

# Vorlesungsskript

LinA II\* SoSe 24

# Inhaltsverzeichnis

<b>1. Eigenwerte und Eigenvektoren .....</b>	<b>3</b>
1.1. Definition und grundlegende Eigenschaften .....	3
1.2. Das charakteristische Polynom .....	7
<b>2. Diagonalisierbarkeit und Normalform .....</b>	<b>16</b>
2.1. Diagonalisierbarkeit .....	16
2.2. Dualräume .....	22
2.3. Zyklische $f$ -invariant Unterräume .....	26
2.4. Die Jordan-Normalform .....	30
<b>3. Euklidische und unitäre Vektorräume .....</b>	<b>39</b>
3.1. Skalarprodukt und Normen .....	39
3.2. Winkel und Orthogonalität .....	44
3.3. Selbstadjungierte Abbildungen .....	50
<b>4. Affine Geometrie .....</b>	<b>56</b>
4.1. Operation einer Gruppe auf einer Menge .....	56
4.2. Affine Räume .....	58
4.3. Lagebeziehungen von affinen Unterräumen .....	62

# Definitionen

## 1 .

---

- 1.1: Eigenwert und Eigenvektor
- 1.2: Eigenwert und Eigenvektor
- 1.7: Eigenraum
- 1.10: Geometrische Vielfachheit
- 1.12: Charakteristisches Polynom
- 1.17: ähnliche Matrizen
- 1.20: Algebraische Vielfachheit

- 4.4: transitive Wirkung
- 4.6: affiner Raum
- 4.13: affiner Unterraum
- 4.16: Aufpunkt und Richtung
- 4.17: schwach

## 2 .

---

- 2.1: Diagonalisierbar
- 2.8: Jordan
- 2.9: Linearform, Dualraum
- 2.12: duale Abbildung
- 2.15: nilpotent vom Grad
- 2.16: equation
- 2.17: Bilinearform
- 2.19: Grad von
- 2.20: Krylov

## 3 .

---

- 3.1: Sesquilinearform
- 3.2: Skalarprodukt
- 3.3: hermitesche Matrix
- 3.6: Norm
- 3.9: orthogonal
- 3.14: Orthogonale und unitäre Matrizen
- 3.16: orthogonale Abbildung
- 3.17: linebreak
- 3.22: adjungierter Endorphismus
- 3.23: selbstadjungiert
- 3.27: positiv definite Matrix

## 4 .

---

- 4.1: Wirkung einer Gruppe
- 4.3: Bahn von

## Wiederholung:

$K$  sei ein beliebiger Körper,  $V$  ein  $n$ -dimensionaler  $K$ -Vektorraum,

$$L(V, V) = \{f : V \rightarrow V \mid f \text{ lin. Abbildung}\}$$

$f \in L(V, V)$  heißt Endomorphismus. Ist  $f \in L(V, V)$ , so läßt sich  $f$  bezüglich einer Basis  $B = \{v_1, \dots, v_n\}$  von  $V$  eindeutig durch eine Matrix

$$A_f^{B,B} = (a_{ij})_{1 \leq i, j \leq n} \in K^{n,n}$$

Es gilt

$$f(v_j) = \sum_{i=1}^n a_{ij} v_i \quad 1 \leq j \leq n$$

Abbildung

$$F : L(V, V) \rightarrow K^{n,n}$$

ist ein Isomorphismus.

Basiswechsel? Basen  $B, C$  von  $V$



(siehe Lem. 5.27, LinA I\*)

Eine zentrale Frage: Sei  $f \in L(V, V)$ , existiert eine Basis  $B = \{v_1, \dots, v_n\}$  von  $V$ , so dass  $A_f^{B,B}$  eine möglichst einfache Form besitzt?

z.B. Diagonalmatrix:

$$A_f^{B,B} = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}$$

Wir werden:

- Endomorphismen charakterisieren, die sich durch eine Diagonalmatrix beschreiben lassen.

Wenn ja: Dann gilt  $f(v_j) = \lambda_j v_j$

$\Rightarrow f$  ist eine Streckung von  $v_j$  um den Faktor  $\lambda_j$ .

- Die Jordan-Normalform herleiten.

# 1. Eigenwerte und Eigenvektoren

Eigenwerte charakterisieren zentrale Eigenschaften linearer Abbildungen. Z.B.

- Lösbarkeit von linearen Gleichungssystemen
  - Eigenschaften von physikalischen Systemen
    - gewöhnliche Differentialgleichungen
    - Eigenschwingungen / Resonanzkatastrophe
- Zerstörung einer Brücke über dem Fluß Maine / Millenium-Bridge London

## 1.1. Definition und grundlegende Eigenschaften

### Definition 1.1: Eigenwert und Eigenvektor (Endomorphismus)

Sei  $V$  ein  $K$ -Vektorraum. Ein Vektor  $v \in V, v \neq 0_V$ , heißt **Eigenvektor** von  $f \in L(V, V)$ , falls  $\lambda \in K$  mit

$$f(v) = \lambda v$$

existiert. Der Skalar  $\lambda \in K$  heißt der **Eigenwert** zum Eigenvektor  $v \in V$ .

### Definition 1.2: Eigenwert und Eigenvektor (Matrix)

Sei  $K$  ein Körper und  $n \in \mathbb{N}$ . Ein Vektor  $v \in K^n, v \neq 0_{K^n}$ , heißt Eigenvektor von  $A \in K^{n,n}$ , falls  $\lambda \in K$  mit

$$Av = \lambda v$$

existiert. Der Skalar  $\lambda \in K$  heißt der Eigenwert zum Eigenvektor  $v \in V$ .

### Bemerkungen:

- In Def 1.1 kann  $\dim(V) = \infty$  sein. Dies ist für viele Definitionen/Aussagen in denen wir Endomorphismen betrachten, der Fall.
- Für  $\dim(V) < \infty$  kann man jedes  $f \in L(V, V)$  eindeutig mit einer Matrix  $A$  identifizieren. Dann: Def 1.2 ist Spezialfall von Def 1.1.

- **Achtung:**  $0 \in K$  kann ein Eigenwert sein:

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0 \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Der Nullvektor  $0 \in V$  ist **nie** ein Eigenvektor.

Für  $\dim(V) = 0$  besitzt  $f$  keinen Eigenvektor für  $f \in L(V, V)$ .

- Ist  $v$  Eigenvektor zum Eigenwert  $\lambda$ , so ist auch  $\alpha v$  für jedes  $\alpha \in K \setminus \{0\}$  ein Eigenvektor

$$f(\alpha v) = \alpha f(v) = \alpha \lambda v = \lambda(\alpha v)$$

Zentrale Frage dieses Kapitels:

Existenz von Eigenwerten? Wenn sie existieren: Weitere Eigenschaften?

**Beispiel 1.3:** Sei  $I \subset \mathbb{R}$  ein offenes Intervall und  $V$  der unendlichdimensionale Vektorraum der auf  $I$  beliebig oft differenzierbaren Funktionen. Ein Endomorphismus  $f \in L(V, V)$  ist gegeben durch

$$f(\varphi) = \varphi' \quad \forall \varphi \in V$$

Die Abbildung  $f$  hat jedes  $\lambda \in \mathbb{R}$  als Eigenwert, da für  $c \in \mathbb{R} \setminus \{0\}$  und die Funktion

$$\varphi(x) := c \cdot e^{\lambda x} \neq 0_V \quad \forall x \in I$$

gilt

$$f(\varphi(x)) = f(c \cdot e^{\lambda x}) = \lambda(c e^{\lambda x}) = \lambda \varphi(x)$$

Hier:  $\varphi'(x) = f(\varphi)$  ist eine gewöhnliche Differentialgleichung.

**Beispiel 1.4:** Wir betrachten die lineare Abbildung  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , welche durch

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

definiert ist. Sei  $x$  ein Eigenvektor, dann gilt

$$\begin{aligned} f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) &= \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ \Leftrightarrow x_2 &= \lambda x_1 \quad \text{und} \quad -x_1 = \lambda x_2 \end{aligned}$$

O.B.d.A:  $x_2 \neq 0$

$$\Rightarrow x_2 = \lambda(-\lambda x_2) = -\lambda^2 x_2 \Rightarrow -\lambda^2 = 1, \quad \lambda \in \mathbb{R} \Rightarrow \lambda^2 \geq 0 \Rightarrow -\lambda^2 \leq 0 \quad \nexists$$

D.h.  $f$  besitzt keinen Eigenwert/-vektor. Für  $f : \mathbb{C}^2 \rightarrow \mathbb{C}^2$  ändert sich dies!  $\Rightarrow$  Die Wahl von  $K$  entscheidet!

**Beispiel 1.5:** Wieder  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , diesmal

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} 2x_2 \\ 2x_1 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}}_{=:A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Dann gilt für  $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ,  $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ,  $v_3 = (-1, 1)$  dass  $f(v_1) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ ,  $f(v_2) = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2 \cdot v_2$  und  $f(v_3) = \begin{pmatrix} 2 \\ -2 \end{pmatrix} = (-2) \cdot v_3$ .



Beobachtung:  $\dim(V) = 2$

zwei Eigenwerte:  $2, -2$ , es existieren keine Weiteren,

zwei Eigenvektoren:  $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ,  $v_3 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ , sind linear unabhängig

**Lemma 1.6:** Es sei  $f \in L(V, V)$  ein Endomorphismus. Eigenvektoren zu paarweise verschiedenen Eigenwerten von  $f$  sind linear unabhängig.

*Beweis:* Es seien  $v_1, \dots, v_m$  Eigenvektoren zu den paarweise verschiedenen Eigenwerten  $\lambda_1, \dots, \lambda_m$  von  $f$ . Beweis durch Induktion:

Induktionsanfang:  $m = 1$ ,  $\lambda_1, v_1 \neq 0 \implies v_1$  lin. unabh.

Induktionsschritt:  $m - 1 \rightarrow m$

Induktionsvoraussetzung: Behauptung gelte für  $m - 1$

Betrachte

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_m v_m = 0 \quad (*) \quad \alpha_m \in K$$

EV, f()

$$\implies \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \dots + \alpha_m \lambda_m v_m = 0$$

(\*) ·  $\lambda_m$

$$\implies \lambda_m \alpha_1 v_1 + \lambda_m \alpha_2 v_2 + \dots + \lambda_m \alpha_m v_m = 0$$

Wir bilden die Differenz aus Zeile 1 und 2

$$\underbrace{(\lambda_1 - \lambda_m)}_{\neq 0} \alpha_1 v_1 + \underbrace{(\lambda_2 - \lambda_m)}_{\neq 0} \alpha_2 v_2 + \dots + \underbrace{(\lambda_{m-1} - \lambda_m)}_{\neq 0} \alpha_{m-1} v_{m-1} = 0$$

$v_1, \dots, v_{m-1}$  lin. unabh.  $\implies \alpha_1 = \alpha_2 = \dots = \alpha_{m-1} = 0$  Einsetzen in (\*) liefert

$$\underbrace{\alpha_m v_m}_{\neq 0} = 0 \implies \alpha_m = 0$$

$\Rightarrow v_1, \dots, v_m$  lin unabh.

□

**Folgerung:** Es gibt höchstens  $n = \dim(V)$  verschiedene Eigenwerte für  $n = \dim(V) < \infty$ .

### Definition 1.7: Eigenraum

Ist  $f \in L(V, V)$  und  $\lambda \in K$ , so heißt

$$\text{Eig}(f, \lambda) = \{v \in V \mid f(v) = \lambda v\}$$

der **Eigenraum** von  $f$  bezüglich  $\lambda$ .

Es gilt:

- $\text{Eig}(f, \lambda) \subseteq V$  ist ein Untervektorraum
- $\lambda$  ist Eigenwert von  $f \iff \text{Eig}(f, \lambda) \neq \{0\}$
- $\text{Eig}(f, \lambda) \setminus \{0\}$  ist die Menge der zu  $\lambda$  gehörenden Eigenvektoren von  $f$ .
- $\text{Eig}(f, \lambda) = \ker(f - \lambda \text{Id})$
- $\dim(\text{Eig}(f, \lambda)) = \dim(V) - \text{rg}(f - \lambda \text{Id})$
- Sind  $\lambda_1, \lambda_2 \in K$  verschiedene Eigenwerte, so ist  $\text{Eig}(f, \lambda_1) \cap \text{Eig}(f, \lambda_2) = \{0\}$

Die letzte Aussage kann verallgemeinert werden zu:

**Lemma 1.8:** Sei  $V$  ein  $K$ -Vektorraum mit  $\dim(V) = n < \infty$  und  $f \in L(V, V)$ . Sind  $\lambda_1, \dots, \lambda_m, m \leq n$ , paarweise verschiedene Eigenwerte von  $f$ , so gilt

$$\text{Eig}(f, \lambda_i) \cap \sum_{\substack{j=1 \\ j \neq i}}^m \text{Eig}(f, \lambda_j) = \{0\} \quad \forall i = 1, \dots, m$$

*Beweis:* Summe von Vektorräumen, vgl. Def 3.32 LinA I.

Sei  $i \in \{1, \dots, m\}$  fest gewählt.

$$v \in \text{Eig}(f, \lambda_i) \cap \sum_{\substack{j=1 \\ i \neq j}}^m \text{Eig}(f, \lambda_j)$$

Also ist

$$v = \sum_{\substack{j=1 \\ j \neq i}}^m v_j \quad \text{für } v_j \in \text{Eig}(f, \lambda_j) \quad \text{für } j \neq i$$

$$\Rightarrow -v + \sum_{\substack{j=1 \\ j \neq i}}^m v_j = 0 \quad \text{Aus Lemma 1.6 folgt damit } v = 0.$$

□



Über die Identifikation von Endomorphismen und Matrizen für  $\dim(V) < \infty$  erhält man:

**Korollar 1.9:** Für ein  $n \in \mathbb{N}$  und einem Körper  $K$  sei  $A \in K^{n,n}$ . Dann gilt für jedes  $\lambda \in K$ , dass

$$\dim(\text{Eig}(A, \lambda)) = n - \text{rg}(A - \lambda I_n)$$

Insbesondere ist  $\lambda \in K$  ein Eigenwert von  $A$ , wenn  $\text{rg}(A - \lambda I_n) < n$  ist.

### Definition 1.10: Geometrische Vielfachheit

Ist  $f \in L(V, V)$  und  $\lambda \in K$  ein Eigenwert von  $f$ , so heißt

$$g(f, \lambda) := \dim(\text{Eig}(f, \lambda)) \quad (> 0)$$

die geometrische Vielfachheit des Eigenwerts  $\lambda$ .

## 1.2. Das charakteristische Polynom

Wir bestimmt man Eigenwerte?

**Lemma 1.11:** Seien  $A \in K^{n,n}$  und  $\lambda \in K$ . Dann ist

$$\det(A - \lambda I_n)$$

ein Polynom  $n$ -ten Grades in  $\lambda$ .

*Beweis:* Mit der Leibniz-Formel folgt,

$$\begin{aligned} \det(\underbrace{A - \lambda I_n}_{\tilde{a}_{ij}}) &= \sum_{\sigma \in S_1} \text{sgn}(\sigma) \cdot \tilde{a}_{1\sigma(1)} \cdot \dots \cdot \tilde{a}_{n\sigma(n)} \\ &= \underbrace{(a_{11} - \lambda) \cdot (a_{22} - \lambda) \cdot \dots \cdot (a_{nn} - \lambda)}_{\substack{\sigma = \text{Id} \\ \in \mathcal{P}_n \text{ in } \lambda}} + \underbrace{\sum_{\substack{\sigma \neq \text{Id} \\ \in \mathcal{P}_{n-2} \text{ in } \lambda}} S}_{S_1} \end{aligned}$$

Weiter gilt:

$$(a_{11} - \lambda) \cdot \dots \cdot (a_{nn} - \lambda) = (-1)^n \lambda^n + (-1)^{n-1} \lambda^{n-1} (a_{11} + \dots + a_{nn}) + \underbrace{S_1}_{\in \mathcal{P}_{n-2} \text{ in } \lambda}$$

Insgesamt: Es existieren Koeffizienten  $a_0, \dots, a_n \in K$  mit

$$\det(A - \lambda I_n) = a_n \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_1 \lambda + a_0$$

$$a_n = (-1)^n$$

$$a_{n-1} = (-1)^{n-1} (a_{11} + \dots + a_{nn})$$

man kann zeigen:  $a_0 = \det(A)$

□

Man nennt  $a_{11} + a_{22} + \dots + a_{nn}$  auch die **Spur** von  $A$ .

