathrm{Hom}(V, W) = \dim V \cdot \dim W$.

Theorem 1. Let V be a vector space over $\mathbb{K}$.

- $\dim V =: n < \infty \Rightarrow \dim V^* = n$
  More precisely: Let $(b_1, \ldots, b_n)$ be a basis of V. Then

  $$b_k^* : b_i \mapsto \delta_{ik} = \begin{cases} 1 & i = k \\ 0 & \text{else} \end{cases}$$

  is a basis of V^* and is called *dual basis*.

- For $v^* \in V^*$ it holds that $v^* = \sum_{k=1}^n \langle v^*, b_k \rangle \cdot b_k^*$.
- If $\dim V = \infty, (b_i)_{i \in I}$ is a basis, then it holds that $(b_k^*)_{k \in I}$ with

  $$\langle b_k^*, b_i \rangle = \delta_{ik}$$

  is *not* a basis of V^*.

Proof.
- Special case of 5.18
  (b_k^*) is linear independent, hence in $\sum_{i=1}^n \lambda_i b_i^* = 0$ all $\lambda_i = 0$.

  $$0 = \left\langle \sum_{i=1}^n \lambda_i b_i^*, b_k \right\rangle = \sum_{i=1}^n \lambda_i \left\langle \underbrace{b_i^*, b_k}_{\delta_{ik}} \right\rangle = \lambda_k \forall k$$

- Let $v \in V$ with $v = \sum_{i=1}^n v_i b_i$. We need to show

  $$\langle v^*, v \rangle \stackrel{!}{=} \left\langle \sum_{k=1}^n \langle v^*, b_w \rangle b_n^*, v \right\rangle$$

  $$\begin{aligned} \left\langle \sum_{k=1}^n \langle v^*, b_k \rangle b_k^*, v \right\rangle &= \sum_{k=1}^n \langle v^*, b_k \rangle \langle b_k^*, v \rangle \\ &= \sum_{k=1}^n \langle v^*, b_k \rangle \left\langle b_k^*, \sum_{i=1}^n v_i b_i \right\rangle \\ &= \sum_{k=1}^n \sum_{i=1}^n \langle v^*, b_k \rangle \underbrace{\langle b_k^*, b_i \rangle}_{\delta_{ki}} \cdot v_i \\ &= \sum_{k=1}^n \langle v^*, b_k \rangle \langle v^*, b_k \rangle \cdot v_k \\ &= \left\langle v^*, \sum_{k=1}^n v_k b_k \right\rangle \\ &= \langle v^*, v \rangle \end{aligned}$$

- (To be done in the practicals) Consider the functional

  $$\langle v^*, b_i \rangle = 1 \Rightarrow v^* \notin L((v_i^*)_{i \in I})$$

$\square$

1.4 Remark and a definition for bilinearity

The mapping $V^* \times V \to \mathbb{K}$ is linear in v (with fixed v^*) with $(v^*, v) \mapsto \langle v^*, v\rangle$ is linear in v^* (with fixed v). Such a mapping is called *bilinear*.

A mapping $F : V_1 \times \ldots \times V_n \to W$ is called *multilinear* (n-linear) if it is linear in every component. Formally:

$$F(v_1, \ldots, v_{k-1}, \lambda v_k' + \mu v_k'', v_{k+1}, \ldots, v_n)$$

$$= \lambda F(v_1, \ldots, v_{k-1}, v_k', v_{k+1}, \ldots, v_n) + \mu F(v_1, \ldots, v_k'', v_{k+1}, \ldots, v_n)$$

1.5 Example

$$V = \mathbb{K}[x] \text{ polynomials}$$

Basis: $\left\{x^k \,\middle|\, k \in \mathbb{N}_0\right\}$ and $\dim V = \aleph_0$

Every $v^* \in V^*$ is uniquely defined by $a_k \coloneqq \left\langle v^*, x^k\right\rangle$

$$(a_k)_{k\in\mathbb{N}_0}$$

$V^* \cong \mathbb{K}[[t]]$ are the formal power series

$$= \left\{\sum_{k=0}^{\infty} a_k t^k \,\middle|\, a_k \in \mathbb{K}\right\}$$

$$\lambda \sum_{k=0}^{\infty} a_k t^k + \mu \sum_{k=0}^{\infty} b_k t^k = \sum_{k=0}^{\infty} (\lambda a_k + \mu b_k) t^k$$

(Compare with Taylor series $f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$))

$$\left\langle \sum_{k=0}^{\infty} a_k t^k, \sum_{k=0}^{n} b_k x^k \right\rangle =: \sum_{k=0}^{n} a_k b_k \text{ is well-defined}$$

$$\to \mathbb{K}[x]^* \cong \mathbb{K}[[t]]$$

1.6 Example

$$C[0,1] \text{ continuous functions}$$

Example:

Example 1.

$$x \in [0,1] \qquad \delta_x : C[0,1] \to \mathbb{R}$$
$$f \mapsto f(x)$$
$$\langle \delta_x, f\rangle = f(x)$$
$$\langle \delta_x, f\rangle = f(x)$$
$$I(f) = \int_0^1 f(x) \,\mathrm{d}x \text{ is linear}$$

$$\langle I_g, f\rangle = \int_0^1 f(x) g(x) \,\mathrm{d}x$$
$$g \in C[0,1] \text{ is fixed}$$
$$\Rightarrow I_g \in C[0,1]$$
$$\langle I_g, \lambda f_1 + \mu f_2\rangle' = \int_0^1 (\lambda f_1(x) + \mu f_2(x)) g(x) \,\mathrm{d}x$$
$$= \lambda \int_0^1 f_1(x) g(x) \,\mathrm{d}x + \mu \int_0^1 f_2(x) g(x) \,\mathrm{d}x$$

This also works with non-continuous g (it suffices to have g integratable). (Compare with measure theory and Riesz' theorem)

Does there exist some g such that $f(x) = \langle \delta_x, f\rangle = \int_0^1 f(t) g(t) \,\mathrm{d}t$. (Compare with Dirac's δ function and Schwartz/Sobder theory)

$$V^{**} = (V^*)^* \cong V \text{ if } \dim V < \infty$$

Lemma 1. Let V be a vector space over $\mathbb{K}$. It requires that $\dim V < \infty$ and the Axiom of Choice holds.

- $v \in V \setminus \{0\} \Leftrightarrow \bigvee_{v^* \in V^*} \langle v^*, v\rangle \neq 0$

- $\bigwedge_{v \in V} v = 0 \Leftrightarrow \bigwedge_{v^* \in V^*} \langle v^*, v \rangle = 0$

Proof. Addition v to a basis B of V: Define $v^* \in V^*$ by

$$\langle v^*, b \rangle = \begin{cases} 1 & b = v \\ 0 & b \neq v \end{cases} \quad \text{for } b \in B$$

□

Theorem 2. Let V be a vector space over $\mathbb{K}$.

- The map $\iota : V \to V^{**} := (V^*)^*$ is called *bidual space*.

$$\langle \iota(v), v^* \rangle := \langle v^*, v \rangle$$

  is linear and injective.

- if $\dim V < \infty$, then isomorphism.

Proof.
- Linearity

$$\iota(\lambda v + \mu w) \stackrel{!}{=} \lambda \iota(v) + \mu \iota(w)$$

  must hold in every point $v^* \in V^*$:

$$\begin{aligned} \langle \iota(\lambda v + \mu w), v^* \rangle &= \langle v^*, \lambda v + \mu w \rangle \\ &= \lambda \langle v^*, v \rangle + \mu \langle v^*, w \rangle \\ &= \lambda \langle \iota(v), v^* \rangle + \mu \langle \iota(w), v^* \rangle \\ &= \langle \lambda \iota(v) + \mu \iota(w), v^* \rangle \end{aligned}$$

  Is it injective? Let $v \in \ker \iota$.

$$\langle \iota(v), v^* \rangle = 0 \quad \forall v^* \in V^*$$

$$\Rightarrow \langle v^*, v \rangle = 0 \quad \forall v^* \in V^*$$

$$\xRightarrow{\text{Lemma 1}} v = 0$$

- Follows immediately, because the dimension is equal.

□

Definition 1. Let V, W be vector spaces over $\mathbb{K}$. $f \in \operatorname{Hom}(V, W)$. We define $f^T \in \operatorname{Hom}(W^*, V^*)$ using $f^T(w^*) \in V^*$ via

$$\left\langle f^T(w^*), v \right\rangle = \langle w^*, f(v) \rangle = w^*(f(v)) = w^* \circ f(v)$$

$$f^T(w^*) = w^* \circ f \text{ is linear} \Rightarrow f^T(w^*) \in V^*$$

V to W (with f) and W to $\mathbb{K}$ (with w^*).

f^T is called *transposed map.*

Example 2. (See practicals) Let $\dim V = n$ and $\dim W = m$ with $B \subseteq V$ and $C \subseteq W$ as bases and dual bases $B^* \subseteq V^*$ and $C^* \subseteq W^*$

$$\Phi_{B^*}^{C^*}(f^T) = \Phi_C^B(f)^T \quad \text{transposition of matrices}$$

This lecture took place on 2nd of March 2016 (Franz Lehner).

2 Determinants

Leibnitz 1693 (3×3 matrices)
Seki Takukazu 1685 (most general version)
Gauß 1801 (“determinant”)
Cayley 1845 (on matrices)

$n = 2$

$$\begin{aligned} ax + by &= e \\ cx + dy &= f \end{aligned}$$

$$\begin{array}{cc|c} a & b & e \\ c & d & f \end{array}$$

1. Case 1: $a \neq 0$ (multiply first row $-\frac{a}{b}$ times second row)

$$\begin{array}{cc} a & b \\ c & d \\ \hline a & b \\ 0 & d - \frac{bc}{a} \end{array}$$

Unique solution:

$$d - \frac{bc}{a} \neq 0$$

2. Case 2: $c \neq 0$ (multiple second row $-\frac{a}{c}$ times first row)

$$\begin{array}{cc} a & b \\ c & d \\ \hline 0 & b - \frac{ad}{c} \\ c & d \end{array}$$

Unique solution:

$$b - \frac{ad}{c} \neq 0$$

This gives us

$$ad - bc \neq 0$$

Definition 2.

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

is called *determinant* of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

2.1 Properties of determinants

- The determinant is bilinear in the columns and rows.

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (v, w)$$

where v and w are column vectors of A.

$$\det(\lambda v_1 + \mu v_2, w) = \lambda \det(v_1, w) + \mu \det(v_2, w)$$

$$\det(v, \lambda w_1 + \mu w_2) = \lambda \det(v, w_1) + \mu \det(v, w_2)$$

$$\det(\lambda v_1 + \mu v_2, w) = \begin{vmatrix} \lambda a_1 + \mu a_2 & b \\ \lambda c_1 + \mu c_2 & d \end{vmatrix}$$

$$= (\lambda a_1 + \mu a_2)d - (\lambda c_1 + \mu c_2)b$$
$$= \lambda(a_1 d - c_1 b) + \mu(a_2 d - c_2 b)$$
$$= \lambda \begin{vmatrix} a_1 & b \\ c_1 & d \end{vmatrix} + \mu \begin{vmatrix} a_2 & b \\ c_2 & d \end{vmatrix}$$

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & c \\ b & d \end{vmatrix}$$

- $\det(v, v) = 0$.

$$\begin{vmatrix} a & a \\ c & c \end{vmatrix} = ac - ac = 0$$

- $$\det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \det(e_1, e_2) = 1$$

Theorem 3. The properties 1–3 of determinants (see above) characterize the determinant.

Let $\varphi : \mathbb{K}^2 \times \mathbb{K}^2 \to \mathbb{K}$

- bilinear
- $\bigwedge_{v \in \mathbb{K}^2} \varphi(v, v) = 0$
- $\varphi(e_1, e_2) = 1$. Then it holds that $\varphi = \det$.

Proof. To show: $\varphi(v, w) = \det(v, w) \forall v, w \in \mathbb{K}^2$

$$v = \underbrace{ae_1 + ce_2}_{\begin{pmatrix} a \\ c \end{pmatrix}} \qquad w = \underbrace{be_1 + de_2}_{\begin{pmatrix} b \\ d \end{pmatrix}}$$

$$\begin{aligned} \varphi(v, w) &= \varphi(ae_1 + ce_2, be_1 + de_2) \\ &= a\varphi(e_1, be_1 + de_2) + c \cdot \varphi(e_2, be_1 + de_2) \\ &= ad \underbrace{\varphi(e_1, e_2)}_{=1} + \underbrace{ab\varphi(e_1, e_1)}_{=0} + cb\varphi(e_2, e_1) + cd \underbrace{\varphi(e_2, e_2)}_{=0} \end{aligned}$$

□

Lemma 2. From (i) bilinearity and (ii) $\bigwedge_{v\in\mathbb{K}^2}\varphi(v,v)=0$ it follows that

$$\bigwedge_{v,w\in\mathbb{K}^2}\varphi(v,w)=-\varphi(w,v)$$

$$\begin{aligned}0 \overset{\text{(ii)}}{=} \varphi(v+w,v+w) &\overset{\text{(i)}}{=} \varphi(v,v)+\varphi(v,w)+\varphi(w,v)+\varphi(w,w)\\ &\overset{\text{(ii)}}{=} \varphi(v,w)+\varphi(w,v)\end{aligned}$$

2.2 Geometric interpretation of the determinant

Consider an area with w defining its breath and v its depth (hence the area spanning vectors). Let e_1 and e_2 be the spanning vectors of a rectangle corresponding to the parallelogram. $\det(v,w)$ is the surface of the spanned parallelogram. The sign defines the orientation of the pair (v,w).

$$\det(e_1,e_2)=1 \qquad \det(e_2,e_1)=-1$$

There are surfaces where the surface is infinite if you follow a vector in some direction:

- Möbius strip
- Klein's bottle (named after Felix Klein)

$$A=|\,v\,|\cdot h$$

Consider Figure 1. h is the length of the projection of w to $v^\perp$.

$$v=\begin{pmatrix}a\\b\end{pmatrix}\to\vec{n}=\begin{pmatrix}-b\\a\end{pmatrix}$$

$$\langle\begin{pmatrix}c\\d\end{pmatrix},\begin{pmatrix}-b\\a\end{pmatrix}\rangle=ad-bc$$

Second proof. $A(v,w)$ satisfies properties (i)—(iii).

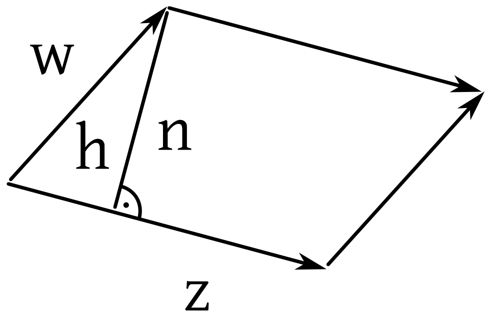

Figure 1: Parallelogram

- Property (iii) follows immediately (the area of unit vectors in two dimensions is 1).
- Property (ii) follows immediately (the area of two vectors in the same direction is 0).

Property (i) defines the linearity in v

1. If v,w are linear dependent, then $A(v,w)=0$ (one is a multiple of the other)
2. $n\in\mathbb{N}$ with $A(nv,w)=nA(v,w)$
3. For $\tilde{v}=n\cdot v$:

$$A(\tilde{v},w)=n\cdot A(\frac{\tilde{v}}{n},w)$$

$$\Rightarrow A(\frac{\tilde{v}}{n},w)=\frac{1}{n}A(\tilde{v},w)$$

$$\begin{aligned} A(nv, w) &= nA(v, w) \\ A(\frac{1}{n}v, w) &= \frac{1}{n}A(v, w) \\ A(\frac{m}{n}v, w) &= \frac{m}{n}A(v, w) \\ A(-v, w) &= -A(v, w) \end{aligned}$$

From continuity it follows that $A(\lambda u, w) = \lambda A(v, w)$ for $\lambda \in \mathbb{R}$. Analogously $A(v, \lambda w) = \lambda A(v, w)$.

4. The sum is given with

$$A(v + w, w) = A(v, w)$$

Compare with Figure 2, where area(2) + area(3) = area(2) + area(1).

$$\begin{aligned} A(\lambda v + \mu w, w) &= A(\lambda v + \mu w, \frac{1}{\mu}\mu w) \\ &= \frac{1}{\mu}A(\lambda v + \mu w, \mu w) \\ &= \frac{1}{\mu}A(\lambda v, \mu w) \\ &= A(\lambda v, w) \end{aligned}$$

General case: v, w are linear independent and therefore basis of $\mathbb{R}^2$. Besides that, v_1 and v_2 are arbitrary.

$$\begin{aligned} v_1 &= \lambda_1 v + \mu_1 w \\ v_2 &= \lambda_2 v + \mu_2 w \end{aligned}$$

$$\begin{aligned} A(v_1 + v_2, w) &= A(\lambda_1 v + \mu_1 w + \lambda_2 v + \mu_2 w, w) \\ &= A((\lambda_1 + \lambda_2)v + (\mu_1 + \mu_2)w, w) \\ &= A((\lambda_1 + \lambda_2)v, w) \\ &= (\lambda_1 + \lambda_2)A(v, w) \\ &= A(\lambda_1 v, w) + A(\lambda_2 v, w) \end{aligned}$$

$$A(\lambda_1 v + \mu_1 w, w) + A(\lambda_2 v + \mu_2 w, w) = A(v_1, w) + A(v_2, w)$$

Additivity follows.

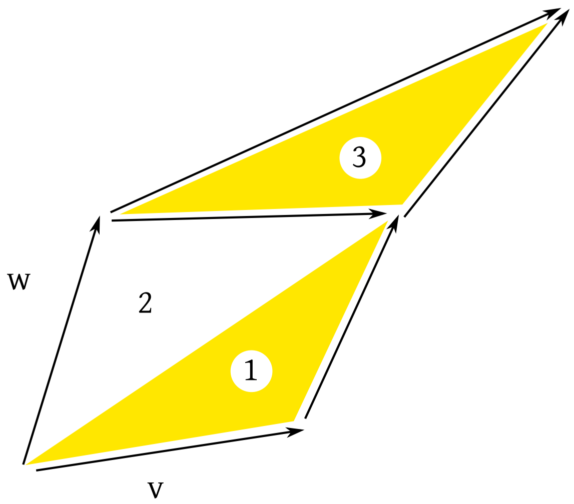

Figure 2: Translation of area 1 to area 3.

□

Definition 3. Let $\dim V = n$. A *determinant form* is a map

$$\triangle : V^n \to \mathbb{K}$$

with properties:

1.

$$\bigwedge_{\lambda} \bigwedge_{k} \bigwedge_{a_1,\dots,a_n \in V} \triangle(a_1, \dots, a_{k-1}, \lambda a_k, a_{k+1}, \dots, a_n) = \lambda \triangle(a_1, \dots, a_k, \dots, a_n)$$

2.
$$\bigwedge_{k} \bigwedge_{\substack{a_1,\dots,a_n \\ a'_k, a''_k}} \triangle(a_1, \dots, a_{k-1}, a'_k + a''_k, a_{k+1}, \dots, a_n)$$
$$:= \triangle(a_1, \dots, a_{k-1}, a'_k + a''_k, a_{k+1}, \dots, a_n)$$

3.
$$\triangle(a_1, \dots, a_n) = 0$$

if $\bigvee_{k \neq l} a_k = e_l$ if $\triangle \neq 0$, i.e. $\triangle$ is non-trivial.

Multilinearity is defined by the first two properties. Multilinearity means linearity in a_k if $a_1, \dots, a_{k-1}, a_{k+1}, \dots, a_n$ gets fixed.

Theorem 4.
$$\dim V = n$$
$$\triangle : V^n \to \mathbb{K} \text{ is determinant form}$$

Then,

4.
$$\bigwedge_{\lambda \in \mathbb{K}} \bigwedge_{i \neq j} \triangle(a_1, \dots, a_{i-1}, a_i + \lambda a_j, a_{i+1}, \dots, a_n) = \triangle(a_1, \dots, a_i, \dots, a_n)$$

"Addition of λa_j to a_i does not change $\triangle$"

5.
$$\bigwedge_{i > j} \triangle(a_1, \dots, a_{j-1}, a_i, a_{j+1}, \dots, a_{i-1}, a_j, a_{i+1}, \dots, a_n)$$
$$= -\triangle(a_1, \dots, a_j, \dots, a_i, \dots, a_n)$$

"Exchanging a_i with a_j inverts the sign"

Proof. 4.
$$\triangle(a_1, \dots, a_i + \lambda a_j, \dots, a_n)$$

Without loss of generality: $i < j$. From properties 1 and 2 it follows that:
$$= \triangle(a_1, \dots, a_i, a_j, a_n) + \lambda \triangle(a_1, \dots, a_j, a_j, \dots, a_k)$$

Oh, a_j occurs twice! Once at index i and once at index j.
$$= 0$$

due to property 3.

5.
$$\begin{aligned} 0 &\overset{\text{property 3}}{=} \triangle(a_1, \dots, a_{i-1}, a_i + a_j, \dots, a_{j-1}, a_i + a_j, \dots, a_n) \\ &= \triangle(a_1, \dots, a_{i-1}, \mathbf{a_i}, \dots, a_{j-1}, \mathbf{a_i}, \dots, a_n) = \mathbf{0} \\ &+ \triangle(a_1, \dots, a_{i-1}, \mathbf{a_i}, \dots, a_{j-1}, \mathbf{a_j}, \dots, a_n) \\ &+ \triangle(a_1, \dots, a_{i-1}, \mathbf{a_j}, \dots, a_{j-1}, \mathbf{a_i}, \dots, a_n) \\ &+ \triangle(a_1, \dots, a_{i-1}, \mathbf{a_j}, \dots, a_{j-1}, \mathbf{a_j}, \dots, a_n) = \mathbf{0} \\ &\Rightarrow \delta \end{aligned}$$

□

Definition 4. A permutation of order n is a bijective mapping $\pi : \{1, \dots, n\} \to \{1, \dots, n\}$.
$$\sigma_n = \text{ set of all permutations}$$

Remark 1. Notation: We write the elements in the first row and their images in the second row.

Definition 5. σ_n constitutes (in terms of composition) a group with neutral element id, the so-called symmetric group.

In the previous course (Theorem 1.40) we have proven: Compositions of bijective functions are bijective.

Remark 2. For $n \geq 3$, σ_n is non-commutative

Theorem 5.
$$|\,\sigma_n\,| = n!$$

Remark 3. These are "a lot"!

Example 3.
$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$

$$\begin{pmatrix}1 & 2 & 3 & 4\\ 4 & 1 & 3 & 2\end{pmatrix}^{-1} = \begin{pmatrix}1 & 2 & 3 & 4\\ 2 & 4 & 3 & 1\end{pmatrix}$$

Definition 6. A *transposition* is a permutation of the structure

$$\tau = \tau_{ij} : \begin{array}{c} i \mapsto j \\ j \mapsto i \\ k \mapsto h \end{array} \quad \text{if } k \notin \{i, j\}$$

Then $\tau_{ij}^{-1} = \tau_{ij}$, hence $\tau_{ij}^2 = \mathrm{id}$.

Theorem 6. σ_n is generated by transpositions. With other words, every permutation π can be represented as composition of transpositions

$$\pi = \tau_1 \circ \ldots \circ \tau_k$$

Proof.

$$\pi = \begin{pmatrix}1 & 2 & \ldots & n\\ \pi(1) & \pi(2) & \ldots & \pi(n)\end{pmatrix}$$

If $\pi = \mathrm{id}$,

$$\pi = \pi \quad \tau \coloneqq \mathrm{id}$$

If $\pi \neq \mathrm{id}$,

$$k_1 = \min\{k \mid k \neq \pi(k)\}$$

1.

$$\tau_1 = \tau_{k_1 \pi(k_1)}$$

$$\pi_1 = \tau_1 \circ \pi = \begin{pmatrix}1 & \ldots & k_1-1 & k_1 & \ldots\\ 1 & \ldots & k_1-1 & k_1 & \ldots\end{pmatrix}$$

Example: Consider $\begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6 & 7\\ 1 & 3 & 5 & 4 & 7 & 6 & 2\end{pmatrix}$.

$$k_1 = 2$$

$$\tau_1 = \tau_{23}$$

$$\pi_1 = \tau_1 \circ \pi = \begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6 & 7\\ 1 & 2 & 5 & 4 & 7 & 6 & 3\end{pmatrix}$$

2.

$$k_2 = \min\{k \mid k \neq \pi_1(k)\} > k_1$$

$$\tau_2 = \tau_{k_2, \pi(k_2)}$$

And so on and so forth. $k_j > k_{j-1}$ ends after $\leq n$ steps.

$$\tau_k \circ \tau_{k-1} \circ \ldots \circ \tau_1 \circ \pi = \mathrm{id}$$

$$\Rightarrow \pi = \tau_1 \circ \tau_2 \circ \ldots \circ \tau_k$$

Regarding the example:

$$k_2 = 3$$

$$\tau_2 = \tau_{35}$$

$$\pi_2 = \tau_2 \circ \pi_1 = \tau_2 \circ \tau_1 \circ \pi = \begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6 & 7\\ 1 & 2 & 3 & 4 & 7 & 6 & 5\end{pmatrix}$$

$$k_3 = 5 \quad \tau_3 = \tau_{57}$$

$$\Rightarrow \pi = \tau_{23} \circ \tau_{35} \circ \tau_{57}$$

□

Definition 7. An *inversion* of π is a pair (i, j) such that $i < j$ with $\pi(i) > \pi(j)$. Let F_π be the set of inversions of π.

$$f_\pi \coloneqq |\, F_\pi \,|$$

$$\mathrm{sign}(\pi) \coloneqq (-1)^{f_\pi} \eqqcolon (-1)^\pi$$

Example 4.

$$\pi = \begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6 & 7\\ 1 & 3 & 5 & 4 & 7 & 6 & 2\end{pmatrix}$$

$$F_\pi = \{(2,7), (3,4), (3,7), (4,7), (5,6), (5,7), (6,7)\}$$

$$f_\pi = 7 \qquad \mathrm{sign}(\pi) = -1$$

This lecture took place on 7th of March 2016 (Franz Lehner).

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Recall: Determinant form:

1. $\triangle(a_1, \dots, \lambda a_k, \dots, a_n) = \lambda \triangle(a_1, \dots, a_n)$
2. $\triangle(a_1, \dots, a_k' + a_k'', \dots, a_n) = \triangle(a, \dots, a_k', \dots, a_n) + \triangle(a_1, \dots, a_k'', \dots, a_n)$
3. $\triangle(a_1, \dots, a_k, \dots, a_l, \dots, a_n) = 0$ if $a_k = a_l$

Conclusions:

4. $\triangle(a_1, \dots, a_k + \lambda a_l, \dots, a_n) = \triangle(a_1, \dots, a_n)$ if $k \neq l$
5. $\triangle(a_1, \dots, a_k, \dots, a_l, \dots a_n) = -\triangle(a_1, \dots, a_l, \dots, a_k, \dots a_n)$

$$\triangle(a_{\pi(1)}, \dots, a_{\pi(n)}) = (-1)^k \triangle(a_1, \dots, a_n)$$

Decompose $\pi = \tau_1 \circ \ldots \circ \tau_k \circ \tau_{12} \circ \tau_{12}$. This decomposition is not distinct (k is distinct $\mod 2$)

$$\pi \in \sigma_n \qquad \text{permutation}$$

$$F_\pi = \{(i,j) \mid i < j, \pi(i) > \pi(j), \text{ inversions } \}$$

$$f_\pi = |\, F_\pi \,|$$

$$\operatorname{sign}(\pi) \coloneqq (-1)^{f_\pi} \eqqcolon (-1)^\pi$$

Theorem 7.

- $\bigwedge_{\pi \in \sigma_n} \operatorname{sign}(\pi) = \prod_{1 \leq i < j \leq n} \frac{\pi(j) - \pi(i)}{j - i}$
- For transposition τ it holds that $\operatorname{sign}(\tau) = -1$

Proof.

- Every pair $\{i, j\}$ occurs in the enumerator exactly once.

$$\frac{\prod_{i<j} \pi(j) - \pi(i)}{\prod_{i<j} (j - i)}$$

Denominator: $j > i$, positive. Enumerator: positive if $\pi(j) > \pi(i)$, negative if $\pi(i) > \pi(j)$.

- $$\tau = \begin{pmatrix} 1 & \dots & k & \dots & l & \dots & n \\ 1 & \dots & l & \dots & k & \dots & n \end{pmatrix}$$

$$F_\tau(\underbrace{(k, k+1), (k, k+2), \dots, (k, l-1)}_{\text{inversions with } k,\ l-k \text{ times}}, (k, l), \underbrace{(k+1, l), \dots, (l-1, l)}_{l-k-1 \text{ times}})$$

Example:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 1 & 2 & 3 & 8 & 5 & 6 & 7 & 4 & 9 & 10 \end{pmatrix}$$

Yields 7 inversions (8 needs to be repositioned with 3 transpositions, 4 needs to be repositions with 4 transpositions).