### Definition 1.12: Charakteristisches Polynom

Sei  $A \in K^{n,n}$  und  $\lambda \in K$ . Dann heißt das Polynom  $n$ -ten Grades

$$P_A(\lambda) := \det(A - \lambda I_n)$$

das charakteristische Polynom zu  $A$ .

**Lemma 1.13:** Sei  $A \in K^{n,n}$  und  $\lambda \in K$ . Der Skalar  $\lambda$  ist genau dann Eigenwert von  $A$ , wenn

$$P_A(\lambda) = 0$$

gilt.

*Beweis:* Die Gleichung

$$Av = \lambda v \iff Av - \lambda v = 0 \iff (A - \lambda I_n)v = 0$$

hat genau eine Lösung  $v \in V, v \neq 0$ , wenn  $\text{rg}(A - \lambda I_n) < n$ , vgl. Satz 6.3 aus LinA I. Dies ist genau dann der Fall, wenn

$$\det(A - \lambda I_n) = 0, \text{ vgl. D10 aus LinA I}$$

□

**Beispiel 1.14:** Eigenwerte und -vektoren von

$$A = \begin{pmatrix} 3 & 8 & 16 \\ 0 & 7 & 8 \\ 0 & -4 & -5 \end{pmatrix}$$

Regel von Sarrus liefert

$$\begin{aligned}
P_A(\lambda) &= \begin{pmatrix} 3-\lambda & 8 & 16 \\ 0 & 7-\lambda & 8 \\ 0 & -4 & -5-\lambda \end{pmatrix} \\
&= (3-\lambda)(-35-7\lambda+5\lambda+\lambda^2+32) \\
&= (3-\lambda)[(7-\lambda)(-5-\lambda)-8(-4)]-8(0-0)+16(0-0) \\
&= (3-\lambda)(\lambda^2-2\lambda-3) = (3-\lambda)(\lambda+1)(\lambda-3)
\end{aligned}$$

$\Rightarrow$  Eigenwerte sind  $\lambda = 3$  und  $\lambda = -1$

Zugehörige Eigenvektoren?

$\lambda = -1$ :

$$\begin{aligned}
Av &= -v \Leftrightarrow (A + I_3)v = 0 \\
\begin{pmatrix} 4 & 8 & 26 \\ 0 & 8 & 8 \\ 0 & -4 & -4 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}$$

LGS lösen:  $\Rightarrow v_2 = -v_3, v_1 = -2v_3$

Damit ist z.B.:  $w_1 = (2, 1, -1)^\top$  Eigenvektor.

$\lambda = 3$ :

$$\begin{aligned}
(A - 3I_3)v &= 0 \Leftrightarrow \\
\begin{pmatrix} 0 & 8 & 16 \\ 0 & 4 & 8 \\ 0 & -4 & -8 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} &= 0 \in \mathbb{R}^3 \Leftrightarrow v_2 + 2v_3 = 0
\end{aligned}$$

Damit sind z.B.:  $w_2 = (1, 2, -1)^\top, w_3 = (-1, 2, -1)$  Eigenvektoren.

$\lambda = -1$ : einfache Nullstelle und  $\dim(\text{Span}(w_1)) = 1$  passt zu  $\text{rg}(A - (-1)I_n) = 2$  und  $\dim(\text{Eig}(A_1 - 1)) = 3 - 2 = 1$ .

$\lambda = -3$ : doppelte Nullstelle und  $\dim(\text{Span}(w_2, w_3)) = 2$  passt zu  $\text{rg}(A - 3I_n) = 1$  und  $\dim(\text{Eig}(A, 3)) = 3 - 1 = 2$

**Lemma 1.15:** Sei  $A \in K^{n,n}$ . Dann gilt

$$p_A(\cdot) = p_{A^\top}(\cdot)$$

D.h. eine Matrix und ihre Transponierte haben die gleichen Eigenwerte.

*Beweis:*

$$p_A(\lambda) = \det(A - \lambda I_n) \stackrel{\text{D12}}{=} \det((A - \lambda I_n)^\top) = \det(A^\top - \lambda I_n) = p_{A^\top}(\lambda)$$

□

**Achtung:** Die Eigenwerte bleiben gleich, aber nicht die Eigenvektoren.

**Beispiel 1.16:** Für die Matrix  $A$  aus Bsp. 1.14 gilt

$$A^\top = \begin{pmatrix} 3 & 0 & 0 \\ 8 & 7 & -4 \\ 16 & 8 & -5 \end{pmatrix} \Rightarrow \det(A^\top - \lambda I_n) = (3 - \lambda)[(7 - \lambda)(-5 - \lambda) + 4 \cdot 8] \\ = -(\lambda - 3)^2(\lambda + 1)$$

Aber

$$\begin{pmatrix} 3 & 0 & 0 \\ 8 & 7 & -4 \\ 16 & 8 & -5 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 6 \\ 27 \\ 45 \end{pmatrix} \neq (-1) \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$$

Man kann ausrechnen:

$$\tilde{w}_1 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \text{ EV zu EW } -1, \tilde{w}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \tilde{w}_3 \text{ EV zu EW } 3$$

Übertragung auf Endomorphismen?

$$p_f(\lambda) \quad f \in L(V, V), B \text{ Basis} \Rightarrow \exists! A_f^{B,B}, C \text{ Basis} \Rightarrow \exists! A_f^{C,C}$$

$$p_{A_f^{B,B}}(\lambda) \stackrel{?}{=} p_{A_f^{C,C}}(\lambda)$$

### Definition 1.17: ähnliche Matrizen

Zwei Matrizen  $A, B \in K^{n,n}$  heißen **ähnlich**, wenn es eine Matrix  $T \in \text{GL}_n(K)$  gibt, so dass  $A = TBT^{-1}$  gilt.

Man kann leicht beweisen, dass die Ähnlichkeit von Matrizen eine Äquivalenzrelation auf der Menge der quadratischen Matrizen ist.

Mit  $\det(A^{-1}) \stackrel{\text{D11}}{=} (\det(A))^{-1}$  folgte für zwei ähnliche Matrizen  $A$  und  $B$ , dass

$$\det(A) = \det(TBT^{-1}) = \det(T) \det(B) \det(T^{-1}) = \det(B)$$

**Beispiel 1.18:** Sei  $f \in L(\mathbb{R}^3, \mathbb{R}^3)$ , d.h.  $V = \mathbb{R}^3$ , gegeben durch

$$f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} x_1 \\ -4x_1 + 7x_2 \\ 3x_1 + 5x_2 + 3x_3 \end{pmatrix}$$

Wir betrachten für den  $\mathbb{R}^3$  die Basen

$$E = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\},$$

$$B = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\},$$

$$C = \left\{ \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} \right\}$$

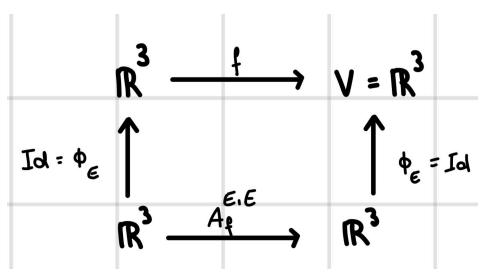
Für darstellende Matrix von  $f$  bezüglich der Standardmatrix  $E$  erhalten wir aus Satz 5.18, LinA I,

$$f(e_j) = \sum_{i=1}^3 a_{ij} e_i \quad \forall j \in \{1, 2, 3\}$$

dass

$$A_f^{E,E} = \begin{pmatrix} 1 & 0 & 0 \\ -4 & 7 & 0 \\ 3 & 5 & 3 \end{pmatrix}$$

Das zugehörige kommutative Diagramm ist gegeben durch



Für die Basis  $B$  erhalten wir

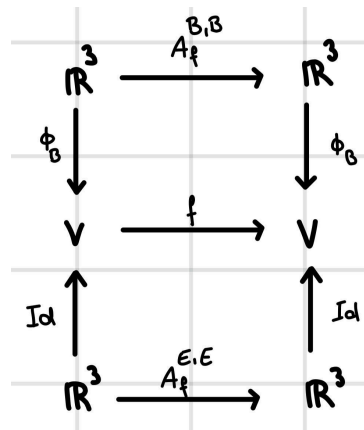
$$f\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ -4 \\ 3 \end{pmatrix} = 5\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (-7)\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 3\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$f\left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 3 \\ 8 \end{pmatrix} = (-2)\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (-5)\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 8\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$f\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 3 \\ 11 \end{pmatrix} = (-2)\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (-8)\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 11\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\Rightarrow A_f^{B,B} = \begin{pmatrix} 5 & -2 & -2 \\ -7 & -5 & -8 \\ 3 & 8 & 11 \end{pmatrix}$$

Herleitung bezüglich Matrizen?



Koordinatenabbildung  $\Phi_B$ ?

Abbildung vom  $\mathbb{R}^3$  + Standardbasis  $E$  in den  $V (= \mathbb{R}^3)$  + Basis  $B$ .

$$\Phi_B = (e_i) = v_i \quad \text{für} \quad B = \{v_1, v_2, v_3\}$$

$$\Rightarrow A_{\Phi_B}^{E,B} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Damit folgt insgesamt:

$$A_f^{B,B} = (A_{\Phi_B}^{E,B})^{-1} I_n A_f^{E,E} I_n^{-1} A_{\Phi_B}^{E,B} = (A_{\Phi_B}^{E,B})^{-1} A_f^{E,E} \underbrace{A_{\Phi_B}^{E,B}}_{\in \text{GL}_n(\mathbb{R})}$$

$$\Rightarrow A_f^{B,B} \text{ und } A_f^{E,E} \text{ sind ähnlich}$$

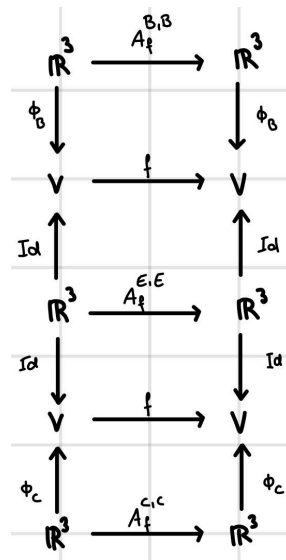
Für die Basis  $C$  erhalten wir

$$\begin{aligned} f\left(\begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}\right) &= \begin{pmatrix} 0 \\ 0 \\ -3 \end{pmatrix} = 3 \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} + 0 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} \\ f\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) &= \begin{pmatrix} 1 \\ -4 \\ 3 \end{pmatrix} = (-3) \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} + 1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 4 \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} \\ f\left(\begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}\right) &= \begin{pmatrix} 0 \\ -7 \\ -5 \end{pmatrix} = 5 \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} + 0 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 7 \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} \end{aligned}$$

Als Darstellungsmatrix erhält man

$$A_f^{C,C} = \begin{pmatrix} 3 & -3 & 5 \\ 0 & 1 & 0 \\ 0 & 4 & 7 \end{pmatrix}$$

Als Matrizenmultiplikation



Darstellung von  $\Phi_C$ ?  $\Phi_C(e_i) = w_i$  für  $C = \{w_1, w_2, w_3\}$

$$A_{\Phi_C}^{E,C} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix}$$

$$A_f^{C,C} = (A_{\Phi_C}^{E,C})^{-1} I_n A_f^{E,E} I_n^{-1} A_{\Phi_C}^{E,C} = (A_{\Phi_C}^{E,C})^{-1} A_f^{E,E} A_{\Phi_C}^{E,C}$$

Also auch:  $A_f^{C,C}$  ist ähnlich zu  $A_f^{E,E}$ .

Alternativ:

$$\begin{aligned} A_f^{C,C} &= (A_{\Phi_C}^{E,C})^{-1} I_n I_n^{-1} A_{\Phi_B}^{E,B} A_f^{B,B} (A_{\Phi_B}^{E,B})^{-1} I_n A_{\Phi_C}^{E,C} \\ &= \underbrace{(A_{\Phi_C}^{E,C})^{-1} A_{\Phi_B}^{E,B} A_f^{B,B} (A_{\Phi_B}^{E,B})^{-1}}_{\in \text{GL}_n(\mathbb{R})} A_{\Phi_C}^{E,C} \end{aligned}$$

Jetzt allgemein:  $f \in L(V, V)$ ,  $\dim(V) < \infty$ ,  $B, C$  seien Basen von  $V \Rightarrow$

$$A := A_f^{B,B} \quad \tilde{A} := A_f^{C,C}$$

und es existiert  $T \in \text{GL}_n(K)$  als Basistransformationsmatrix, so dass

$$\tilde{A} = T A T^{-1}$$

Dann gilt

$$\begin{aligned} p_{\tilde{A}}(\lambda) &= \det(\tilde{A} - \lambda I_n) = \det(T A T^{-1} - \lambda T T^{-1}) \\ &= \det(T(A - \lambda I_n)T^{-1}) \\ &= \det(T) \det(A - \lambda I_n) \det(T^{-1}) \\ &= p_A(\lambda) \end{aligned}$$

D.h. für einen Endomorphismus ist das charakteristische Polynom der zugehörigen Darstellungsmatrix unabhängig von der Wahl der Basis!

Damit ist es sinnvoll, für  $f \in L(V, V)$ ,  $\dim(V) < \infty$ ,

$$p_f(\cdot) := p_A(\cdot)$$

für  $A$  als Darstellungsmatrix  $A_f^{B,B}$  für eine Basis  $B$ .

**Lemma 1.19:** Sei  $V$  ein  $K$ -Vektorraum mit  $\dim(V) = n < \infty$  und  $f \in L(V, V)$ . Dann sind folgende Aussagen äquivalent:

1.  $\lambda \in K$  ist ein Eigenwert von  $f$ .
2.  $\lambda \in K$  ist ein Eigenwert der Darstellungsmatrix  $A_f^{B,B}$  für eine gewählte  $B$  von  $V$ .

Des weiteren gilt auch. Für zwei ähnliche  $A$  und  $B$  gilt  $p_A(\lambda) = p_B(\lambda)$

$$A, B \text{ ähnlich} \implies p_A(\lambda) = p_B(\lambda)$$

z.B.

$$A = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$p_A(\lambda) = (1 - \lambda)^2 = p_B(\lambda), \text{ aber für jedes } T \in \text{GL}_2(\mathbb{R}) \text{ gilt}$$

$$TBT^{-1} = TT^{-1} = I \neq A \text{ also } A, B \text{ nicht ähnlich}$$

Weitere Beobachtung: Aus Lemma 1.13 und Lemma 1.19 folgt, dass die Eigenwerte von  $f \in L(V, V)$  die Nullstellen des charakteristischen Polynoms der Matrix  $A_f^{B,B}$  für eine Basis  $B$  ist. Dies gilt **nicht** i.a. für Darstellungsmatrizen  $A_f^{B,C}$  für  $B \neq C$ .

### Definition 1.20: Algebraische Vielfachheit

Sei  $V$  ein  $K$ -Vektorraum mit  $\dim(V) = n < \infty$ . Ist  $f \in L(V, V)$  und  $\tilde{\lambda}$  ist Eigenwert von  $f$  hat das charakteristische Polynom  $p_f(\lambda)$  die Form

$$p_f(\lambda) = (\lambda - \tilde{\lambda})^d \cdot \tilde{p}(\lambda)$$

für ein  $\tilde{p}(\cdot) \in \mathbb{K}[\lambda]$  mit  $\tilde{p}(\tilde{\lambda}) \neq 0$ , so nennt man  $d$  die **algebraische Vielfachheit** von  $\tilde{\lambda}$  und bezeichnet sie  $a(f, \tilde{\lambda})$ .

**Lemma 1.21:** Seien  $V$  ein  $K$ -Vektorraum,  $\dim(V) = n < \infty$ , und  $f \in L(V, V)$ . Für Eigenwert  $\tilde{\lambda}$  von  $f$  gilt

$$g(f, \tilde{\lambda}) \leq a(f, \tilde{\lambda})$$



*Beweis:* Ist  $\tilde{\lambda}$  EW von  $f$  mit der geometrischen Vielfachheit  $m := g(f, \tilde{\lambda})$ , so gibt es nach Def. 1.10 zu  $\tilde{\lambda}$   $m$  linear unabhängige Eigenvektoren  $v_1, \dots, v_m \in V$ .

Gilt  $m = n = \dim(V)$  sind  $\{v_1, \dots, v_m\}$  schon Basis von  $V$ .

Gilt  $m < n$ , so folgt aus dem Basisergänzungssatz (Satz 3.21, LinA I), dass man  $\{v_1, \dots, v_m\}$  zu einer Basis  $\{v_1, \dots, v_m, v_{m+1}, \dots, v_n\} =: B$  ergänzen. Wegen  $f(v_j) = \tilde{\lambda}v_j, 1 \leq j \leq m$ , gilt

$$A_f^{B,B} = \begin{pmatrix} \tilde{\lambda}I_n & A_1 \\ 0 & A_2 \end{pmatrix} \in K^{n,n}$$

für zwei Matrizen  $A_1 \in K^{m,n-m}, A_2 \in K^{n-m,n-m}$ .

Mit D9 aus LinA I folgt

$$p_f(\lambda) = (\tilde{\lambda} - \lambda)^m \cdot \det(A_2 - \lambda I_{n-m,n-m})$$

$\Rightarrow$  EW  $\tilde{\lambda}$  ist mindestens  $m$ -fache Nullstelle von  $p_f(\lambda)$ . Für  $m = n \Rightarrow A_f^{B,B} = \tilde{\lambda}I_n \Rightarrow p_f(\lambda) = (\tilde{\lambda} - \lambda)^n$

□

## 2. Diagonalisierbarkeit und Normalform

### 2.1. Diagonalisierbarkeit

#### Definition 2.1: Diagonalisierbar

Sei  $V$  ein  $K$ -Vektorraum mit  $\dim(V) = n < \infty$ . Ein  $f \in L(V, V)$  heißt **diagonalisierbar**, wenn es eine Basis  $B$  von  $V$  gibt, so dass  $A_f^{B,B}$  eine Diagonalmatrix ist. D.h. es existieren  $\lambda_1, \dots, \lambda_n \in K$  mit

$$A_f^{B,B} = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix} \in K^{n,n}$$

Entsprechend nennen wir eine Matrix  $A \in K^{n,n}$  **diagonalisierbar**, wenn es eine Matrix  $T \in \text{GL}_n(K)$  und eine Diagonalmatrix  $D \in K^{n,n}$  gibt mit

$$A = TDT^{-1}$$

D.h.  $A$  ist ähnlich zu einer Diagonalmatrix.

**Satz 2.2:** Sei  $V$  ein  $K$ -Vektorraum mit  $\dim(V) = n < \infty$  und  $f \in L(V, V)$ . Dann sind folgende Aussagen äquivalent:

1.  $f$  ist diagonalisierbar
2. Es gibt eine Basis  $B$  von  $V$  bestehend aus Eigenvektoren von  $f$ .
3. Das charakteristische Polynom  $p_f(\cdot)$  zerfällt in  $n$  Linearfaktoren über  $K$ , d.h.

$$p_f(\lambda) = (\lambda - \lambda_1) \cdot \dots \cdot (\lambda - \lambda_n)$$

mit Eigenwerten  $\lambda_1, \dots, \lambda_n \in K$  für  $f$  und für jeden Eigenwert  $\tilde{\lambda}$  gilt  $a(f, \tilde{\lambda}) = g(f, \tilde{\lambda})$ .

*Beweis:*

“1  $\Rightarrow$  2”:  $f$  diagonalisierbar  $\Rightarrow \exists \{v_1, \dots, v_n\} = B$  Basis von  $V$ ,  $\lambda_1, \dots, \lambda_n \in K$ :

$$\tilde{A} := A_f^{B,B} = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix} \quad f(v_j) = \sum_{i=1}^n a_{ij} v_i$$

$\Rightarrow f(v_j) = \lambda_j v_j, 1 \leq j \leq n, v_j \neq 0. \Rightarrow$  Damit sind  $\lambda_1, \dots, \lambda_n$  Eigenwerte von  $f$  mit zugehörigen Eigenvektoren  $v_1, \dots, v_n. \Rightarrow 2$ .

“2  $\Rightarrow$  1”: Ist  $B = \{v_1, \dots, v_n\}$  eine Basis von  $V$  bestehend aus Eigenvektoren, so gibt es zugehörige Eigenwerte  $\lambda_1, \dots, \lambda_n$  mit  $f(v_j) = \lambda_j v_j, 1 \leq j \leq n \Rightarrow$

$$A_f^{B,B} = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}$$

“2  $\Rightarrow$  3”: Sei  $B = \{v_1, \dots, v_n\}$  eine Basis von Eigenvektoren,  $\lambda_1, \dots, \lambda_n$  seien die zugehörigen Eigenwerte  $\Rightarrow$

$$\begin{aligned} p_f(\lambda) &= p_{A_f^{B,B}}(\lambda) = \det(A_f^{B,B} - \lambda I_n) \\ &= (\lambda_1 - \lambda) \cdot (\lambda_2 - \lambda) \cdot \dots \cdot (\lambda_n - \lambda) \end{aligned}$$

$\Rightarrow p_f(\cdot)$  zerfällt in Linearfaktoren. Verschiedene Eigenwerte  $\tilde{\lambda}_1, \dots, \tilde{\lambda}_k, k \leq n$ . Der Eigenwert  $\tilde{\lambda}_i$  besitzt die algebraische Vielfachheit  $m_j := a(f, \tilde{\lambda}_j)$  genau dann, wenn er  $m_j$ -mal auf den Diagonalen von  $A_f^{B,B}$  steht. Dies ist genau dann der Fall, wenn  $m_j$  Eigenvektoren zu  $\tilde{\lambda}_j$  in  $B$  enthalten sind. Diese sind linear unabhängig  $\Rightarrow$

$$1. \dim(\text{Eig}(f, \tilde{\lambda}_j)) = g(f, \tilde{\lambda}_j) \geq m_j = a(f, \tilde{\lambda}_j)$$

$$2. \text{ Lemma 1.21: } g(f, \tilde{\lambda}_j) \leq a(f, \tilde{\lambda}_j)$$

$$1 \wedge 2 \Rightarrow g(f, \tilde{\lambda}_j) = a(f, \tilde{\lambda}_j)$$

“3  $\Rightarrow$  2”: Seien  $\tilde{\lambda}_1, \dots, \tilde{\lambda}_k, k \leq n$  die paarweise verschiedenen Eigenwerte von  $f$ . Wir wissen:  $p_f(\cdot)$  zerfällt in Linearfaktoren,  $a(f, \tilde{\lambda}_j) = g(f, \tilde{\lambda}_j), 1 \leq j \leq n$ .

$$\dim(V) = n = \sum_{j=1}^k a(f, \tilde{\lambda}_j) = \sum_{j=1}^k g(f, \tilde{\lambda}_j) = \sum_{j=1}^k \dim(\text{Eig}(f, \tilde{\lambda}_j))$$

Es gilt (Lemma 1.8):

$$\text{Eig}(f, \tilde{\lambda}_j) \cap \sum_{i=1}^k \text{Eig}(f, \tilde{\lambda}_i) = 0 \quad \forall j = 1, \dots, k$$

Dann folgt (Lemma 3.31, (2), Lemma 3.35, Satz 3.14) (direkte Summe,  $U \subset V$  UVR  $\Rightarrow \dim(U) \leq \dim(V)$ ,  $U = V \Leftrightarrow \dim(U) = \dim(V)$ , Basis  $\Leftrightarrow$  eindeutige Darstellung), dass die zu  $\tilde{\lambda}_1, \dots, \tilde{\lambda}_n$  linear unabhängigen Eigenvektoren, die jeweils eine Basis von  $\text{Eig}(f, \tilde{\lambda}_j), 1 \leq j \leq k$ , eine Basis von  $V$  bilden.

□

In Verbindung mit Lemma 1.6 folgt unmittelbar:

**Korollar 2.3:** Sei  $V$  ein  $K$ -Vektorraum mit  $\dim(V) = n < \infty$  und  $f \in L(V, V)$  mit  $n$  paarweise verschiedenen Eigenwerten, dann ist  $f$  diagonalisierbar.

**Bemerkung:** Das Kriterium der  $n$  paarweise verschiedenen Eigenwerte ist nicht notwendig z.B.  $V = K^n$ ,  $B = E$  Standardbasis

$$f : \text{Id} : K^n \rightarrow K^n, \implies A_f^{E,E} = I_n \implies 1n\text{-facher Eigenwert}$$

**Beispiel 2.4:** Fortsetzung von Bsp. 1.14

$$A = \begin{pmatrix} 3 & 8 & 16 \\ 0 & 7 & 8 \\ 0 & -4 & -5 \end{pmatrix}, \text{EW: } -1, 3$$

$$w_1 = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} \text{ EV zu } -1, \quad w_2 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \quad w_3 = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} \text{ EV zu } 3$$

$\implies \exists$  Basis von Eigenvektoren  $\xrightarrow{\text{Satz 2.2}} A$  ist diagonalisierbar

$$p_A(\lambda) = (3 - \lambda)(\lambda + 1)(\lambda - 3)$$

$$a(f, -1) = 1 = g(f, -1)$$

$$a(f, 3) = 2 = g(f, 3)$$

$T \in \text{GL}_n(\mathbb{R})$  so, dass  $T^{-1}AT = D$ ?

Die zu  $B = \{w_1, w_2, w_3\}$  gehörende Koordinatentransformation  $\Phi_B$  ist gegeben durch

$$A_{\Phi_B}^{E,B} = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 2 \\ -1 & -1 & -1 \end{pmatrix}$$

Dann gilt: Für  $f \in L(\mathbb{R}^3, \mathbb{R}^3)$  mit

$$A_f^{E,E} = A \quad A_f^{B,B} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = D$$

Mit Basiswechsel von  $A$  zu  $D$

$$D = (A_{\Phi_B}^{E,B})^{-1} \underbrace{A A_{\Phi_B}^{E,B}}_{=T}$$

**Beispiel 2.5:** Nicht jeder Endomorphismus bzw. jede Matrix ist diagonalisierbar. Bsp. 1.4:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \overbrace{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}^A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad p_f(\lambda) = \lambda^2 + 1$$

D.h. über  $\mathbb{R}$  zerfällt  $p_f(\cdot)$  nicht in Linearfaktoren.

## Ein weiteres Beispiel

$$A = \begin{pmatrix} 5 & 10 & 7 \\ 0 & -3 & -3 \\ 0 & 3 & 3 \end{pmatrix}$$

$\Rightarrow p_A(\lambda) = (5 - \lambda)\lambda^2 \Rightarrow p_A(\cdot)$  zerfällt in Linearfaktoren.  $a(f, \lambda_i), g(f, \lambda_i)$  für  $\lambda_1 = 5, \lambda_2 = 0$ . Lemma 1.21:  $g(f, \lambda_i) \leq a(f, \lambda_i) \Rightarrow g(f, 5) = 1 = a(f, 5)$ ,  $a(f, 0) = 2, g(f, 0) \geq 1$  Ein Eigenvektor zu  $\lambda = 0$  sind

$$w_1 = \begin{pmatrix} 3 \\ -5 \\ 5 \end{pmatrix} \Rightarrow g(f, 0) = 1 < 2 = a(f, 0)$$

$\Rightarrow f$  nicht diagonalisierbar.

Mit Satz 2.2 erhält man einen Algorithmus zur Überprüfung, ob ein gegebenes  $f \in L(V, V)$  (bzw.  $A \in K^{n,n}$ ) diagonalisierbar ist:

1. Bestimme mit einer Basis  $B$  von  $V$  die Darstellungsmatrix  $A = A_f^{B,B}$
2. Bestimme für  $A$  das charakteristische Polynom  $p_A(\cdot)$  (Determinantenberechnung)
3. Zerfällt  $p_A(\cdot)$  in Linearfaktoren über  $K$ ? Nein:  $f$  nicht diagonalisierbar. Ja: Seien  $\lambda_i, 1 \leq i \leq k \leq n = \dim(V)$  die paarweise verschiedene Eigenwerte von  $f$ .

Für  $i = 1, \dots, k$

1. Bestimme eine Basis von  $\text{Eig}(f, \lambda_i)$
2. Prüfe, ob  $a(f, \lambda_i) = g(f, \lambda_i)$

Gilt  $a(f, \lambda_i) = g(f, \lambda_i)$  für alle  $i \in \{1, \dots, k\}$ . Nein:  $f$  ist nicht diagonalisierbar. Ja:  $f$  ist diagonalisierbar.

**Beispiel 2.6:** Fischer/Springborn

Betrachtet wird: Masse aufgehängt an einer Feder. Zur Zeit  $t = 0$  in Position  $y(0) = \alpha$  und ausgelenkt in senkrechter Richtung mit Geschwindigkeit  $\beta = \dot{y}(0)$

$y(t) \hat{=}$  Position der Masse zum Zeitpunkt  $t$



Dieses System wird durch die gewöhnliche Differentialgleichungen

$$\ddot{y} + 2\mu\dot{y} + \omega^2 y = 0, \quad y(0) = \alpha, \dot{y}(0) = \beta$$

## Umschreiben

$$\dot{y}_0 = y_1$$

$$\dot{y}_1 = -\omega^2 y_0 - 2\mu y_1$$

mit  $y_0 = y, \ddot{y}_0 = \ddot{y}, y_0(0) = \alpha, y_1(0) = \beta$ .

$$\dot{\tilde{y}} := \begin{pmatrix} \dot{y}_0 \\ \dot{y}_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\omega^2 & -2\mu \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}$$

$$p_A(\lambda) = \lambda^2 + 2\mu\lambda + \omega^2$$

mit den potentiellen Nulstellen

$$\lambda = -\mu \pm \sqrt{\mu^2 - \omega^2}$$

Man unterscheidet drei Fälle:

- $0 \leq \mu < \omega$ , d.h.  $\mu^2 - \omega^2 < 0 \implies$  schwache Dämpfung
- $\mu = \omega$ , d.h.  $\mu^2 = \omega^2 \implies$  aperiodischer Fall  $\implies a(A, -\mu) = 2$ ,  
 $\dim(\text{Eig}(A, -\mu)) = 1$ ,  $A$  nicht diagonalisierbar
- $\mu > \omega$ , d.h.  $\mu^2 > \omega^2$ , starke Dämpfung

Eine solche Eigenwertanalyse kann auch nutzen, um das Langzeitverhalten von Lösungen von gewöhnlichen DGL zu bestimmen.



**Satz 2.7:** Sei  $V$  ein  $K$ -Vektorraum mit  $\dim(V) = n < \infty$  und  $f \in L(V, V)$ . Dann sind folgende Aussagen äquivalent:

1. Das charakteristische Polynom  $p_f(\cdot)$  zerfällt über  $K$  in Linearfaktoren.
2. Es gibt eine Basis  $B$  von  $V$ , so dass  $A_f^{B,B}$  eine obere Dreiecksmatrix ist, d.h.

$$A_f^{B,B} = \begin{pmatrix} 1 & \dots & * \\ \vdots & \ddots & \vdots \\ 0 & \dots & * \end{pmatrix}$$

und  $f$  ist damit **triangulierbar**.

*Beweis:* Beweis von Satz 14.18 im Liesen/Mehrmann

□

Nun ist das Ziel:

Bestimmung einer Basis  $B$  von  $V$ , so dass  $A_f^{B,B}$  eine obere Dreiecksmatrix ist, die möglichst nah an einer Diagonalmatrix ist und von der geometrischen Vielfachheiten der Eigenwerte abgelesen werden können.

D.h.  $p_f(\cdot)$  zerfällt in Linearfaktoren mit den Eigenwerten  $\lambda_1, \dots, \lambda_k$  (notwendig, Satz 2.7) und wir wollen eine Basis  $B$  bestimmen, so dass  $A_f^{B,B}$  Diagonalblockgestalt hat mit

$$A_f^{B,B} = \begin{pmatrix} J_1(\lambda_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & J_m(\lambda_m) \end{pmatrix}$$

wobei jeder Diagonalblock die Form

$$J_i(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & \lambda_i \end{pmatrix} \in K^{d_i, d_i} \quad (*)$$

### Definition 2.8: Jordan-Block

Sei  $V$  ein  $K$ -Vektorraum mit  $\dim(V) = n < \infty$ ,  $f \in L(V, V)$  und  $\lambda_i$  ein Eigenwert von  $f$ . Eine Matrix der Form  $(*)$  heißt **Jordan-Block** der Größe  $d_i$  zum Eigenwert  $\lambda_i$ .