□

$$\operatorname{sign}(\pi) = \prod_{i<j} \frac{\pi(j) - \pi(i)}{j - i} \qquad \binom{n}{2} \text{ factors}$$

$$\operatorname{sign}(\tau) = -1$$

Theorem 8.

1. $\operatorname{sign}(\mathrm{id}) = 1$
2. $\operatorname{sign}(\pi \circ \sigma) = \operatorname{sign}(\pi) \cdot \operatorname{sign}(\sigma)$, hence

$$\operatorname{sign} \sigma_n \to (\{+1, -1\}, \cdot)$$

is a group homomorphism. (In general: A group homomorphism $h : G \to (\mathcal{T}, \cdot)$ is called *character*)

3. $\operatorname{sign}(\pi^{-1}) = \operatorname{sign}(\pi)$

Remark 4.

$$\mathcal{T} = \{z \in \mathbb{C} \,|\, |\,z\,| = 1\}$$

Torus with multiplication is a group.

$$|\,z_1 \cdot z_2\,| = |\,z_1\,| \cdot |\,z_2\,| = 1$$

Proof. 1. trivial

2.

$$\begin{aligned}\operatorname{sign}(\pi \cdot \sigma) &= \prod_{i<j} \frac{\pi \circ \sigma(j) - \pi \circ \sigma(i)}{j-i} \\ &= \underbrace{\prod_{i<j} \frac{\pi(\sigma(j)) - \pi(\sigma(i))}{\sigma(j) - \sigma(i)}}_{=\operatorname{sign}(\pi)} \cdot \underbrace{\prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j-i}}_{\operatorname{sign}(\sigma)}\end{aligned}$$

3. Group homomorphism!

□

Corollary 1.

- If $\pi = \tau_1 \circ \tau_2 \circ \ldots \circ \tau_k$, product of transpositions

$$\Rightarrow \operatorname{sign}(\pi) = (-1)^k$$

- $\mathfrak{a}_n := \ker(\operatorname{sign}) = \{\pi \in \sigma_n \,|\, \operatorname{sign}(\pi) = 1\}$

"even permutations", "alternating group"

$$|\,\mathfrak{a}_n\,| = \frac{n!}{2}$$

Corollary 2.

$$\triangle : V^k \to \mathbb{K} \text{ determinant form}$$

then it holds that

$$\bigwedge_{\pi \in \sigma_n} \bigwedge_{a_1,\ldots,a_n \in V} \triangle(a_{\pi(1)}, \ldots, a_{\pi(n)}) = \operatorname{sign}(\pi) \cdot \triangle(a_1, \ldots, a_n)$$

Proof.

- If $\pi = \tau_{kl}$ transposition $\xrightarrow{\text{Theorem 4}}$ $\triangle(a_{\tau(1)}, \ldots, a_{\pi(n)}) = -\triangle(a_1, \ldots, a_n) = \operatorname{sign}(\tau_{kl}) \cdot \triangle(a_1, \ldots, a_n)$
- If $\pi = \tau_1 \circ \ldots \circ \tau_k = \tau_1 \circ \tilde{\pi}, \tilde{\pi} = \tau_2 \circ \ldots \circ \tau_k$

$$\begin{aligned}&\triangle(a_{\tau_1 \circ \tilde{\pi}(1)}, \ldots, a_{\tau_1 \circ \tilde{\pi}(n)}) \\ &= -\triangle(a_{\tilde{\pi}(1)}, \ldots, a_{\tilde{\pi}(n)}) \\ &= (-1)^2 \cdot \triangle(a_{\tilde{\pi}(1)}, a_{\tilde{\pi}(n)}) \\ &\to (-1)^k \cdot \triangle(a_1, \ldots, a_n)\end{aligned}$$

□

Theorem 9 (Leibnitz' definition of $\det(A)$)**.** Let $B = (b_1, \ldots, b_n)$ be the basis of V. $a_1, \ldots, a_n \in V$ with coordinates

$$\Phi_B(a_j) = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix}$$

$$A := [a_{ij}]_{i,j=1,\ldots,n} = [\Phi_B(a_1), \Phi_B(a_2), \ldots, \Phi_B(a_n)]$$

Then it holds that for every determinant form $\triangle : V^k \to \mathbb{K}$:

$$\triangle(a_1, \ldots, a_n) = \det(A) \cdot \triangle(b_1, \ldots, b_n)$$

where

$$\det(A) := \sum_{\pi \in \sigma_n} \operatorname{sign}_{\mathbb{K}} \pi a_{\pi(1),1} a_{\pi(2),2} \ldots a_{\pi(n),n}$$

is the determinant of A

Example 5. Example ($n = 2$):

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

$$\operatorname{sign}\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = 1$$

$$\operatorname{sign}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = -1$$

Proof.

$$a_j = \sum_{j=1}^{n} a_{ij} b_i$$

$$\begin{aligned}\triangle(a_1,\dots,a_n) &= \triangle\left(\sum_{i_=1}^{n} a_{i,1}b_i, \sum_{i_2=1}^{n} a_{i_2,2}b_i,\dots,\sum_{i_n=1}^{n} a_{i_n,n}b_i\right)\\ &= \sum_{i_1=1}^{n} a_{i,1}\sum_{i_2=1}^{n} a_{i_2,2}\dots\sum_{i_n=1}^{n} a_{i_n,n}\underbrace{\triangle(b_i,b_{i_2},\dots,b_{i_n})}_{=0 \text{ if some } i_k=i_l}\end{aligned}$$

So summands with equal indices disappear. It holds that $\sum_{i_1,\dots,i_n}$ such that $i_1,\dots,i_n$ are different. Hence every value from $\{1,\dots,n\}$ occurs exactly once. This is the set of all permutations π $(i_j = \pi(j))$

$$= \sum_{\pi\in\sigma_n} a_{\pi(1)1}a_{\pi(2)2}\dots a_{\pi(n)n}\underbrace{\triangle(b_{\pi(1)},\dots,b_{\pi(n)})}_{\text{sign}(\pi)\cdot\triangle(b_1,\dots,b_n)}$$

□

Corollary 3. A determinant form is *uniquely* defined on a basis $(b_1,\dots,b_n)$ by the value $\triangle(b_1,\dots,b_n)$. Especially $\triangle$ is nontrivial,

$\Leftrightarrow$ $\triangle(b_1,\dots,b_n)\neq 0$ on some basis.

$\Leftrightarrow$ $\triangle(b_1,\dots,b_n)\neq 0$ in every basis $b_1,\dots,b_n$.

Let $\triangle(b'_1,\dots,b'_n)=0$ for some other basis, represent $b_1,\dots,b_n$ in basis $b'_1,\dots,b'_n$

$$b_j = \sum a_{ij}b'_i \Rightarrow \triangle(b_1,\dots,b_n) = \det(A)\cdot\triangle(b'_1,\dots,b'_n) = 0$$

$$\triangle(a_1,\dots,a_n) = \det(A)\cdot\triangle(b_1,\dots,b_n)$$

Theorem 10. Let $B=(b_1,\dots,b_n)$ be a basis of V over $\mathbb{K}$. $c\in\mathbb{K}$. For $a_1,\dots,a_n\in V$, let $A=[\Phi_B(a_1),\dots,\Phi_B(a_n)]$. Then

$$\triangle(a_1,\dots,a_n) = c\cdot\det(A)$$

defines a determinant form, specifically the unique determinant form with value

$$\triangle(b_1,\dots,b_n) = c$$

Proof. The 3 properties of a determinant form:

1.

$$\begin{aligned}\triangle(a_1,\dots,\lambda a_k,\dots,a_n) &= c\cdot\det[\Phi_B(a_1),\dots,\lambda\cdot\Phi_B(a_k),\dots,\Phi_B(a_n)]\\ &= c\cdot\sum_{\pi\in\sigma_n}\operatorname{sign}\pi\cdot a_{\pi(1),1}a_{\pi(2),2}\dots\lambda a_{\pi(k),k}\dots a_{\pi(n),n}\\ &= \lambda\cdot c\cdot\sum_{\pi\in\sigma_n}\operatorname{sign}\pi\cdot a_{\pi(1),1}a_{\pi(2),2}\dots a_{\pi(n),n}\\ &= \lambda\cdot\triangle(a_1,\dots,a_n)\end{aligned}$$

2.

$$\begin{aligned}&\triangle(a_1,\dots,a'_k+a''_k,\dots,a_n)\\ &= c\cdot\det[\Phi_B(a_1),\dots,\Phi_B(a'_k)+\Phi_B(a''_k),\dots,\Phi_B(a_n)]\\ &= c\cdot\sum_{\pi\in\sigma_n}\operatorname{sign}\pi\cdot a_{\pi(1),1}\cdot a_{\pi(2),2}\cdot\dots\left(a'_{\pi(k),k}+a''_{\pi(k),k}\right)\cdot\dots\cdot a_{\pi(n),n}\\ &= c\cdot\sum_{\pi\in\sigma_n}\operatorname{sign}\pi\cdot a_{\pi(1),1}\cdot\dots\cdot a'_{\pi(k),k}\dots a_{\pi(n),n}\\ &+ c\cdot\sum_{\pi\in\sigma_n}\operatorname{sign}(\pi)a_{\pi(1),1}\dots a''_{\pi(k),k}\dots a_{\pi(n),n}\\ &= \triangle(a_1,\dots,a'_k,\dots,a_n)+\triangle(a_1,\dots,a''_k,\dots,a_n)\end{aligned}$$

3. Let $a_k=a_l$ for $k<l$. Show that $\triangle(a_1,\dots,a_n)=0$

$$\tau_{kl} = \text{ transposition exchanging } k \text{ and } l$$

$$\sigma_n = \mathfrak{a}_n\dot{\cup}(\mathfrak{a}_n\cdot\tau_{kl})$$

Claim: $\{\pi\mid\operatorname{sign}\pi=-1\}=\{\pi\circ\tau_{kl}\mid\operatorname{sign}\pi=+1\}$

$\supseteq$ If $\operatorname{sign}\pi=+1\Rightarrow\operatorname{sign}(\pi\circ\tau_{kl})=\underbrace{\operatorname{sign}\pi}_{+1}\cdot\underbrace{\operatorname{sign}\tau_{kl}}_{-1}=-1$

$\subseteq$ If $\operatorname{sign}\pi=-1\Rightarrow\operatorname{sign}(\pi\circ\tau_{kl})=+1\Rightarrow\pi=\underbrace{(\pi\circ\tau_{kl})}_{\in\mathfrak{a}_n}\circ\tau_{kl}\in\mathfrak{a}_n\cdot\tau_{kl}$

$$
\begin{aligned}
\triangle(a_1,\dots,a_n) &= c\cdot \sum_{\pi\in\sigma_n=\mathfrak{a}_n\cup\mathfrak{a}_n\cdot\tau_{kl}} \operatorname{sign}(\pi) a_{\pi(1),1}\dots a_{\pi(n),n}\\
&= \underbrace{c\cdot\sum_{\pi\in\mathfrak{a}_n} a_{\pi(1),1}\dots a_{\pi(n),n}}_{\text{even}}\\
&\quad \underbrace{-\sum_{\pi\in\mathfrak{a}_n} a_{\pi\circ\tau_{kl}(1),1}\dots a_{\pi\circ\tau_{kl}(k),k}\dots a_{\pi\circ\tau_{ul}(l),l}\dots a_{\pi\circ\tau_{kl}(n),n}}_{\text{odd}}\\
&= c\cdot\sum_{\pi\in\mathfrak{a}_n} a_{\pi(1),1}\dots, a_{\pi(n),n}\\
&\quad -\sum_{\pi\in\mathfrak{a}_n} a_{\pi(1),1}\dots \underbrace{a_{\pi(l),k}}_{a_{\pi(l),l}}\dots \quad a_{\pi(k),k}\ \underbrace{a_{\pi(k)l}}_{\text{because } a_k=a_l}\dots a_{\pi(n),n}
\end{aligned}
$$

What we did:

(a) $a_{\pi(l),k}=a_{\pi(l),l}$ and $a_{\pi(k),l}=a_{\pi(k),k}$ because $a_k=a_l$

(b) exchange factors

$$
\begin{aligned}
&= c\sum_{\pi\in\mathfrak{a}_n} a_{\pi(1),1}\dots a_{\pi(k),k}\dots a_{\pi(l),l}\dots a_{\pi(n),n}\\
&- c\sum_{\pi\in\mathfrak{a}_n} a_{\pi(1),1}\dots a_{\pi(k),k}\dots a_{\pi(l),l}\dots a_{\pi(n),n}\\
&= 0
\end{aligned}
$$

Value for $(b_1,\dots,b_n)$

$$a_{ij}=\delta_{ij}\Rightarrow A=I$$

$$\det(I)=\sum_{\pi\in\sigma_n}\operatorname{sign}\pi\cdot\delta_{\pi(1),1}\dots\delta_{\pi(n),n}=+1$$

for all $\pi(j)=j$ otherwise 0.

$$\Rightarrow \pi=\text{id is the only summand}$$

$$\triangle(b_1,\dots,b_n)=\det(I)\cdot c=c$$

$\square$

Remark 5. "$\mathfrak{a}_n$ is the subgroup of index 2" denoted $[\sigma_n:\mathfrak{a}_n]=2$

You might be familiar with:

$$\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$$

$$[\mathbb{Z}:n\mathbb{Z}]=n$$

Theorem 11 (Summary).
- The set of determinant forms $\triangle : V^n\to\mathbb{K}$ constructs a one-dimensional vector space, $\Lambda^n V$
- There exists a non-trivial determinant form with $\triangle(b_1,\dots,b_n)=1$

This lecture took place on 9th of March 2016 (Franz Lehner).

Revision:

$$\triangle : V^n\to\mathbb{K}$$

$$\triangle(a_1,\dots,a_n)=\det A\cdot\triangle(b_1,\dots,b_n)$$

$$\phi_B(a_j)=\begin{pmatrix}a_{1j}\\ \vdots\\ a_{nj}\end{pmatrix}$$

$$\det A=\sum_{\pi\in\sigma_n}\operatorname{sign}\pi\cdot a_{\pi(1),1}\dots a_{\pi(n),n}$$

$$\triangle(v_1,\dots,v_n)\neq 0\Leftrightarrow v_1,\dots,v_n \text{ linear independent } (\Leftrightarrow \text{basis})$$

Theorem 12.

$$\det(A\cdot B)=\det(A)\cdot\det(B)$$

Lemma 3. Let V,W be vector spaces over $\mathbb{K}$ with $\dim V=\dim W=n$.

$$\triangle : W^n\to\mathbb{K}$$

$$f:V\to W$$

$$\Rightarrow f^n : V^n \to W^n \overset{\triangle}{\to} \mathbb{K}$$
$$(v_1, \ldots, v_n) \mapsto (f(v_1), \ldots, f(v_n))$$

Then $\triangle^f : V^n \to \mathbb{K}$

$$\triangle^f(v_1, \ldots, v_n) = \triangle(f(v_1), \ldots, f(v_n))$$

is a determinant form in V.

Proof. 1.

$$\begin{aligned}\triangle f(v_1, \ldots, \lambda v_k, \ldots, v_n) &= \triangle(f(v_1), \ldots, f(\lambda v_k), \ldots, f(v_n)) \\ &= \lambda\triangle(f(v_1), \ldots, f(v_n)) \\ &= \lambda \cdot \triangle^f(v_1, \ldots, v_n)\end{aligned}$$

2.

$$\begin{aligned}&= \triangle^f(v_1, \ldots, v_k', +v_k'', \ldots, v_n) \\ &= \triangle(f(v_1), \ldots, f(v_k' + v_k''), \ldots, f(v_n)) \\ &= \triangle(f(v_1), \ldots, f(v_k') + f(v_k''), \ldots, f(v_n)) \\ &= \triangle(f(v_1), \ldots, f(v_k'), \ldots, f(v_n)) + \triangle(f(v_1), \ldots, f(v_k''), \ldots, f(v_n)) \\ &= \triangle^f(v_1, \ldots, v_k', \ldots, v_n) + \triangle^f(v_1, \ldots, v_k'', \ldots, v_n)\end{aligned}$$

3.

$$\begin{aligned}\triangle^f(v_1, \ldots, v_k, \ldots, v_l, \ldots, v_n) \qquad & v_k = v_l \Rightarrow f(v_k) = f(v_l) \\ &= \triangle(f(v_1), \ldots, f(v_k), \ldots, f(v_l), \ldots, f(v_n)) \\ &= 0\end{aligned}$$

□

Corollary 4 (Conclusions for $V = W$)**.**

$$\triangle : V^n \to \mathbb{K}$$

non-trivial determinant form

$$f : V \to V$$
$$\Rightarrow \triangle^f \text{ is a determinant form}$$

$$\dim \bigwedge^n \bigvee = 1 \Rightarrow \bigvee_{c_f \in \mathbb{K}} \triangle^k = c_f \cdot \triangle$$

$$c_f =: \det f \text{ is called determinant of } f$$

Corollary 5. Let V, $\triangle$ and f be like above.

1. For every basis $B = (b_1, \ldots, b_n)$ it holds that
$$\triangle^f(b_1, \ldots b_n) = \triangle(f(b_1), \ldots, f(b_n)) = \det(f) \cdot \triangle(b_1, \ldots, b_n)$$
$$\det(f) = \frac{\triangle(f(b_1), \ldots, f(b_n))}{\triangle(b_1, \ldots, b_n)}$$
2. with $a_j = f(b_j)$ it holds that
$$\det \Phi_B^B(f) = \det(f)$$
$$A = \Phi_B^B(f)$$
$a_{ij} =$ i-th coordinate of $f(b_j)$ and $s_j(A) = \Phi_B(f(b_j))$.

Theorem 13. Let $f : V \to V$ be an isomorphism $\Leftrightarrow \det(f) \neq 0$.

Proof. Let f be an isomorphism.

$$\begin{aligned}&\Leftrightarrow (f(b_1), \ldots, f(b_n)) \text{ is basis} \\ &\Leftrightarrow \triangle(f(b_1), \ldots, f(b_n)) \neq 0 \\ &\Leftrightarrow \det(f) \cdot \triangle(b_1, \ldots, b_n) \\ &\Leftrightarrow \det(f) \neq 0\end{aligned}$$

□

Theorem 14. Let $f, g : V \to V$ be linear.

$$\Rightarrow \det(f \circ g) = \det(f) \cdot \det(g)$$

Remark 6. We show: $f \circ g$ is isomorphism $\Leftrightarrow$ f and g are isomorphisms.

If f, g are invertible, then $f \circ g$ are invertible.

1. $$(f \circ g)^{-1} = g^{-1} \circ f^{-1}$$

2. Attention! This is wrong, if $\dim = \infty$! For example: $\delta : (x_1, x_2, \dots) \mapsto (0, x_1, x_2, \dots)$ over $\mathbb{K}^\infty$ is injective, but not surjective!

$$S_L : (x_1, x_2, \dots) = (x_2, x_3, \dots)$$

is not injective, but surjective.

$$S_L \circ S_R = I$$
$$S_R \circ S_L - I - P_1$$

If $f \circ g$ is bijective, then g is injective and f surjective.

$$\xrightarrow{\dim < \infty} g \text{ bijective}, f \text{ bijective}$$

Proof. Case distinction:

$\det(f \circ g) = 0$

$$\begin{aligned}
&\xleftrightarrow{\text{Theorem 13}} f \circ g \text{ is not bijective} \\
&\Leftrightarrow f \text{ is not bijective or g not bijective} \\
&\Leftrightarrow \det(f) = 0 \vee \det(g) = 0 \\
&\Leftrightarrow \det(f) \cdot \det(g) = 0
\end{aligned}$$

$\det(f \circ g) \neq 0$

$$\begin{aligned}
&\Leftrightarrow f \circ g \text{ is bijective} \\
&\Rightarrow g \text{ bijective} \\
&\Rightarrow \triangle^g \text{ non-trivial}
\end{aligned}$$

Let $(b_1, \dots, b_n)$ be a basis of V, then $\triangle$ is non-trivial determinant.

$$\begin{aligned}
\det(f \circ g) &= \frac{\triangle(f \circ g(b_1) \dots, f \circ g(b_n))}{\triangle(b_1, \dots, b_n)} \\
&= \frac{\triangle(f(g(b_1)), \dots, f(g(b_n)))}{\triangle(g(b_1), \dots, g(b_n))} \cdot \frac{\triangle(g(b_1), \dots, g(b_n))}{\triangle(b_1, \dots, b_n)} \\
&= \frac{\triangle((b_1'), \dots, f(b_n'))}{\triangle(b_1', \dots, b_n')} \cdot \frac{\triangle(g(b_1), \dots, g(b_n))}{\triangle(b_1, \dots, b_n)} \\
&= \det(f) \cdot \det(g)
\end{aligned}$$

$b_i' = g(b_i)$ are also a basis, because g is bijective.

□

Corollary 6. Let $A, B \in \mathbb{K}^{n \times n}$.

1. $\det(A \cdot B) = \det(A) \cdot \det(B)$
2. A is regular $\Rightarrow \det(A^{-1}) = \frac{1}{\det(A)}$
3. $\det(A) = 0 \Leftrightarrow \operatorname{rank}(A) < n$
4. $\det(A^t) = \det(A)$

Proof. 1. A first proof follows from Theorem 14.
A second proof approach is:

$$A = [s_1, \dots, s_n] \qquad \text{column vectors}$$

$$A \cdot B = \left[\sum_{i_1=1}^n s_{i_1} \cdot b_{i_1,1}, \sum_{i_2=1}^n s_{i_2} b_{i_2,2}, \dots, \sum_{i_n=1}^n s_{i_n} b_{i_n,n} \right]$$

Select determinent form such that $\triangle(e_1, \dots, e_n) = 1$.

$$\det(A \cdot B) = \triangle \left(\sum_{i_1=1}^n s_{i_1} b_i, \dots, \sum_{i_n=1}^n s_{i_n} b_{i_n,n} \right)$$

From multilinearity it follows that

$$\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_n=1}^{n} b_{i_1,1}b_{i_2,2}\cdots b_{i_n,n}\triangle(s_{i_1},\dots,s_{i_n})$$

If two indices satisfy $i_k = i_l \Rightarrow \triangle = 0$.

$$\Rightarrow \sum_{\text{different indices}} = \sum_{\text{permutations}}$$

$$= \underbrace{\sum_{\pi\in\sigma_n} b_{\pi(1),1}b_{\pi(2),2}\cdots b_{\pi(n),n}}_{\det(B)} \underbrace{\triangle(s_{\pi(1)},\dots,s_{\pi(n)})}_{\operatorname{sign}(\pi)\underbrace{\triangle(s_1,\dots,s_n)}_{\det(A)}}$$

$$= \det A\cdot\det B$$

Be aware that $\det(B)$ also includes $\operatorname{sign}(\pi)$ from the right-hand side.

2.
$$A\cdot A^{-1} = I \Leftrightarrow \det(A\cdot A^{-1}) = \det I = 1$$

$$\det(A\cdot A^{-1}) \stackrel{1.}{=} \det(A)\cdot\det(A^{-1})$$

3. $\det(A) = 0$ and $\det(A) = \det(f_A)$.

$$\Leftrightarrow f_A \text{ is not bijective } \Leftrightarrow \operatorname{rank}(A) < n$$

4.

$$\begin{aligned}\det(A^T) &= \sum_{\pi\in\sigma_n}\operatorname{sign}(\pi)a^T_{\pi(1),1}\dots a^T_{\pi(n),n}\\ &= \sum_{\pi\in\sigma_n}\operatorname{sign}(\pi)a_{1,\pi(1)}\dots a_{n,\pi(n)}\\ &= \sum_{\pi\in\sigma_n}\operatorname{sign}\pi a_{\pi^{-1}(1),1}\dots a_{\pi^{-1}(n),1}\\ &= \sum_{\rho}\operatorname{sign}\rho^{-1}a \qquad\qquad \rho = \pi^1\end{aligned}$$

For fixed π:

$$\prod_{j=1}^{n} a_{j,\pi(j)} = \prod_{k=1}^{n} a_{\pi^{-1}(k),k}$$

$$\pi(j) = 1 \Leftrightarrow j = \pi'(1)$$

$$\pi(j) = k \Leftrightarrow j = \pi'(k)$$

$$\sum_{\pi}\operatorname{sign}\pi a_{\pi^{-1}(1),1}\dots a_{\pi^{-1}(n),n}$$

$$= \sum\operatorname{sign}(\rho^{-1})a_{\rho(1),1}\dots a_{\rho(n),n} = \sum_{\rho}\operatorname{sign}(\rho)a_{\rho(1),1}\dots a_{\rho(n),n} = \det A$$

Remark:

$$\sigma_n \to \sigma_n \text{ is bijective}$$

$$\pi \mapsto \pi^{-1}$$

$$\operatorname{sign}(\rho) = (-1)^k \text{ where } \rho = \tau_1,\dots,\tau_k$$

$$\Rightarrow \rho^{-1} = \tau_k\circ\dots\circ\tau_n$$

$$\operatorname{sign}\rho^{-1} = (-1)^k$$

□

Remark 7 (Determination of determinants)**.** dim ≤ 3

For $n = 2$:

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

For $n = 3$:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix} = \sum_{\pi\in\sigma_3}\operatorname{sign}(\pi)a_{\pi(1),1}a_{\pi(2),2}a_{\pi(3),3}$$

General linear group:

$$\begin{aligned}\mathrm{GL}(n,\mathbb{K}) &= \text{ group of invertible matrices}\\ &= \left\{A\in\mathbb{K}^{n\times n}\,\middle|\,\det(A)\neq 0\right\}\\ \mathrm{SL}(n,\mathbb{K}) &= \text{ special linear group}\\ &= \left\{A\in\mathbb{K}^{n\times n}\,\middle|\,\det(A)=1\right\}\end{aligned}$$

σ_3 is a coxeter group.

$$\sigma_3 = \langle\tau_{12},\tau_{23}\rangle$$

Is created by two transpositions.

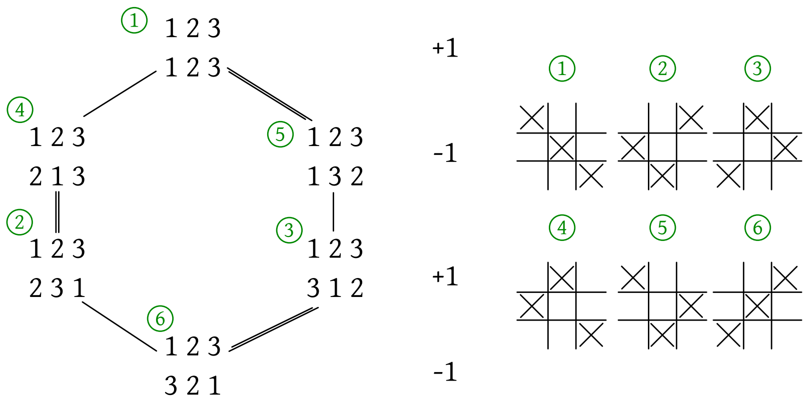

Figure 3: Sign of a permutation

$$= a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23} - a_{31}a_{22}a_{13}$$

corresponding to (1) + (2) + (3) + (4) + (5) + (6) in Figure 3.

Remark 8 (Rule of Sarrus)**.** Compare with Figure 4.

You write the first two columns next to right side of the matrix. You add up all 3 diagonals (the product of their values) from top left diagonally to the right bottom and subtract all 3 diagonals from left bottom to the top right.

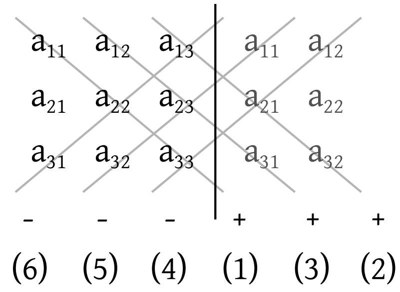

Figure 4: Rule of Sarrus visualized

The rule of Sarrus does not hold for $n = 4$!

Example 6.

$$\det\begin{vmatrix}1 & 2 & 5\\ 2 & 5 & 14\\ 5 & 14 & 42\end{vmatrix} = 1\cdot5\cdot42 + 2\cdot14\cdot5 + 5\cdot2\cdot14 - 5\cdot5\cdot5 - 14\cdot14\cdot1 - 2\cdot2\cdot42$$

$$= 14(1\cdot5\cdot3 + 2\cdot5 + 5\cdot2 - 14 - 2\cdot2\cdot3) - 125 = 14\cdot9 - 125 = 1$$

It turns out, if we use Catalan numbers, we always end up with determinant 1.