Wegen der Bedeutung der Jordan-Normalform gibt es zahlreiche Herleitungen mit unterschiedlichen mathematischen Hilfsmitteln.

Hier: Beweis über die Dualitätstheorie basierend auf einer Arbeit von V. Pt  $\bar{a}$  k (1956)

## 2.2. Dualräume

### Definition 2.9: Linearform, Dualraum

Sei  $V$  ein  $K$ -Vektorraum. Eine Abbildung  $f \in L(V, K)$  heißt **Linearform**. Den  $K$ -Vektorraum  $V^* := L(V, K)$  nennt man **Dualraum**.

Gilt  $\dim(V) = n < \infty$  so folgt aus Satz 5.18 LinA I, dass  $\dim(V^*) = n$  gilt. Ist  $B = \{v_1, \dots, v_n\}$  eine Basis von  $V$  und  $C = \{1\}$  eine Basis des  $K$ -Vektorraum  $K$ , dann gilt für

$$f(v_i) = \mu_i \in K \quad \text{für } f \in V^*, \text{ d.h. } f : V \rightarrow K,$$

für  $i = 1, \dots, n$  und damit

$$A_f^{B,C} = (\mu_1, \dots, \mu_n) \in K^{1,n}$$

**Beispiel 2.10:** Sei  $V$  der  $\mathbb{R}$ -Vektorraum der auf dem Intervall  $[0, 1]$  stetigen, reellwertigen Funktionen und  $a \in [0, 1]$ . Dann sind

$$g_1 : V \rightarrow \mathbb{R}, \quad g_1(f) := \int_0^1 f(x) dx$$

$$g_2 : V \rightarrow \mathbb{R}, \quad g_2(f) := f(a)$$

Linearformen auf  $V$ .

Basis des Dualraums?

**Satz 2.11:** Sei  $V$  ein  $K$ -Vektorraum mit  $\dim(V) = n < \infty$  und  $B = \{v_1, \dots, v_n\}$  eine Basis von  $V$ . Dann gibt es genau eine Basis  $B^* = \{v_1^*, \dots, v_n^*\}$  von  $V^* = L(V, K)$  für die

$$v_i^*(v_j) = \delta_{ij} \quad i, j = 1, \dots, n$$

gilt. Diese Basis heißt die zu  $B$  duale Basis.

*Beweis:* Lemma 4.10: LinA I. Es gibt eine lineare Abbildung  $v_i^*$  für die  $v_i^*(v_j) = \delta_{ij}$  für  $j = 1, \dots, n$ , für  $i = 1, \dots, n$ . Noch zu zeigen:  $v_i^*$  sind Basis von  $V^*$ . Wir wissen schon:  $\dim(V^*) = n$ . Also: Es reicht zu zeigen:  $\{v_i^*\}_{i=1, \dots, n}$  linear unabhängig. Seien  $\mu_i \in K$  so, dass

$$\sum_{i=1}^n \mu_i v_i^* = 0 \in V^* = L(V, K)$$

Dann gilt:



$$0_K = 0_{V^*}(v_j) = \sum_{i=1}^n \mu_i v_i^*(v_j) = \mu_j \quad j = 1, \dots, n$$

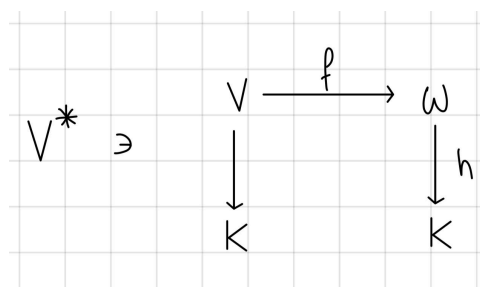
□

**Definition 2.12: duale Abbildung**

Seien  $V$  und  $W$  zwei  $K$ -Vektorräume mit den zugehörigen Dualräumen  $V^*$  und  $W^*$ . Für  $f \in L(V, W)$  heißt

$$f^* : W^* \rightarrow V^*, \quad f^*(h) = h \circ f$$

die zu  $f$  **duale Abbildung**.



Seien  $U \subseteq V$  und  $Z \subseteq V^*$  zwei Unterräume. Dann heißt die Menge

$$U^0 := \{h \in V^* \mid h(u) = 0 \text{ für alle } u \in U\}$$

**Annihilator** von  $U$  und die Menge

$$Z^0 := \{v \in V \mid z(v) = 0 \text{ für alle } z \in Z\}$$

**Annihilator** von  $Z$ .

Man kann sich überlegen:

- Die Mengen  $U^0 \subseteq V^*$  und  $Z^0 \subseteq V$  sind Untervektorräume von  $V^*$  bzw  $V$
- Es gilt für  $f \in L(V, V)$

$$(f^k)^* = (f^*)^k$$

Des Weiteren besitzt die duale Abbildung folgende Eigenschaften:

**Lemma 2.13:** Sind  $V, W$  und  $X$  drei  $K$ -Vektorräume. Dann gilt

1. Ist  $f \in L(V, W)$ , dann ist die duale Abbildung  $f^*$  linear, d.h.  $f^* \in L(W^*, V^*)$
2. Ist  $f \in L(V, W)$  und  $g \in L(W, X)$ , dann ist  $(g \circ f)^* \in L(X^*, V^*)$  und es gilt  $(g \circ f)^* = f^* \circ g^*$
3. Ist  $f \in L(V, W)$  bijektiv, dann ist  $f^* \in L(W^*, V^*)$  bijektiv und es gilt  $(f^*)^{-1} = (f^{-1})^*$

*Beweis:* ÜB



**Lemma 2.14:** Sei  $V$  ein endlichdimensionaler Vektorraum,  $f \in L(V, V)$ ,  $f^* \in L(V^*, V^*)$  und  $U \subseteq V$ , sowie  $W \subseteq V^*$  zwei Vektorräume. Dann gilt:

1.  $\dim(V) = \dim(W) + \dim(W^0)$
2. Ist  $f$  nilpotent vom Grad  $m$ , dann ist die duale Abbildung  $f^*$  ebenfalls nilpotent vom Grad  $m$ .
3. Ist  $W \subseteq V^*$  ein  $f^*$ -invarianter Vektorraum, dann ist  $W^0$  ein  $f$ -invarianter Unterraum.

*Beweis:* ÜA



### Definition 2.15: nilpotent vom Grad $m$

Sei  $\{0\} \neq V$  ein  $K$ -Vektorraum. Man nennt  $f \in L(V, V)$  **nilpotent**, wenn ein  $m \in \mathbb{N}$  existiert, so dass  $f^m = 0 \in L(V, V)$  gilt. Gilt für dieses  $m$ , dass  $f^{m-1} \neq 0 \in L(V, V)$ , so heißt  $f$  **nilpotent vom Grad  $m$**  und  $m$  ist der **Nilpotenzindex** von  $f$ .

### Definition 2.16: $f$ -invarianter Unterraum

Sei  $V$  ein  $K$ -Vektorraum mit  $\dim(V) = n < \infty$ ,  $U \subseteq V$  ein Unterraum und  $f \in L(V, V)$ . Gilt  $f(U) \subseteq U$ , d.h. ist  $f(u) \in U$  für alle  $u \in U$ , so nennt man  $U$  einen  $f$ -invarianten Unterraum von  $V$ .

**Definition 2.17: Bilinearform**

Seien  $V$  und  $W$  zwei  $K$ -Vektorräume. Eine Abbildung  $a : V \times W \rightarrow K$  heißt Bilinearform, wenn

1.  $a(\cdot, w) : V \rightarrow K$  für alle  $w \in W$  eine lineare Abbildung ist und
2.  $a(v, \cdot) : W \rightarrow K$  für alle  $v \in V$  eine lineare Abbildung ist

Eine Bilinearform  $a(\cdot, \cdot)$  heißt **nicht ausgeartet** in der ersten Variable, wenn aus

$$a(v, w) = 0 \quad \text{für alle } w \in W$$

folgt, dass  $v = 0$  ist. Eine Bilinearform heißt nicht ausgeartet in der zweiten Variable, wenn aus

$$a(v, w) = 0 \quad \text{für alle } v \in V$$

folgt, dass  $w = 0$  ist. Falls  $a(\cdot, \cdot)$  in beiden Variablen nicht ausgeartet ist, so nennt man  $a(\cdot, \cdot)$  eine **nicht ausgeartete Bilinearform** und die Räume  $V, W$  ein **duales Paar von Räumen** oder **duales Raumpaar** bezüglich  $a(\cdot, \cdot)$ . Ist  $V = W$ , so heißt  $a(\cdot, \cdot)$  eine **Bilinearform auf  $V$** . Eine Bilinearform  $a(\cdot, \cdot)$  auf  $V$  heißt **symmetrisch**, wenn  $a(v, w) = a(w, v)$  für alle  $v, w \in V$ , ansonsten heißt  $a(\cdot, \cdot)$  unsymmetrisch.

**Bemerkung:** Damit  $V, W$  ein duales Raumpaar für eine nicht ausgeartete Bilinearform bilden können, muss  $\dim(V) = \dim(W)$  gelten.

**Lemma 2.18:** Sei  $V$  ein endlichdimensionaler  $K$ -Vektorraum,  $f \in L(V, V)$ ,  $f^* \in L(V^*, V^*)$  die duale Abbildung zu  $f$ ,  $U \subseteq V$  und  $W \subseteq V^*$  zwei Untervektorräume. Ist die Bilinearform

$$a : U \times W \rightarrow K, (v, h) \mapsto h(v)$$

nicht ausgeartet ist, d.h. sind  $U$  und  $W$  ein duales Raumpaar bezüglich dieser Bilinearform, so ist

$$V = U \oplus W^0$$

*Beweis:* Sei  $u \in U \cap W^0$ . Dann gilt  $h(u) = 0$  für alle  $h \in W$ . Weil  $U, W$  ein duales Raumpaar bzgl.  $a(\cdot, \cdot)$  bilden, folgt  $u = 0$ . Außerdem  $\dim(U) = \dim(W)$  gelten. Damit folgt aus Lemma 2.14, 1., dass

$$\begin{aligned} \dim(V) &= \dim(W) + \dim(W^0) \\ &= \dim(U) + \dim(W^0) \end{aligned}$$

$$\Rightarrow V = U \oplus W^0$$

□

## 2.3. Zyklische $f$ -invariant Unterräume

Jetzt: Genauere Analyse der Struktur von Eigenräumen

**Beispiel:** Ist  $V$  ein  $K$ -Vektorraum,  $f \in L(V, V)$  und  $\lambda \in K$  ein Eigenwert von  $f$ , so ist  $\text{Eig}(f, \lambda)$  ein  $f$ -invarianter Unterraum, da: Für  $v \in \text{Eig}(f, \lambda)$  gilt  $f(v) = \lambda v \in \text{Eig}(f, \lambda)$ .

Sei  $V$  ein  $K$ -Vektorraum mit  $\dim(V) = n < \infty$  und  $f \in L(V, V)$ . Ist  $v \in V \setminus \{0\}$ , so existiert ein eindeutig definiertes  $m = m(f, v) \in \mathbb{N}$ , sodass die Vektoren

$$v, f(v), f(f(v)), \dots, f^{m-1}(v)$$

linear unabhängig, die Vektoren

$$v, f(v), \dots, f^m(v)$$

jedoch linear abhängig sind. Wegen  $\dim(V) = n$ , muss  $m \leq n$  gelten!

### Definition 2.19: Grad von $v$

Die eindeutig definiert Zahl  $m(f, v) \in \mathbb{N}$  heißt Grad von  $v$  bezüglich  $f$ .

$$0 \neq v, f(v), f^2(v), \dots, f^{m-1}(v) \text{ lin. unabh.}$$

$$v, f(v), \dots, f^m(v) \text{ lin. abh.}$$

$\implies$  Grad  $m$  von  $v$ ,  $m \in \mathbb{N}$ .

### Bemerkungen:

- Der Vektor  $v = 0 \in V$  ist lin. abhängig. Deswegen muss man  $v \neq 0$  fordern oder  $m \in \mathbb{N} \cup \{0\}$ .
- Der Grad von  $0 \neq v \in V$  ist gleich 1, genau dann wenn  $v, f(v)$  linear abhängig sind. Das ist genau dann der Fall wenn  $v$  ein Eigenvektor von  $f$  ist. Damit folgt auch: Ist  $v \in V$  kein Eigenvektor von  $f$  und  $v \neq 0$ , so ist der Grad von  $v$  also  $m(v, f) \geq 2$ .

### Definition 2.20: Krylov-Raum

Sei  $V$  ein  $K$ -Vektorraum mit  $\dim(V) = n < \infty$ ,  $f \in L(V, V)$ ,  $v \in V$  und  $j \in \mathbb{N}$ . Der Unterraum

$$\mathcal{K}_j(f, v) := \text{Span}\{v, f(v), f^2(v), \dots, f^{j-1}(v)\} \subseteq V$$

heißt **j-ter Krylov-Raum** von  $f$  und  $v$ .

Alexei Krylov (russischer Schiffsbauingenieur und Mathematiker, 1863-1945). Krylov-Räume spielen auch eine wichtige Rolle für das CG-Verfahren (Conjugate Gradients).

**Lemma 2.21:** Sei  $V$  ein  $K$ -Vektorraum mit  $\dim(V) = n < \infty$  und  $f \in L(V, V)$ . Dann gilt:

1. Hat  $0 \neq v \in V$  den Grad  $m$  bzgl.  $f$ , so ist  $\mathcal{K}_m(f, v)$  ein  $f$ -invarianter Unterraum und es gilt:

$$\text{Span } \{v\} = \mathcal{K}_1(f, v) \subset \mathcal{K}_2(f, v) \subset \dots \subset \mathcal{K}_m(f, v) = \mathcal{K}_{m+j}(f, v)$$

für alle  $j \in \mathbb{N}$ .

2. Hat  $0 \neq v \in V$  den Grad  $m$  bzgl.  $f$  und ist  $U \subseteq V$  ein  $f$ -invarianter Unterraum, so dass  $v \in U$ , so ist

$$\mathcal{K}_m(f, v) \subseteq U$$

D.h. betrachtet man alle  $f$ -invarianten Unterräume von  $V$ , die  $v$  enthalten, so ist  $\mathcal{K}_m(f, v)$  derjenige mit der kleinsten Dimension.

3. Gilt für  $v \in V$ , dass  $f^{m-1}(v) \neq 0$  und  $f^m(v) = 0$  für ein  $m \in \mathbb{N}$ , dann ist

$$\dim(\mathcal{K}_j(f, v)) = j \quad \text{für } j = 1, \dots, m$$

*Beweis:*

1. ÜA
2. Sei  $U \subseteq V$  ein  $f$ -invarianter Unterraum mit  $v \in U$ . Dann gilt  $f^j(v) \in U$  für  $j = 1, \dots, m-1$ . Da  $v$  den Grad  $m$  hat, sind  $v, f(v), \dots, f^{m-1}(v)$  linear unabhängig.  $\implies \mathcal{K}_m(f, v) \subseteq U$  und  $\dim(\mathcal{K}_m(f, v)) = m \leq \dim(U)$
3. Seien  $\mu_0, \dots, \mu_{m-1} \in K$  so gewählt, dass

$$0 = \mu_0 v + \mu_1 f(v) + \dots + \mu_{m-1} f^{m-1}(v)$$

gilt. Anwendung  $f^{m-1}$

$$0 = \mu_0 f^{m-1}(v) + \mu_1 f^m(v) = \mu_0 \underbrace{f^{m-1}(v)}_{\neq 0}$$

$$\implies \mu_0 = 0$$

Für  $m > 1$  kann man dieses Argument induktiv für  $f^{m-j}$ ,  $j = 2, \dots, m$ , anwenden und erhält damit

$$\mu_1 = \dots = \mu_{m-1} = 0$$

$\implies$  Beh.

□

**Beobachtungen:** Hat  $v$  den Grad  $m$  bzgl.  $f$  gilt

- $\mathcal{K}_j(f, v)$  ist für  $j < m$  kein  $f$ -invarianter Unterraum, da  $0 \neq f(f^{j-1}(v)) = f^j(v) \notin \mathcal{K}_j(f, v)$

- wie oben gezeigt, bilden die Vektoren  $v, f(v), \dots, f^{m-1}(v)$  eine Basis von  $\mathcal{K}_m(f, v)$ .  
Wendet man  $f$  auf ein Element dieser Basis an, d.h.  $f^{k+1}(v)$ ,  $k = 0, \dots, m-1$ , so erhält man für  $k = m-1$   $f^m(v)$  als Linearkombination von  $v, f(v), \dots, f^{m-1}(v) \Rightarrow f^m(v) \in \mathcal{K}_m(f, v)$ . Deswegen wird  $\mathcal{K}_m(f, v)$  auch **zyklische invarianter Unterraum** zu  $v$  von  $f$  genannt.

**Lemma 2.22:** Sei  $\{0\} \neq V$  ein  $K$ -Vektorraum. Ist  $f \in L(V, V)$  nilpotent vom Grad  $m$ , so gilt  $m \leq \dim(V)$ .

*Beweis:* Nach Definition existiert ein  $v \in V$  mit  $f^{m-1}(v) \neq 0$  und  $f^m(v) = 0$ . Lemma 2.21 sichert, dass  $v, f(v), \dots, f^{m-1}(v)$  linear unabhängig  $\Rightarrow m \leq \dim(V)$ . □

**Beobachtung:** Sei  $V$  ein  $K$ -Vektorraum und  $f \in L(V, V)$ . Ist  $U \subseteq V$  ein  $f$ -invarianter Unterraum, so gilt für die Einschränkung von  $f$  auf  $U$ , d.h.

$$f|_U : U \rightarrow U, \quad u \mapsto f(u),$$

dass  $f|_U \in L(U, U)$ .

### Satz 2.23: Fittingzerlegung

Sei  $V$  ein endlichdimensionaler  $K$ -Vektorraum und  $f \in L(V, V)$ . Dann existieren  $f$ -invariante Unterräume  $U \subseteq V$  und  $W \subseteq V$ , so dass gilt:

1.  $V = U \oplus W$
2.  $f|_U \in L(U, U)$  ist bijektiv
3.  $f|_W \in L(W, W)$  ist nilpotent

*Beweis:*  $v \in \ker(f)$ . Dann gilt wegen der Linearität von  $f$ , sodass  $f^2(v) = f(f(v)) \stackrel{f(v)=0}{=} 0 \Rightarrow \ker(f) \subseteq \ker(f^2)$

Induktiv zeigt man:

$$\{0\} \subseteq \ker(f) \subseteq \ker(f^2) \subseteq \ker(f^3) \subseteq \dots$$

Da  $\dim(V) < \infty$ , muss es eine kleinste Zahl  $m \in \mathbb{N} \cup \{0\}$  geben, so dass  $\ker(f^m) = \ker(f^{m+j})$  für alle  $j \in \mathbb{N}$ . Damit sehen wir

$$U = \operatorname{im}(f^m) \quad \text{und} \quad W = \ker(f^m)$$

Zeige:  $U$  und  $W$  sind  $f$ -invariant. Sei  $u \in U$ . Dann existiert  $w \in V$  mit  $f^m(w) = u \Rightarrow f(u) = f(f^m(w)) = f^m(f(w)) \in U$ .

Sei  $w \in W$ . Dann gilt

$$f^m(f(w)) = f(f^m(w)) = 0 \Rightarrow f(w) \in W$$

Also existieren  $f$ -invariante Unterräume  $U \subseteq V$  und  $W \subseteq V$ .

1. Es gilt  $U + W \subseteq V$ . Die Dimensionsformel für lineare Abbildungen (Satz 4.16, LinA I) liefert für  $f^m$ , dass

$$\dim(V) = \dim(U) + \dim(W)$$

$$\text{Ist } v \in U \cap W \implies \exists w \in V : v = f^m(w) (v \in U)$$

$$v \in W \implies 0 = f^m(v) = f^m(f^m(v)) = f^{2m}(v)$$

$$\text{Es gilt } \ker(f^m) = \ker(f^{2m}) \implies v = f^m(v) = 0$$

$$\implies V = U \oplus W$$

2. Sei  $v \in \ker(f|_U) \subseteq U$ . Dann existiert ein  $w \in V$ , so dass  $f^m(w) = v$  gilt.  $\implies$   
 $0 = f(v) = f(f^m(w)) = f^{m+1}(w)$ . Mit  
 $\ker(f^m) = \ker(f^{m+1}) \implies w \in \ker(f^m) \implies v = f^m(w) = 0 \implies f$  injektiv.

Aus der Dimensionsformel folgt, dass  $f$  surjektiv ist.

3. Sei  $v \in W$ . Dann gilt

$$0 = f^m(v) = (f|_W)^m(v)$$

$$\implies (f|_W)^m = 0 \in L(W, W), \text{ d.h. } (f|_W)^m \text{ ist die Nullabbildung} \implies f|_W \text{ nilpotent.}$$

**Satz 2.24:** Sei  $V$  ein endlichdimensionaler  $K$ -Vektorraum,  $f \in L(V, V)$  nilpotent vom Grad  $m$ ,  $v \in V$  ein beliebiger Vektor mit  $f^{m-1}(v) \neq 0$  und  $h \in V^*$  mit  $h(f^{m-1}(v)) \neq 0$ . Dann sind  $v$  und  $h$  vom Grad  $m$  bzgl.  $f$  und  $f^*$ . Die beiden Räume  $\mathcal{K}_m(f, v)$  bzw.  $\mathcal{K}_m(f^*, h)$  sind zyklisch  $f$ - bzw.  $f^*$ -invariante Unterräume von  $V$  bzw.  $V^*$ . Sie bilden ein duales Raumpaar bzgl. der Bilinearform

$$a : \mathcal{K}_m(f, v) \times \mathcal{K}_m(f^*, h) \rightarrow K, \quad (\bar{v}, \bar{h}) \mapsto \bar{h}(\bar{v})$$

und es gilt

$$V = \mathcal{K}_m(f, v) \oplus (\mathcal{K}_m(f^*, h))^0$$

Hierbei ist  $\mathcal{K}_m(f^*, h)^0$  ein  $f$ -invarianter Unterraum von  $V$ .

*Beweis:* Für  $v \in V$  gilt  $f^{m-1}(v) \neq 0$ . Lemma 2.20  $\implies \mathcal{K}_m(f, v)$   $m$ -dimensionaler zyklischer  $f$ -invarianter Unterraum von  $V$ . Für  $V^*$  gilt

$$0 \neq h(f^{m-1}(v)) = (f^*)^{m-1}(h)(v)$$

Dann ist  $0 \neq (f^*)^{m-1}(h) \in L(V^*, V^*)$ .  $f$  nilpotent von Grad  $m \implies$  (Lemma 2.14)  $f^*$  nilpotent von Grad  $m \implies$

$$(f^*)^m(h) = 0 \in L(V^*, V^*)$$

$\implies$  (Lemma 2.20)  $\mathcal{K}_m(f^*, h)$  ist  $m$ -dimensionaler zyklischer  $f^*$ -invarianter Unterraum von  $V^*$ .

Nun zu zeigen:  $\mathcal{K}_m(f, v), \mathcal{K}_m(f^*, h)$  sind ein duales Raumpaar. Sei

$$\bar{v} = \sum_{i=0}^{m-1} \mu_i f^i(v) \in \mathcal{K}_m(f, v)$$

so gewählt, dass

$$\bar{h}(\bar{v}) = a(\bar{v}, \bar{h}) = 0 \quad \forall \bar{h} \in \mathcal{K}_m(f^*, h)$$

Zeige induktiv, dass  $\mu_k = 0, k = 0, \dots, m-1$ . Wegen  $((f^*)^{m-1}(h)) \in \mathcal{K}_m(f^*, h)$  folgt

$$0 = ((f^*)^{m-1}(h))(\bar{v}) = h(f^{m-1}(\bar{v})) = \sum_{i=0}^{m-1} \mu_i h(f^{m-1+i}(v)) = \underbrace{\mu_0 h(f^{m-1}(v))}_{\neq 0} \\ \Rightarrow \mu_0 = 0$$

Sei nun  $\mu_0 = \dots = \mu_{k-1} = 0$  für ein  $k \in \{1, \dots, m-2\}$ . Wegen  $(f^*)^{m-1-k}(h) \in \mathcal{K}_m(f^*, h)$  folgt aus der Darstellung von  $\bar{v}$ , dass

$$0^{(*)} = ((f^*)^{m-1-k}(h))(\bar{v}) = h(f^{m-1-k}(\bar{v})) = \sum_{i=0}^{m-1} \mu_i h(f^{m-1-k+i}(v)) = \underbrace{\mu_k h(f^{m-1}(v))}_{\neq 0} = \mu_k \\ \Rightarrow \bar{v} = 0$$

$\Rightarrow a(., .)$  ist nicht ausgeartet in der ersten Komponente. Analog zeigt man, dass  $a(., .)$  auch in der zweiten Komponente nicht ausgeartet ist  $\Rightarrow \mathcal{K}_m(f, v)$  und  $\mathcal{K}_m(f^*, h)$  sind ein duales Raumpaar.

Mit Lemma 2.18:  $V = \mathcal{K}_m(f, v) \oplus (\mathcal{K}_m(f^*, h))^0$

Mit Lemma 2.14, 3:  $(\mathcal{K}_m(f^*, h))^0$  ist  $f$ -invarianter UR von  $V$ .

□

(zyklisch  $f$ -invarianter UR:  $v, f(v), f^2(v), \dots$ )

## 2.4. Die Jordan-Normalform

**Satz 2.25:** Sei  $V$  ein endlichdimensionaler  $K$ -Vektorraum und  $f \in L(V, V)$ . Ist  $\lambda \in K$  ein Eigenwert von  $f$ , dann gibt es  $f$ -invariante Unterräume  $U \subset V$  und  $\{0\} \neq W \subseteq V$ , so dass

1.  $V = U \oplus W$
2. die Abbildung  $f|_U - \lambda \text{Id}_U$  ist bijektiv und
3. die Abbildung  $f|_W - \lambda \text{Id}_W$  ist nilpotent

Des Weiteren ist  $\lambda$  kein Eigenwert von  $f|_U$ .

*Beweis:* Wir definieren

$$g := f - \lambda \text{Id}_V \in L(V, V)$$



Satz 2.23:  $\exists g$ -invariante UR  $U \subseteq V$  und  $W \subseteq V$ :

$$V = U \oplus W, \quad g|_U \text{ bijektiv, } g|_W \text{ nilpotent}$$

Annahme:  $\{0\} = W \implies V = U$

$$\implies g|_U = g|_V = g \text{ bijektiv}$$

$\lambda$  ist Eigenwert von  $f \implies \exists 0 \neq v : f(v) = \lambda v$

$$\implies g(v) = f(v) - \lambda v = \lambda v - \lambda v = 0$$

$$\implies \ker(g) \supseteq \{0, v\} \neq \{0\} \nmid g \text{ bijektiv}$$

$$\implies U \subset V$$

Annahme:  $\lambda$  ist Eigenwert von  $f|_U$

$$\implies \exists 0 \neq v \in U : f(v) = \lambda v$$

$$\implies g|_U(v) = f(v) - \lambda v = \lambda v - \lambda v = 0 \nmid g|_U \text{ bijektiv}$$

□

**Beispiel 2.26:** Wir betrachten  $V = \mathbb{R}^5$ , die Standardbasis  $E$  und  $f \in L(V, V)$  gegeben durch

$$A = \begin{pmatrix} -3 & -1 & 4 & -3 & -1 \\ 1 & 1 & -1 & 1 & 0 \\ -1 & 0 & 2 & 0 & 0 \\ 4 & 1 & -4 & 5 & 1 \\ -2 & 0 & 2 & -2 & 1 \end{pmatrix}$$

Dann gilt

$$p_f(\lambda) = p_A(\lambda) = (\lambda - 1)^4(\lambda - 2)^1$$

$$\implies \text{EW: } 1, 2 \quad a(f, 1) = 4 \quad a(f, 2) = 1$$

$\implies p_A(\cdot)$  zerfällt in Linearfaktoren

$\lambda_1 = 1$  : Es gilt  $\ker(g_1^3) = \ker(g_1^4)$  für  $g_1 := f - \lambda_1 \text{Id}_V$

$$\implies m_1 = 3$$

$$U_1 = \text{Span} \left\{ \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ -2 \end{pmatrix} \right\} \quad W_1 = \text{Span} \left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

$\lambda_2 = 2$  : Für  $g_2 = f - \lambda_2 \text{Id}_V$  gilt  $\ker(g_2) = \ker(g_2^2)$

$$\implies m_2 = 1$$

$$U_2 = \text{Span} \left\{ \begin{pmatrix} -5 \\ 1 \\ -1 \\ 4 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ -1 \\ 0 \\ -4 \\ 2 \end{pmatrix}, \begin{pmatrix} -3 \\ 1 \\ 0 \\ 3 \\ -2 \end{pmatrix} \right\}, W_2 = \text{Span} \left\{ \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ -2 \end{pmatrix} \right\}$$

Beobachtung:  $\dim(W_1) = a(f, \lambda_1)$ ,  $\dim(W_2) = a(f, \lambda_2)$

**Satz 2.27:** Sei  $V$  ein endlichdimensionaler  $K$ -Vektorraum und  $f \in L(V, V)$ . Ist  $\lambda \in K$  ein Eigenwert von  $f$ , dann existieren für den Unterraum  $W$  aus Satz 2.25 Vektoren  $w_1, \dots, w_k \in W$  und  $d_1, \dots, d_k \in \mathbb{N}$ , so dass

$$W = \mathcal{K}_{d_1}(f, w_1) \oplus \mathcal{K}_{d_2}(f, w_2) \oplus \dots \oplus \mathcal{K}_{d_k}(f, w_k)$$

Des Weiteren gibt es eine Basis  $B$  von  $W$ , so dass

$$A_{f|_W}^{B,B} = \begin{pmatrix} J_{d_1}(\lambda) & & 0 \\ & \ddots & \\ 0 & & J_{d_k}(\lambda) \end{pmatrix}$$

*Beweis:* Sei wie in Satz 2.25  $g := f - \lambda \text{Id}_V$  und  $g_1 := g|_W$  nilpotent vom Grad  $d_1$ . Dann gilt  $1 \leq d_1 \leq \dim(W)$ .

Sei  $w_1 \in W$  ein Vektor mit  $g_1^{d_1-1}(w_1) \neq 0$ . Wegen  $g^{d_1}(w_1) = 0$

$\Rightarrow g_1^{d_1-1}(w_1)$  ist ein Eigenvektor von  $g_1$  zum Eigenwert 0.

Lemma 2.21, 3, liefert, dass die  $d_1$  Vektoren

$$\{w_1, g(w_1), \dots, g_1^{d_1-1}(w_1)\}$$

linear unabhängig sind. Außerdem ist  $W_1 := \mathcal{K}_{d_1}(g_1, w_1)$  ein  $d_1$ -dimensionaler zyklischer  $g_1$ -invarianter UR von  $W$ . Also ist

$$B_1 := \{g_1^{d_1-1}(w_1), g_1^{d_1-2}(w_1), \dots, g_1(w_1), w_1\}$$

eine Basis von  $\mathcal{K}_{d_1}(g_1, w_1) = W_1$  und

$$A_{g_1|_{W_1}}^{B_1, B_1} = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} = J_{d_1}(0) \in K^{d_1, d_1}$$

Per Definition gilt  $A_{g_1|_{W_1}}^{B_1, B_1} = A_{g|_{W_1}}^{B_1, B_1}$ . Ist  $d_1 = \dim(W)$ : siehe unten  $\bullet \bullet \bullet$ .

Sei nun  $d_1 < \dim(W)$ . Satz 2.25 sichert, dass es für  $g_1 \in L(W, W)$  einen  $g_1$ -invarianten Unterraum  $\tilde{W} \neq \{0\}$  mit  $W = W_1 \oplus \tilde{W}$  gibt.

Die Abbildung  $g_2 := g_1|_{\tilde{W}}$  ist nilpotent vom Grad  $\lambda_2$  mit  $1 \leq d_2 \leq d_1$ .