Lemma 4. Let A be an upper triangular matrix, hence $a_{ij} = 0 \forall i > j$. Then it holds that $\det A = a_{11}a_{22}\dots a_{nn}$.

Proof.

$$\det A = \sum_{\pi\in\sigma_n} \operatorname{sign}\pi a_{\pi(1),1}\dots a_{\pi(n),n}$$

it must hold that

$$\pi(j)\le j \qquad \forall j$$

$$\Rightarrow \pi(1)=1, \pi(2)=2, \dots, \pi(n)=n$$

The only permutation which contributes something is the identity. And sign id $=$ 1, hence

$$= 1\cdot a_{11}a_{22}\dots a_{nn}$$

Lemma 5 (Elementary row and column transformations)**.**

$$A = [a_{ij}] \in \mathbb{K}^{n \times n}$$

1.
$$s_i = \begin{bmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{bmatrix} \text{ column vectors}$$

$$\Rightarrow \det[as_1, \ldots, s_i + \lambda s_j, \ldots, s_n] = \det(A) \qquad i \neq j$$

2. Let $z_i = [a_{i_1}, \ldots, a_{i_n}]$ rows of A.

$$\det \begin{bmatrix} z_1 \\ \vdots \\ z_i + \lambda z_j \\ \vdots \\ z_n = \end{bmatrix} = \det A \qquad \text{for } i \neq j$$

Proof. 1. compare with determinant form

2. $\det A = \det A^T$

□

Example 7.

$$\begin{vmatrix} 1 & 2 & 5 \\ 2 & 5 & 14 \\ 5 & 14 & 42 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 5 \\ 0 & 1 & 4 \\ 0 & 4 & 17 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 5 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{vmatrix} = 1 \cdot 1 \cdot 1 = 1$$

This lecture took place on 14th of March 2016 (Franz Lehner).

Lemma 6. Recall: The following operations do not change the determinant:

- $\triangle(s_1, \ldots, s_i + \lambda s_j, \ldots, s_n) = \triangle(s_1, \ldots, s_n)$
  Addition of a multiple of a column (or row) to another

□

- Gauss-Jordan operations (elementary row/column transformations)

Example 8.

$$\begin{vmatrix} 1 & 0 & 3 & -2 \\ 2 & 6 & 4 & 1 \\ 3 & 3 & -1 & -1 \\ -1 & 2 & 4 & 1 \end{vmatrix} \rightsquigarrow \begin{vmatrix} 1 & 0 & 3 & -2 \\ 0 & 6 & -2 & 5 \\ 0 & 3 & -10 & 5 \\ 0 & 2 & 7 & -1 \end{vmatrix} \rightsquigarrow \frac{1}{3}\frac{1}{2} \begin{vmatrix} 1 & 0 & 3 & -2 \\ 0 & 6 & -2 & 5 \\ 0 & 6 & -20 & 10 \\ 0 & 6 & 21 & -3 \end{vmatrix}$$

We multiplied the third row times 2 and the fourth row times 3. Be aware that this way we avoided fractions in the matrix.

$$\rightsquigarrow \frac{1}{6} \begin{vmatrix} 1 & 0 & 3 & -2 \\ 0 & 6 & -2 & 5 \\ 0 & 0 & -18 & 5 \\ 0 & 0 & 23 & -8 \end{vmatrix} \cdot \frac{23}{18} = \frac{1}{6} \begin{vmatrix} 1 & 0 & 3 & -2 \\ 0 & 6 & -2 & 5 \\ 0 & 0 & -8 & 5 \\ 0 & 0 & 0 & -8 + 5\frac{23}{18} \end{vmatrix}$$

Even though we have a fraction $\frac{1}{6}$ at the front, our result will remain to be integral (i.e. without decimal points).

Triangular matrix:

$$\frac{1}{6} \cdot 1 \cdot 6 \cdot (-18) \cdot \left(-8 + \frac{5 \cdot 23}{18}\right)$$
$$= -(-18 \cdot 8 + 5 \cdot 23) = -(-144 + 115) = 29$$

Lemma 7. 1.

$$\left|\begin{array}{c|ccc} a_{11} & * & \ldots & * \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right| = a_{11} \cdot \det B$$

2.

$$\left|\begin{array}{ccc|c} & & & 0 \\ & B & & 0 \\ & & & \vdots \\ \hline * & \ldots & * & a_{nn} \end{array}\right| = \det B \cdot a_{nn}$$

Proof.

$$\det A = \sum_{\pi \in \sigma_n} (-1)^{\pi} a_{\pi(1),1} \ldots a_{\pi(2),2}$$

2.

$$\begin{aligned} a_{\pi(n),n} &= 0 \text{ except when } \pi(n) = n \\ &= \sum_{\pi\in\sigma_n} (-1)^\pi a_{\pi(1),1} \dots a_{\pi(n),n} \\ &= \sum_{\rho\in\sigma_{n-1}} (-1)^\rho a_{\rho(1),1} \dots a_{\rho(n-1),n-1} a_{\rho(n),n} \qquad = \det B \cdot a_{nn} \end{aligned}$$

□

Definition 8. Let $A \in \mathbb{K}^{n\times n}$.

$$1 \leq k, l \leq n$$

$A_{k,l}$ (dimension $(n-1)\times(n-1)$) which is generated by A if you cancel out row k and column l.

$$\begin{vmatrix} a_{1,1} & \dots & a_{1,l-1} & a_{1,l+1} & \dots & a_{1,n} \\ a_{2,1} & \dots & a_{2,l-1} & a_{2,l+1} & \dots & a_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{k-1,1} & \dots & a_{k-1,l-1} & a_{k-1,l+1} & \dots & a_{k-1,n} \\ a_{k+1,1} & \dots & a_{k+1,l-1} & a_{k+1,l+1} & \dots & a_{k+1,n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,l-1} & a_{n,l+1} & \dots & a_{n,n} \end{vmatrix}$$

Theorem 15 (Generative theorem of Laplace (dt. Entwicklungssatz von Laplace))**.** Let $A \in K^{n\times n}$, then it holds that

$$\det(A) = \sum_{k=1}^n a_{k,l} \cdot (-1)^{k+l} \cdot \det A_{k,l}$$

Generation to l-th column.

$$\det A = \sum_{l=1}^n a_{k,l} \cdot (-1)^{k+l} \cdot \det A_{k,l}$$

Generation to k-th row.

Proof. l-th column is

$$a_l = \sum_{k=1}^n a_{kl} e_k$$

$$\begin{aligned} \det(A) &= \triangle(a_1, \dots, a_n) \\ &= \triangle(a_1, \dots, a_{l-1}, \sum_{k=1}^n a_{kl} e_k, \dots, a_n) \\ &= \sum_{k=1}^n a_{kl} \triangle(a_1, \dots, a_{l-1}, e_k, \dots a_n) \\ &= \sum_{k=1}^n a_{kl} \begin{vmatrix} a_{11} & \dots & a_{1,l-1} & 0 & a_{1,l+1} & \dots & a_{1,n} \\ & & & \vdots & & & \\ & & & 0 & & & \\ \vdots & & & 1 & & & \vdots \\ & & & 0 & & & \\ & & & \vdots & & & \\ a_{n1} & \dots & a_{n,l-1} & 0 & a_{n,l+1} & \dots & a_{n,n} \end{vmatrix} \end{aligned}$$

where 1 is given on the k-th row and the l-th column which is e_k.

We exchange the l-th column with the $(l-1)$-th, then $(l-2)$-th and so on and so forth ... This requires $(l-1)$ transpositions.

$$\sum_{k=1}^n a_{kl} (-1)^{l-1} \begin{vmatrix} 0 & a_{11} & \dots & a_{1,l-1} & a_{1,l-1} & \dots & a_{1,n} \\ \vdots & a_{21} & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots \\ \vdots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & a_{n1} & \dots & a_{n,l-1} & a_{n,l-1} & \dots & a_{n,n} \end{vmatrix}$$

where 1 is given on the k-th row.

Exchange k-th and $(k-1)$-th row, then $(k-2)$-th and so on and so forth ... This requires $k-1$ transpositions.

$$= \sum_{k=1}^{n} a_{kl}(-1)^{k-1+l-1} \begin{vmatrix} 1 & \\ \hline 0 & \\ \vdots & \\ \dots & A_{k,l} \\ \vdots & \\ 0 & \end{vmatrix} = \sum_{l=1}^{n} a_{k,l}(-1)^{k+l} \det A_{k,l}$$

□

Example 9.

$$\begin{vmatrix} 1 & 2 & 5 \\ 2 & 5 & 14 \\ 5 & 14 & 42 \end{vmatrix} = 1 \cdot \begin{vmatrix} 5 & 14 \\ 14 & 42 \end{vmatrix} - 2 \cdot \begin{vmatrix} 2 & 14 \\ 5 & 42 \end{vmatrix} + 5 \cdot \begin{vmatrix} 2 & 5 \\ 5 & 4 \end{vmatrix}$$

+	−	+
−	+	−
+	−	+

where the top right + refers to the third summand (submatrix) and the top middle − refers to the second summand (submatrix).

$= (5 \cdot 42 - 14 \cdot 14) - 2 \cdot (2 \cdot 42 - 5 \cdot 14) + 5 \cdot (2 \cdot 14 - 5 \cdot 5) = 14 - 2 \cdot 14 + 5 \cdot 3 = 1$

Theorem 16. A is invertible iff $\det A \neq 0$.

Let $A \in K^{n \times n}$, $\hat{A} := [\hat{a}_{kl}]_{k,l=1,\dots,n}$ is the *complementary matrix* or *adjoint matrix.*

$$\hat{a}_{kl} = (-1)^{k+l} \det A_{lk}$$

Then

$$A^{-1} = \frac{1}{\det A} \cdot \hat{A}$$

Proof. Show that $B := \hat{A} \cdot A = \det A \cdot I$.

$$b_{k,l} = \sum_{j=1}^{n} \hat{a}_{kj} a_{jl} = \sum_{j=1}^{n} (-1)^{k+j} \det A_{jk} a_{jl}$$

Case $k = l$

$$b_{kk} = \sum_{j=1}^{n} (-1)^{k+j} a_{jk} \det A_{jk} = \det A \text{ (Laplace generation to } k\text{-th column)}$$

Case $k \neq l$ Without loss of generality $k < l$.

$$0 = \det \begin{bmatrix} a_{11} & \dots & a_{1l} & \dots & a_{1l} & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nl} & \dots & a_{nl} & \dots & a_{nn} \end{bmatrix}$$

We replace the k-th column (left column with a_{1l} in the middle) by the l-th column (right column with a_{1l} in the middle).

Laplace generation by k-th column:

$$= \sum_{j=1}^{n} a_{jl} \det \begin{bmatrix} a_{11} & \dots & 0 & \dots & a_{1l} & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & 1 & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \dots & 0 & \dots & a_{nl} & \dots & a_{nn} \end{bmatrix}$$

Similar to Laplace:

$$= \sum_{j=1}^{n} a_{jl}(-1)^{j+l} \det A_{jk} = \sum_{j=1}^{n} a_{jl} \hat{a}_{kj} = b_{kl}$$

□

Example 10 (Cayley 1855)**.** Cayley considered it as partial derivations:

$$\frac{1}{\triangledown} \begin{vmatrix} \partial_a \triangledown & \partial_c \triangledown \\ \partial_b \triangledown & \partial_d \triangledown \end{vmatrix}$$

Consider $n = 2$:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Consider $n = 3$:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}^{-1} = \frac{1}{\det A} \begin{bmatrix} \begin{vmatrix} a_{22} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} \\ \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \end{bmatrix}$$

Example 11.

$$\begin{bmatrix} 1 & 2 & 5 \\ 2 & 5 & 14 \\ 5 & 14 & 42 \end{bmatrix} = \begin{bmatrix} \begin{vmatrix} 5 & 14 \\ 14 & 42 \end{vmatrix} & -\begin{vmatrix} 2 & 5 \\ 14 & 42 \end{vmatrix} & \begin{vmatrix} 2 & 5 \\ 5 & 14 \end{vmatrix} \\ -\begin{vmatrix} 2 & 14 \\ 5 & 42 \end{vmatrix} & \begin{vmatrix} 1 & 5 \\ 5 & 42 \end{vmatrix} & -\begin{vmatrix} 1 & 5 \\ 2 & 14 \end{vmatrix} \\ \begin{vmatrix} 2 & 5 \\ 5 & 14 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 5 & 14 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 2 & 5 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} 14 & -14 & 3 \\ -14 & 17 & -4 \\ 3 & -4 & 1 \end{bmatrix}$$

$$\begin{vmatrix} 5 & 14 \\ 14 & 42 \end{vmatrix} = 5 \cdot 3 \cdot 14 - 14 \cdot 14 = 14$$

$$\begin{vmatrix} 2 & 5 \\ 14 & 42 \end{vmatrix} = 2 \cdot 3 \cdot 14 - 5 \cdot 14 = 14$$

Theorem 17 (Arnold's hypothesis)**.** "No theorem in mathematics is named after it's original author"

Proof. No proof provided here. $\square$

Theorem 18 (Cramer's rule)**.** Originally by McLansin (1748) based on work by Leibniz (1678) and reformulated by G. Cramer (1750).

A regular $n \times n$ matrix with column vectors $a_1, \dots, a_n \in \mathbb{K}^n$.

Then the unique solution to the equation system $Ax = b$ is given by

$$x_i := \frac{\triangle(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n)}{\triangle(a_1, \dots, a_n)} = \frac{\det(a_1, \dots, b, \dots, a_n)}{\det A}$$

Its complexity is given by $n + 1$ determinants!

Proof.

$$b = \sum_{j=1}^{n} b_j e_j$$

$$x = A^{-1} b - \frac{1}{\det A} \hat{A} \cdot b$$

$$x_i = \frac{1}{\det A} \sum_{j=1}^{n} \hat{a}_{ij} b_j = \frac{1}{\det A} \sum_{j=1}^{n} (-1)^{i+j} \det(A_j) b_j$$

$$= \frac{1}{\det A} \sum_{j=1}^{n} \triangle(a_1, \dots, a_{i-1}, \dots, a_{j-1}, e_j, a_{j+1}, \dots, a_n) \cdot b_j$$

$$= \frac{1}{\det A} \triangle(a_1, \dots, a_{i-1}, \underbrace{\sum_{j=1}^{n} b_j e_j}_{=b}, \dots, a_n)$$

$\square$

Example 12.

$$2x_1 + 2x_2 = 7$$

$$x_1 - 3x_2 = 0$$

$$A = \begin{bmatrix} 2 & 2 \\ 1 & -3 \end{bmatrix} \qquad b = \begin{bmatrix} 7 \\ 0 \end{bmatrix}$$

$$\det A = -8 \qquad x_1 = \frac{\begin{vmatrix} 7 & 2 \\ 0 & -3 \end{vmatrix}}{-8} = \frac{21}{8} \qquad x_2 = \frac{\begin{vmatrix} 2 & 7 \\ 1 & 0 \end{vmatrix}}{-8} = \frac{7}{8}$$

Remark 9.

- in higher dimensions ($n \geq 4$) Cramer's rule is disallowed.
  1. too computationally intense
  2. numerically unstable (small errors have large effects)
- Anyways, still useful for theoretical considerations
  1. the map $A \mapsto \det A$ is C^∞ (polynomial!) (this denotes infinite differentiability)

2. The set of invertible matrices in $\mathbb{R}^{n\times n}$ is open, because if $\det A \neq 0$, then also $\det \tilde{A} \neq 0$ as long as $|\, a_{ij} - \tilde{a}_{ij} \,| < \delta$.
3. The solution of the equation system $Ax = b$, for invertible A, depends continuously and differentiable on A and b:

$$x_i = \underbrace{\frac{1}{\det A}}_{\text{continuous as long as } \det A \neq 0} \underbrace{\hat{A}b}_{\text{polynomial}}$$

4. The map $\mathrm{GL}(n,\mathbb{R}) \to \mathrm{GL}(n,\mathbb{R})$

$$A \mapsto A^{-1}$$

is continuous.

$$A^{-1} = \frac{1}{\det A} \cdot \hat{A}$$

So $\mathrm{GL}(n,\mathbb{R})$ is a Lie group.

This lecture took place on 16th of March 2016 (Franz Lehner).

3 Inner products

Descartes introduced "La Géometrie" (1637).

Definition 9. The length of a vector in $\mathbb{R}^2/\mathbb{R}^3$ is:

$$\left\| \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right\| = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

Definition 10 (Scalar product)**.**

$$\cos\theta = \cos(2\pi - \theta)$$

The scalar product is defined as

$$\langle a, b\rangle = \|a\| \cdot \|b\| \cdot \cos\theta$$

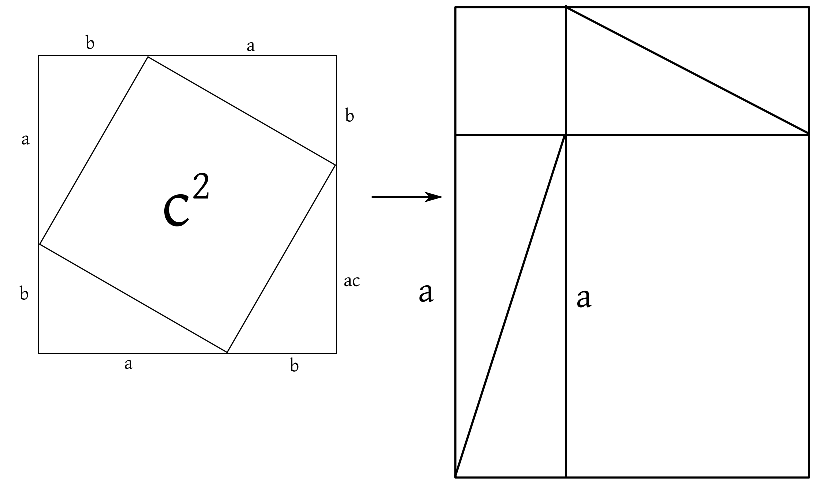

Figure 5: Pythagorian proof of $c^2 = a^2 + b^2$

Theorem 19. The following properties hold:

- $\|\lambda \cdot a\| = |\,\lambda\,| \cdot \|a\|$
- $\|a + b\| \leq \|a\| + \|b\|$ (triangle inequality)
- $\langle a, a\rangle = \|a\|^2 \geq 0$
- $\langle a, a\rangle = 0 \Leftrightarrow a = 0$
- $\langle a, b\rangle = 0 \Leftrightarrow a = 0 \vee b = 0$

$$\langle a, b\rangle > 0 \Leftrightarrow \text{ acute angle}$$

$$\langle a, b\rangle < 0 \Leftrightarrow \text{ obtuse angle}$$

Theorem 20.

$$\langle a, b\rangle = \langle b, a\rangle \tag{1}$$
$$\langle \lambda a, b\rangle = \lambda \langle a, b\rangle \tag{2}$$
$$\langle a+b, c\rangle = \langle a, c\rangle + \langle b, c\rangle \tag{3}$$

So it actually describes a bilinear map.

Proof.
- immediate
- $\lambda > 0$ immediate
  $\lambda < 0$ Angle θ becomes $\pi - \theta$.
  $$\cos(\pi-\theta) = -\cos\theta$$
  $$\langle \lambda a, b\rangle = |\,\lambda\,| \cdot \|a\| \cdot \|b\| \cos(\pi - \theta) = -|\,\lambda\,| \cdot \|a\| \cdot \|b\| \cdot \cos\theta = \lambda \langle a, b\rangle$$
- Let $b = e, \|e\| = 1$.
  $$\langle a, e\rangle = \|a\| \cdot \cos\theta$$
  $$\langle a+b, c\rangle = \|c\| \left\langle a+b, \frac{c}{\|c\|}\right\rangle - \|c\| \left(\left\langle a, \frac{c}{\|c\|}\right\rangle + \left\langle b, \frac{c}{\|c\|}\right\rangle \right) = \langle a, c\rangle + \langle b, c\rangle$$
  Compare with Figure 6.

□

Theorem 21.
$$\left\langle \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \right\rangle = a_1 b_1 + a_2 b_2 + a_3 b_3$$

Proof.
$$\begin{aligned} \langle a, b\rangle &= \langle a_1 e_1 + a_2 e_2 + a_3 e_3, b\rangle \\ &= a_1 \langle e_1, b\rangle + a_2 \langle e_2, b\rangle + a_3 \langle e_3, b\rangle \\ &= a_1 b_1 + a_2 b_2 + a_3 b_3 \\ \langle e_i, b\rangle &= \langle e_i, b_1 e_1 + b_2 e_2 + b_3 e_3\rangle \\ &= b_1 \langle e_i, e_1\rangle + b_2 \langle e_i, e_2\rangle + b_3 \langle e_i, e_3\rangle \\ &= b_1 \delta_{i1} + b_2 \delta_{i2} + b_3 \delta_{i3} \\ &= b_i \end{aligned}$$

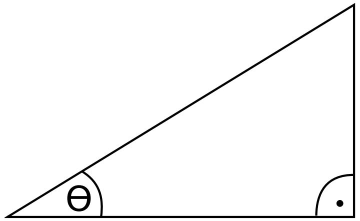

Figure 6: $\langle a+b, c\rangle = \langle a, c\rangle + \langle b, c\rangle$

with dim $\langle e_i, e_j\rangle = \delta_{ij}$. □

Example 13 (Law of cosines)**.**

$$a^2 + b^2 = c^2 + 2ab\cos\gamma$$

Compare with Figure 7.

$$\begin{aligned} \|c\|^2 &= \langle a-b, a-b\rangle \\ &= \langle a, a\rangle - \langle a, b\rangle - \langle b, a\rangle + \langle b, b\rangle \\ &= \|a\|^2 - 2 \cdot \|a\|\|b\| \cos\gamma + \|b\|^2 \end{aligned}$$

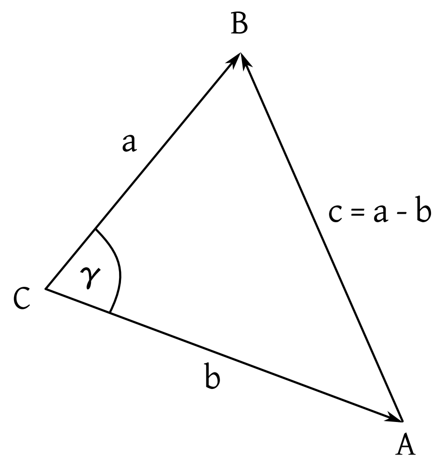

Figure 7: Law of cosines

Theorem 22. Theorem by Thales TODO: image

$$\begin{aligned}\langle a-b,-a-b\rangle &= \|a-b\|\|a+b\|\cos\theta\\ \langle a-b,-a-b\rangle &= -\langle a-b,a+b\rangle\\ &= -\left(\langle a,a\rangle - \langle b,a\rangle + \langle a,b\rangle - \langle b,b\rangle\right)\\ &= -\left(\|a\|^2 - \|b\|^2\right)\\ &= 0\\ &\Rightarrow \theta = \frac{\pi}{2}\end{aligned}$$

Remark 10. How do we find the normal vector?

$$\vec{n} = \begin{pmatrix} a_2 \\ -a_1 \end{pmatrix}$$

$$\left\langle \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} a_2 & -a_1 \end{pmatrix} \right\rangle = a_1a_2 - a_2a_1 = 0$$

Definition 11 (Outer product)**.** “Outer product”, “cross product” or “vector product”

TODO: image missing

This is only available in $\mathbb{R}^3$.

Let $a, b \in \mathbb{R}^3$, then $a \times b$ is the vector with properties:

- $\|a \times b\| = \|a\| \cdot \|b\| \cdot \sin\theta$
  This corresponds to the are of a parallelogram.

  $$\|b\| \cdot \sin\theta = \text{ heigh of a parallelogram}$$

- $a \times b \perp a, b$

  $$\begin{aligned}\langle a \times b, a\rangle = 0\\ \langle a \times b, b\rangle = 0\end{aligned}$$

- $(a, b, a \times b)$ are clockwise (consider a screw coming out of Figure)

$$a \times b = 0 \Leftrightarrow a = 0 \vee b = 0 \vee a, b \text{ are linear dependent}$$

Theorem 23. 1. $b \times a = -a \times b$ (counter-clockwise)

2. $(\lambda a) \times b = \lambda \cdot a \times b = a \times (\lambda b)$
3. $(a + b) \times c = a \times c + b \times c$

So it is bilinear in $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$

Proof.

$$a \times c, b \times c, (a + b) \times c \in E$$

Let $a', b', (a + b)'$ be the projection of a, b and $a + b$ in the plane.

TODO: image missing

1. $$(a + b)' = a' + b'$$

   Projection of the sum = sum of projections.

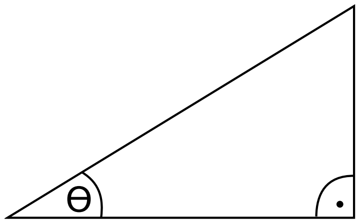

Figure 8: Theorem 23, third statement

2.
$$a \times c = a' \times c$$
$$\|a' \times c\| = \|a'\| \cdot \|c\|$$

$$\begin{aligned}\|a \times c\| &= \|a\| \cdot \|c\| \cdot \sin\theta \\ &= \|a'\| \cdot \|c\|\end{aligned}$$

$$\|a'\| = \|c\| \cdot \sin\theta$$

and they have the same direction.

TODO: image missing

3.
$$(a' + b') \times c = c' \times c + b' \times c$$

From above:

TODO: image missing

$$\|a' \times c\| = \|c\| \cdot \|a'\|$$

So this operation is linear.

$$\begin{aligned}(a+b) \times c &\overset{2}{=} (a+b)' \times c \\ &\overset{1}{=} (a' + b') \times c \\ &\overset{3}{=} (a' \times c + b' \times c) \\ &\overset{2}{=} a \times c + b \times c\end{aligned}$$

□

Corollary 7. The cross product is a map $x : \mathbb{R}^3 \to \mathbb{R}^3 \to \mathbb{R}^3$ with properties:

- bilinear
- anti-symmetric

- "chiral", namely

$$e_1 \times e_2 = e_3$$
$$e_2 \times e_3 = e_1$$
$$e_3 \times e_1 = e_2$$

Corollary 8.

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{bmatrix} = \begin{bmatrix} \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \\ -\begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} \\ \begin{vmatrix} a_2 & b_2 \\ a_2 & b_2 \end{vmatrix} \end{bmatrix}$$

$$\overset{\text{Laplace}}{=} \begin{vmatrix} a_1 & b_1 & e_1 \\ a_2 & b_2 & e_2 \\ a_3 & b_3 & e_3 \end{vmatrix}$$

Formally, matrices in a vector of values are disallowed, but as far as it boils down to addition, this is fine.

Proof.

$$\begin{aligned}
&(a_1e_1 + a_2e_2 + a_3e_3) \times (b_1e_1 + b_2e_2 + b_3e_3) \\
&= a_1b_1e_1 \times e_1 + a_1b_2e_1 \times e_2 + a_1b_3e_1 \times e_3 \\
&+ a_2b_1e_2 \times e_1 + a_2b_2e_2 \times e_2 + a_2b_3e_3 \times e_3 \\
&+ a_3b_1e_3 \times e_1 + a_3b_2e_3 \times e_2 + a_3b_3e_3 \times e_3 \\
&= a_1b_2e_3 + a_1b_3(-e_2) + a_2b_1(-e_3) + a_2b_3e_1 + a_3b_1e_2 + a_3b_2(-e_1) \\
&= (a_2b_3 - a_3b_2)e_1 + (a_3b_1 - a_1b_3)e_2 + (a_1b_2 - a_2b_1)e_3
\end{aligned}$$

□

Theorem 24 (Scalar triple product)**.** The three-dimensional parallelogram is called "Spat" in German (compare with Figure 9).

$$\langle a \times b, c \rangle = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \text{volume of spanned 3-dimensional parallelogram}$$

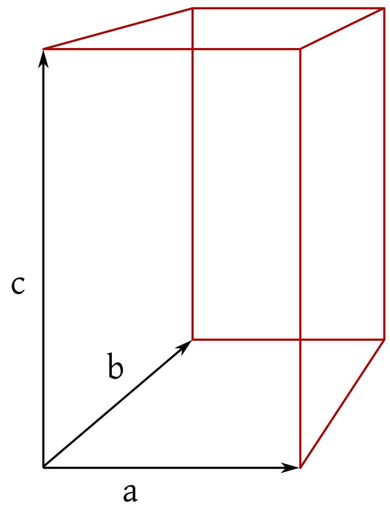

Figure 9: Three-dimensional parallelogram

$\|a \times b\|$ is the area of the parallelogram. $\langle a \times b, c \rangle = \|a \times b\| \cdot \|c\| \cdot \cos\theta$ where $\|c\| \cdot \cos\theta$ is the height of the 3-dimensional parallelogram.

$$\langle a \times b, c \rangle = \left\langle \begin{pmatrix} \begin{vmatrix} a_1 & b_2 \\ a_3 & b_3 \end{vmatrix} \\ -\begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} \\ \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \end{pmatrix}, \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \right\rangle$$

$$\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \cdot c_1 - \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} \cdot c_2 + \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \cdot c_3 = \text{Laplace generated by third column}$$

Example 14. Given a plane in parameter representation:

$$E = \{v_0 + \lambda a + \mu b \mid \lambda, \mu \in \mathbb{R}\}$$

Find $\alpha_1, \alpha_2, \alpha_3$ and β with ("implicit representation")

$$E = \{x \mid \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 = \beta\}$$

$$\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix} = a \times b$$

TODO: image missing

$$\beta = \langle v_0, a \times b\rangle$$

In the following chapters we always consider $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.

Definition 12. An inner product over a vector space in $\mathbb{R}$ or $\mathbb{C}$ is a map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ with properties:

- $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle \quad \forall x, y, z \in V$
- $\langle \lambda x, y\rangle = \lambda \langle x, y\rangle \quad \forall x, y \in V \forall \lambda \in \mathbb{K}$
- $\langle y, x\rangle = \overline{\langle x, y\rangle} \quad \forall x, y \in V$

where $\overline{\langle x, y\rangle}$ denotes the complex conjugate. Especially $\langle x, x\rangle \in \mathbb{R} \forall x \in V$.

An inner product is called

positive semidefinite if $\langle x, x\rangle \geq 0 \quad \forall x$

positive definite if $\langle x, x\rangle > 0 \quad \forall x \neq 0$

negative semidefinite if $\langle x, x\rangle \leq 0 \quad \forall x$

negative definite if $\langle x, x\rangle < 0 \quad \forall x \neq 0$

indefinite if $\exists x : \langle x, x\rangle > 0 \land \exists y : \langle y, y\rangle < 0$

Definition 13. Scalar product if $\mathbb{K} = \mathbb{R}$
Hermitian product (or unitary product) if $\mathbb{K} = \mathbb{C}$

Quadratic form if $\mathbb{K} = \mathbb{R}$
Hermitian form if $\mathbb{K} = \mathbb{C}$

Lemma 8.
- $\langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle$
- $\langle x, \lambda y\rangle = \overline{\lambda} \langle x, y\rangle$
- $\langle x, 0\rangle = 0$

Linear in x and anti-linear in y!

Sesquilinear.

This lecture took place on 11th of April 2016 (Franz Lehner).

Scalar product.

1. Not bilinear, but sequilinear
   - $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$
   - $\langle \lambda x, y\rangle = \lambda \langle x, y\rangle$
   - $\langle x, y\rangle = \overline{\langle y, x\rangle}$ (hermetian)
2. $\langle , \rangle$ is called positive definite, if

$$\langle x, x\rangle > 0 \qquad \forall x \neq 0$$

≥ 0	positive semidefinite
< 0	negative definite
$\leq$	negative semidefinite
$\neq$	indefinite

3.1 Examples

- $\mathbb{R}^n$:

$$\left\langle \begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}, \begin{bmatrix}y_1\\ \vdots\\ y_n\end{bmatrix} \right\rangle = \sum_{x=1}^{n} x_i y_i = x^t y$$

- $\mathbb{C}^n$

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i} = x^t \overline{y}$$

- $A = [a_j]_{j=1,\dots,n}$ because $\langle x, y\rangle_A = x^t A y$ is complex.

Exercise: is symmetrical if and only if $A = A^t$, hence $a_{ij} = a_{ji}$.
Exercise: is hermetian if and only if $a_{ij} = \overline{a_{ji}}$ (A is hermitian)

dim $= \infty$.

$$\mathbb{R}^\infty : \langle x, y\rangle = \sum_{i=1}^{\infty} x_n y_n$$

Development on $l^2 = \left\{ (x_i) \,\middle|\, \sum_{|\,x_i\,|^2} < \infty \right\}$ where l stands for Lebeque.

$\Rightarrow$ Hilbert space.

$$V = C([a,b], \mathbb{C})$$

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

3.2 Norm

Definition 14. A *norm* on a vector space V over $\mathbb{R}$ or $\mathbb{C}$ is a mapping $\|\,\| \to [0, \infty[$ with properties

N1. $\|X\| \geq 0 \forall X$, $\|X\| = 0 \Leftrightarrow X = 0$

N2. $\|\lambda X\| = \|\lambda\| \cdot \|X\|$ (homogeneous)

N3. $\|X + Y\| \leq \|X\| + \|Y\|$ (triangle inequality)

Remark 11. A norm induces a metric.

$$d(x, y) = \|x - y\|$$

The induced metric satisfies

$$d(x + z, y + z) = d(x, y)$$

Example 15. • In $\mathbb{R}^n/\mathbb{C}^n$

$$\|X\|_\infty = \max(\|X_1\|, \dots, \|X_n\|)$$

The *euclidean norm* is given by:

$$\|X\|_2 = \left(\sum \|X\|^2\right)^{\frac{1}{2}}$$

The L^1-norm is given by (compare it with possible paths in a grid)

$$\|X\|_1 = \sum_{i=1}^{n} \|X_i\|$$

- Analogously for $V = C[a, b]$

$$\|f\|_\infty = \sup_{x \in [a,b]} \|f(x)\|$$

$$\|f\|_2 = \left(\int_a^b \|f(x)\|^2\,dx \right)^{\frac{1}{2}}$$

$$\|f\|_1 = \int_a^b \|f(x)\|\,dx$$

Theorem 25. Let $\langle,\rangle$ be a positive definite scalar product in V. Then $\|X\| = \sqrt{\langle x, x\rangle}$ defines a norm in V.

Proof. N1.

$$\langle x, x\rangle \geq 0 \Rightarrow \sqrt{\ } \text{ is well-defined in } \mathbb{R}^+$$

$$\|X\| = 0 \Leftrightarrow \langle x, x\rangle = 0 \xRightarrow{\text{positive definite}} X = 0$$

N2.

$$\|\lambda \cdot X\| = \sqrt{\langle \lambda X, \lambda X\rangle} = \sqrt{\lambda\overline{\lambda}\,\langle X, X\rangle} = \|\lambda\|\sqrt{\langle X, X\rangle} = \|\lambda\| \cdot \|X\|$$

because $\langle x, y\rangle = \overline{\langle \lambda y, x\rangle} = \overline{\lambda \langle y, x\rangle} = \overline{\lambda} \cdot \overline{\langle y, x\rangle} = \overline{\lambda} \cdot \langle x, y\rangle$.

$\square$

Lemma 9 (Cauchy-Bunjakovsky-Schwarz Inequality)**.** Cauchy (1789–1857), Bunjakovsky (1804–1880), Schwarz (1843–1921)

For a positive definite scalar product, the following inequality holds:

$$\|\langle x, y\rangle\| \leq \|x\| \cdot \|y\|$$

Equality holds if and only if x, y are linear independent.

Lemma 10. Cauchy (in "Cours d'Analyse", 1815)

$$|\sum_{i=1}^{n} x_i \overline{y}_i| \leq \left(\sum \|x_i\|^2\right)^{\frac{1}{2}} \left(\sum \|y\|^2\right)^{\frac{1}{2}}$$

Bunjakovsky (1859)

$$|\int_a^b f(x)\overline{g(x)}\, dx| \leq \left(\int_a^b \|f(x)\|^2\, dx\right)^{\frac{1}{2}} \cdot \left(\int_a^b \|g(x)\|^2\, dx\right)^{\frac{1}{2}}$$

Schwarz (1882), abstract

Lagrange (17??)

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{m}(x_i y_i - x_j y_i)^2 &= \sum_{x_i^2 y_j^2} -2\sum_{i,j} x_i y_j x_j y_i + \sum_{i,j} x_j^2 y_i^2 \\ &= 2\left(\sum x_i^2\right)\left(\sum x_i^2\right)\left(\sum_j y_j^2\right) - 2\left(\sum_{i=1}^{n} x_i \cdot y_i\right)^2\end{aligned}$$

$$\Rightarrow \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{j=1}^{m} y_j^2\right) = \left(\sum_{i=1}^{n} x_i \cdot y_i\right)^2 + \frac{1}{2}\sum_{i,j}(x_i y_j - x_j y_i)^2 \geq \left(\sum_{i=1}^{n} x_i y_i\right)^2$$

$h = 3$

$$\|X\|^2\|y\|^2 = \|\langle x, y\rangle\|^2 + \|x \cdot y\|^2$$

A geometrical proof is left as an exercise to th reader. □

General proof. **Case 1:** $y = 0$ trivial, $\langle x, y\rangle = 0$

Case 2: $y \neq 0$

$$\begin{aligned}0 \leq \langle x - \lambda y, x - \lambda y\rangle &= \langle x, x\rangle - \langle x, \lambda y\rangle - \langle \lambda y, x\rangle + \langle \lambda y, \lambda y\rangle \\ &= \langle x, x\rangle - \overline{\lambda}\langle x, y\rangle - \lambda \underbrace{\langle y, x\rangle}_{=\lambda\overline{\langle x,y\rangle}} - \|\lambda\|^2 \langle y, y\rangle\end{aligned}$$

holds for all λ. Especially:

$$\lambda = \frac{\langle x, y\rangle}{\langle y, y\rangle}$$

$$0 \leq \langle x, x\rangle - \frac{\overline{\langle x, y\rangle}}{y, x} \cdot \langle x, y\rangle - \frac{\langle x, y\rangle}{\langle y, x\rangle}\overline{\langle y, x\rangle} + \frac{|\ \langle x, y\rangle\ |^2}{\langle x, y\rangle^2}\langle y, y\rangle$$

$$= \langle x, x\rangle - \frac{|\ \langle x, y\rangle\ |^2}{\langle y, y\rangle}$$

$$\langle x, x\rangle \cdot \langle y, y\rangle \geq |\ \langle x, y\rangle\ |^2$$

Equality $\Rightarrow \|x - \lambda y\|^2 = 0 \Rightarrow x = \lambda y \Rightarrow$ linear independent. Inequality if $x = \lambda y$, $|\langle x, y\rangle| = |\langle \lambda y, y\rangle| = |\lambda|\|y\|^2 = \|x\| \cdot \|y\| = \|\lambda y\| \cdot \|y\|$.

The triangle inequality can be proven this way:

$$\begin{aligned}\|x + y\|^2 &= \langle x + y, x + y\rangle \\ &= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &= \|x\|^2 + 2\Re\langle x, y\rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| - \|y\|)^2\end{aligned}$$

□

Remark 12.

$$\|X\|_p = \left(\sum_{i=1}^{n}\|x_i\|^2\right)^{\frac{1}{p}}$$

with $1 \leq p < \infty$ is the L^p-norm

$\Rightarrow$ Höldische Ungleichung

$$|\sum x_i y_i| \leq \left(\sum |\, x_i \,|^p\right)^{\frac{1}{p}} \cdot \left(\sum |\, y_i \,|^2\right)^{\frac{1}{q}}$$

where q such that $\frac{1}{p} + \frac{1}{q} = 1$.

Theorem 26. Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$ with an inner product $\langle,\rangle$. Let $B = \{b_1, \dots, b_n\}$ be the basis of V.

Then there exists exactly one hermetian matrix $A \in \mathbb{K}^{n \times m}$ such that

$$\langle x, y \rangle = \Phi_B(x)^t A \overline{\Phi_B(y)}$$

then $\langle,\rangle$ is positive definite, A is regular.

Proof. Let $x = \sum_{i=1}^n \xi_i b_i$ and $y = \sum_{i=1}^n \eta_i b_i$.

$$\Phi_B(x) = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

$$\Phi_B(y) = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix}$$

$$\begin{aligned} \langle x, y \rangle &= \left\langle \sum_{i=1}^n \xi_i b_i, \sum_{j=1}^n \eta_j b_j \right\rangle \\ &= \sum_{i=1}^n \sum_{i=1}^n \xi_i \overline{\eta_j} \underbrace{\langle b_i, b_j \rangle}_{=:a_{ij}} \\ &= \sum_{i,j} \xi_i a_{ij} \overline{\eta_j} = \xi^t A \overline{\eta} \\ a_{ji} &= \langle b_j, b_i \rangle = \overline{b_i, b_j} = \overline{a_{ij}} \end{aligned}$$

$\Rightarrow$ A is regular.

It suffices to show that $\ker A = \{0\}$. Let $A\xi = 0 \Rightarrow \xi^t A \xi = 0 \Rightarrow \sum \xi a_i = 0$. And also $\xi^t A \xi = \langle \sum \xi b_i, \sum \xi b_i \rangle$ for all $\xi_i = 0$. $\square$

Definition 15. Let $A \in \mathbb{C}^{n \times n}$. Then the matrix

$$A^* := \overline{A^t}$$

$$(A^*)_{ij} = \overline{a_{ji}}$$

is the conjugate matrix to A (german: adjungiert).

A is called *self-conjugate* if $A = A^*$, *symmetrical* if $K = \mathbb{R}$ and *hermitian* if $K = \mathbb{C}$.

A is called positive/negative semidefinite/definite or indefinite if the inner product

$$\langle \xi, \eta \rangle_A := \xi^t A \overline{\eta}$$

has the corresponding property, hence A is positive positive definite if $\xi^t A \overline{\xi} > 0 \quad \forall \xi \neq 0$.

$$\langle x, x \rangle > 0 \quad x$$

We want to determine how "positive" a given matrix is.

Analogously to the rank, we consider: Every rank is equivalent to some matrix of the form

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

$$\exists P, Q \in \mathrm{GL}(n, \mathbb{K}) : PAQ = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

Definition 16. Two matrices $A, B \in \mathbb{C}^{m \times n}$ are called congruent if

$$\exists C \in \mathrm{GL}(n, \mathbb{C}) : C^* A C = B$$

(strong condition for equivalence)

Theorem 27. Every hermetian matrix is congruent to the diagonal matrix

$$D = \operatorname{diag}(+1, \dots, +1, -1, \dots, -1, 0, \dots 0)$$

Remark 13.
1. If $A \geq 0$ and C is arbitrary. Then $C^*AC \geq 0$. (A is positive semidefinite)
2. If $A > 0$ and $C \in \mathrm{GL}(n, \mathbb{K}) \Rightarrow C^*AC > 0$ (A is positive definite)

Theorem 28 (Sylvester's law of inertia)**.** Let $A \in \mathbb{C}^{n\times n}$ be a hermitian matrix and $C \in \mathrm{GL}(n, \mathbb{C})$.

Let $C^*AC = \operatorname{diag}(+1, \dots, +1, -1, \dots, -1, 0, \dots, 0)$. Then the number of $+1, -1$ and 0 is defined distinctly.

Definition 17. Let $A \in \mathbb{C}^{n\times n}$ be hermitian congruent to $\operatorname{diag}(\underbrace{+1, \dots, +1}_{r}, \underbrace{-1, \dots, -1}_{s}, 0, \dots, 0)$.

That means $\operatorname{ind}(A) := r$ is called *index of A* and $\operatorname{sign}(A) := r - s$ is called *signature of A*.

$$r + s = \operatorname{rank}(A)$$

A is positive definite if and only if $r = n$.

This lecture took place on 13th of April 2016 (Franz Lehner).

Theorem 29. Every Hermitian matrix ($A = A^*$) is congruent to $D = (+1, \dots, +1, -1, \dots, -1)$.

Constructive proof by induction. $n = 1$ Let $A = [a_{11}]$.

Find: c_{11} such that $\overline{c_{11}}a_{11}c_{11} = +1, -1, -0$

$$|\, c_{11}\,|^2 a_{11} = \pm 1, 0$$

$$c_{11} = \begin{cases} \frac{1}{\sqrt{|\, a_n \,|}} & \text{if } a_{11} \neq 0 \\ 1 & \text{if } a_{11} = 0 \end{cases}$$

$n - 1 \to n$ Basic idea:

$$\begin{bmatrix} 1 & \to 0 & \dots & 0 \\ \downarrow & \ddots & & \vdots \\ \vdots & & & \vdots \\ 0 & \dots & \dots & \ddots \end{bmatrix}$$

Create 0 in first column and row.

Case 1 $A = 0 \Rightarrow C = I$

Case 2 $a_n = 0$

Case 2a

$$\exists j : a_{jj} \neq 0, \qquad C = T_{(1j)} = C^* \Rightarrow (C^*AC)_{11} = a_{jj} \neq 0 \Rightarrow \text{ case 3}$$

Case 2b All $a_{jj} = 0$, $\exists i, j : a_{ij} \neq 0$

$$C = I + E_{ij} \cdot e^{i\theta} \text{ such that } e^{-i\theta} a_{ij} = |\, a_{ij} \,|$$

$$i \to \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Example 16.

$$A = \begin{bmatrix} 0 & 1 & i \\ 1 & 0 & 1 \\ -i & 1 & 0 \end{bmatrix}$$

Case 2b (cont.)

$$a_{12} \neq 0 \qquad C_1 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$A' = C_1^* A C_1 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & i \\ 1 & 0 & 1 \\ i & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & i \\ 1 & 2 & 1+i \\ -i & 1-i & 0 \end{bmatrix}$$

$\Rightarrow$ Case 2a, $a_{22} \neq 0$

$$C_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$C_2^* A' C_2 = \begin{bmatrix} 2 & 1 & 1+i \\ 1 & 0 & i \\ 1-i & -i & 0 \end{bmatrix}$$

$\Rightarrow$ Case 3

$$\begin{aligned} C^* AC &= \left[\left(I + E_{ji} e^{-i\theta}\right) A \left(I + E_{ij} e^{i\theta}\right)\right]_{ij} \\ &= \left[A + E_{ji} e^{-\theta} A + A E_{ij} e^{i\theta} + E_{ji} A E_{ij}\right]_{ij} \\ &= \underbrace{a_{ji}}_{=0} + e^{i\theta} a_{ij} + \underbrace{a_{ji} e^{i\theta}}_{=\overline{a_{ij} e^{-i\theta}}} + \underbrace{a_{ii}}_{=0} \\ &\overset{\text{by selection of } \theta}{=} |\, a_{ij} \,| \cdot 2 \end{aligned}$$

$\Rightarrow A'_{ji} \neq 0 \Rightarrow$ Case 2a
$\Rightarrow$ Case 3: $a_{11} \neq 0$
We generate zeroes.

Case 3: $a_{11} \neq 0$

$$C = \begin{bmatrix} 1 - \frac{a_{12}}{a_{11}} & -\frac{a_{13}}{a_k} & \dots & \dots & -\frac{a_{1n}}{a_k} \\ \vdots & \ddots & & & \vdots \\ \vdots & & 1 & & \vdots \\ \vdots & & & \ddots & \vdots \\ \vdots & \dots & & \dots & 1 \end{bmatrix}$$

$$A'' = C_2^* A C_2 = \begin{bmatrix} 2 & 1 & 1+i \\ 1 & 0 & i \\ 1-i & -i & 0 \end{bmatrix} \Rightarrow \text{case 3}$$

$$\begin{aligned} C_3^* A'' C_3 &= \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ -\frac{1-i}{2} & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1+i \\ 1 & 0 & i \\ 1-i & -i & 0 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1+i}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 1 & 1+i \\ 0 & -\frac{1}{2} & \frac{-1+i}{2} \\ 0 & \frac{-1-i}{2} & -1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{-i+2}{2} \\ 0 & \frac{-1-i}{2} & -1 \end{bmatrix} \end{aligned}$$

□

TODO

This lecture took place on 18th of April 2016 (Franz Lehner).

Revision: A is positive definite. $A = A^*$.

$$\bigwedge_{x \neq 0} x^t A x > 0 \Leftrightarrow \operatorname{index} A = n$$

$$A \mathrel{\widehat{=}} D = \begin{bmatrix} +1 & & & & & \\ & +1 & & & & \\ & & -1 & & & \\ & & & -1 & & \\ & & & & 0 & \\ & & & & & 0 \end{bmatrix}$$

where r is the number of $+1$ and s is the number of -1.

Hence

$$\bigvee_{C \in \mathrm{GL}(n, \mathbb{C})}$$

index $A = r$ and sign $A = r - s$.

Remark 14. A matrix is called *non-negative* if all $a_{ij} \geq 0$.

We denote $A \geq 0$.

$A < 0$.

$A \prec 0$ if sign $A = -n$.

Indefinite:

$$\begin{cases} r > 0 & \operatorname{index} A \neq 0 \\ s > 0 & \operatorname{index} A - \operatorname{sign} A \neq 0 \end{cases} \operatorname{index} A \cdot (\operatorname{index} A - \operatorname{sign} A) \neq 0$$

Remark 15. The *minors of a matrix* are defined as

$$[A]_{I,J} = \begin{vmatrix} a_{i_1,j_1} & a_{i_1,j_2} & \dots & a_{i_1,j_r} \\ \vdots & \ddots & \ddots & \vdots \\ a_{i_r,j_1} & a_{i_r,j_2} & \dots & a_{i_r,j_r} \end{vmatrix}$$

$$I = \{i_1 < i_2 < \ldots < i_r\}$$
$$J = \{j_1 < j_2 < \ldots < j_r\}$$

Theorem 30 (Fundamental minor criterion)**.**

$$A > 0 \Leftrightarrow \begin{vmatrix} a_{11} & \ldots & a_{ir} \\ \vdots & & \vdots \\ a_{r1} & \ldots & a_{rr} \end{vmatrix} > 0 \quad \text{for } r = 1, 2, \ldots, n$$

$$\Rightarrow A_r = \begin{bmatrix} a_{11} & \ldots & a_{ir} \\ \vdots & & \vdots \\ a_{r1} & \ldots & a_{rr} \end{bmatrix}$$

are all defined positively.

$$\{\xi^t A_t\} = \begin{bmatrix} \xi \\ \overline{0_{n-r}} \end{bmatrix} A \begin{bmatrix} \xi \\ \overline{0_{n-r}} \end{bmatrix} > 0 \qquad \text{if } \xi \neq 0$$

Lemma 11. 4. $A > 0 \Rightarrow \det A > 0$ hence

$$C^* A C = I$$

where C is invertible.

$$\Rightarrow |\det(C)|^2 \cdot \det A = 1$$

Proof. Induction: all submatrices A_r are positive definite.

IB $r = 1$**:** $A_1 = [a_{11}]$ is positive definite, because $a_{11} = \det[a_n] > 0$

IS $r \to r + 1$**:** Assume $A_{r-1} > 0$ and $\det A_r > 0$, then $A_{r-1} \mathrel{\widehat{=}} I_{r-1}$

$$\Rightarrow \bigvee_{C_{r-1} \in \mathrm{GL}(r-1, \mathbb{C})} C^*_{r-1} A_{r-1} C_{r-1} = I_{r-1}$$

$$A'_r = \begin{bmatrix} C^*_{r-1} & \\ & 1 \end{bmatrix} \cdot A_r \cdot \begin{bmatrix} C_{r-1} & \\ & 1 \end{bmatrix} = \begin{bmatrix} I_{r-1} & & & a_{1r} \\ & & & a_{2r} \\ & & & \vdots \\ \overline{a_{1r}} & \ldots & \overline{a_{2r}} & a_{rr} \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & & & -a_{ir} \\ & 1 & & \vdots \\ & & \ddots & \vdots \\ & & & -a_{r-1,r} \\ & & & 1 \end{bmatrix}$$

$$C^* A'_r C = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ -\overline{a_{i,r}} & -\overline{a_{2,r}} & \ldots & -\overline{a_{r-1,r}} & 1 \end{bmatrix}.$$

$$\begin{bmatrix} 1 & & & & a_{1,r} \\ & 1 & & & a_{2,r} \\ & & \ddots & & \vdots \\ & & & 1 & a_{r-1,r} \\ \overline{a_{1,r}} & \overline{a_{2,r}} & \ldots & \overline{a_{r-1,r}} & a_{r,r} \end{bmatrix} \cdot \begin{bmatrix} 1 & & & -a_{1,r} \\ & 1 & & \vdots \\ & & \ddots & -a_{r-1,r} \\ & & & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & & & a_{1,r} \\ & \ddots & & \vdots \\ & & 1 & a_{r-1,r} \\ 0 & \ldots & 0 & \underbrace{a_{r,r} - \sum_{j=1}^{r-1} |a_{1,r}|^2}_{=\tilde{a}} \end{bmatrix} \cdot \begin{bmatrix} 1 & & & -a_{1,r} \\ & 1 & & \vdots \\ & & \ddots & -a_{r-1,r} \\ & & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \tilde{a} \end{bmatrix}$$

$$\det A'_r = |\det C_{r-1}|^2 \cdot \det A_r > 0$$

$$\det C^* A'_r C = |\det C|^2 \cdot \det A'_r > 0$$

$$\begin{bmatrix} 1 & & \\ & \ddots & \\ & & \frac{1}{\sqrt{\tilde{a}}} \end{bmatrix} C^* A'_r C \begin{bmatrix} 1 & & \\ & \ddots & \\ & & \frac{1}{\sqrt{\tilde{a}}} \end{bmatrix} = I_r \Rightarrow A_r \mathrel{\widehat{=}} I_r \Rightarrow A_r > 0$$

□

In the following, we will only consider positive definite inner products. Consider $(V, \langle , \rangle)$ and choose a basis $(b_1, \dots, b_n)$.

$$A = [\langle b_i, b_j \rangle]$$

$$\Rightarrow A > 0 \ ?$$

We have already shown: Cauchy-Bunjakowsky-Schwarz:

$$|\ \langle x, y \rangle\ | \le \|X\| \cdot \|Y\|$$

where

$$\|X\| = \sqrt{\langle X, X \rangle}$$

$\Rightarrow$ is a norm.

Definition 18. David Hilbert (1862–1943) $\to$ Hilbert's 23 problems (1900)

1. A vector space V with positive definite scalar product is called
   - Euclidean space ($K = \mathbb{R}$)
   - unitary space ($K = \mathbb{C}$)
   - (pre-)Hilbert space ($\dim = \infty$)
2. An element $v \in V$ is called *normed* if $\|v\| = 1$.

   $$v \neq 0 \Rightarrow \frac{v}{\|v\|} \text{ is normed}$$

3. Let $v, w \neq 0$, then the angle $\angle(v, w)$ is exactly $\arccos \frac{\Re(\langle v, w \rangle)}{\|v\| \cdot \|w\|}$.

   $$\arccos : [-1, 1] \to [0, \pi]$$

4. Two vectors v, w are called *orthogonal* ($v \perp w$) if

   $$\langle v, w \rangle = 0$$

   hence, $v = 0 \vee w = 0 \vee \varphi = \frac{\pi}{2}$.

Theorem 31. In $(V, \langle , \rangle)$ is holds that

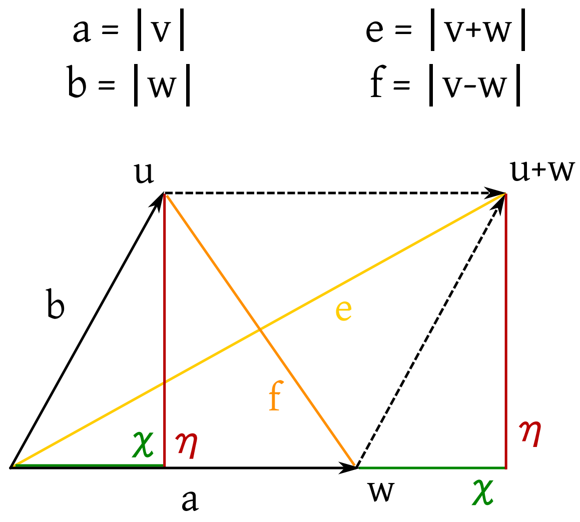

Figure 10: Norm addition illustrated in a parallelogram

1. $\|v + w\|^2 = \|v\|^2 + \|w\|^2 + 2\|v\|\|w\| \cos \varphi$ (Cosine theorem)
2. If $v \perp w$, then $\|v + w\|^2 = \|v\|^2 + \|w\|^2$ (Pythagorean theorem)
3. $\|v + w\|^2 + \|v - w\|^2 = 2(\|v\|^2 + \|w\|^2)$ (parallelogram equation)

   Compare with Figure 10.

$$\xi^2 + \eta^2 = b^2$$
$$(a + \xi)^2 + \eta^2 = e^2$$
$$(a - \xi)^2 + \eta^2 = f^2$$
$$\underbrace{(a + \xi)^2 + (a - \xi)^2}_{2(a^2+\xi^2+\eta^2)=2(a^2+b^2)} + 2\eta^2 = e^2 + f^2$$

Example 17 (Counterexample)**.** $\|x\|_1 = |\,x_1\,| + \dots + |\,x_n\,|$ does not satisfy the third property.

Remark 16. It is possible to show (von Neumann): If a norm satisfies the parallelogram equation, it originates from a scalar product.