Wiederholung der Konstruktion:

$\exists w_2 \in \tilde{W} : g_2^{d_2-1}(w_1) \neq 0, \dots, W_2 := \mathcal{K}_{d_2}(g_2, w_2) \dots$  UR von  $\tilde{W} \subseteq W$ ,

$$B_2 := \{g_2^{d_2-1}(w_2), g_2^{d_2-2}(w_2), \dots, g_2(w_2), w_2\}$$

$$A_{g|_{W_2}}^{B_2, B_2} = A_{g_2|_{W_2}}^{B_2, B_2} = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}$$

Nach  $k \leq \dim(W)$  Schritten muss diese Konstruktion abbrechen und es gilt

$$\begin{aligned} W &= \mathcal{K}_{d_1}(g_1, w_1) \oplus \mathcal{K}_{d_2}(g_2, w_2) \oplus \dots \oplus \mathcal{K}_{d_k}(g_k, w_k) \\ &= \mathcal{K}_{d_1}(g, w_1) \oplus \mathcal{K}_{d_2}(g, w_2) \oplus \dots \oplus \mathcal{K}_{d_k}(g, w_k) \end{aligned}$$

Vereinigt man die Basen  $B_1, \dots, B_k$  zu einer Basis  $B$  von  $W$  (direkte Summe!), so erhält man

$$A_{g|_W}^{B, B} = \begin{pmatrix} A_{g|_{W_1}}^{B_1, B_1} & & 0 \\ & \ddots & \\ 0 & & A_{g|_{W_k}}^{B_k, B_k} \end{pmatrix} = \begin{pmatrix} J_{d_1}(0) & & 0 \\ & \ddots & \\ 0 & & J_{d_k}(0) \end{pmatrix}$$

Jetzt: Übertragung auf  $f = g + \lambda \text{Id}_V$ . Man kann sich leicht überlegen, dass jeder  $g$ -invariante Unterraum von  $V$  auch  $f$ -invariant ist und damit gilt:

$$\mathcal{K}_{d_i}(f, w_i) = \mathcal{K}_{d_i}(g, w_i) \quad \text{für } i = 1, \dots, k$$

$$\stackrel{\text{ÜA}}{\implies} W = \mathcal{K}_{d_1}(f, w_1) \oplus \dots \oplus \mathcal{K}_{d_k}(f, w_k)$$

Für  $j \in \{1, \dots, k\}$  und  $0 \leq l \leq d_j - 1$  ist

$$\begin{aligned} f(g^l(w_j)) &= g(g^l(w_j)) + \lambda g^l(w_j) \\ &= \lambda g^l(w_j) + \underbrace{g^{l+1}(w_j)}_{=0, l=d_j-1} \end{aligned}$$

$$\implies A_{f|_W}^{B, B} = \begin{pmatrix} A_{f|_{W_1}}^{B_1, B_1} & & 0 \\ & \ddots & \\ 0 & & A_{f|_{W_k}}^{B_k, B_k} \end{pmatrix} = \begin{pmatrix} J_{d_1}(\lambda) & & 0 \\ & \ddots & \\ 0 & & J_{d_k}(\lambda) \end{pmatrix}$$

□

**Beispiel 2.28:** Fortsetzung von Bsp 2.26

$$A = \begin{pmatrix} -3 & -1 & 4 & -3 & -1 \\ 1 & 1 & -1 & 1 & 0 \\ -1 & 0 & 2 & 0 & 0 \\ 4 & 1 & -4 & 5 & 1 \\ -2 & 0 & 2 & -2 & 1 \end{pmatrix} \quad \text{EW: } \begin{aligned} \lambda_1 &= 1, a(f, \lambda_1) = 4 = \dim(W_1) \\ \lambda_2 &= 2, a(f, \lambda_2) = 1 = \dim(W_2) \end{aligned}$$

$\lambda_1 = 1$ :  $g_1^1|_{W_1}$  nilpotent vom Grad  $\lambda_1^1 = 3$  und  $1 < d_1^1 < \dim(W_1)$

Erinnerung:  $g_1^1 = f - \lambda_1^1 I_d$ . Für  $w_1^1 = (0 \ 0 \ 0 \ 0 \ 1)^T$  ist  $(g_1^1)^2(w_1^1) = (1 \ 0 \ 1 \ 0 \ 0)^T \neq 0$  und  $(g_1^1)^3(w_1^1) = 0 \in V = \mathbb{R}^5$ .

Mit Lemma 2.21:

$$\{w_1, (g_1^1)^1(w_1^1), (g_1^1)^2(w_1^1)\} = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$\implies \text{Span}\{w_1, g_1^1(w_1^1), (g_1^1)^2(w_1^1)\} = \mathcal{K}_3(g_1^1, w_1^1)$$

$d_1^1 < \dim(W_1) \implies$  es existiert zu  $W_{11} := \mathcal{K}(g_1^1, w_1^1)$  ein  $\tilde{W}_1 \neq \{0\}$  mit  $W_1 = W_{11} \oplus \tilde{W}_1$ .

Zum Beispiel:  $w_2^1 = (1 \ -1 \ 1 \ 0 \ 1)^T \implies$

$$w_2^1, w_1^1, g_1^1(w_1^1), (g_1^1)^2(w_1^1) \quad \text{lin. unab.}$$

$$\tilde{W}_1 := \text{Span}\{w_2^1\} \cap \mathcal{K}_3(g_1^1, w_1^1) = \{0\}$$

Es gilt  $g_2^1 := g_1^1|_{\tilde{W}_1}$  nilpotent vom Grad 1

$$\implies d_2^1 = 1 \quad W_1 = \mathcal{K}_3(g_1^1, w_1^1) \oplus \mathcal{K}_1(g_2^1, w_2^1)$$

Weitherhin kann man nachrechnen

$$\mathcal{K}_3(f, w_1^1) = \text{Span}\{w_1, g_1^1(w_1^1), (g_1^1)^2(w_1^1)\} = \mathcal{K}_3(g_1^1, w_1^1)$$

$$\mathcal{K}_1(f, w_2^1) = \text{Span}\{w_2^1\} = \mathcal{K}_1(g_2^1, w_2^1)$$

$\lambda_2 = 2$ :

$$g_1^2|_{W_2} \text{ nilpotent vom Grad } \lambda_1^2 = 1$$

$$\lambda_1^2 = \dim(W_2)$$

$$w_1^2 = (0 \ 1 \ 2 \ 3 \ -2)^T \neq 0$$

$$(g_1^2)^1(w_1^2) = 0 \in V \implies W_2 = \mathcal{K}_1(f, w_1^2)$$

$$A_{f|_{W_1}}^{B^1, B^1} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad A_{f|_{W_2}}^{B^2, B^2} = (2)$$

$$\stackrel{\text{Ziel:}}{\implies} A_f^{B, B} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & \\ 0 & 0 & 1 & \\ & & & 1 \\ 0 & & & & 2 \end{pmatrix}$$

**Satz 2.29:** Sei  $V$  ein  $K$ -Vektorraum mit  $\dim(V) < \infty$  und  $f \in L(V, V)$ . Ist  $\lambda \in K$  ein Eigenwert von  $f$ , dann gilt für die  $d_j$ ,  $1 \leq j \leq k$  aus Satz 2.27, dass

$$a(f, \lambda) = \dim(W) = d_1 + \dots + d_k$$

$$g(f, \lambda) = k$$

*Beweis:* Für den Unterraum  $U$  aus Satz 2.23/2.25 ist die Abbildung  $f|_U = (f - \lambda \text{Id})|_U$  bijektiv  $\implies \lambda$  ist kein Eigenwert von  $f|_U$ . Daraus erhält man

$$a(f, \lambda) = \dim W = d_1 + \dots + d_k$$

Zur Bestimmung von  $g(f, \lambda)$  sei  $v \in W$  ein beliebiger Vektor. Dann ist

$$v = \sum_{j=1}^k \sum_{l=0}^{d_j-1} \mu_{jl} g^l(w_j)$$

und es gilt

$$\begin{aligned} f(v) &= \sum_{j=1}^k \sum_{l=0}^{d_j-1} \mu_{jl} f(g^l(w_j)) = \sum_{j=1}^k \sum_{l=0}^{d_j-1} \mu_{jl} g^l(w_j) + \sum_{j=1}^k \sum_{l=0}^{d_j-1} \mu_{jl} g^{l+1}(w_j) \\ &= \lambda v + \underbrace{\sum_{j=1}^k \sum_{l=0}^{d_j-2} \mu_{jl} g^{l+1}(w_j)}_{=0} \quad \text{lin. unab.} \end{aligned}$$

$$v \in \text{Eig}(f, \lambda) \iff \mu_{jl} = 0, 1 \leq j \leq k, 0 \leq l \leq d_j - 2$$

$$\iff v = \sum_{j=1}^k \mu_j g^{d_j-1}(w_j)$$

Für  $v \neq 0$  muss mindestens ein Koeffizient  $\mu_j \neq 0$ ,  $j = 1, \dots, k$ . Daraus folgt

$$\text{Eig}(f, \lambda) = \text{Span} \underbrace{\{g^{d_1-1}(w_1), \dots, g^{d_k-1}(w_k)\}}_{\text{lin. unab. wegen direkter Summe}}$$

**Beispiel 2.30:** Fortsetzung von Bsp. 2.28. Es gilt

$$\text{Eig}(f, 1) = \text{Span} \left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ -1 \end{pmatrix} \right\} \Rightarrow g(f, 1) = 2$$

$$\lambda_1 = 1: a(f, 1) = 4 = 3 + 1 = d_1^1 + d_2^1, g(f, 1) = 2 = k$$

$$\lambda_2 = 2: a(f, 2) = 1 = d_1^2, g(f, 2) = 1$$

**Fazit:** Für einen Eigenwert  $\lambda$  zu  $f \in L(V, V)$  gilt:

- Die geometrische Vielfachheit des Eigenwert  $\lambda$  ist gleich der Anzahl der Jordanblöcke zu diesem Eigenwert in der entsprechenden Darstellungsmatrix

$$A_f^{B,B} = \begin{pmatrix} J_{d_1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{d_k}(\lambda_m) \end{pmatrix}$$

- Die algebraische Vielfachheit des Eigenwert  $\lambda$  ist gleich der Summe der Dimensionen der zugehörigen Jordanblöcke
- In jedem Unterraum  $\mathcal{K}_{d_j}(f, w_j)$  gehört genau ein Eigenvektor und seine Vielfachheiten.

Was gilt für weitere Eigenwerte?

Ist  $\tilde{\lambda} \neq \lambda$  ein weiterer Eigenwert von  $f$ , dann ist  $\tilde{\lambda}$  auch ein Eigenwert der Einschränkung  $f|_U \in L(U_\lambda, U_\lambda)$

$\Rightarrow$  Man kann die Sätze 2.25-2.29 auf  $f|_{U_\lambda}$  anwenden. Damit erhält man

- $U_\lambda = X \oplus Y$
- $f|_X - \tilde{\lambda} \text{Id}_X$  ist bijektiv
- $f|_Y - \tilde{\lambda} \text{Id}_Y$  ist nilpotent
- Der UVR  $Y$  ist die direkte Summe von Krylovräumen
- Es gibt eine Darstellungsmatrix von  $f|_Y$  bestehend aus Jordanblöcken

Da man dieses Argument für alle paarweise verschiedene Eigenwerte von  $f$  anwenden kann, erhält man.

**Satz 2.31:** Sei  $V$  ein endlichdimensionaler  $K$ -Vektorraum und  $f \in L(V, V)$ . Zerfällt das charakteristische Polynom  $p_f(\cdot)$  in Linearfaktoren, so gibt es eine Basis  $B$  von  $V$  für welche die Darstellungsmatrix in Jordan-Normalform ist, d.h.

$$A_f^{B,B} = \begin{pmatrix} J_{d_1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{d_k}(\lambda_m) \end{pmatrix}$$

*Beweis:* s.o.

□

Marie Ennemond Jordan (fr. Mathematiker, 1838-1922) gab diese Form 1870 an. Zwei Jahre vor Jordan bewies Karl Weierstraß (dt. Mathematiker, 1815-1897) ein Resultat, aus dem die JNF folgt.

**Beispiel 2.32:**

$$A = \begin{pmatrix} -3 & -1 & 4 & -3 & -1 \\ 1 & 1 & -1 & 1 & 0 \\ -1 & 0 & 2 & 0 & 0 \\ 4 & 1 & -4 & 5 & 1 \\ -2 & 0 & 2 & -2 & 1 \end{pmatrix} \rightsquigarrow J = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 & \\ & & & 1 \\ 0 & & & & 2 \end{pmatrix} = A_f^{B,B}$$

$$B = \left\{ (g_1^1)^2(w_1^1), g_1^1(w_1^1), w_1^1, w_2^1, w_1^2 \right\}$$

Für

$$S = \begin{pmatrix} 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 1 & -2 \end{pmatrix}$$

gilt  $J = S^{-1}AS$ . Also  $J$  ähnlich zu  $A$ .

Für  $f \in L(V, V)$  hatten wir:

- $f$  ist diagonalisierbar  $\iff$ 
  - $p_f(\cdot)$  zerfällt in Linearfaktoren
  - $\forall$  EW  $\lambda$  von  $f : a(f, \lambda) = g(f, \lambda)$
- zerfällt  $p_f(\cdot)$  in Linearfaktor  $\implies \exists$  Basis  $B : A_f^{B,B}$  in JNF

**Folgerung:** Existiert eine Darstellungsmatrix in Jordan-Normalform:  $f$  ist diagonalisierbar  $\iff d_i = 1 \forall i \in \{1, \dots, k\}$

**Frage:** Wann zerfällt  $p_f(\cdot)$  in Linearfaktoren?

Fundamentalsatz der Algebra:

Jedes Polynom  $p \in P[t]$  über  $\mathbb{C}$  mit einem Grad größer 0 hat mindestens eine Nullstelle.

*Beweis:* Liesen, Mehrmann, Kapitel 15, braucht substantiell Hilfsmittel aus der Analysis.

Damit folgt unmittelbar:

**Korollar 2.33:** Jedes Polynom  $p \in P[t]$  über  $\mathbb{C}$  zerfällt in Linearfaktoren, d.h. es gibt  $a, \lambda_1, \dots, \lambda_n \in \mathbb{C}$  mit  $n = \text{grad}(p)$  und

$$p(t) = a(t - \lambda_1)(t - \lambda_2) \dots (t - \lambda_n)$$

Daraus folgt direkt:

**Korollar 2.34:** Sei  $V$  ein endlichdimensionaler  $\mathbb{C}$ -Vektorraum. Dann besitzt jedes  $f \in L(V, V)$  eine Jordan-Normalform.

Matrix-Version:

**Korollar 2.35:** Sei  $K$  ein Körper und  $A \in K^{n,n}$ , so dass das charakteristische Polynom  $p_A(\cdot)$  in Linearfaktoren zerfällt. Dann ist  $A$  ähnlich zu einer Matrix  $J$  in Jordan-Normalform.

Ist die Jordan-Normalform eindeutig bestimmt?

**Satz 2.36:** Sei  $V$  ein  $K$ -Vektorraum mit  $\dim(V) = n < \infty$ . Bestitzt  $f \in L(V, V)$  eine Jordan-Normalform, so ist diese bis auf die Reihenfolge der Jordanblöcke eindeutig bestimmt.

*Beweis:* sehr technisch, z.B. Liesen, Mehrmann Satz 16.12, Fischer/Springborn, Abschnitt 5.7.

□

Alternativer Beweis für die JNF über Hauptvektoren und Haupträume, vgl. Fischer/Springborn, Abschnitt 5.5.

Damit: Für Bsp. 2.32 wären

$$\begin{pmatrix} 2 & & \\ & J_3(1) & \\ & & 1 \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} 2 & & \\ & 1 & \\ & & J_3(1) \end{pmatrix}$$

alternative JNF. Jordanblöcke bleiben gleich. D.h. Satz 2.36 rechtfertigt den Namen "Normalform".



### 3. Euklidische und unitäre Vektorräume

Jetzt:  $V$  Vektorraum über  $\mathbb{R}$  oder  $\mathbb{C}$  mit  $\dim(V) < \infty$ .

Damit: Definition eines Skalarproduktes und Verallgemeinerung von Begriffen aus der Geometrie für  $\mathbb{R}^2$  bzw.  $\mathbb{R}^3$ . Dies beinhaltet auch Orthogonalität und orthonormale Basen.