Definition 19. Let $(V, \langle,\rangle)$ be a vector space with scalar product. A family $(v_i)_{i\in I} \subseteq V$ is called

orthogonal if $\bigwedge_{i\neq j} \langle v_i, v_j\rangle = 0$

orthonormal if $\bigwedge_{i,j} \langle v_i, v_j\rangle = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$

orthonormal basis if it is a basis and orthonormal

Example 18. $(e_1, \dots, e_n)$ in $\mathbb{K}^n$ is orthonormal basis in regards of the standard scalar product.

1. $\langle e_i, e_j\rangle = \delta_{ij}$
2. $$\int_0^1 \sin(2\pi mx)\sin(2\pi nx)\,dx = \delta_{mn}\cdot 2$$
$$\int_0^1 \sin(2\pi nx)\cos(2\pi nx)\,dx = 0$$
$$\int_0^1 \cos(2\pi mx)\cos(2\pi nx)\,dx = \delta_{mn}\cdot 2$$
$$\{1\} \cup \left\{ \frac{\sin(2\pi nx)}{\sqrt{2}} \;\middle|\; n \in \mathbb{N}\right\} \cup \left\{ \frac{\cos(2\pi nx)}{\sqrt{2}} \;\middle|\; n \in \mathbb{N}\right\}$$
where
$$\langle f, g\rangle = \int_0^1 f(x)g(x)\,dx$$
is orthonormal in $C[0,1]$.
This is the spanned linear subspace.
The result are the so-called trigonometric polynomials.
$$f(x) = \sum_{n=0}^{\infty} a_n \cos(2\pi nx) + \sum_{n=1}^{\infty} b_n \sin(2\pi nx)$$

Theorem 32. Let $(v_i)_{i\in I} \subseteq V$, $v_i \neq 0$.

1. $(v_i)_{i\in I}$ is orthogonal $\Leftrightarrow \left(\frac{v_i}{\|v_i\|}\right)_{i\in I}$ is orthonormal.
2. If $(v_i)_{i\in I}$ is orthogonal, then $(v_i)_{i\in I}$ is linear independent.

Proof. 1. trivial

2. Let $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ and $\lambda_1 v_{i_1} + \ldots + \lambda_k v_{i_k} = 0$, then all $\lambda_j = 0$.
$$\begin{aligned} 0 &= \langle 0, v_{ij}\rangle \\ &= \langle \lambda_1 v_{i_1} + \ldots + \lambda_k v_{i_k}, v_{ij}\rangle \\ &= \lambda_1 \langle v_{i_1}, v_{ij}\rangle + \lambda_2 \langle v_{i_2}, v_{ij}\rangle + \ldots + \lambda_k \langle v_{i_k}, v_{i_j}\rangle \qquad = \lambda_j \|v_{ij}\|^2 \\ &\Rightarrow \lambda_j = 0 \qquad \text{for } j = 1, \dots, k \end{aligned}$$

□

Theorem 33. Let $B = (b_1, \dots, b_n)$ be a orthonormal basis (ONB) of a finite-dimensional vector space V over $\mathbb{K}$. Let $v, w \in V$ with

$$\Phi_B(v) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \qquad \Phi_B(w) = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$$

Then it holds that

1. $\bigwedge_{i\in\{1,\dots,n\}} \lambda_i = \langle v, b_i\rangle$
2. $\langle v, w\rangle = \sum_{i=1}^n \lambda_i \overline{\mu_i}$

Proof. 1. $\langle v, b_i\rangle = \left\langle \sum_{j=1}^n \lambda_j b_j, b_i\right\rangle = \sum_{j=1}^n \lambda_j \left\langle \underbrace{b_j, b_i}_{\delta_{ji}} \right\rangle = \lambda_i$

2. $$\langle v, w\rangle = \Phi_B(v)^t A \overline{\Phi_B(w)} = \Phi_B(v)^t \overline{\Phi_B(w)} = \sum_{i=1}^n \lambda_i \overline{\mu_i}$$

$$a_{ij} = \langle b_i, b_j \rangle = \delta_{ij}$$

□

Definition 20. $(V, \langle , \rangle)$. $M \subseteq V$ be a subset. Then

$$M^\perp := \left\{ v \in V \,\middle|\, \bigwedge_{u \in M} \langle v, u \rangle = 0 \right\}$$

is called *orthogonal complement* of M. For $v \in V$, let $v^\perp := \{v\}^\perp$.

Theorem 34. Let $(V, \langle , \rangle)$ and $M, N \subseteq V$.

1. $M^\perp$ is a subspace.
2. $M \subseteq N \Rightarrow N^\perp \subseteq M^\perp$.
$$(M_1 \cup M_2)^\perp = M_1^\perp \cap M_2^\perp$$
3. $\{0\}^\perp = V$
4. $V^\perp = \{0\}$
5. $M \cap M^\perp \subseteq \{0\}$
6. $M^\perp = \mathcal{L}(M)^\perp$
7. $M \subseteq (M^\perp)^\perp$

Proof. 1.
$$u^\perp = \{v \,|\, \langle v, u \rangle = 0\}$$
$$T_u : \substack{V \to \mathbb{K} \\ v \mapsto \langle v, u \rangle} \text{ is linear}$$
$$\{v \,|\, \langle v, u \rangle = 0\} = \{v \,|\, T_u(v) = 0\} = \ker T_u \text{ is subspace}$$
$$M^\perp = \bigcap_{u \in M} u^\perp \text{ is intersection of subspaces}$$

2.
$$N^\perp = \bigcap_{u \in N} u^\perp \subseteq \bigcap_{u \in M} u^\perp = M^\perp$$
$$(M_1 \cup M_2)^\perp = \bigcap_{u \in M_1 \cup M_2} u^\perp - \bigcap_{u \in M_1} u^\perp \cap \bigcap_{u \in M_2} u^\perp = M_1^\perp \cap M_2^\perp$$

3. trivial
4.
$$V^\perp = V \cap V^\perp = \{0\}$$
5. $v \in M \cap M^\perp \Rightarrow \langle v, v \rangle = 0 \Rightarrow v = 0$
6.
$$\mathcal{L}(M)^\perp \subseteq M^\perp \qquad \text{(because of 2.)}$$
Show that: $M^\perp \subseteq \mathcal{L}(M)^\perp$: Let $v \in M^\perp$, $u \in \mathcal{L}(M)$ Then
$$\exists u_1, \ldots, u_n \in M \exists \lambda_1, \ldots, \lambda_n \in \mathbb{K} : u = \lambda_1 u_1 + \ldots + \lambda_n u_n$$
$$\Rightarrow \langle v, u \rangle = \left\langle v, \sum_{i=1}^n \lambda_i u_i \right\rangle = \sum_{i=1}^n \overline{\lambda_i} \langle v, u_i \rangle = 0$$
7. Show: Let $v \in M$, then $\bigwedge_{u \in M^\perp} \langle v, u \rangle = 0$
$$\bigwedge_{u \in M^\perp} \langle v, u \rangle = \bigwedge_{u \in M^\perp} \langle u, v \rangle = 0$$

□

This lecture took place on 20th of April 2016 (Franz Lehner).

Theorem 35. Let $M^\perp = \left\{ v \,\middle|\, \bigwedge_{u \in M} u \perp v \right\}$ is subspace.

6. $M^\perp = L(M)^\perp$
2. $M \subseteq N \Rightarrow N^\perp \subseteq M^\perp$
3. $0^\perp = V$

4. $V^\perp = \{0\}$
5. $M \cap M^\perp \subseteq \{0\}$

$$M \subseteq (M^\perp)^\perp$$

Corollary 9. If $U \subseteq V$ is a subspace of V, then the sum $U + U^\perp$ is direct.

$$(U+U^\perp)^\perp \overset{6.}{=} (U \cup U^\perp)^\perp = U^\perp \cap (U^\perp)^\perp \overset{5.}{=} \{0\}$$

From $(U+U^\perp)^\perp = \{0\}$, $U + U^\perp = V$ follows only in finite dimensions.

Example 19.

$$V = e^2 = \left\{ (\xi_n)_n \,\middle|\, \sum_{n=1}^{\infty} |\,\xi_n\,|^2 < \infty \right\}$$

$$U = \mathcal{L}((e_i)_{i\in\mathbb{N}}) \neq V = \{(\xi_n)_n \,|\, \xi_n = 0 \text{ for almost all } n\}$$

$$U^\perp = \left\{ x = (\xi_n)_{n\in\mathbb{N}} \,\middle|\, \underbrace{\langle x, e_i\rangle}_{=\xi_i} = 0 \forall i \right\} = \{0\}$$

$$V = (U^\perp)^\perp \neq U \qquad U = U + U^\perp \neq V, U^\perp = \{0\}$$

Practicals:

$$U + U^\perp = V \Leftrightarrow U = (U^\perp)^\perp$$

In the following we always assume: $V = U \dot{+} U^\perp$.

$\rightarrow$ projection: every vector has a unique decomposition.

$$x = u + v$$

$$u \in U \qquad v \in U^\perp \text{ such that } u \perp v$$

Definition 21. Let V be a vector space. A subset $K \subseteq V$ is called convex if

$$\bigwedge_{x,y\in\mathbb{K}} \bigwedge_{\lambda\in[0,1]} x + \lambda(y - x) \in \mathbb{K}$$

$(1-\lambda)x + \lambda y$ is called *convex combination.*

Informally: A set is convex if all elements of the path between two points of the set are inside the set.

Example 20. 1. Let $(V, \|\|)$ be a normed space. Then

$$B(0,1) - \{x \,|\, \|x\| < 1\} \text{ is convex}$$

$$x, y \in B(0,1), \lambda \in [0,1] : \|(1-\lambda)x + \lambda y\| \leq (1-\lambda)\|x\| + \lambda\|y\| < (1-\lambda) + \lambda > 1$$

2. Subspaces are convex.
3. Translations and scalar multiples of convex sets are convex
   - Linear manifolds
   - $B(\times, r)$ is convex.
4. $K \subseteq V$ is convex, $f : V \to W$ is linear $\Rightarrow f(\mathbb{K})$ is convex (the proof is left as an exercise).

Remark 17. What does optimization mean?

Given an audio file with data. We want to approximate these data, but the maximum size of the data is defined. So we optimize the data such that the file size is decreased.

Formally: Find $x \in K$ with $\|X\| = \min$.

Remark 18. Consider $l^1 : \|x\| = |\,x_1\,||\,x_2\,|$. The unit circle is a square rotated by 45°.

If we expand this unit circle to our desired K (a straight line like $f(x) = -x$), the intersection of K and this expanded unit circle yields infinitely many points.

Theorem 36. Let $(V, \langle,\rangle)$ be a vector space with scalar product. $K \subseteq V$ is convex, $x \in V$, $y_0 \in K$.

DFASÄ:

1. $\bigwedge_{y\in K} \|x - y_0\| \leq \|x - y\|$
2. $\bigwedge_{y\in K} \Re\,\langle x - y_0, y - y_0\rangle \leq 0$
3. $\bigwedge_{y\in K\setminus\{y_0\}} \|x - y_0\| < \|x - y\|$

Remark 19. If K is a linear manifold, then (2.) is equivalent to:

2'. $\bigwedge_{y\in K} \langle x - y_0, y - y_0\rangle = 0$

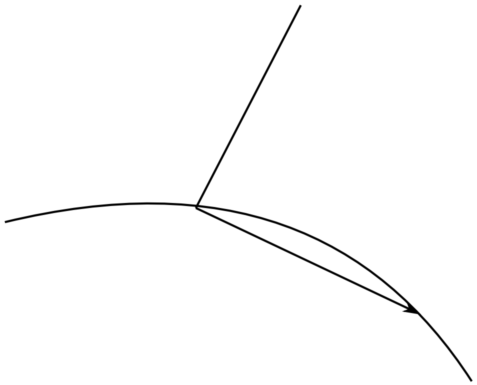

Figure 11: Norm

Proof. **1.** $\to$ **2.** Let $y \in K$. $0 < \varepsilon < 1$. Compare with Figure 11.

$$\Rightarrow y_\varepsilon = (1-\varepsilon)y_0 + \varepsilon y_1 \in K$$

$$\Rightarrow \|x - y_0\| \leq \|x - y_\varepsilon\|$$
$$y_\varepsilon = y_0 + \varepsilon(y - y_0)$$

$$\begin{aligned}
0 &\leq \|x - y_\varepsilon\|^2 - \|x - y_0\|^2 \\
&= \|x - y_0 - \varepsilon(y - y_0)\|^2 - \|x - y_0\|^2 \\
&= \|x - y_0\|^2 + \varepsilon^2\|y - y_0\|^2 - 2\Re\langle x - y_0, y - y_0\rangle\varepsilon - \|x - y_0\|^2 \\
&= \varepsilon(\varepsilon\|y - y_0\|^2 - 2\Re\langle x - y_0, y - y_0\rangle) \\
\varepsilon \to 0 &\Rightarrow -2\Re\langle x - y_0, y - y_0\rangle \geq 0
\end{aligned}$$

If $\Re x - y_0, y - y_0 > 0$, then $y \neq y_0$. Choose

$$\varepsilon < \frac{2\Re\langle x - y_0, y - y_0\rangle}{\|y - y_0\|}$$

This leads to a contradiction.

2. $\to$ **3.** Let $y \in K \setminus \{y_0\}$.

$$\begin{aligned}
\|x - y\|^2 &= \|x - y_0 - (y - y_0)\|^2 \\
&= \|x - y_0\|^2 + \|y_0 - y\|^2 - 2\Re\langle x - y_0, y - y_0\rangle \\
&\geq \|x - y_0^2\| + \|y_0 - y\|^2 \\
&> \|x - y_0\|^2
\end{aligned}$$

3. $\to$ **1.** trivial

If $K = U$ is a subspace.

2.

$$y - y_0 \in U \Leftrightarrow y \in U$$

$$\begin{aligned}
&\bigwedge_{y \in U} \Re\left\langle x - y_0, \underbrace{y - y_0}_{=: z \in U}\right\rangle \leq 0 \\
\Leftrightarrow &\bigwedge_{z \in U} \Re\langle x - y_0, z\rangle \leq 0 \\
\Rightarrow &\bigwedge_{z \in U} \Re\langle x - y_0, -z\rangle \leq 0 \\
\Leftrightarrow &\bigwedge_{z \in U} \Re\langle x - y_0, z\rangle \geq 0 \\
\Rightarrow &\bigwedge_{z \in U} \Re\langle x - y_0, z\rangle = 0 \\
\Rightarrow &\bigwedge_{z \in U} \Re\langle x - y_0, iz\rangle = 0 \\
\Rightarrow &\bigwedge_{z \in U} \Re(-i\,\langle x - y_0, z\rangle) = 0
\end{aligned}$$

$$\Re(-i(a+ib)) = b$$
$$\Rightarrow \bigwedge_{z\in U} \Im \langle x - y_0, z\rangle = 0$$
$$\Rightarrow \bigwedge_{z\in U} \langle x - y_0, z\rangle = 0 \Rightarrow x - y_0 \in U^\perp$$

□

Corollary 10. Let $(V, \langle,\rangle)$ be a vector space with a scalar product.

1. If $K \subseteq V$ is convex, then the optimization problem

$$\begin{cases} \|x-y\| = \min! \\ y \in K \end{cases}$$

has at most one solution.

2. If $U \subseteq V$ is a subspace, $x \in V$, then there exists at most one point $y_0 \in U$ such that $x - y_0 \in U^\perp$.

$\Rightarrow$ the sum $U + U^\perp$ is direct.

Definition 22. Let $(V, \langle,\rangle)$ is a vector space with scalar product. Let $U \subseteq V$ a subspace with $V = U \dot{+} U^\perp$.

Let's recognize that

$$V = U \dot{+} W$$
$$\bigwedge_{x} \bigvee_{\substack{u\in U\\ w\in W}} = u + w$$

Then $\pi_U : V \to V$ and $\pi_{U^\perp} : V \to V$ such that

$$\bigwedge_{x\in V} \pi_U(x) \in U \wedge \pi_{U^\perp}(x) \in U^\perp$$

are called *orthogonal projections* to U and $U^\perp$.

Compare with Figure 12.

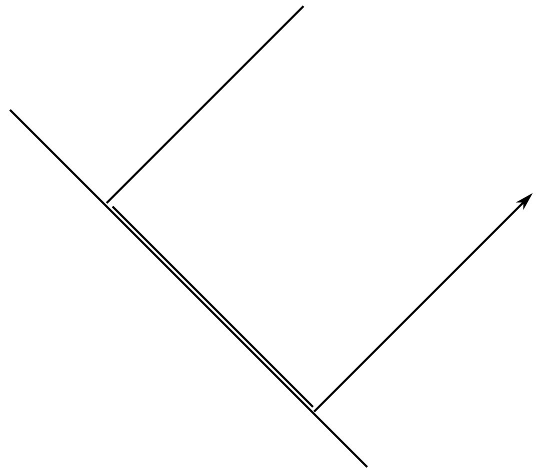

Figure 12: Orthogonal projections

Theorem 37 (Revision of direct sums of vector spaces)**.**

1. $x \in U \Leftrightarrow \pi_U(x) = x \Leftrightarrow \pi_{U^\perp}(x) = 0$
2. $x \in U^\perp \Leftrightarrow \pi_U(x) = 0 \Leftrightarrow \pi_{U^\perp}(x) = x$
3. $\pi_{U^\perp} = \mathrm{id} - \pi_U$
4. $\pi_U \circ \pi_U = \pi_U$
5. π_U is linear

Theorem 38. Let $V = U \dot{+} U^\perp$.

1. $\bigwedge_{x,y\in V} \langle x, \pi_U(y)\rangle = \langle \pi_U(x), y\rangle = \langle \pi_U(x), \pi_U(y)\rangle$
2. $\bigwedge_{x\in V} \|\pi_U(x)\| \leq \|x\|$ and $\|\pi_U(x)\| = \|X\| \Leftrightarrow x \in U$

Proof. 1.

$$
\begin{aligned}
x &= \pi_U(x) + \pi^{U^\perp}(x) \\
y &= \pi_U(y) + \pi_{U^\perp}(y) \\
\langle x, \pi_U(y)\rangle &= \langle \pi_U(x) + \pi_{U^\perp}(x), \pi_U(y)\rangle \\
&= \langle \pi_U(x), \pi_U(y)\rangle + \left\langle \underbrace{\pi_{U^\perp}(x)}_{\in U^\perp}, \underbrace{\pi_U(y)}_{\in U} \right\rangle \\
\langle \pi_U(x), y\rangle &= \langle \pi_U(x), \pi_U(y)\rangle
\end{aligned}
$$

2.

$$
\begin{aligned}
\|x\|^2 &= \|\pi_U(x) + \pi_{U^\perp}(x)\|^2 \\
\text{Pythagorean theorem} &= \|\pi_U(x)\|^2 + \|\pi_{U^\perp}(x)\|^2 \\
&\geq \|\pi_U(x)\|^2 \\
\text{equality} &\Leftrightarrow \pi_{U^\perp}(x) = 0 \Leftrightarrow x \in U
\end{aligned}
$$

□

Definition 23. Jørgen Pedersen Gram (1850–1916)

Let $(V, \langle,\rangle)$ is a vector space with scalar product. Let $v_1, \dots, v_m \in V$.

Then the matrix is called

$$
\mathrm{Gram}(v_1, \dots, v_m) := \begin{bmatrix} \langle v_1, v_1\rangle & \langle v_1, v_2\rangle & \dots & \langle v_1, v_m\rangle \\ \langle v_2, v_1\rangle & \langle v_2, v_2\rangle & \dots & \langle v_2, v_m\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle v_m, v_1\rangle & \langle v_m, v_2\rangle & \dots & \langle v_m, v_m\rangle \end{bmatrix} \in \mathbb{K}^{m\times m}
$$

Gram's matrix of tuple $(v_1, \dots, v_m)$

Remark 20.

$$
\begin{aligned}
V &= \mathbb{R}^n \qquad (\mathbb{C}^n) \\
\langle v_i, v_j\rangle &= v_i^t v_j \\
\rightsquigarrow G &= V^t \overline{V} \\
V &= \begin{pmatrix} V_1 & V_2 & \dots & V_m \\ \vdots & \vdots & & \vdots \end{pmatrix}
\end{aligned}
$$

Theorem 39. Let $(V, \langle,\rangle)$ be a vector space with a scalar product. $v_1, \dots, v_m \in V$.

1. $G = \mathrm{Gram}(v_1, \dots, v_m)$ is hermitian and positive semidefinite. Furthermore it holds that
$$\xi^t \cdot G \cdot \overline{\xi} = \|\sum_{i=1}^m \xi_i v_i\|^2$$
2. $$\xi \in \ker(G) \Leftrightarrow \sum_{i=1}^m \overline{\xi_i} v_i = 0$$
3. G is positive definite iff G is regular iff $v_1, \dots, v_m$ are linear independent.

Proof. 1.
$$g_{ij} = \langle v_i, v_j\rangle = \overline{\langle v_j, v_i\rangle} = \overline{g_{ji}}$$

$\Rightarrow$ G is Hermitian.

$$
\begin{aligned}
\xi^t G \overline{\xi} &= \sum_{i,j=1}^m \xi_i \langle v_i, v_j\rangle \overline{\xi_j} \\
&= \left\langle \sum_{i=1}^m \xi_i v_i, \sum_{j=1}^m \xi_j v_j \right\rangle \\
&= \|\sum_{i=1}^m \xi_i v_i\|^2
\end{aligned}
$$

2. $\Rightarrow$ Let $\xi \in \ker G$.

$$G \cdot \xi = 0 \Rightarrow \underbrace{\overline{\xi}^t G \xi}_{=\|\sum \overline{\xi_i v_i}\|^2} = 0$$

$$\Rightarrow \sum_{i=1}^m \overline{\xi_i} v_i = 0$$

$\Leftarrow$ Let $\sum_{i=1}^m \xi_i v_i = 0$.

$$(G \cdot \xi)_i = \sum_{j=1}^m \langle v_i, v_j \rangle \xi_j = \left\langle v_i, \underbrace{\sum_{j=1}^m \overline{\xi_j}\, v_j}_{=0} \right\rangle = 0$$

holds for all $i = 1, \dots, m$.

$$\Rightarrow G \cdot \xi = 0 \Rightarrow \xi \in \ker G$$

3.

$$\begin{aligned} G > 0 &\Leftrightarrow \xi^t G \overline{\xi} > 0 \quad \forall \xi \neq 0 \\ &\Leftrightarrow \left\|\sum \xi_i v_i\right\|^2 > 0 \quad \forall \xi \neq 0 \\ &\Leftrightarrow \sum \xi_i v_i \neq 0 \quad \forall \xi \neq 0 \\ &\Leftrightarrow v_1, \dots, v_m \text{ is linear independent} \\ &\Leftrightarrow \ker G = \{0\} \\ &\Leftrightarrow G \text{ regular} \end{aligned}$$

□

This lecture took place on 25th of April 2016 (Franz Lehner).

Revision:

- Approximation of a x (outside) in a convex set by computing the orthogonal line from x to the border of the convex set intersecting the set at y_0 ($\|x-y_0\|$ min, $y_0 \in K$)

$$\bigwedge_{y \in K} \Re\langle x - y_0, y - y_0 \rangle \leq 0$$

- $u = \pi_u(x)$ is the unique element $u \in U$ such that $x - u \in U^\perp$. U is defined such that $U \dot{+} U^\perp = V \Leftrightarrow U = (U^\perp)^\perp$. π_U is linear.

Theorem 40. Consider $v_1, \dots, v_m \in V$.

$$\operatorname{Gram}(v_1, \dots, v_m) = [\langle v_2, v_1 \rangle]_{i,j=1}^m$$

$$\xi^t G \xi = \left\|\sum \xi_i v_i\right\|^2$$

G is positive definite iff $v_1, \dots, v_m$ are linear independent.

Theorem 41. Let $(V, \langle,\rangle)$ be a vector space with scalar product. $U \subseteq V$ is subspace with $\dim U < \infty$. $(u_1, \dots, u_n)$ is basis of U. $G = \operatorname{Gram}(u_1, \dots, u_m)$ is positive definite and regular. Then the orthogonal projection to U for $x \in V$ is given by

$$\pi_U(x) = \sum_{j=1}^m \eta_j u_j$$

where $\overline{\eta} = G^{-1}\xi$.

$$\xi = \begin{pmatrix} \langle x, u_1 \rangle \\ \vdots \\ \langle x, u_m \rangle \end{pmatrix}$$

(wrt. claiming $u_1, \dots, u_m$ is orthogonal normal basis and Gram matrix provides correction to achieve that)

Proof. Let $u = \sum_{j=1}^m \eta_j u_j$. Because

$$\bigwedge_{y \in U} \langle x - y_0, y - y_0 \rangle = 0$$

holds, it holds that $x - u \in U^\perp = \underbrace{\{u_1, \dots, u_m\}}_{\text{basis of } U'}^{\perp}$.

Hence we need to show:

$$\langle u_i, u \rangle = \langle u_i, x \rangle \qquad \text{for } i = 1, \dots, m$$

$$\begin{aligned} \langle u_i, u \rangle = \left\langle u_i, \sum_{j=1}^m \eta_j u_j \right\rangle &= \sum_{j=1}^m \underbrace{\langle u_i, u_j \rangle}_{=G_{ij}} \overline{\eta}_j \\ = (G\overline{\eta})_i = \overline{\xi}_i &= \overline{\langle x_i, u_i \rangle} = \langle u_i, x \rangle \\ &\Rightarrow \langle x - u, u_i \rangle = 0 \quad \forall i \\ &\Rightarrow x - u \in U^\perp \end{aligned}$$

□

Example 21. Determine the polynomial $p(x)$ of degree ≤ 2 such that

$$\int_0^1 |\, t^3 - p(t)\,|^2 \, dt = \text{ min!}$$

where $V = \mathbb{R}[x]$ and $K = U = \mathbb{R}_2[x] = \mathcal{L}(1, x, x^2)$.

$$\langle f, g \rangle = \int_0^1 f(t)g(t)\, dt$$

Find: Orthogonal projection of x^3 to $\mathcal{L}(1, x, x^2)$

Step 1: Gram matrix G

$$G_{ij} = \int_0^1 t^{i-1}t^{j-1}\, dt = \int_0^1 t^{i+j-2}\, dt = \frac{1}{i+j-1}$$

$$G = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix} \qquad \text{“Hilbert matrix”}$$

Hankel matrix: $a_{ij}a_{i'j'}$ if $i + j = i' + j'$

$$\begin{pmatrix} h_1 & h_2 & h_3 & \dots \\ h_2 & h_3 & h_4 & \dots \\ h_3 & h_4 & h_5 & \dots \vdots \end{pmatrix}$$

“momentum problem”

Our solution:

$$p(t) = \sum_{i=1}^{3} \eta_i t^{i-1}$$

$$\eta = G^{-1}\xi \qquad \xi_i = \langle x^3, u_i \rangle = \int_0^1 t^3 t^{i-1}\, dt = \frac{1}{3+i}$$

$$\eta = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{4} \\ \frac{1}{5} \\ \frac{1}{6} \end{bmatrix} = \begin{bmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{bmatrix} \begin{bmatrix} \frac{1}{4} \\ \frac{1}{5} \\ \frac{1}{6} \end{bmatrix} = \begin{bmatrix} -\frac{1}{20} \\ -\frac{3}{5} \\ \frac{3}{2} \end{bmatrix}$$

Solution:

$$p(x) = \frac{1}{20} - \frac{3}{5}x + \frac{3}{2}x^2$$

$n \times n$ Hilbert matrix:

$$\det H_n = \frac{c_n^4}{c_{2n}} \sim \frac{(2\pi)\overbrace{l_n}^{\text{Stirling}}}{4^{n^2}\sqrt[4]{n}}$$

$$c_n = \prod_{i=1}^{n-1} i!$$

“Almost” not invertible. So this matrix is actually a good test matrix for numerical algorithms.

Corollary 11. Let $U \subseteq V$ be a subspace and $(u_1, \dots, u_m)$ be an orthonormal basis. Then it holds that

- $\bigwedge_{v \in V} \pi_u(v) = \sum_{i=1}^{m} \langle v, u_i \rangle\, u_i$
- $\bigwedge_{v \in V} \sum_{i=1}^{m} |\, \langle v, u_i \rangle\, |^2 \leq \|v\|^2$, “Bessel’s inequality”
  $\bigwedge_{v \in V} \sum_{i=1}^{m} |\, \langle v, u_i \rangle\, |^2 = \|v\|^2 \Leftrightarrow v \in U$, “Parseval’s identity”

Proof. Immediate. $G = I$, hence $\eta_i = \xi_i$. □

Remark 21. So if $u_1, \dots, u_m$ is an orthonormal basis of U, then we can immediately determine $\pi_U(v)$. Can we fine an orthonormal basis?

Theorem 42 (Gram–Schmidt process)**.** Given $v_1, \dots, v_m$ is linear independent. Determine orthonormal system $u_1, \dots, u_m$ such that $\mathcal{L}(u_1, \dots, u_m) = \mathcal{L}(v_1, \dots, v_m)$.

Let $(V, \langle , \rangle)$ be vector space with scalar product. Let $(v_1, \dots, v_n)$ be linear independent.

Then there exists a orthonormal system $(u_1, \dots, u_m) \subseteq V$ such that $\mathcal{L}(u_1, \dots, u_m) = \mathcal{L}(v_1, \dots, v_m)$. Find $u \in U_k$ such that $v_{k+1} - u \in U_k^{\perp}$.

Inductively,

$$u_1 = \frac{v_1}{\|v_1\|} \qquad U_k = \mathcal{L}(v_1, \dots, v_k) = \mathcal{L}(u_1, \dots, u_k)$$

and for $k = 2, \dots, m$: Let

$$\tilde{u_k} = v_k - \underbrace{\sum_{j=1}^{k-1} \langle v_k, u_j \rangle u_j}_{\pi_{U_{k-1}}(v_k)}$$

$$U_k = \frac{\tilde{U_k}}{\|\tilde{U_k}\|}$$

Proof. **Case** $k = 1$ $v_1 \neq 0$ because they are linear independent.

$$\rightarrow u_1 = \frac{v_1}{\|v_1\|}$$

Case $k - 1 \rightarrow k$ $u_1, \dots, u_{k-1}$ is orthonormal sysm with $\mathcal{L}(u_1, \dots, u_{k-1}) = \mathcal{L}(v_1, \dots, v_{k-1})$. $v_k \notin \mathcal{L}(u_1, \dots, u_{k-1}) =: U_{k-1}$.

$$\begin{aligned}
\tilde{u_k} &= v_k - \sum_{j=1}^{k-1} \langle v_k, u_j \rangle u_j \\
&\overset{\text{Theorem ??}}{=} v_k - \pi_{U_{k-1}}(v_k) \in U_{k-1}^{\perp} \\
&\Rightarrow \tilde{u_k} \perp u_1, \dots, u_{k-1}, \tilde{u_k} \neq 0 \text{ because } v_k \text{ is linear independent of } u_1, \dots, u_{k-1} \\
u_k &= \frac{\tilde{u_k}}{\|\tilde{u_k}\|} \rightarrow (u_1, \dots, u_k) \text{ is an orthonormal system}
\end{aligned}$$

Immediate:

$$\tilde{u_k} \in \mathcal{L}(u_1, \dots, u_{k-1}, u_k) = \mathcal{L}(v_1, \dots, v_k)$$

$$v_k = \tilde{u_k} + \sum \lambda_j u_j \in \mathcal{L}(u_1, \dots, u_{k-1}, \tilde{u_k}) = \mathcal{L}(u_1, \dots, u_k)$$

□

Example 22. Let $V = \mathbb{R}^3$ with an inner product.

$$\langle x, y \rangle = x^t A y \text{ with } A = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 2 \end{pmatrix}$$

Find an orthonormal basis.

1. We orthogonalize the canonical basis e_1, e_2 and e_3.

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \|v_1\| = v_1^t A v_1 = 1$$

$$u_1 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}$$

$$v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \tilde{u_2} = v_2 - \overbrace{\langle v_2, u_1 \rangle}^{=a_{21}} u_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - (-1) \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$\langle u_1, \tilde{u_2} \rangle = 0$$

$$\|\tilde{u}_2\|^2 = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} A \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 1 - 1 - 1 + 3 = 2$$

$$u_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\begin{aligned}
\tilde{u}_3 &= v_3 - \langle v_3, u_1 \rangle u_1 - \frac{\langle v_3, \tilde{u}_2 \rangle \tilde{u}_2}{\|\tilde{u}_2\|^2} \\
&= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - 1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - 0 \cdot \tilde{u}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}
\end{aligned}$$

Solution:

$$\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, TODO \right)$$

Theorem 43 (Alternative approach to build an orthogonal projection (in $\mathcal{C}^n$))**.** Determine orthonormal basis $(u_1, \dots, u_m)$ of subspace $U \subseteq \mathcal{C}^n$, then

$$P = \sum_{i=1}^{m} u_i u_i^* = \sum_{i=1}^{m} u_i \overline{u_i^t}$$

Hence given $v \in \mathbb{C}^n$,

$$P \cdot v = \sum_{i=1}^{m} u_i \underbrace{u_i^* v}_{=\langle v, u_i \rangle}$$

Example 23. Considering Exercise 21 again.

$$V = \mathbb{R}[x] \qquad U = \mathcal{L}(1, x, x^2)$$

Orthonormal basis:

$$\|v_1\|^2 = \int_0^1 1^2 \, dt = 1$$

$$u_1 = 1$$

$$\begin{aligned}\tilde{u}_2 &= v_2 - \langle v_2, u_1 \rangle \, u_1 \\ &= \left(\langle x, 1 \rangle = \int_0^1 t \cdot 1 \, dt = \frac{1}{2} \right) \\ &= x - \langle x, 1 \rangle \cdot 1 \\ &= x - \frac{1}{2}\end{aligned}$$

$$\|\tilde{u}_2\|^2 = \int_0^1 (t - \frac{1}{2})^2 \, dt = \frac{(t - \frac{1}{2})^3}{3}\bigg|_0^1 = \frac{1}{12}$$

$$u_2 = \sqrt{12} \cdot \left(x - \frac{1}{2} \right)$$

$$\tilde{u}_3 = x^2 - \langle x^2, 1 \rangle \cdot 1 - \left\langle x^2, x - \frac{1}{2} \right\rangle \left(x - \frac{1}{2} \right) \cdot 12 = x^2 - x + \frac{1}{6}$$

$$\|\tilde{u}_3\|^2 = \int_0^1 \left(t^2 - t + \frac{1}{6} \right)^2 dt = \frac{1}{180}$$

$$\pi_U(x^3) = \langle x^3, 1 \rangle \cdot 1 + \left\langle x^3, x - \frac{1}{2} \right\rangle \left(x - \frac{1}{2} \right) \cdot 12 + \left\langle x^3, x^2 - x + \frac{1}{6} \right\rangle \left(x^2 - x + \frac{1}{6} \right) \cdot 180$$

Remark 22. Let V be a vector space with a scalar product. $\dim V < \infty$.

- Every subspace U has a unique orthogonal complement $U^\perp$ such that $V = U \oplus U^\perp$.
- $\pi_U : V \to U$ is an orthogonal projection with the property:

$$\|\pi_U(v)\| \leq \|v\|$$

- We can determine an orthonormal basis of U, then

$$\pi_U(v) = \sum_{i=1}^{n} \langle v, u_i \rangle \, u_i$$

$\Rightarrow$ Bessel's inequality

Remark 23 (Revision)**.** $\operatorname{Hom}(V, \mathbb{K}) = V^*$ dual space. The map

$$V \times V^* \to \mathbb{K}$$

$$(x, f) \mapsto \langle f, x \rangle = f(x)$$

is bilinear.

Consider scalar product:

$$V \times V \to \mathbb{K}$$
$$(x, y) \mapsto \langle x, y \rangle$$

is sesquilinear.

Every $y \in V$ induces $f_y \in V^*$, $f_y(x) = \langle x, y \rangle$.

Consider:

$$V \to V^*$$
$$y \mapsto f_y$$

is an anti-linear embedding.

Why? Injectivity: $f_y = 0 \Rightarrow f_y(x) = 0 \quad \forall x$. Especially, $f_y(y) = \langle y, y \rangle = \|y\|^2 = 0 \Rightarrow y = 0$.

Theorem 44 (Riesz representation theorem)**.** Frigyes Riesz (1880–1956) Let $(V, \langle,\rangle)$ is vector space with scalar product. $\dim V < \infty$.

Then the map

$$V \to V^*$$

$$y \mapsto f_y : V \to \mathbb{K} \text{ with } x \mapsto \langle x, y\rangle$$

is an antilinear isomorphism.

Proof. Injectivity has already been shown. □

This lecture took place on 27th of April 2016 (Franz Lehner).

Remark 24 (Revision)**.**

$$U \subseteq V$$

How to determine $\pi_U : V \to U$? Gram-Schmidt process or construct orthonormal basis $(u_1, u_2, \dots, u_n)$ with $\pi_u(v) = \sum_{i=1}^m \langle v_i u_i\rangle u_i$

$$\|\pi_u(v)\| \leq \|v\|$$

$$\sum_{i=1}^{w} | \langle v_i, u_i\rangle |^2 \leq \|v\|^2$$

Bessel or equivalently $v \in U$ Parseval (*1755)

Theorem 45.

$$V^* = \operatorname{Hom}(V, \mathbb{K}) \overset{\sim}{-} V$$

The map $V \to V^*$ with $y \mapsto f_y : V \to \mathbb{K}$ with $x \mapsto \langle x, y\rangle$ is anti-linear[1] and bijective.

$$f_{\lambda y + \mu z} = \overline{\lambda} f_y + \overline{\mu} f_z$$

Remark 25. We have already shown that $y \mapsto f_y$ is injective.

Remark 26 (Surjectivity)**.** Given $f \in V^*$. Find $y \in V$ such that $f = f_y$.

[1] In $\mathbb{R}$ linear, in $\mathbb{C}$ it does make a difference

Theorem 46. Let $(u_1, \dots, u_n)$ be an orthonormal basis of V. Let $y = \sum_{i=1}^n \overline{f(u_i)} u_i$.

$$\begin{aligned} f_y(x) = \langle x, y\rangle &= \left\langle x, \sum_{i=1}^n \overline{f(u_i)} u_i \right\rangle \\ &= \sum_{i=1}^n f(u_i) \langle x, u_i\rangle \\ &= f\left(\sum_{i=1}^n \langle x, u_i\rangle u_i\right) \\ &= f(x) \\ \Rightarrow f_y &= f \end{aligned}$$

Theorem 47 (Second Riesz representation theorem)**.**

$$C[0,1]^* = \text{ space of measures in } [0,1]$$

Proof. The proof would take about 4 months of lectures. □

Example 24.

- $$\begin{aligned} v = 0 &\Leftrightarrow \bigwedge_{w \in V} \langle v, w\rangle = 0 \\ &\Leftrightarrow \bigwedge_{f \in V^*} f(v) = 0 \end{aligned}$$

  or

  $$v_1 = v_2 \Leftrightarrow \bigwedge_{w \in V} \langle v_1, w\rangle = \langle v_2, w\rangle$$

- $$\begin{aligned} \|v\| = \sqrt{\langle v, v\rangle} &= \sup\{| \langle v, w\rangle | \mid \|w\| \leq 1\} \\ &= \sup\{| f(v) | \mid f \in V^*, \|f\| \leq 1\} \end{aligned}$$

Remark 27 (Reminder)**.**

$$f \in \operatorname{Hom}(V, W) \qquad f : V \xrightarrow{f} W \xrightarrow{w^*} \mathbb{K}$$

$$f^T : W^* \to V^*$$
$$f^T(w^*) = w^* \circ f \quad \in V^*$$
$$\overset{w^*}{\underset{\mathbb{K}}{\downarrow}} \to v^*$$
$$V^* \cong V$$

Theorem 48 (Theorem with a definition)**.** Let $(V, \langle \cdot, \cdot \rangle_V)$ and $(W, \langle \cdot, \cdot \rangle_W)$ with vector spaces with scalar product. Let $\dim V < \infty$ and $\dim W < \infty$.

Let $g \in \operatorname{Hom}(V, W)$.

1. For every $w \in W$ the map
$$v \mapsto \underbrace{\langle g(v), w \rangle}_{\text{linear}} = f_W \circ g(v)$$

2. We get a unique $u \in V$,
$$\bigwedge_{w \in W} \dot{\bigvee}_{u \in V} \bigwedge_{v \in V} \langle g(v), w \rangle = \langle v, u \rangle$$
We denote $g^*(w) := u$.

3. The map $g^* : W \to V$ is linear and is called *conjugate map*.

4. The map $\operatorname{Hom}(V, W) \to \operatorname{Hom}(W, V)$ is an antilinear involution, hence $g^{**} = g$.

Proof. 1. $\langle g(v), w \rangle = f_w \circ g(v)$ is linear $\Rightarrow f_w \circ g \in V^*$.

2. From Riesz' representation theorem it follows that
$$\dot{\bigvee}_{u \in V} f_w \circ g = f_u$$
$$\bigwedge_{v \in V} \langle g(v), w \rangle = \langle v, u \rangle =: \langle v, g^*(w) \rangle$$
if $\bigwedge_{v \in V} \langle v, u_1 \rangle = \langle v, u_2 \rangle$ and by Exercise 24 $u_1 = u_2$.

3. g^* is linear.
$$\bigwedge_{v \in V} \langle v, g^*(\lambda w_1 + \mu w_2) \rangle \overset{!}{=} \langle v, \lambda g^*(w_1) + \mu g^*(w_2) \rangle$$
$$\begin{aligned}
\langle v, g^*(\lambda w_1 + \mu w_2) \rangle &= \langle g(v), \lambda w_1 + \mu w_2 \rangle \\
&= \overline{\lambda} \langle g(v), w_1 \rangle + \overline{\mu} \langle g(v), w_2 \rangle \\
&= \overline{\lambda} \langle v, g^*(w_1) \rangle + \overline{\mu} \langle v, g^*(w_2) \rangle \\
&= \langle v, \lambda g^*(w_1) \rangle + \langle v, \mu g^*(w_2) \rangle \\
&= \langle v, \lambda g^*(w_1) + \mu g^*(w_2) \rangle
\end{aligned}$$

4. Consider $\operatorname{Hom}(V, W) \to \operatorname{Hom}(W, V)$ and $g \mapsto g^*$ is antilinear. We need to show: $(\lambda g + \mu h)^* = \overline{\lambda} g^* + \overline{\mu} h^*$.
We need to show:
$$\bigwedge_{v \in V} \bigwedge_{w \in W} \langle v, (\lambda g + \mu h)^*(w) \rangle = \left\langle v_1 (\overline{\lambda} g^* + \overline{mu} g^*)(w) \right\rangle$$
$$\begin{aligned}
\langle v, (\lambda g + \mu h)^*(w) \rangle &= \langle (\lambda g + \mu h)(v), w \rangle \\
&= \lambda \langle g(v), w \rangle + \mu \langle h(v), w \rangle \\
&= \lambda \langle v, g^*(w) \rangle + \mu \langle v, h^*(w) \rangle \\
&= \left\langle v, \overline{\lambda} g^*(w) + \overline{\mu} h^*(w) \right\rangle \\
&= \left\langle v, (\overline{\lambda} g^* + \overline{\mu} h^*)(w) \right\rangle
\end{aligned}$$
Remember, $g^* : W \to V$ and $g^{**} : V \to W$.
$$\begin{aligned}
\langle g^{**}(v), w \rangle &= \overline{\langle w, g^{**}(v) \rangle} \\
&= \overline{\langle g^*(w), v \rangle} \\
&= \langle v, g^*(w) \rangle \\
&= \langle g(v), w \rangle
\end{aligned}$$
$$\overset{\text{Exercise } 24}{\Rightarrow} \bigwedge_{v \in V} g^{**}(v) = g(v)$$
$$\Rightarrow g^* * = g$$

$\square$

Remark 28. If $\dim V = \infty$, then the conjugate map must not exist! Consider this example:

$$V = \mathbb{R}[x]$$

$$\langle f, g \rangle = \int_0^1 f(t)g(t)\,dt$$

Polynomials are differentiable.

$$D : V \to V$$

$$p(x) \mapsto p'(x)$$

$D^*(x)$ does not exist!

$$\left\langle x^n, \underbrace{D^*(x^1)}_{=:q(x)} \right\rangle = \langle Dx^n, x \rangle = \int_0^1 nt^{n-1}t\,dt = \frac{n}{n+1}$$

$$\sup_{t\in[0,1]} |\, q(t)\,| \leq M$$

$$= \int_0^1 t^n q(t)\,dt \Rightarrow |\, \langle x^n, D^*x \rangle \,| \leq \int_0^1 t^n |\, q(t)\,|\,dt \leq \frac{M}{n+1}$$

This is a contradiction.

How do we fix this? We use the wrong space. $\mathbb{R}[x]$ is replaced with $\mathcal{L}^2$, which is not differentiable. This is the major topic for Numerics / Computational Mathematics and discussed further here.

Theorem 49. Let $B \subseteq V$ and $C \subseteq W$ be orthonormal bases. $B = (b_1, \ldots, b_m)$ and $C = (c_1, \ldots, c_n)$.

$$f \in \operatorname{Hom}(V, W) \qquad \to f^* \in \operatorname{Hom}(W, V)$$

$$\Phi_B^C(f^*) = \Phi_C^B(f)^* = \overline{\Phi_C^B(f)^t}$$

The most difficult part about this theorem is understanding it. Proving is trivial.

Proof.

$$A = \Phi_C^B(f) \qquad (a)_{ij} = \Phi_C(f(b_j))_i \overset{\text{orthonormal}}{=} \langle f(b_j), c_i \rangle$$

$$\tilde{A} = \Phi_B^C(f^*)$$

$$\tilde{a}_{ij} = \Phi_B(f^*(c_j))_i = \langle f^*(c_j), b_i \rangle = \overline{\langle b_i \rangle, f^*(c_j)}$$

$$= \overline{\langle f(b_i), c_j \rangle} = \overline{a_{ji}} \Rightarrow \tilde{A} = \overline{A^t}$$

□

Theorem 50. Properties of the conjugate:

$$U \xrightarrow{f} V \xrightarrow{g} W \qquad \text{and} \qquad W \xrightarrow{g^*} V \xrightarrow{f^*} U$$

Consider $\dim U < \infty$, $\dim V < \infty$ and $\dim W < \infty$ (otherwise kernelspace not necessarily closed).

- $(g \circ f)^* = f^* \circ g^*$
- $f^{**} = f$
- $\ker f = \operatorname{im}(f^*)^\perp$
- $\operatorname{im}(f) = \ker(f^*)^\perp$
- f is injective $\Leftrightarrow$ f^* is surjective
- f is surjective $\Leftrightarrow$ f^* is injective

Proof. 1.

$$\bigwedge_{u\in U} \bigwedge_{w\in W} \langle u, (g \circ g)^*(w) \rangle_U = \langle g(f(u)), w \rangle_W$$
$$= \langle f(u), g^*(w) \rangle_V$$
$$= \langle u, f^*(g^*(w)) \rangle_U$$

From Example 24 it follows that $(g \circ f)^* = f^* \circ g^*$.

2. We have already shown that.

3. Let $u \in \ker(f)$. We need to show that:

$$\bigwedge_{v \in V} u + f^*(v)$$

$$\langle u, f^*(v)\rangle = \left\langle \underbrace{f(u)}_{=0}, v \right\rangle = 0$$

First we show "$\supseteq$": Let $u \in \operatorname{im}(f^*)^\perp$. Show that $f(u) = 0$. From Example 24 it follows that it suffices to show $\langle f(u), v\rangle = 0 \quad \forall v \in V$.

$$\langle f(u), v\rangle - \langle u, f^*(v)\rangle = 0$$

because $u \perp \operatorname{im}(f^*)$.

4. Apply the third property to f^*.

$$\ker(f^*) = \operatorname{im}(f^{**})^\perp = \operatorname{im}(f)^\perp$$

$$\Rightarrow \ker(f^*)^\perp = \operatorname{im}(f)^{\perp\perp} = \operatorname{im}(f)$$

5. f is injective

$$\Leftrightarrow \ker(f) = \{0\} \overset{\text{3rd property}}{\Leftrightarrow} \operatorname{im}(f^*) = U \Leftrightarrow f^* \text{ surjective}$$

6. Like the proof for the fifth property, but applied to f^*

$\square$

Definition 24. 1. A linear map $f : V \to V$ is called *self-adjoint*, if

$$f^* = f$$

for matrices $A = A^*$, $\mathbb{K} = \mathbb{R}$ and $A = A^t$ symbolically.

2. $f \in \operatorname{Hom}(V, W)$ is called *unitary* (linear isometry), if

$$\bigwedge_{x,y \in V} \langle f(x), f(y)\rangle_W = \langle x, y\rangle_V$$

Remark 29. • Unitary maps are isometric and therefore injective

$$f(x) = 0 \Rightarrow \|f(x)\|^2 = 0 = \langle f(x), f(x)\rangle = \langle x, x\rangle = \|x\|^2$$

• $\dim V < \infty$.

$$f : V \to V \text{ unitary } \Rightarrow f \text{ invertible}$$

$$\varphi^{-1} = \varphi^*$$

Mostly unitary operators are defined by this relation $\varphi^{01} = \varphi^*$.

• If $\dim V = \infty$ is defined, then linear isometries are not necessarily invertible.

Proof. • Immediate.

• f is injective, so bijective.

$$\bigwedge_{x,y \in V} \langle x, y\rangle = \langle f(x), f(y)\rangle = \langle f^*(f(x)), y\rangle$$

From Exercise 24 it follows that

$$\bigwedge_x x = f^*(f(x))$$

$$\Rightarrow f^* \circ f = \mathrm{id} \Rightarrow f^* = f^{-1}$$

• Example: Consider $V = l^2 = c_{00}$ (space of finite sequences) $= \{(\xi_1, \xi_2, \ldots, \xi_n, 0, 0, \ldots) \mid n \in \mathbb{N}, \xi_i \in \mathbb{K}\}$.

$$S : V \to V$$

$$(\xi_1, \xi_2, \ldots) \mapsto (0, \xi_1, \xi_2, \ldots) \text{ is linear isometry}$$

$$\langle Sx, y\rangle = \langle (0, \xi_1, \xi_2, \ldots), (\eta_1, \eta_2, \ldots)\rangle$$

$$= \sum_{i=2}^{\infty} \xi_{i-1}\overline{\eta_i} = \sum_{i=1}^{\infty} \xi_i \cdot \overline{\eta_{i+1}} = \langle (\xi_1, \xi_2, \ldots), (\eta_2, \eta_3, \ldots)\rangle = \langle x, S^*(y)\rangle$$

$$\to S^* : V \to V$$

$$(\eta_1, \eta_2, \ldots) \mapsto (\eta_2, \eta_3, \ldots)$$

$$S^* \circ S = \mathrm{id}$$
$$S \circ S^* = \mathrm{id} - P_1 \neq \mathrm{id}$$
$$P1 : V \to V$$
$$(\xi_1, \xi_2, \dots) \mapsto (\xi_1, 0, 0, \dots)$$

So S is isometry, but not invertible (only works in infinity).

□

Definition 25. A matrix $U \in \mathbb{C}^{n \times n}$ is called *unitary*, if

$$U^* U = I$$

A matrix $U \in \mathbb{R}^{n \times n}$ is called *orthogonal* if

$$U^T U = I$$

This lecture took place on 2nd of May 2016 (Franz Lehner).

$$f : V \to W \text{ is unitary } (= \text{linear isometry})$$
$$\langle f(x), f(y) \rangle_W = \langle x, y \rangle_V$$

Definition 26. A matrix $U \in \mathbb{C}^{n \times n}$ is called *unitary* if $U^* U = I$ (i.e. $U^* = U^{-1}$). A matrix $R \in \mathbb{C}^{n \times n}$ is called *orthogonal* if $U^+ U = I$ (i.e. $U^+ = U^{-1}$).

Definition 27. Let $T \in \mathbb{C}^{n \times n}$. DFASÄ:

1. T is unitary.
2. $\bigwedge_{x \in \mathbb{C}^n} \|T_x\| = \|x\|$
3. $\bigwedge_{x,y \in \mathbb{C}^n} \Re\langle Tx, Ty \rangle = \Re\langle x, y \rangle$
4. $\bigwedge_{x,y \in \mathbb{C}^n} \langle Tx, Ty \rangle = \langle x, y \rangle$
5. The columns of T are orthogonal to each other. They satisfy the properties of a orthogonal normal basis.

Proof. **1.** $\to$ **2.** Let $T^* = T^{-1}$.

$$\Rightarrow \|Tx\|^2 = \langle Tx, Tx \rangle = \langle x, T^* Tx \rangle = \langle x, x \rangle = \|x\|^2$$

2. $\to$ **3.** Polarization:

$$\|x+y\|^2 - \|x-y\|^2 = \|x\|^2 + \|y\|^2 + 2\Re\langle x, y \rangle - (\|x\|^2 + \|y\|^2 - 2\Re\langle x, y \rangle) = 4\Re\langle x, y \rangle$$
$$\|T(x+y)\|^2 - \|T(x-y)\|^2 = \|Tx + Ty\|^2 - \|Tx - Ty\|^2 = \dots = 4\Re\langle Tx, Ty \rangle$$

Because $\|x+y\|^2 - \|x-y\|^2 = \|T(x+y)\|^2 - \|T(x-y)\|^2$, the cosine between Tx and Tx is the same like between x and y.

3. $\to$ **4.**

$$\Re Tx, Ty = \Re x, y \qquad \forall x, y$$
$$\Rightarrow \forall x, iy$$
$$\Re\langle Tx, T(iy) \rangle = \Re(-i \langle Tx, Ty \rangle) = \Im(\langle Tx, Ty \rangle)$$
$$\Re\langle Tx, T(iy) \rangle = \Re\langle x, iy \rangle = \Im(\langle x, y \rangle)$$

4. $\to$ **5.** The columns of T are

$$u_i = T \cdot e_i \qquad (n \text{ times})$$

Hence it suffices to show that u_i are an orthonormal system.

$$\langle u_i, u_j \rangle = \langle Te_i, Te_j \rangle = \langle e_i, e_j \rangle = \delta_{ij}$$

5. $\to$ **1.**

$$(T^* T)_{ij} = u_i^* u_j = \langle u_j, u_i \rangle = \delta_{ij}$$
$$\Rightarrow T^* T = I$$

□

Definition 28. Let (X, d) be a metric space. Consider (X', d'). $f : X \to X'$ is called *isometry* if $d'(f(x), f(y)) = d(x, y)$. Normed spaces are metric spaces.

$$d(x, y) = \|x - y\|$$

Isometry between metric spaces:

$$\bigwedge_{x,y \in V} \|f(x) - f(y)\| = \|x - y\|$$

Example 25. Translation is non-linear. Rotation around x is linear iff $x = 0$. Reflection along some axis g is linear iff $0 \in g$. Compare with Figure **??**.

$$U = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

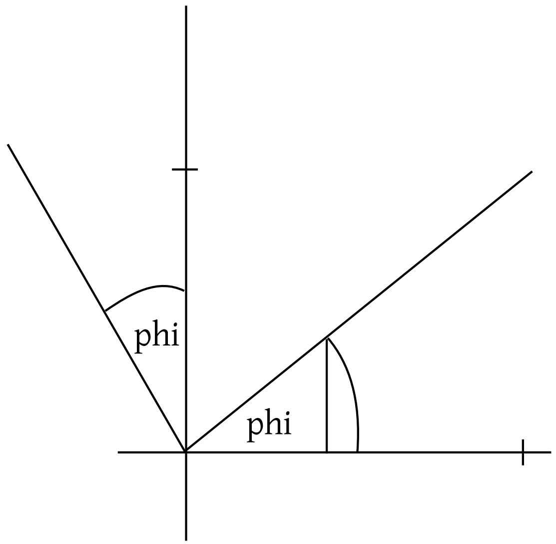

Figure 13: Rotation in $\mathbb{R}^2$

Remark 30. Newton considered motion (compare with Figure 14). We derive componentwise.

$$\hat{x}(t) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \cdot x(t)$$

$$\hat{x} = \alpha x$$

$$\frac{dx}{dt} = \alpha x$$

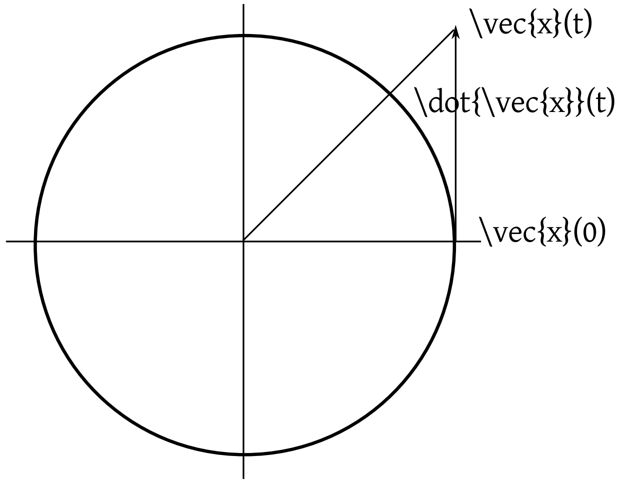

Figure 14: Motion

$$\frac{dx}{x} = \alpha dt$$

$$\int \frac{dx}{x} \int \alpha\, dt$$

$$\log(x) = \alpha t \Rightarrow x = e^{\alpha t} \cdot e^c + c$$

$$x(0) = e^c$$

$$\Rightarrow x(t) = e^{\begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} t} \cdot x(0)$$

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \sum_{n=0}^{\infty} \frac{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^n}{n!} t^n$$

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^3 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^4 = I$$

This holds for any $M^{4i}, M^{4i+1}, M^{4i+2}$ and M^{4i+3} respectively.

$$e^{i\varphi} = \sum_{n=0}^{\infty} \frac{(i\varphi)^n}{n!} = \sum_{n=0}^{\infty} \frac{(-1)^n \varphi^{2n}}{2n!} + \sum_{n=0}^{\infty} (-1)^n i \frac{\varphi^{2n+1}}{2(n+1)!} = \cos\varphi + i \sin\varphi$$

$$e^{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \varphi} = I \cos\varphi + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \sin\varphi = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

Remark 31. $\boxed{\text{Sophus Lie}}$

Unitary matrices build a group. This defines the field of Lie groups.