#### 3.1. Skalarprodukt und Normen

Für  $K = \mathbb{R}$  werden wir Bilinearformen (Def. 2.17) verwenden. Für  $K = \mathbb{C}$  benötigen wir

##### Definition 3.1: Sesquilinearform

Seien  $V$  und  $W$  zwei  $\mathbb{C}$ -Vektorräume. Man nennt eine Abbildung

$$s : V \times W \rightarrow \mathbb{C}, (v, w) \mapsto s(v, w)$$

**Sesquilinearform** auf  $V \times W$ , wenn für alle  $v, v_1, v_2 \in V$ ,  $w, w_1, w_2 \in W$  und  $\lambda \in \mathbb{C}$  gilt

1.  $s(v_1 + v_2, w) = s(v_1, w) + s(v_2, w)$  und  $s(\lambda v, w) = \lambda s(v, w)$   
 $\hat{=}$   $s(., .)$  ist linear in der ersten Komponente
2.  $s(v, w_1 + w_2) = s(v, w_1) + s(v, w_2)$  und  $s(v, \lambda w) = \bar{\lambda} s(v, w)$

Ist  $V = W$ , so heißt  $s$  Sesquilinearform auf  $V$ . Eine Sesquilinearform auf  $V$  nennt man hermitesch, wenn

$$s(v, w) = \overline{s(w, v)} \quad \forall v, w \in V$$

##### Definition 3.2: Skalarprodukt

Sei  $V$  ein  $K$ -Vektorraum. Eine Abbildung

$$\langle ., . \rangle : V \times V \rightarrow K, (v, w) \mapsto \langle v, w \rangle$$

nennt man **Skalarprodukt** oder **inneres Produkt** auf  $V$ , wenn gilt

1. Ist  $K = \mathbb{R}$ , so ist  $\langle ., . \rangle$  eine symmetrische Bilinearform
2. Ist  $K = \mathbb{C}$ , so ist  $\langle ., . \rangle$  eine hermitesche Sesquilinearform
3.  $\langle ., . \rangle$  ist positiv definit, d.h. es gilt

$$\langle v, v \rangle \geq 0 \quad \forall v \in V$$

$$\langle v, v \rangle = 0 \iff v = 0 \in V$$

Ein  $\mathbb{R}$ -Vektorraum mit einem Skalarprodukt nennt man **euklidischen Vektorraum** und einen  $\mathbb{C}$ -Vektorraum mit einem Skalarprodukt **unitären Vektorraum**.

### Bemerkungen:

- Für alle  $v \in V$  gilt  $\langle v, v \rangle \in \mathbb{R}^+$  unabhängig von  $K = \mathbb{R}$  oder  $K = \mathbb{C}$
- Ein Unterraum eines euklidischen (unitären) Vektorraums ist wieder ein euklidischer (unitärer) Vektorraum.

#### Definition 3.3: hermitesche Matrix

Für eine Matrix  $A = (a_{ij}) \in \mathbb{C}^{m,n}$  ist die hermitesch transponierte von  $A$  definiert als

$$A^H = (\bar{a}_{ji}) \in \mathbb{C}^{n,m}$$

Gilt  $A = A^H$ , so heißt  $A$  **hermitesche Matrix**.

Ist  $A \in \mathbb{R}^{m,n}$ , so  $A^H = A^T$ . Für eine hermitesche Matrix  $A$  gilt  $a_{ii} = \bar{a}_{ii} \Rightarrow a_{ii} \in \mathbb{R}$ .

**Beispiel 3.4:** Man kann leicht nachrechnen:

- Für  $V = \mathbb{R}^n$  ist

$$\langle v, w \rangle := v^T w = \sum_{i=1}^n v_i w_i$$

ein Skalarprodukt. Es ist das Standardskalarprodukt im  $\mathbb{R}^n$ .

- Für  $V = \mathbb{C}^n$  ist

$$\langle v, w \rangle := w^H v = \sum_{i=1}^n \bar{w}_i v_i$$

ein Skalarprodukt. Es ist das Standardskalarprodukt im  $\mathbb{C}^n$ .

- Für  $V = K^{m,n}$  ist

$$\langle A, B \rangle := \text{Spur} \underbrace{(B^H A)}_{\in K^{n,n}} = \sum_{i=1}^n \left( \sum_{j=1}^m \bar{b}_{ij} a_{ji} \right)$$

- Auf dem Vektorraum der auf dem Intervall  $[0, 1]$  stetigen, reellwertigen Funktionen ist

$$\langle f, g \rangle := \int_0^1 f(x)g(x)dx$$

ein Skalarprodukt.

#### Lemma 3.5: Cauchy-Schwarz-Ungleichung

Ist  $V$  ein euklidischer oder unitärer Vektorraum, so gilt

$$|\langle v, w \rangle|^2 \leq \langle v, v \rangle \cdot \langle w, w \rangle \quad \forall v, w \in V$$

wobei das Gleichheitszeichen genau dann gilt, wenn  $v$  und  $w$  linear abhängig sind.

*Beweis:* Für  $w = 0$  folgt die (Un-)gleichung.

Für  $w \neq 0$  definiert man

$$\lambda := \frac{\langle v, w \rangle}{\langle w, w \rangle}$$

Dann folgt

$$\begin{aligned} 0 &\leq \langle v - \lambda w, v - \lambda w \rangle = \langle v, v \rangle - \bar{\lambda} \langle v, w \rangle - \lambda \langle w, v \rangle - \lambda \cdot (-\bar{\lambda}) \langle w, w \rangle \\ &= \langle v, v \rangle - \frac{\overline{\langle v, w \rangle}}{\langle w, w \rangle} \langle v, w \rangle - \frac{\langle v, w \rangle}{\langle w, w \rangle} \langle w, v \rangle + \frac{|\langle v, w \rangle|^2}{(\langle w, w \rangle)^2} \langle w, w \rangle \\ &= \langle v, v \rangle - \frac{|\langle v, w \rangle|^2}{\langle w, w \rangle} \\ &\implies |\langle v, w \rangle|^2 \leq \langle v, v \rangle \cdot \langle w, w \rangle \end{aligned}$$

“=”:

$$\begin{aligned} 0 &= \langle v - \lambda w, v - \lambda w \rangle \\ &\iff v - \lambda w = 0 \iff v = \lambda w \iff w = \lambda^{-1} v \end{aligned}$$

□

$$\langle v, \lambda w \rangle = \bar{\lambda} \langle v, w \rangle$$

Deshalb:

Die Cauchy-Schwartzsche Ungleichung ist ein sehr wichtiges Instrument der Analysis, z.B. für Approximationsfehler.

Nächstes Ziel: Vektoren  $v \in V$  eine Länge zuzuordnen  $\rightarrow$  Norm als Verallgemeinerung des Betrags

Für die reellen Zahlen:  $|\cdot| : \mathbb{R} \rightarrow \mathbb{R}^+, x \mapsto |x|$  mit

- $|\lambda x| = |\lambda| \cdot |x| \quad \forall \lambda \in \mathbb{R}, \forall x \in \mathbb{R}$
- $|x| \geq 0 \quad \forall x \in \mathbb{R}, |x| = 0 \iff x = 0$
- $|x + y| \leq |x| + |y| \quad \forall x, y \in \mathbb{R}$

**Definition 3.6: Norm**

Sei  $V$  ein  $K$ -Vektorraum. Eine Abbildung

$$\|\cdot\| : V \rightarrow \mathbb{R}, \quad v \mapsto \|v\|$$

nennt man Norm auf  $V$ , wenn für alle  $v, w \in V$  und  $\lambda \in K$  gilt:

- sie ist homogen, d.h.

$$\|\lambda v\| = |\lambda| \cdot \|v\|$$

- sie ist positiv definit, d.h:

$$\|v\| \geq 0, \quad \|v\| = 0 \iff v = 0 \in V$$

- sie erfüllt die Dreiecksungleichung, d.h.

$$\|v + w\| \leq \|v\| + \|w\|$$

Einen  $K$ -Vektorraum, auf dem eine Norm definiert ist, nennt man **normierten Raum**.

**Beispiel 3.7:** Man kann leicht nachrechnen:

- Ist  $\langle \cdot, \cdot \rangle$  das Standardskalarprodukt auf  $\mathbb{R}^m$  oder  $\mathbb{C}^m$ , dann definiert

$$\|v\| := \langle v, v \rangle^{\frac{1}{2}} = (v^T v)^{\frac{1}{2}} \quad \text{bzw.} \quad = (v^H v)^{\frac{1}{2}}$$

eine Norm auf  $\mathbb{R}^m$  bzw.  $\mathbb{C}^m$ . Sie wird **euklidische Norm** genannt

- Für  $V = K^{m,n}$  ist

$$\|A\|_F := (\text{Spur}(A^H A))^{\frac{1}{2}} = \left( \sum_{i=1}^n \left( \sum_{j=1}^m |a_{ji}|^2 \right) \right)^{\frac{1}{2}}$$

eine Norm. Sie wird Frobeniusnorm genannt. Es gilt  $\|A\|_F = \|A^H\|_F$  für alle  $A \in K^{m,n}$ .

- Auf dem Vektorraum der auf dem Intervall  $[0, 1]$  stetigen, reellwertigen Funktionen ist

$$\|f\| := \langle f, f \rangle^{\frac{1}{2}} = \left( \int_0^1 (f(x))^2 dx \right)^{\frac{1}{2}}$$

eine Norm. Sie wird  $L_2$ - oder  $L^2$ -Norm genannt.

- Sei  $p \in \mathbb{R}, p \geq 1$  und  $V = K^n$ . Dann definiert

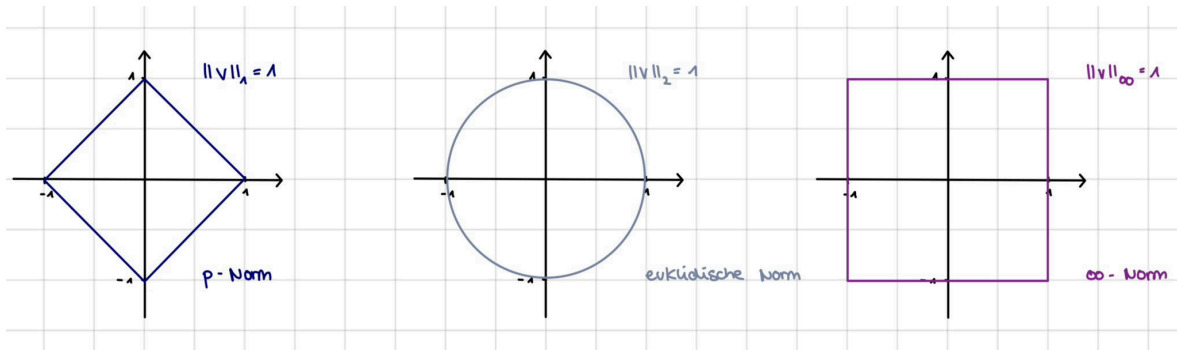
$$\|v\|_p = \left( \sum_{i=1}^n |v_i|^p \right)^{\frac{1}{p}}$$

eine Norm im  $K^n$ . Sie wird  $p$ -Norm genannt. Für  $n = 2$  erhält man die euklidische Norm.

Für  $p \rightarrow \infty$  erhält man die sogenannte  $\infty$ -Norm

$$\|v\|_{\infty} := \max_{1 \leq i \leq n} |v_i|$$

Je nach Situation kann es einem erheblichen Unterschied bedeuten, welche Norm betrachtet wird. Für  $V = \mathbb{R}^2$ :



- Die  $p$ -Norm auf  $K^{m,n}$  ist definiert durch

$$\|A\|_p := \sup_{0 \neq v \in \mathbb{K}^n} \frac{\|Av\|_p}{\|v\|_p}$$

$\|A\|_p$  ist die durch die  $p$ -Norm induzierte Matrix-Norm.

Man kann zeigen:

- Supremum wird angenommen
- $\|A\|_p = \max_{\substack{v \in K^n \\ \|v\|_p = 1}} \|Av\|_p$

Man kann zeigen:

$$\|A\|_1 = \max_{1 \leq j \leq m} \sum_{i=1}^n |a_{ij}| \quad (\text{Spaltensummennorm})$$

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^m |a_{ij}| \quad (\text{Zeilensummennorm})$$

**Korollar 3.8:** Sei  $V$  ein  $K$ -Vektorraum mit einem Skalarprodukt. Dann ist die Abbildung

$$\|\cdot\| : V \rightarrow \mathbb{R}, \quad v \mapsto \|v\| := (\langle v, v \rangle)^{\frac{1}{2}}$$

eine Norm auf  $V$ . Man nennt sie die durch das Skalarprodukt induzierte Norm.

*Beweis:*

1. Homogenität: (Es gilt mit  $\operatorname{Re}(z) \leq |z| \forall z \in \mathbb{C}$ )<sup>(\*)</sup>)

$$\|\lambda v\|^2 = \langle \lambda v, \lambda v \rangle = \lambda \bar{\lambda} \langle v, v \rangle = |\lambda|^2 \langle v, v \rangle$$

2. Positive Definitheit:

$$\begin{aligned}\langle v, v \rangle &\geq 0 \implies \|v\| \geq 0 \\ \langle v, v \rangle &= 0 \iff v = 0, \\ &\iff \|v\| = 0\end{aligned}$$

$$3. \|v + w\| \leq \|v\| + \|w\|$$

$$\begin{aligned}\|v + w\|^2 &= \langle v + w, v + w \rangle = \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle \\ &= \langle v, v \rangle + \langle v, w \rangle + \overline{\langle v, w \rangle} + \langle w, w \rangle \\ &= \langle v, v \rangle + 2 \operatorname{Re}(\langle v, w \rangle) + \langle w, w \rangle \\ &\stackrel{(*)}{\leq} \langle v, v \rangle + 2|\langle v, w \rangle| + \langle w, w \rangle \\ &\stackrel{\text{CSU}}{\leq} \langle v, v \rangle + 2\langle v, v \rangle \langle w, w \rangle + \langle w, w \rangle \\ &= \|v\|^2 + 2\|v\| \|w\| + \|w\|^2 \\ &= (\|v\| + \|w\|)^2 \\ &\implies \sqrt{\phantom{x}} \|v + w\| \leq \|v\| + \|w\|\end{aligned}$$

## 3.2. Winkel und Orthogonalität

In  $\mathbb{R}^2$  bzw.  $\mathbb{R}^3$  ist der von zwei Vektoren eingeschlossene Winkel anschaulich klar.  
Übertragung auf allgemeine Vektorräume?

Zunächst:  $V = \mathbb{R}^2$ , Standardskalarprodukt  $\langle v, w \rangle = w^T v$  und der damit induzierten Norm.

Aus Cauchy-Schwartz folgt:

$$-1 \leq \frac{\langle v, w \rangle}{\|v\| \cdot \|w\|} \quad \forall v, w \in \mathbb{R}^2 \setminus \{0\}$$

D.h. dieser Quotient ist gleich  $\cos(\theta)$  für ein  $\theta \in [0, \pi]$ . Diesen nennt man den zwischen  $v$  und  $w$  eingeschlossenen Winkel.

$$\frac{\langle v, w \rangle}{\|v\| \cdot \|w\|} = \cos(\theta) \quad \rightarrow \quad \angle(v, w) := \arccos \frac{\langle v, w \rangle}{\|v\| \cdot \|w\|}$$

Passt das zur “üblichen” Winkeldefinition?