Remark 32 (Reflection)**.**

$$U = \begin{bmatrix} \cos 2\varphi & \cos(2\varphi - \frac{\pi}{2}) \\ \sin 2\varphi & \sin(2\varphi - \frac{\pi}{2}) \end{bmatrix} = \begin{bmatrix} \cos 2\varphi & \sin 2\varphi \\ \sin 2\varphi & -\cos 2\varphi \end{bmatrix}$$

Theorem 51. The following sets are groups:

$$\mathcal{O}(n) = \left\{ U \in \mathbb{R}^{n \times n} \,\middle|\, U^t U = I \right\} \qquad \text{orthogonal group}$$

$$\mathcal{U}(n) = \left\{ U \in \mathbb{U}^{n \times n} \,\middle|\, U^t U = I \right\} \qquad \text{unitary group}$$

$$S\mathcal{O}(n) = \{ U \in \mathcal{O}(n) \mid \det U = 1 \} \qquad \text{special orthogonal group}$$

$$S\mathcal{U}(n) = \{ U \in \mathcal{U}(n) \mid \det U = 1 \} \qquad \text{special unitary group}$$

"Classical" Lie groups.

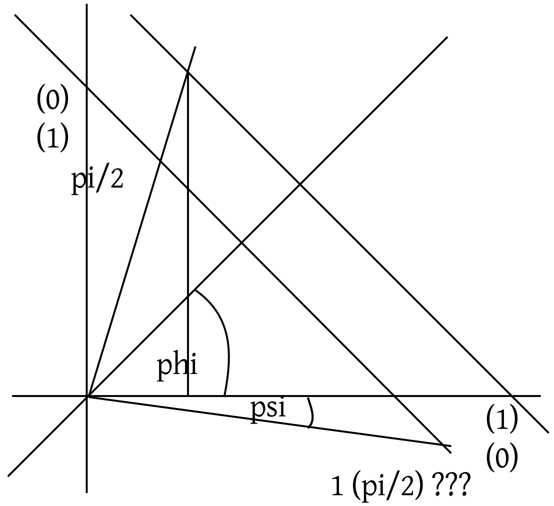

Figure 15: Reflection in R2

Theorem 52. For $U \in \mathcal{U}(n)$ it holds that $|\det U| = 1$.

$$U^* = U^{-1}$$

$$\begin{aligned}\det U^* &= \det \overline{U^t}\\ &= \overline{\det U^t}\\ &= \overline{\det U}\\ &= \det U^{-1}\\ &= \frac{1}{\det U}\\ &\Rightarrow \overline{\det U}\cdot \det U = 1\end{aligned}$$

Remark 33.

$$\mathcal{O}(n) = \{\det(U) = 1\} \cup \{\det(U) = -1\}$$

Example 26 ($\mathcal{O}(2)$)**.** Rotation:

$$\begin{bmatrix}\cos(\varphi) & -\sin(\varphi)\\ \sin\varphi & \cos\varphi\end{bmatrix} \qquad \det = 1$$

Reflection:

$$\begin{bmatrix}\cos(2\varphi) & \sin(2\varphi)\\ \sin 2\varphi & -\cos 2\varphi\end{bmatrix} \qquad \det = -1$$

Orthogonal:

$$U = \begin{bmatrix}a & b\\ c & d\end{bmatrix}$$

$$\begin{cases} a^2 + c^2 &= 1\\ b^2 + d^2 &= 1\\ ab + cd &= 0\end{cases}$$

$$U = \begin{bmatrix}a & b\\ c & d\end{bmatrix} = \begin{bmatrix}\cos\varphi & \cos\varphi\\ \sin\varphi & \sin\varphi\end{bmatrix}$$

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

$$\psi = \varphi + (k + \frac{1}{2})\pi$$

This sum equation can be derived from Euler:

$$e^{i(\alpha + i\beta)} = e^{i\alpha}e^{i\beta}$$

$$\begin{aligned}\cos(\alpha+\beta) + i\sin(\alpha+\beta) &= (\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta)\\ &= \cos\alpha\cos\beta - \sin\alpha\sin\beta + i(\sin\alpha\cos\beta + \cos\alpha\sin\beta)\end{aligned}$$

$$\begin{aligned}\cos\psi &= \cos(\varphi + (k + \frac{1}{2})\pi)\\ &= \cos\varphi\underbrace{\cos(k+\frac{1}{2})}_{=0}\pi - \sin\varphi\underbrace{\sin(k+\frac{1}{2})}_{\varepsilon:=\pm 1}\pi\\ &= -\varepsilon\sin\varphi\\ \sin\psi &= \sin(\varphi + (k+\frac{1}{2})\pi)\\ &= \sin\varphi\underbrace{\cos(k+\frac{1}{2})}_{=0}\pi + \cos\varphi\sin(k+\frac{1}{2})\pi = \varepsilon\cos\varphi\end{aligned}$$

Remark 34.

$$U = \begin{bmatrix}\cos\varphi & -\sin\varphi\cdot\varepsilon\\ \sin\varphi & \cos\varphi\cdot\varepsilon\end{bmatrix} = \begin{bmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{bmatrix}\cdot\begin{bmatrix}1 & 0\\ 0 & \varepsilon\end{bmatrix}$$

$$\det U = \varepsilon$$

U is either a rotation (if $\det U = 1$) or rotation with reflection (if $\det U = -1$).

Remark 35.

$$S\mathcal{O}(2) = \text{ rotations } \cong \mathcal{T} = \left\{e^{i\varphi} \,\middle|\, \varphi \in [0, 2\pi[\right\}$$

$$e^{i\varphi}e^{i\psi} = e^{i(\varphi+\psi)}$$

Remark 36. $\boxed{\text{William R. Hamilton (1805–1865)}}$

$$S\mathcal{U}(2) = \{q \in H \mid \|q\| = 1\}$$

Defined the Hamilton operator. Extension to $\mathbb{R}$ ("quaternions"):

$$H = \{a_0 + a_1 i + a_2 j + a_3 k \mid a_0, a_1, a_2, a_3 \in \mathbb{R}\}$$

$$i^2 = j^2 = k^2 = -1$$
$$ij = k \qquad jk = i \qquad ki = j$$
$$ji = -1 \qquad kj = -i \qquad ik = -j$$

Almost a field (inverse elements, but not commutative). A skew field.

Octonionen (inverse elements, but not associative).

4 Polynomials and Algebras

Definition 29. $\mathbb{K}$ is a field. A $\mathbb{K}$-Algebra is a vector space $\mathcal{A}$ over $\mathbb{K}$ with multiplication: $* : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ such that

1. $u * (b + c) = u * b + u * c$
2. $(a + b) * c = a * c + b * c$
3. $\lambda \cdot (a * b) = (\lambda \cdot a) * b = a * (\lambda \cdot b)$

where λ is an algebra.

Remark 37. If it furthermore holds that

$$a * (b * c) = (a * b) * c$$

then $\mathcal{A}$ is associative.

If it furthermore holds that

$$a * b = b * a$$

then $\mathcal{A}$ is called commutative.

Example 27.
- $(\mathbb{K}, +, * = \cdot)$ associative, commutative
- $\mathrm{Hom}(V, V) = \mathrm{End}(V)$

  $$f * g := f \circ g$$

  non-commutative, associative algebra.

  This is isomorphic to $(\mathbb{K}^{n \times n}, +, \cdot)$ where $\cdot$ is matrix multiplication.

  Hadamond- or *Schur product*:

  $$[a_{ij}][b_{ij}] = [a_{ij}, b_{ij}]$$

- $$\mathcal{C}[0, 1]$$
  $$(f * g)(t) = f(t) \cdot g(t)$$

  Convolution:

  $$(f * g)(t) = \int_0^1 f(t - s) g(s)\, ds$$

- Consider $(\mathbb{R}^3, +, \times)$ with cross product. $a \times b = -b \times a$.

  $$(a \times b) \times c \neq a \times (b \times c)$$

  Non-associative, non-commutative.

- Consider $(\mathbb{K}^{n \times n}, +, [])$.

  $$[A, B] = A \cdot B - B \cdot A$$

  Lie product or commutator product. Non-commutative, non-associative. From this, the *Jacobi identity* follows.

- $$= A = \mathbb{K}^{n \times n}_{\text{symm}} = \left\{ A \,\middle|\, A = A^t \right\}$$
  $$A * B = \frac{AB + BA}{2}$$

  Jordan product, associative, commutative.

Definition 30.

$$\mathbb{K}^\infty = \{(a_0, a_1, a_2, \ldots) \,|\, a_i \in \mathbb{K}\}$$

Vector of all sequences.

$$P_{\mathbb{K}} = \{(a_0, a_1, \ldots, a_n, 0, 0, \ldots) \,|\, a_i \in \mathbb{K}, n \in \mathbb{N}\}$$

Subspace of finite sequences. Basis of $P_{\mathbb{K}} : (e_i)_{i \geq 0}$.

$$(a_n) * (b_n) = (c_n)$$
$$(c_n) = \sum_{k=0}^{n} a_k b_{n-k} \qquad \text{(Cauchy product)}$$

Theorem 53. 1. $(P_{\mathbb{K}}, *)$ is an associative, commutative $\mathbb{K}$-Algebra with one element $(1, 0, 0, \ldots) = e_0$. $P_{\mathbb{K}} = \mathbb{K}[x]$.

$$x^k := e_k$$

$$e_i * e_j = e_{i+j}$$

$$x^0 = 1 \qquad \text{the one element}$$

2. $$\underbrace{\mathbb{K}[[x]]}_{\mathbb{K}(x)} = \left\{ \sum_{k=0}^{\infty} a_k x^k \,\middle|\, a_k \in \mathbb{K} \right\}$$

is a formal power series. Defines a commutative, associative algebra.

This lecture took place on 4th of May 2016 (Franz Lehner).

Remark 38. What is $\log(-1)$?

$$e^{i\varphi} = \cos\varphi + i \cdot \sin\varphi$$

$$e^{\log -1} = -1$$

$$\Rightarrow e^{i\pi} = -1$$

$$\Rightarrow \log(-1) = i\pi$$

This is ambiguous.

$$\sqrt{-1} = \pm i$$

$$\sqrt[3]{i} = 1, e^{2\frac{\pi}{i}}3, e^{-2\frac{\pi i}{3}}$$

$$\sqrt{e^{ix}} = e^{\frac{ix}{2}}$$

Riemann replaced the complex plane with a plane which consists of two planes. If you follow the unit circle one rotation, you end up at the other plane.

Remark 39. Consider $(P_{\mathbb{K}}, *)$.

$$P_{\mathbb{K}} = \{(a_0, a_1, \ldots, a_n, 0, \ldots) \,|\, a_k \in \mathbb{K}, n \in \mathbb{N}_0\}$$

with $(a_i) * (b_j) = (c_k)$ is an associative and commutative $\mathbb{K}$-algebra.

$$c_k = \sum_{j=i}^{k} a_j b_{k-j} = \sum_{j=0}^{k} a_{k-j} b_j$$

$$e_i * e_j = e_{i+j}$$

$$x^i := e_i$$

Let $a_i = 0$ for $i > m$, $b_j = 0$ for $j > n$. Compare with Figure 16.

$$c_k = \sum_{i=0}^{k} a_i b_{k-i} = \sum_{i=0}^{m} a_i \underbrace{b_{k-i}}_{=0} = 0$$

$$k > m + n \qquad i < m$$

$$k - i > m + n - i$$

$$k - i > n$$

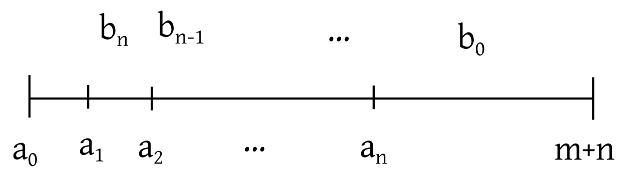

Figure 16: relation of a_i and b_i

$$\deg(p(x) \cdot q(x)) \leq \deg(p(x)) + \deg(q(x))$$

$$p(x) = a_0 + a_1 x + \ldots + a_m x^m$$

$$\deg(p(x)) = \max\{i \,|\, a_i \neq 0\}$$

Distributive law:

$$(a*(b+c))_k = \sum_{i=0}^{k} a_i(b_{k-i}+c_{k-i}) = \sum_{i=0}^{k} a_i b_{k-i} + \sum_{i=0}^{k} a_i c_{k-i} = (a*b)_k + (a*c)_k$$

This also works for $(a_0, a_1, \ldots)$ arbitrary sequences. $(a*b)_k = \sum_{i=0}^{k} a_i b_{k-i}$ is finite for all k. Polynomials (= finite sequences) form a subalgebra.

Definition 31.

$$X^0 = (1, 0, \ldots, 0) = 1$$
$$X^k = (0, \ldots, 0, 1, 0, \ldots)$$
$$X^k \cdot X^l = X^{k+l}$$

We write $\mathbb{K}[x]$ instead of $P_{\mathbb{K}}$.

$$p(x) = \sum_{i=0}^{m} a_i x^i$$

$$\partial p(x) = \deg(p(x)) = \max\{i \mid a_i \neq 0\}$$

We need to define $\deg(0) = -\infty$.

Lemma 12. 1. $\deg(p(x)\cdot q(x)) = \deg(p(x)) + \deg(q(x))$

2. $\mathbb{K}[x]$ is zero divisor free.

$$p(x)q(x) = 0 \Rightarrow p(x) = 0 \vee q(x) = 0$$

Counterexamples for zero divisor freedom:

$$\mathbb{Z}_n, n \notin \mathbb{P} : n = pq \Rightarrow p \neq 0 \mod n \wedge q \neq 0 \mod n$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = 0$$

Definition 32. Every polynomial $p(x) \in \mathbb{K}[x]$ induces a function.

$$p : \mathbb{K} \to \mathbb{K}$$

$$\alpha \mapsto p(\alpha) = \sum_{k=0}^{m} a_k \alpha^k$$

$$(\lambda p + \mu q)(\alpha) = \lambda p(\alpha) + \mu q(\alpha)$$
$$(p \cdot q)(\alpha) = p(\alpha)\cdot q(\alpha)$$

$$\mathbb{K}[x] \to \mathbb{K}^{\mathbb{K}}$$

is an algebra homomorphism.

Remark 40. Is it injective? If $|\mathbb{K}| < \infty$, it is not injective.

$$\dim \mathbb{K}[x] = \infty \qquad \dim \mathbb{K}^{\mathbb{K}} = |\mathbb{K}|$$

$$p(x) = (x-\xi_1)(x-\xi_2)\ldots(x-\xi_n)$$

has degree n.

From this we can see the difference between a polynomial and a polynomial function.

Example 28. Every function $f : \mathbb{K} \to \mathbb{K}$ is a polynomial function, hence there exists some polynomial $p(x) \in \mathbb{K}[x]$ such that $p(\xi) = f(\xi)$ for all $\xi \in \mathbb{K}$.

Definition 33. A map $\psi : \mathcal{A} \to \mathcal{B}$ between two $\mathbb{K}$-algebras $\mathcal{A}$ and $\mathcal{B}$ is called *algebra homomorphism* if Ψ is linear and multiplicative.

$$\bigwedge_{a,b\in\mathcal{A}} \psi(a *_{\mathcal{A}} b) = \psi(a) *_{\mathcal{B}} \psi(b)$$

Example 29.

- $\mathbb{K}[x] \to \mathbb{K}^{\mathbb{K}}$ with $p(x) \mapsto$ polynomial function
- For all $\alpha \in \mathbb{K}$,
$$\psi_\alpha : \mathbb{K}[x] \to \mathbb{K}$$
$$p(x) \mapsto p(\alpha)$$
is algebra homomorphism.
- $\mathbb{K} \to \mathbb{K}[x]$ with $\alpha \to \alpha \cdot \mathbb{1}$. Embedding is algebra homomorphism.

Theorem 54 (Insertion theorem)**.** Let $\mathcal{A}$ be an associative algebra over $\mathbb{K}$ with one-element $1_{\mathcal{A}}$.

- $$\Rightarrow: L : \mathbb{K} \to \mathcal{A}$$
  $$\alpha \mapsto \alpha \cdot \mathbb{A}_{\mathcal{A}}$$
  is an algebra homomorphism.

- For every $a \in \mathcal{A}$ is a map
  $$\psi_a : \mathbb{K}[x] \to \mathcal{A}$$
  $$\sum_{k=0}^{n} c_k x^k \mapsto \sum_{k=0}^{n} c_k a^k$$
  where $a^0 := 1_{\mathcal{A}}$ and $a^{k+1} = a * a^k$, the *unique* algebra homomorphism $\psi : \mathbb{K}[x] \to \mathcal{A}$ with the property $\psi(x) = a$.

- Every algebra homomorphism $\psi : \mathbb{K}[x] \to \mathcal{A}$ has this structure.

Proof. If $a = \psi(x)$, then $a^k = \psi(x)^k = \psi(x^k)$. If $\psi(x)$ is known, then $\psi(x^k)$ is defined for all k. So $\psi(p(x))$ is defined for all $p(x) \in \mathbb{K}[x]$. This follows because they represent a basis and by the Fortsetzungssatz. $\square$

Remark 41. Linearity of ψ_a will be shown in the practicals. Multiplicativity:

$$\psi_a(p(x) \cdot q(x)) \stackrel{!}{=} \psi_a(p(x)) * \psi_a(q(x))$$

Let $p(x) = \sum_{i=0}^{m} \alpha_i \cdot x^i$ and $q(x) = \sum_{j=0}^{n} \beta_j x^j$.

$$p(x) \cdot q(x) = \sum_{k=0}^{m+n} \gamma_k x^k \qquad \gamma_k = \sum_{i=0}^{k} \alpha_i \beta_{k-i}$$

$$\psi_a(p(x) \cdot q(x)) = \sum_{k=0}^{m+n} \gamma_k a^k$$

$$\begin{aligned}\psi_a(p(x)) * \psi_a(q(x)) &= \left(\sum_{i=0}^{m} \alpha_i a^i\right) \cdot \left(\sum_{j=0}^{n} \beta_j a^j\right) \\ &= \sum_{i=0}^{m} \sum_{j=0}^{n} \alpha_i \beta_j a^{i+j} \\ &= \sum_{k=0}^{m+n} \underbrace{\sum_{i,j \geq 0} \alpha_i \beta_j}_{\sum_{i=0}^{k} \alpha_i \beta_{k-i} = \gamma_k} a^k\end{aligned}$$

Remark 42 (Notation)**.**
$$\psi_a(p(x)) =: p(a)$$

Example 30.
- $\mathcal{A} = \mathbb{K}$
  $$\psi_\alpha(p(x)) = p(\alpha)$$
- $\mathcal{A} = \mathrm{Hom}(V, V)$
  $$L : \mathbb{K} \to \mathrm{Hom}(V, V)$$
  $$\lambda \mapsto \lambda \cdot \mathrm{id}$$
  $$f^0 = \mathrm{id}$$
  $$f^k = \underbrace{f \circ f \circ \ldots \circ f}_{kx} \Rightarrow \psi_f \left(\sum_{k=0}^{n} \alpha_k x^k\right) = \sum_{k=0}^{n} \alpha_k f^k$$
- $\mathcal{A} = \mathbb{K}^{n \times n}$
  $$\psi_A(p(x)) = p(A) = \sum_{k=0}^{n} \alpha_k A^k$$

Remark 43. $\mathbb{K}[x]$ is a free associative algebra with a generator over $\mathcal{A}$. Every map $f : \{X\} \to \mathcal{A}$ has a unique continuous to an algebra homomorphism.

$$\psi : \mathbb{K}[x] \to \mathcal{A}$$

$\mathbb{K}[x]$ is the smallest algebra over $\mathbb{K}$ which contains x.

For two generators?

$$\begin{array}{ccc} f:\{x,y\} & \to & \mathcal{A} \\ \downarrow & & \\ \psi:\mathbb{K}\langle x,y\rangle & \to & \mathcal{A} \end{array}$$

is a non-commutative polynomial in x, y.

Analogously: free group, free monoid and every vector space is free over its basis.

Definition 34. Let $p(x) \in \mathbb{K}[x]$. A root of $p(x)$ is some $\xi \in \mathbb{K}$ such that $p(\xi) = 0$.

$$\Leftrightarrow p(x) \in \ker(\psi_\xi)$$

Example 31.

$$p(x) = a_0$$

No non-trivial roots.

$$p(x) = a_0 + a_1x + a_2x^2$$

The solution equation was found 2000 BC.

$$p(x) = a_0 + a_1x + a_2x^2 + a_3x^3$$

Equation of Cardano.

$\boxed{\text{Gerdama Cardano (1501–1576)}}$
"Ars Magna" (1545)

$\boxed{\text{Niccolo Tartaglia (1499-1557)}}$

Niccolo Tartaglia found the solution to Cardano's equation. But actually $\boxed{\text{Siphione del Forzzo (1465–1526)}}$ found the equation first and forwarded it to his student Antonio Fiore. This was also the first time someone reasoned about complex numbers. They did not get explicitly defined.

$\deg(p(x)) = 4$ (L. Ferrari)

$\deg(p(x)) \geq 5$ (1826, Abel)

Remark 44 (Cadano and Tartaglia formula)**.** Originally: cub pi6 reb eq̄lis 20

$$x^3 + 6x = 20$$

Approach: $x = u + v$.

$$\begin{aligned} (u+v)^3 + 6(u+v) &= 20 \\ u^3 + 3u^2v + 3uv^2 + v^3 + 6(u+v) &= 20 \\ = u^3 + v^3 + (3uv+6)(u+v) &= 20 \end{aligned}$$

We choose v such that $3uv + 6 = 0$.

$$\Rightarrow uv = -2$$

$$\Rightarrow u^3v^3 = -18$$

$$u^3 + v^3 = 20$$

Let $a = u^3$ and $b = v^3$.

$$a \cdot b = -8 \wedge a + b = 20 \Rightarrow a(20-a) = -8$$

$$\begin{aligned} a^2 - 20a - 8 &= 0 \\ a = \frac{20 \pm \sqrt{400+32}}{2} &= 10 \pm \sqrt{108} \\ \Rightarrow u^3 &= 10 + \sqrt{108} \\ v^3 &= 10 - \sqrt{108} \\ x = u + v &= \sqrt[3]{10+\sqrt{108}} + \sqrt[3]{10-\sqrt{108}} \end{aligned}$$

Theorem 55 (Division with remainder)**.** Let $p(x), q(x) \in \mathbb{K}[x]$ and $q(x) \neq 0$. Then there exists exactly one $s(x), r(x) \in \mathbb{K}[x]$ such that $\deg(r(x)) < \deg(q(x))$ and $p(x) = s(x) \cdot q(x) + r(x)$.

Compare this with natural numbers and the extended euclidean algorithm.

$$m \in \mathbb{Z}, n \in \mathbb{N} \Rightarrow \exists! a, b : m = a \cdot n + b \text{ with } 0 \leq b < n$$

Proof. Complete induction over $\deg(p(x))$.

Case 1: $\deg(p(x)) < \deg(q(x))$

$$\Rightarrow p(x) = 0 \cdot q(x) + p(x)$$

is unique.

Case 2: $\deg(p(x)) \geq \deg(q(x))$

$$p(x) = \sum_{k=0}^{m} a_k x^k \qquad q(x) = \sum_{l=0}^{n} b_l x^l$$

$m \geq n$. Let $p_1(x) = p(x) - \frac{a_m}{b_n} x^{m-n} \cdot q(x)$.

$$= \sum_{k=0}^{m} a_k x^k - \sum_{l=0}^{n} \frac{a_m}{b_n} b_l x^{m-n+l}$$
$$= \sum_{k=0}^{m-1} a_k x^k - \sum_{l=0}^{n-1} \frac{a_m}{b_n} b_l x^{m-n+l}$$

$\deg(p_1(x)) < \deg(p_2(x))$.

Induction hypothesis $\Rightarrow p_1(x) = s_1(x) \cdot q(x) + r_1(x)$

$$\Rightarrow p(x) = \left(\frac{a_m}{b_n} x^{m-n} + s_1(x)\right) q(x) + r_1(x)$$
$$= p_1(x) + \frac{a_m}{b_n} x^{m-n} \cdot q(x)$$

□

Example 32.

$$\begin{array}{rl}
3x^5 - x^4 + 2x^3 + x^2 + 1 & : x^2 - 3x + 1 = 3x^3 + 8x^2 + 23x + 62 \\
-(3x^5 - 9x^4 + 3x^3) & \\
\hline
0 + 8x^4 - x^3 + x^2 + 1 & \\
0 - (8x^4 - 24x^3 + 8x^2) & \\
\hline
0 + 0 + 23x^3 - 7x^2 + 1 & \\
0 + 0 - 23x^3 - 69x^2 + 23x & \\
\hline
0 + 0 + 62x^2 - 23x + 1 & \\
0 + 0 + 62x^2 - 186x + 62 & \\
\hline
0 + 0 + 0 + 163x - 61 = r(x) &
\end{array}$$

This lecture took place on 9th of May 2016 (Franz Lehner).

Remark 45. An Euclidean ring is a ring in which the extended Euclidean algorithm works.

Theorem 56. Let $p(x), q(x) \in \mathbb{K}[x]$ and $q(x) \neq 0$. Then there exists exactly one $s(x)$ and $r(x)$ such that $\deg r(x) < \deg q(x)$.

$$p(x) = r(x) \cdot q(x) + r(x)$$

Definition 35. $q(x)$ divides $p(x)$ if $\exists s(x) : p(x) = s(x) \cdot q(x)$ (i.e. division happens without remainder).

Theorem 57. Consider $q(x) = x - \xi$.

$$\Rightarrow p(x) = s(x) \cdot (x - \xi) + r$$
$$\Rightarrow p(\xi) = r$$

Corollary 12. ξ is a root of $p(x) \Leftrightarrow x - \xi$ is divisor of $p(x)$.

Theorem 58 (Horner schema)**.** Let $p(x) \in \mathbb{K}[x]$ and $\lambda \in \mathbb{K}$. Determine $p(\lambda)$.

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$$
$$p(\lambda) = a_n \lambda^n + a_{n-1} \lambda^{n-1} + \ldots + a_1 \lambda + a_0$$

Remark 46. Naively it requires a quadratic number of evaluations (for λ, per $\lambda - 1$, per $\ldots = \frac{n(n+1)}{2} \approx n^2$). Using binary composition it takes a logarithmic times linear number of evaluations (for λ, for $\lambda - 1, \ldots$). However, it also works with n multiplications.

$$p(x) = (a_n \lambda^{n-1} + a_{n-1} \lambda^{n-2} + \ldots + a_1) \lambda + a_0$$
$$= ((a_n \lambda^{n-2} + a_{n-1} \lambda^{n-3} + \ldots + a_2) \lambda + a_1) \lambda + a_0$$

Example 33.

$$\begin{array}{lr}
p(x) = a_3 \lambda^3 + a_2 \lambda^2 + a_1 \lambda + a_0 & a_3 \\
(a_3 \lambda^2 + a_2 \lambda + a_1) \lambda + a_0 & a_3 \lambda + a_2 \\
((a_3 \lambda + a_2) \lambda + a_1) \lambda + a_0 & ((a_3 \lambda + a_2) \lambda + a_1) \lambda + a_0
\end{array}$$

Algorithm:

$$\xi_n = a_n \text{ for } k = n-1, \dots, 0$$
$$\xi_k = \lambda \xi_{k+1} + a_k$$
$$\Rightarrow p(\lambda) = \xi_0$$

We evaluate:

$$p(x) = 3x^5 - x^4 + 2x^3 + x^2 + 1$$

Let $\lambda = 5$.

$$\begin{aligned}
\xi_5 &= 3 \\
\xi_4 &= 5 \cdot 3 - 1 = 14 \\
\xi_3 &= 5 \cdot 14 + 2 = 72 \\
\xi_2 &= 5 \cdot 72 + 1 = 361 \\
\xi_1 &= 5 \cdot 36 + 0 = 1805 \\
\xi_0 &= 5 \cdot 1805 + 1 = 9026
\end{aligned}$$

Compare with division:

$$\begin{array}{rl}
3x^5 - x^4 + 2x^3 + x^2 + 1 & : x - 5 = 3x^4 + 14x^3 + 72x^2 + 361x + 1805 \\
3x^5 - 15x^4 & \\
\hline
0 + 14x^4 + 2x^3 + x^2 + 1 & \\
0 + 14x^4 - 70x^3 & \\
\hline
0 + 0 + 72x^3 + x^2 + 1 & \\
0 + 0 + 72x^3 - 360x^2 & \\
\hline
0 + 0 + 0 + 361x^2 + 1 & \\
0 + 0 + 0 + 361x^2 - 1805x & \\
\hline
0 + 0 + 0 + 0 + 1805x + 1 & \\
0 + 0 + 0 + 0 + 1805x - 5 \cdot 1805 & \\
\hline
0 + 0 + 0 + 0 + 0 + 9026 &
\end{array}$$

The Horner scheme is equivalent to division, but written down more efficiently.

Definition 36. A polynomial $p(x) \in \mathbb{K}[x]$ is called *reducible* if $\exists p_1(x) \exists p_2(x) \in \mathbb{K}[x] : \deg p_1(x), \deg p_2(x) < \deg p(x)$ and $p(x) = p_1(x) \cdot p_2(x)$ (hence there exists a non-trivial divisor). Otherwise $p(x)$ is called irreducible.

Remark 47.
- Constant and linear polynomials are irreducible.
- Irreducible polynomials of degree ≥ 2 have no roots (otherwise $x - \xi$ is divisor!)

Example 34.

- $x^2 - 2$ is irreducible in $\mathbb{K} = \mathbb{Q}$. Is reducible in $\mathbb{K} = \mathbb{R}$ with $(x - \sqrt{2})(x + \sqrt{2})$.
- $x^2 + 1$ is irreducible in $\mathbb{K} = \mathbb{Q}$ and $\mathbb{K} = \mathbb{R}$ and reducible in $\mathbb{K} = \mathbb{C}$ with $(x - i)(x + i)$.
- $x^2 + x + 1 \in \mathbb{Z}_2[x] \Rightarrow$ irreducible.

$$x^3 + x + 1 \text{ is irreducible}$$

$$x^5 + x + 1 = (x^2 + x + 1)(x^3 + x^2 + 1) \text{ is reducible}$$

Remark 48. How about explicitly defining fields such that $x^2 + x + 1$ is reducible?

Consider $(\sqrt{-1})^2 + 1 = 0$. In which field does $x^2 + x + 1 \in \mathbb{Z}_2[x]$ have a root?

Let α be a "number" such that $\alpha^2 + \alpha + 1 = 0$.

$$\Rightarrow \alpha^2 = \alpha + 1$$

$$\alpha^3 = \alpha(\alpha + 1) = \alpha^2 + \alpha = \alpha + 1 + \alpha = 1$$

$$\begin{aligned}
(a \cdot \alpha + b)(c \cdot \alpha + d) &= ac \cdot \underbrace{\alpha^2}_{=\alpha+1} (b \cdot c + a \cdot d) \cdot \alpha + bd \\
&= (ac + bc + ad)\alpha + (ac + bd)
\end{aligned}$$

$$\Rightarrow \{a \cdot \alpha + b \,|\, a, b \in \mathbb{Z}_2\} =: \mathrm{GF}(2^2) = \mathrm{GF}(4)$$

is a ring and even a field[2]. For the practicals you need to determine the inverse $\alpha^{-1} = \alpha^2$.

For all $p \in \mathbb{P}$ (and $k \in \mathbb{N}$) you need to consider a different Galois field $\mathrm{GF}(p^k)$ (is a field of order p^k).

[2] GF stands for Galois field

Theorem 59. Fundamental theorem of algebra $\mathbb{C}$ is algebraically closed, hence every polynomial $p(x) \in \mathbb{C}[x]$ has a root $\xi \in \mathbb{C}$. Corollaries:

1. $p(x) \in \mathbb{C}[x]$ is irreducible $\Leftrightarrow \deg p(x) \leq 1$
2. Every $p(x) \in \mathbb{C}[x]$ has a factor ring
$$p(x) = (x - \xi_1)(x - \xi_2) \dots (x - \xi_n)$$
where $\xi_i \in \mathbb{C}$ and $n = \deg p(x)$.

Proof. No proof given.

This theorem cannot be proven with means of algebra. You need to study theory of complex functions.

$$f'(z_0) = \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

Theorem of Lionville: Every complex differentiable function is unbounded. □

Theorem 60. In arbitrary fields it holds that: every polynomial has (except for the order) a unique factorization $p(x) = p_n(x) \dots p_k(x)$ to irreducible factors.

Theorem 61. Let $p(x)$ and $q(x) \in \mathbb{K}[x] \setminus \{0\}$. Then there exists a unique monic polynomial (ie. leading coefficient 1) of maximum degree denoted $\gcd(p(x), q(x))$ such that
$$\gcd(p, q) \mid p(x) \wedge \gcd(p, q) \mid q(x)$$
Then it holds that *all* common divisors of $p(x)$ and $q(x)$ divide $\gcd(p(x), q(x))$.

Proof. Let $g(x)$ be a polynomial of maximum degree, which divides $p(x)$ and $q(x)$.
$$\Rightarrow p(x) = f(x) \cdot g(x) \qquad q(x) = h(x) \cdot g(x)$$
Let $d(x)$ be a common divisor, then $\deg d(x) \leq \deg g(x)$. We apply division to retrieve $g(x) = s(x) \cdot d(x) + r(x)$ and $\deg r(x) < \deg d(x)$.
$$p(x) = \tilde{f}(x) \cdot d(x), \qquad q(x) = \tilde{h}(x) \cdot d(x)$$

It holds that
$$f(x) \cdot g(x) = \tilde{f}(x) \cdot d(x)$$
$$f(x) \cdot g(x) = f(x)\,(s(x)d(x) + r(x))$$
$$\Big(\tilde{f}(x) - f(x) \cdot s(x)\Big) \cdot d(x) = r(x)$$
$$\deg(\text{LHS}) > \deg(\text{RHS})$$
$$\Rightarrow \tilde{f}(x) = f(x) \cdot s(x) = 0 \Rightarrow r(x) = 0 \Rightarrow d(x) \mid g(x)$$
□

Remark 49. Only one unique greatest common divisor can exist. Proof is given in the practicals.

If $p(x) = s(x) \cdot q(x) + r(x)$
$$\Rightarrow \gcd(p(x), q(x)) = \gcd(q(x), r(x))$$
just like for integers.

Definition 37. A root ξ of a polynomial $p(x)$ has multiplicity m if $(x - \xi)^m \mid p(x)$ but $(x - \xi)^{m+1} \not| p(x)$.

Remark 50. In the practicals we will see that the roots with multiplicity ≥ 2 are the roots of the greatest common divisor $p(x)$ and $q(x)$.
$$(x^n)' = nx^{n-1}$$

5 Eigenvalues and Eigenvectors

Our goal is to find $f \in \text{End}(V) = \text{Hom}(V, V)$ of a basis B such that $\Phi_B^B(f)$ has the simplest possible structure or find (for a given matrix A) a matrix T such that
$$TAT^{-1}$$
has the simplest possible structure ("equivalence transformation").

Compare it with

$$\Phi_C^B(f) = \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 \end{bmatrix}$$

$$UAV = \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 \end{bmatrix}$$

Definition 38. Let V be a vector space over $\mathbb{K}$. Let $f \in \operatorname{End} V$. $\lambda \in \mathbb{K}$ is called eigen value of f if there exists $v \in V \setminus \{0\} : f(v) = \lambda \cdot v$. v is called eigen vector of a vector value λ.

$$\operatorname{spectrum}(f) := \{\lambda \mid \lambda \text{ is eigenvalue of } f\}$$

is called *spectrum* of f.

Lemma 13.

$$\eta_\lambda = \{v \in V \mid f(v) = \lambda v\} = \ker(\lambda \cdot \mathrm{id} - f)$$

is a subspace and is called Eigenspace of f for eigenvalue λ.

$$v \in \eta_\lambda \Leftrightarrow f(v) = \lambda \cdot v$$
$$\Leftrightarrow \lambda \cdot v - f(v) = 0$$
$$\Leftrightarrow (\lambda \cdot \mathrm{id} - f)(v) = 0$$
$$\Leftrightarrow v \in \ker(\lambda \cdot \mathrm{id} - f)$$

Example 35.

- Let $f = c \cdot \mathrm{id}$

$$\Rightarrow f(v) = c \cdot v \qquad \text{for all } v \in V$$
$$\Rightarrow \operatorname{spectrum}(f) = \{c\}$$
$$\lambda = c \qquad \eta_\lambda = V$$

- Let B be a basis of V

$$f : V \to V$$
$$b_i \mapsto \lambda_i \cdot b_i$$

and continuation theorem: $f(\sum \alpha_i b_i) = \sum \alpha_i \lambda_i b_i$.

$$\Leftrightarrow \Phi_B^B(f) = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

$$\Rightarrow \operatorname{spectrum}(f) = \{\lambda_1, \dots, \lambda_n\}$$
$$\eta_\lambda = \mathcal{L}\left(b_i \,|\, \lambda_i = \lambda\right)$$

Let λ be an Eigenvalue.

$$\Rightarrow \exists v = \sum \alpha_i b_i : f\left(\sum \alpha_i b_i\right) = \lambda \cdot \sum_{i=1}^n \alpha_i b_i$$
$$\sum \alpha_i b_i = \sum_{i=1}^n \alpha_i \lambda_i b_i$$
$$\sum (\lambda - \lambda_i) \alpha_i b_i = 0$$
$$(b_i) \text{ is linear independent} \Rightarrow (\lambda - \lambda_i)\alpha_i = 0 \forall i$$
$$\lambda = \lambda_i \vee \alpha_i = 0$$

-

$$V = C^\infty(\mathbb{R})$$
$$\frac{d}{dx} : C^\infty \to C^\infty$$
$$f \mapsto f' = \frac{df}{dx}$$

Eigenvectors?

$$y' = \lambda y$$
$$\frac{dy}{dx} = \lambda y \Rightarrow \frac{dy}{y} = \lambda dx$$

$$\int \frac{dy}{y} = \lambda \int dx$$
$$\log y = \lambda x + C$$
$$\Rightarrow y = c_1 \cdot e^{\lambda x}$$
$$V = C^\infty(\mathbb{R}, \mathbb{C})$$
$$\Rightarrow \text{spectrum}\left(\frac{d}{dx}\right) = \mathbb{R}$$
$$\eta_\lambda = \mathcal{L}(e^{\lambda x})$$

From e^{iwx} follows Fourier transformation.

- $C^\infty[0, L]$

$$\frac{d^2}{dx^2} : C^\infty[0, L]$$
$$\frac{d^2}{dx^2} e^{\lambda x} = \lambda^2 e^{\lambda x}$$
$$\frac{d^2}{dx^2} e^{i\omega x} = -w^2 e^{i\omega x}$$
$$\to \frac{d^2}{dx^2} \cos \omega x = -\omega^2 \cos \omega x$$
$$\frac{d^2}{dx^2} \sin(\omega x) = -\omega^2 \sin(\omega x)$$

This lecture took place on 11th of May 2016 (Franz Lehner).

$$\frac{d}{dx} e^{\lambda x} = \lambda e^{\lambda x}$$
$$\frac{d^2}{dx^2} e^{\lambda x} = \lambda^2 e^{\lambda x} \qquad \forall \lambda \in \mathbb{R}$$
$$\frac{d^2}{dx^2} \sin(\omega x) = -\omega^2 \sin(\omega x) \qquad \forall w \in \mathbb{R}$$
$$\frac{d^2}{dx^2} \cos(\omega x) = -\omega^2 \cos(\omega x)$$

A lot of applications for eigenvalues can be found. Physics, for example:

$$C_0^\infty[0, L] = \{f \in C^\infty[0, L] \mid f(0) = f(L) = 0\}$$
$$\frac{d^2}{dx^2} C_0^\infty[0, L]$$
$$\sin(\omega x) \qquad \omega L = \pi \cdot k \qquad \Rightarrow \omega = \frac{\pi}{L} k$$
$$\to \sin\left(\frac{\pi}{L} kx\right)$$

Definition 39. Let $A \in \mathbb{K}^{n \times n}$. $\lambda \in \mathbb{K}$ is called *rightsided eigenvalue* if

$$\exists v \in \mathbb{K}^n \setminus \{0\} : A \cdot v = \lambda \cdot v$$

A *leftsided eigenvalue* is given if

$$\exists v \in \mathbb{K}^n \setminus \{0\} : v^* \cdot A = \lambda \cdot v^t$$
$$\Leftrightarrow A^t \cdot v = \lambda \cdot v$$
$$\Leftrightarrow \text{ is right-sided eigenvalue of } A^t$$

Lemma 14 (Leftsided eigenvalue equals rightsided eigenvalue)**.** Let λ be a rightsided eigenvalue.

$$Av - \lambda v = 0$$
$$\ker(\lambda \cdot I - A) \neq \{0\}$$
$$\Leftrightarrow \operatorname{rank}(\lambda I - A) < n$$
$$\Leftrightarrow \operatorname{rank}(\lambda I - A^t) < n$$
$$\Leftrightarrow \lambda \text{ is leftsided eigenvalue of } A$$

Remark 51.
1. Eigenvectors must not be equal.
2. This does not hold if $\dim = \infty$.

$$\delta : (\xi, \xi_2, \ldots) \mapsto (0, \xi, \xi_2, \ldots)$$

From injectivity follows that 0 is not a rightsided eigenvector.

$$\delta^t : (\xi, \xi_2, \ldots) \mapsto (\xi_2, \xi_3, \ldots)$$

has eigenvalue 0: $\delta^t(1,0,0,\dots) = (0,0,\dots)$ Hence,

$$\text{spectrum}(\delta) = \{\lambda \,|\, \lambda I - S \text{ is not invertible}\}$$

just depends on the definition of the spectrum.

Definition 40. Let $A \in \mathbb{K}^{n\times n}$.

$$\begin{aligned}\text{spectrum}_{\mathbb{K}}(A) &= \{\lambda \in \mathbb{K} \,|\, \lambda \text{ is rightsided eigenvalue of } A\}\\ &= \{\lambda \in \mathbb{K} \,|\, \lambda \text{ is leftsided eigenvalue of } A\}\end{aligned}$$

Lemma 15. Let $\dim V = n$, $f \in \text{End}(V)$ and B is basis of V. Then,

$$\text{spectrum}(f) = \text{spectrum}(\Phi_B^B(f))$$

$$f(v) = \lambda v \Leftrightarrow \Phi_B^B(f) \cdot \Phi_B(v) = \lambda \cdot \Phi_B(v)$$

Corollary 13. 1. The spectrum does not depend on the selection of the basis.

2. If T is regular, then $\text{spectrum}(T^{-1}AT) = \text{spectrum}(A)$.

Remark 52. Eigenvectors of $T^{-1}AT$?

$$\begin{aligned}Ax = \lambda x &\Leftrightarrow ATT^{-1}x = \lambda x\\ &\Leftrightarrow T^{-1}ATT^{-1}x = \lambda \cdot T^{-1}x\end{aligned}$$

X is eigenvector of A

$$\Leftrightarrow T^{-1}x \text{ is eigenvector of } T^{-1}AT$$

$$\lambda I - A \text{ is not injective}$$

Theorem 62. Let $A \in \mathbb{K}^{n\times n}$.

1. $\chi_A(\lambda) = \det(\lambda \cdot I - A)$ is a polynomial of degree n.
   $\chi_A(x)$ is called *characteristical polynomial* of A.

2. λ is eigenvalue of A iff λ is a root of $\chi_A(x)$.

Proof. 1.

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \dots & -a1n \\ -a_{21} & \lambda - a_{22} & & \vdots \\ \vdots & & \ddots & \\ -a_{n1} & \dots & & \lambda - a_{nn}\end{vmatrix}$$

$$= \sum_{\pi\in\sigma_n} \dots = (\lambda - a_{11})(\lambda - a_{22})\dots(\lambda - a_{nn})$$

plus polynomial of degree $\leq n-2$. Polynomial in λ

$$= \lambda^n + (\text{polynomial of degree} \leq n-1)$$

2. $\lambda I - A$ is not injective iff $\det(\lambda I - A) = 0$.

$\square$

Example 36.

$$A = \begin{bmatrix} -1 & 1 & 2 \\ -1 & -5 & 2 \\ 2 & -2 & -4\end{bmatrix} \qquad \text{spectrum}(A) = ?$$

$$\det(\lambda I - A) = \begin{vmatrix} \lambda+1 & -1 & -2 \\ 1 & \lambda+5 & -s \\ -2 & 2 & \lambda+4\end{vmatrix} \overset{\substack{\text{Laplace over}\\ \text{1st column}}}{=} \begin{vmatrix} \lambda & -1 & -2 \\ \lambda+6 & \lambda+5 & -2 \\ 0 & 2 & \lambda+4\end{vmatrix}$$

$$\lambda \begin{vmatrix} \lambda+5 & -2 \\ 2\lambda+4 & \end{vmatrix} - (\lambda+6)\begin{vmatrix} -1 & -2 \\ 2 & \lambda+4\end{vmatrix} = \lambda(\lambda^2 + 10\lambda + 30)$$

$$\text{spectrum}_{\mathbb{R}}(A) = \{0\} \qquad \text{spectrum}_{\mathbb{C}}(A) = \left\{0, -5 \pm i\sqrt{5}\right\}$$

$$\lambda_1 = 0$$

$$\lambda_{2,3} = \frac{-10 \pm \sqrt{100-120}}{2} = -5 \pm i\sqrt{5}$$

Theorem 63. Let $A \in \mathbb{K}^{n\times n}$. We denote $[n] \coloneqq \{1,\dots,n\}$.

$$\chi_A(x) = \sum_{k=0}^{n} (-1)^{n-k} c_k(A) x^k$$

$$c_k(A) = \sum_{\substack{J\subseteq[n]\\|J|=n-k}} \underbrace{[A]_{J,J}}_{\text{minors}}$$

with $J = \{j_1 < \dots < j_{n-k}\}$.

$$\begin{vmatrix} a_{j_1j_1} & a_{j_1j_2} & \dots & a_{j_1j_{n-k}} \\ \vdots & \ddots & & \vdots \\ a_{j_{n-k}j_1} & a_{j_{n-k}j_2} & \dots & a_{j_{n-k}j_{n-k}} \end{vmatrix}$$

Sum of all symmetrical minors.

Remark 53. Especially:

$$C_0 = \det A$$
$$C_{n-1} = \sum_{i=1}^{n} a_{ii} = \operatorname{Tr}(A)$$
$$C_n(A) = 1$$

Proof.

$$\det(\lambda I - A) = \sum_{\pi\in\sigma_n} (-1)^{\pi} \prod_{i=1}^{n} (xI - A)_{\pi(i),i}$$

Remark: Determinants are not only for fields, but arbitrary commutative rings defined ("non-conjugate determinant").

$$= \sum_{\pi\in\sigma_n} (-1)^{\pi} \prod_{i=1}^{n} (x\cdot\delta_{\pi(i),i} - a_{\pi(i),i})$$

$$\deg \prod_{i=1}^{n} \left(x\delta_{\pi(i),i} - a_{\pi(i),i}\right) = \#\{i \mid \pi(i) = i\} = \#\operatorname{Fix}(\pi) = \begin{cases} n & \pi = (1) \\ \le n-2 & \pi \ne (1) \end{cases}$$

$\Rightarrow \det \chi_A(x) = n$ and $c_n(A) = 1$ (leading coefficient is one).

$$A = \begin{bmatrix} a_1 & \dots & a_n \\ \vdots & & \vdots \end{bmatrix}$$

a_j is the i-th column of A. $I = (e_1, e_2, \dots, e_n)$.

$$\chi_A(x) = \triangle(xe_1 - a_1, xe_2 - a_2, \dots, xe_n - a_n)$$

$$= \sum_{I\subseteq[n]} \triangle(y_1,\dots,y_n)$$

$$y_i = \begin{cases} x\cdot e_i & i\in I \\ -a_i & i\notin I \end{cases}$$

$$(a+b)^n = \sum_{I\subseteq[n]} a^{|I|} b^{|I^c|}$$

$$\triangle(y_1 \dots y_{k-1}, xe_k, y_{k+1} \dots y_n)$$

$$= k \begin{vmatrix} y_1 & y_2 & \dots & y_{k-1} & 0 & y_{n+1} & \dots & y_n \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & 0 & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & x & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & 0 & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & 0 & \vdots & & \vdots \end{vmatrix} = (-1)^{k-1}(-1)^{k-1}$$

$$\rightsquigarrow = \begin{vmatrix} x & \dots & \dots & \dots & \dots & & & & & \\ 0 & \tilde{y_1} & \tilde{y}_{k-1} & \tilde{y}_{k+1} & \tilde{y}_n \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots & \vdots & & & & \end{vmatrix} = (-1)^{k-1}(-1)^{k-1}$$

$$= x \begin{vmatrix} \tilde{b}_1 & \dots & \tilde{y}_{k-1} & \tilde{y}_{k+1} & \dots & \tilde{y}_n \\ \vdots & & \vdots & \vdots & & \vdots \end{vmatrix}$$

$\tilde{y}_i = y_i$ with crossed-out row k. $(n - |\, I\,|) \times (n - |\, I\,|)$ determinant.

$$= x^{|I|} \underbrace{\triangle(-\tilde{a}_{j_1}, \dots, -\tilde{a}_{j_{1-|I|}})}_{(-1)^{|I^C|}}$$

$$[A]_{I^C}, I^C$$

$J = I^C$ is not crossed-out rows and columns. So the idea was: If there is some x, we can extract it using Laplace. □

Lemma 16.

$$\chi_{T^{-1}AT}(x) = \chi_A(x)$$

similar matrices have the same characteristic polynomial.

Proof.

$$\begin{aligned} \det(\lambda I - T^{-1}AT) &= \det(\lambda \cdot T^{-1} \cdot T - T^{-1}AT) \\ &= \det(T^{-1}(\lambda I - A)T) \\ &= \det(T^{-1}) \cdot \det(\lambda I - A) \cdot \det(T) \\ &= \det(\lambda I - A) \end{aligned}$$

□

Definition 41. Let $A \in \mathbb{K}^{n\times n}$. A is called *diagonizable* if

$$\exists T \in \mathrm{GL}(n, \mathbb{K}) : T^{-1}AT = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

From Lemma 16 it follows that

$$\chi_A(x) = \begin{vmatrix} x-\lambda_1 & & \\ & \ddots & \\ & & x-\lambda_n \end{vmatrix} = \prod_{j=1}^{n}(x-\lambda_i)$$

Lemma 17. *A is diagonizable iff there exists a basis of eigenvectors.*

Proof.

$$B^{-1}AB = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \Leftrightarrow AB = B \begin{bmatrix} -\lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

Let $B = \begin{bmatrix} b_1 & \dots & b_n \\ \vdots & & \vdots \end{bmatrix}$.

$$A \begin{bmatrix} -b_1 & b_2 & \dots & b_n \\ \vdots & \vdots & & \vdots \end{bmatrix} = \begin{bmatrix} b_1 & b_2 & \dots & b_n \\ \vdots & \vdots & & \vdots \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

$$\begin{bmatrix} Ab_1 & Ab_2 & \dots & Ab_n \\ \vdots & \vdots & & \vdots \end{bmatrix} = \begin{bmatrix} \lambda_1 b_1 & \lambda_2 b_2 & \dots & \lambda_n b_n \\ \vdots & \vdots & & \vdots \end{bmatrix} \Leftrightarrow Ab_i = \lambda_i b_i \quad \forall i$$

□

Example 37.

$$A = \begin{bmatrix} -1 & 2 & 4 \\ 4 & -3 & -8 \\ -2 & 2 & 5 \end{bmatrix}$$

Task: Diagonalize!

Step 1: Eigenvalues.

$$\chi_A(\lambda) = \begin{vmatrix} \lambda+1 & -2 & -4 \\ -4 & \lambda+3 & 8 \\ 2 & -2 & \lambda-5 \end{vmatrix} = \begin{vmatrix} \lambda-1 & -2 & -4 \\ \lambda-1 & \lambda+3 & 8 \\ 0 & -2 & \lambda-5 \end{vmatrix}$$

$$= (\lambda-1) \begin{vmatrix} 1 & -2 & -4 \\ 1 & \lambda+3 & 8 \\ 0 & -2 & \lambda-5 \end{vmatrix}$$

$$= (\lambda-1) \begin{vmatrix} 1 & -2 & -4 \\ 0 & \lambda+5 & 12 \\ 0 & -2 & \lambda-5 \end{vmatrix} = (\lambda-1)\,[(\lambda+5)(\lambda-5)+24] = (\lambda-1)(\lambda^2-1) = (\lambda-1)^2(\lambda+1)$$

$$\mathrm{spectrum}(A) = \{\pm 1\}$$

Step 2: Eigenvectors $=\ker(\lambda I - A)$. For $\lambda = +1$ we have,

$$\begin{array}{ccc} 2 & -2 & -4 \\ \\ -4 & 4 & 8 \\ 2 & -2 & -4 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}$$

$$x_1 = x_2 + 2x_3$$

$$\eta_{+1} = \underbrace{\mathcal{L}\left\{\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}2\\0\\1\end{bmatrix}\right\}}_{\text{basis of } \ker(I-A)=\eta_{+1}}$$

For $\lambda = -1$: $\ker(-I - A)$.

$$\begin{array}{ccc} 0 & -2 & -4 \\ -4 & 2 & 8 \\ 2 & -2 & -6 \\ \hline 0 & -2 & -4 \\ \\ 0 & -2 & -4 \\ \hline 0 & 0 & 0 \end{array}$$

$$x_1 = x_2 + 3x_3 = x_3$$

$$x_2 = -2x_3$$

$$\eta_{-1} = \mathcal{L}\begin{pmatrix}1\\-2\\1\end{pmatrix}$$

$$B = \begin{bmatrix}1 & 2 & 1\\ 1 & 0 & -2\\ 0 & 1 & 1\end{bmatrix} \qquad \Rightarrow B^{-1}AB = \begin{bmatrix}1 & & \\ & 1 & \\ & & -1\end{bmatrix}$$

Example 38.

$$A = B^{-1}\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}$$

$$A^2 = B^{-1}\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix} BB^{-1}\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix} B$$

$$= B^{-1}\begin{bmatrix}\lambda_1^2 & & \\ & \dots & \\ & & \lambda_n^2\end{bmatrix} B$$

$$A^k = B^{-1}\begin{bmatrix}\lambda_1^k & & \\ & \dots & \\ & & \lambda_n^k\end{bmatrix} B$$

$$e^A = \sum_{k=0}^{\infty}\frac{A^k}{k!} = \sum_{k=0}^{\infty}\frac{1}{k!}B^{-1}\begin{bmatrix}\lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k\end{bmatrix}\cdot B = B^{-1}\begin{bmatrix}e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n}\end{bmatrix} B$$

$$\to \frac{dx}{dt} = A\cdot x \Rightarrow x = e^{At}\cdot x_0$$

Example 39. $\boxed{\text{Leonardo of Pisa (1170–1250)}}$
"Libera Abaci" (1202)

The Arabic approach to a recursive formula $F_n = F_{n-1} + F_{n-2}$ was too theoretical to him. He described it using reproductive bunnies.

What is an explicit formula for F_n?

$$F_{n+1} = F_n F_{n-1}$$

$$F_n = F_n$$

We represent it as matrix:

$$\begin{pmatrix}F_{n+1}\\F_n\end{pmatrix} = \begin{bmatrix}1 & 1\\1 & 0\end{bmatrix}\begin{bmatrix}F_n\\F_{n-1}\end{bmatrix} = A^n\begin{bmatrix}F_1\\F_0\end{bmatrix} = A^n\begin{bmatrix}1\\1\end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\chi_A(x) = \begin{vmatrix} x-1 & -1 \\ -1 & x \end{vmatrix} = x(x-1) - 1 = x^2 - x - 1$$

$$\lambda_{1,2} = \frac{1 \pm \sqrt{5}}{2}$$

Remark: $\frac{1+\sqrt{5}}{2}$ is the golden ratio.

$$\ldots B = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix}$$

$$B^{-1} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ -1 & \frac{1+\sqrt{5}}{2} \end{bmatrix}$$

$$F_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^{n+1} - \left(\frac{1-\sqrt{5}}{2} \right)^{n+1} \right]$$

Hence,

$$\frac{F_{n+1}}{F_n} \xrightarrow[n\to\infty]{} \frac{1+\sqrt{5}}{2}$$

German keywords

English keywords