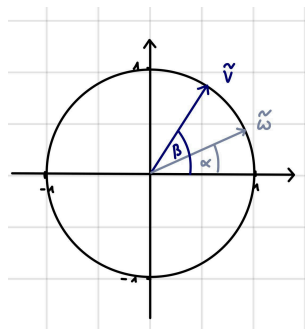
Aufgrund der Eigenschaften des Skalarprodukts folgt

$$\angle(v, w) = \angle(w, v), \quad \angle(\lambda v, w) = \angle(v, w) = \angle(v, \lambda w) \quad \forall \lambda > 0$$

Für  $v \neq 0 \neq w$  und

$$\tilde{v} = \frac{1}{\|v\|} v \quad (\Rightarrow \|\tilde{v}\| = 1) \quad \text{und} \quad \tilde{w} = \frac{1}{\|w\|} w \quad (\Rightarrow \|\tilde{w}\| = 1)$$

gilt  $\angle(v, w) = \angle(\tilde{v}, \tilde{w})$ . Im Einheitskreis erhält man



Also gibt es  $\alpha, \beta \in [0, 2\pi]$  mit

$$\tilde{v} = (\cos \beta, \sin \beta)^T \quad \tilde{w} = (\cos \alpha, \sin \alpha)^T$$

Gilt  $\alpha, \beta \in [0, \pi]$  folgt aus einem Additionstheorem für  $\cos$

$$\begin{aligned} \cos(\beta - \alpha) &= \cos \alpha \cos \beta + \sin \alpha \sin \beta \\ &= \langle \tilde{v}, \tilde{w} \rangle \cdot 1 \cdot 1 \end{aligned}$$

$$\angle(\tilde{v}, \tilde{w}) = \cos(\beta - \alpha)$$

Man kann den Winkel auch über die Gleichung

$$\langle v, w \rangle = \|v\| \cdot \|w\| \cdot \cos(\angle(v, w))$$

definiere. Dann ist auch  $v = 0$  und/oder  $w = 0$  erlaubt. Stehen  $v$  und  $w$  senkrecht aufeinander ( $v \perp w$ )

$$\cos(\angle(v, w)) = \cos\left(\frac{\pi}{2}\right) = 0 \quad \Rightarrow \quad \langle v, w \rangle = 0$$

### Definition 3.9: orthogonal

Sei  $V$  ein endlichdimensionaler euklidischer oder unitärer Vektorraum.

1. Zwei Vektoren  $v, w \in V$  heißen **orthogonal** bezüglich des gegebenen Skalarproduktes  $\langle \cdot, \cdot \rangle$ , wenn gilt  $\langle v, w \rangle = 0$ .
2. Für dieses Skalarprodukt heißt eine Basis  $\{v_1, \dots, v_n\}$  von  $V$  **Orthogonalbasis**, wenn

$$\langle v_i, v_j \rangle = 0 \quad i, j = 1, \dots, n, \quad i \neq j$$

Ist zusätzlich für die induzierte Norm

$$\langle v_i, v_i \rangle^{\frac{1}{2}} = \|v_i\| = 1 \quad i = 1, \dots, n$$

so heißt  $\{v_1, \dots, v_n\}$  **Orthonormalbasis** von  $V$ . ( $\Leftrightarrow \langle v_i, v_j \rangle = \delta_{ij}$ )

**Satz 3.10:** Sei  $V$  ein euklidischer oder unitärer Vektorraum mit  $\dim(V) = n < \infty$ . Sei  $\{v_1, \dots, v_n\}$  eine Basis von  $V$ . Dann existiert eine Orthonormalbasis  $\{w_1, \dots, w_n\}$  von  $V$ .

*Beweis:* Per Induktion über  $n$ .

Induktionsanfang:  $n = 1$

Sei  $v_1 \in V$ ,  $v_1 \neq 0$ . Dann gilt für  $w_1 = \|v_1\|^{-1}v_1$ ,  $\|w_1\| = 1$  und  $\text{Span}\{v_1\} = \text{Span}\{w_1\}$ .  
 $\Rightarrow \{w_1\}$  ONB

Induktionsschritt:  $n \rightarrow n + 1$

Die Aussage gelte für  $n$ . Sei  $\dim(V) = n + 1$  und  $\{v_1, \dots, v_{n+1}\}$  eine Basis von  $V$ . Dann ist  $U = \text{Span}\{v_1, \dots, v_n\}$  ein  $n$ -dimensionaler Unterraum von  $V$ . Nach Induktionsvoraussetzung existiert eine ONB  $\{w_1, \dots, w_n\}$  von  $U$ . D.h.,

$$\text{Span}\{w_1, \dots, w_n\} = \text{Span}\{v_1, \dots, v_n\}$$

Für

$$\tilde{w}_{n+1} = v_{n+1} - \sum_{k=1}^n \langle v_{n+1}, w_k \rangle w_k$$

gilt wegen  $v_{n+1} \notin U$ , dass  $\tilde{w}_{n+1} \neq 0$ . Mit dem Austauschsatz von Steinitz (Satz 2.23, LinA I) folgt für  $w_{n+1} = \|\tilde{w}_{n+1}\|^{-1}\tilde{w}_{n+1}$ , dass

$$V = \text{Span}\{v_1, \dots, v_{n+1}\} = \text{Span}\{w_1, \dots, w_{n+1}\}$$

Für  $j = 1, \dots, n$  erhält man

$$\begin{aligned} \langle w_{n+1}, w_j \rangle &= \langle \|\tilde{w}_{n+1}\|^{-1} \tilde{w}_{n+1}, w_j \rangle \\ \|\tilde{w}_{n+1}\|^{-1} \langle v_{n+1} - \sum_{k=1}^n \langle v_{n+1}, w_k \rangle w_k, w_j \rangle \\ &= \|\tilde{w}_{n+1}\|^{-1} \left( \langle v_{n+1}, w_j \rangle - \sum_{k=1}^n \langle v_{n+1}, w_k \rangle \langle w_k, w_j \rangle \right) \\ \|\tilde{w}_{n+1}\|^{-1} (\langle v_{n+1}, w_j \rangle - \langle v_{n+1}, w_j \rangle) &= 0 \end{aligned}$$

$\Rightarrow \{w_1, \dots, w_{n+1}\}$  sind ONB.

□

Diese Orthogonalisierung ist als Gram-Schmidt-Verfahren bekannt. Jorgen Gram (dänischer Mathematiker, 1850-1916), Erhard Schmidt (deutscher Mathematiker, 1876-1959). Das Verfahren wurde bereits vor Laplace und Cauchy verwendet.

### Algorithmus 3.11: Gram-Schmidt-Verfahren

Gegeben:  $\{v_1, \dots, v_n\}$  als Basis eines euklidischen (unitären) Vektorraums  $V$



1. Setze  $w_1 := \|v_1\|^{-1}v_1$
2. Für  $j = 2, \dots, n$  setze

$$\tilde{w}_j := v_j - \sum_{k=1}^{j-1} \langle v_j, w_k \rangle w_k$$

$$w_j := \|\tilde{w}_j\|^{-1} \tilde{w}_j$$

Die ursprüngliche Basis  $\{v_1, \dots, v_n\}$  hat dann die Darstellung

$$(v_1, \dots, v_n) = (w_1, \dots, w_n) \underbrace{\begin{pmatrix} \|v_1\| & \langle v_1, w_1 \rangle & \dots & \langle v_n, w_1 \rangle \\ 0 & \|\tilde{w}_2\| & \ddots & \vdots \\ \vdots & \vdots & \ddots & \langle v_n, w_{n-1} \rangle \\ 0 & 0 & & \|\tilde{w}_n\| \end{pmatrix}}_{=R}$$

Da alle Diagonaleinträge von  $R$  ungleich 0 sind, ist  $R$  invertierbar. Sei nun  $U$  ein  $m$ -dimensionaler Unterraum von  $\mathbb{R}^n$  oder  $\mathbb{C}^n$  mit dem Skalarprodukt. Wir definieren eine Orthonormalbasis  $\{w_1, \dots, w_m\}$  die Matrix

$$Q = (w_1, \dots, w_m) \in K^{n,m}$$

Damit gilt im reellen Fall

$$\mathbb{R}^{m,m} \ni Q^T Q = (w_i^T w_j)_{i,j=1,\dots,m} = (\delta_{ij})_{i,j=1,\dots,m} = I_m$$

und im komplexen Fall

$$\mathbb{C}^{m,m} \ni Q^H Q = (w_i^H w_j)_{i,j=1,\dots,m} = I_m$$

für  $m = n$ :  $Q^T = Q^{-1}$  bzw.  $Q^H = Q^{-1}$

Umgekehrt gilt: Ist für eine Matrix  $Q \in K^{m,n}$   $Q^T Q = I_m$  bzw.  $Q^H Q = I_m$ , so sind die Spalten von  $Q$  eine ONB bzgl. des Standardskalarproduktes eines  $m$ -dimensionalen Unterraums von  $\mathbb{R}^n$  bzw.  $\mathbb{C}^n$ . Damit gilt:

**Satz 3.12:** Sind  $v_1, \dots, v_m \in K^n$  linear unabhängig, dann gibt es eine Matrix  $Q \in K^{n,m}$  mit orthonormalen Spalten bezüglich des Standardskalarproduktes und eine obere Dreiecksmatrix  $R \in \text{GL}_m(K)$  mit

$$K^{n,m} \ni (v_1, \dots, v_m) = QR$$

als sogenannte  $QR$ -Zerlegung

$QR \rightarrow$  numerische lineare Algebra  $\rightarrow$  kleinste Quadrate-Problem

Die Matrix  $Q$  ist längenerhaltend:

**Lemma 3.13:** Sei  $Q \in K^{m,n}$  eine Matrix mit orthogonalen Spalten bzgl. des Standardskalarproduktes. Dann gilt  $\|v\|_2 = \|Qv\|_2$  für alle  $v \in K^n$ , wobei hier  $\|\cdot\|_2$  die euklidische Norm ist.

*Beweis:*

$$\|v\|_2^2 = \langle v, v \rangle = v^H v = v^H I v = v^H Q^H Q v = \|Qv\|_2^2$$

□

### Definition 3.14: Orthogonale und unitäre Matrizen

- Eine Matrix  $Q \in \mathbb{R}^{n,n}$  heißt **orthogonal**, wenn  $Q^T Q = I_n$  gilt. Wir definieren

$$\mathcal{O}_n(\mathbb{R}) := \{Q \in \mathbb{R}^{n,n} \mid Q \text{ orthogonal}\}$$

- Eine Matrix  $Q \in \mathbb{C}^{n,n}$  heißt **unitär**, wenn  $Q^H Q = I_n$ . Wir definieren

$$U_n(\mathbb{C}) := \{Q \in \mathbb{C}^{n,n} \mid Q \text{ unitär}\}$$

Für orthogonale bzw. unitäre Matrizen gilt

$$\mathbb{R}^{n,n} \ni Q^T Q = I_n \implies Q^T = Q^{-1}, \mathbb{C}^{n,n} \ni Q^H Q = I_n \implies Q^H = Q^{-1}$$

D.h.

**Lemma 3.15:** Die Mengen  $\mathcal{O}_n(\mathbb{R})$  und  $U_n(\mathbb{C})$  bilden Untergruppen von  $\text{GL}_n(\mathbb{R})$  und  $\text{GL}_n(\mathbb{C})$ .

*Beweis:* Hier nur  $\text{GL}_n(\mathbb{R})$

Für  $I_n \in \mathbb{R}^{n,n}$  gilt  $I_n^T I_n = I_n \implies I_n \in \mathcal{O}_n(\mathbb{R}) \implies \mathcal{O}_n(\mathbb{R}) \neq \emptyset$ .

zu zeigen: Gruppeneigenschaften

1. Abgeschlossenheit bzgl. der inneren Verknüpfung

Sind  $Q_1, Q_2 \in \mathcal{O}_n(\mathbb{R})$ . Dann gilt:

$$\begin{aligned} (Q_1 Q_2)^T Q_1 Q_2 &= Q_2^T Q_1^T Q_1 Q_2 = I_n \\ \implies Q_1 Q_2 &\in \mathcal{O}_n(\mathbb{R}) \end{aligned}$$

2. Neutrales Element:  $I_n$

3. Inverses Element:  $Q^{-1} = Q^T$

Jetzt: Übertragung auf Endomorphismen, auch der geometrische Aspekt

**Definition 3.16: orthogonale Abbildung**

Eine Abbildung  $f \in L(V, V)$  heißt **orthogonal** ( $V = \mathbb{R}$ ) bzw. **unitär** ( $V = \mathbb{C}$ ) falls gilt

$$\langle f(v), f(w) \rangle = \langle v, w \rangle \quad \forall v, w \in V$$

**Definition 3.17:**

Wir definieren für einen euklidischen Vektorraum  $V$

$$\mathcal{O}(V) := \{f \in L(V, V) \mid f \text{ orthogonal}\}$$

bzw. für einen unitären Vektorraum  $V$

$$U(V) := \{f \in L(V, V) \mid f \text{ unitär}\}$$

**Lemma 3.18:** Sei  $f \in L(V, V)$  orthogonal bzw. unitär. Dann gilt:

1.  $\|f(v)\| = \|v\| \quad \forall v \in V$  für die durch das Skalarprodukt induzierte Norm
2.  $v \perp w \implies f(v) \perp f(w)$
3.  $f$  ist ein Isomorphismus und  $f^{-1}$  ist ebenfalls orthogonal bzw. unitär.
4. Ist  $\lambda \in K$  ein Eigenwert von  $f$ , so gilt  $|\lambda| = 1$

*Beweis:* 1 und 2 folgt direkt aus der Definition.

3: Injektivität folgt aus 1 + pos. Definitheit der Norm. Surjektivität folgt dann aus der Dimensionsformel. Aus der Surjektivität von  $f$  und  $f$  orthogonal bzw. unitär folgt diese Eigenschaft auch für  $f^{-1}$ . 4: Ist  $\lambda$  ein Eigenwert von  $f$  mit dem Eigenvektor  $v \neq 0$ , so gilt

$$\begin{aligned} \|v\| &= \|f(v)\| = \|\lambda v\| = |\lambda| \|v\| \\ 1 &= |\lambda| \end{aligned}$$

Aus der Definition des Skalarproduktes und orthogonal bzw. unitär folgt

**Korollar 3.19:** Gilt für  $f \in L(V, V)$ , dass

$$\|f(v)\| = \|v\|$$

für die durch das Skalarprodukt induzierte Norm, so ist  $f$  orthogonal bzw. unitär.

Aus diesen Gründen werden orthogonale bzw. unitäre Abbildungen auch Isometrien genannt.

**Satz 3.20:** Sei  $V$  ein euklidischer (unitärer) Vektorraum mit einer Orthonormalbasis  $B = \{v_1, \dots, v_n\}$  und  $f \in L(V, V)$ . Dann gilt:

$$f \in \mathcal{O}(V) \text{ bzw. } f \in U(V) \iff A_f^{B,B} \in \mathcal{O}_n(\mathbb{R}) \text{ bzw. } A_f^{B,B} \in U_n(\mathbb{C})$$

D.h. die Abbildungen

$$\mathcal{O}(V) \rightarrow \mathcal{O}_n(\mathbb{R}), f \mapsto A_f^{B,B} \text{ bzw. } U(V) \rightarrow U_n(\mathbb{C}), f \mapsto A_f^{B,B}$$

sind Isomorphismen.

*Beweis:* Hier nur für  $K = \mathbb{R}$

“ $\implies$ ”:  $f$  orthogonal

Dann gilt wegen der Orthonormalität von  $B$  für  $A_f^{B,B} = (a_{ij})$ , dass

$$\delta_{ij} = \langle v_i, v_j \rangle \stackrel{3.16}{=} \langle f(v_i), f(v_j) \rangle = \left\langle \sum_{l=1}^n a_{li} v_l, \sum_{k=1}^n a_{kj} v_k \right\rangle = \sum_{l=1}^n a_{li} a_{lj}$$

Also:

$$I_n = (A_f^{B,B})^T A_f^{B,B} \implies A_f^{B,B} \in \mathcal{O}_n(\mathbb{R})$$

“ $\impliedby$ ”:  $A_f^{B,B} \in \mathcal{O}_n(\mathbb{R})$ . Für die zugehörige lineare Abbildung  $f$  gilt wegen

$$f(v_i) = \sum_{l=1}^n a_{li} v_l,$$

dass

$$\begin{aligned} \langle f(v_i), f(v_j) \rangle &= \left\langle \sum_{l=1}^n a_{li} v_l, \sum_{k=1}^n a_{kj} v_k \right\rangle = \sum_{l=1}^n a_{li} a_{lj} \stackrel{A_f^{B,B} \in \mathcal{O}_n(\mathbb{R})}{=} \delta_{ij} = \langle v_i, v_j \rangle \\ &\implies f \in \mathcal{O}(V) \end{aligned}$$

□

### 3.3. Selbstadjungierte Abbildungen

Was ist ein adjungierter Endomorphismus?

**Lemma 3.21:** Sei  $V$  ein euklidischer (unitärer) Vektorraum und  $f \in L(V, V)$ . Dann gibt es genau ein  $g \in L(V, V)$  mit

$$\langle f(v), w \rangle = \langle v, g(w) \rangle \quad \forall v, w \in V$$

Ist  $B$  eine Orthonormalbasis von  $V$ , so gilt

$$A_g^{B,B} = \left( A_f^{B,B} \right)^H$$

*Beweis:* Hier nur für  $K = \mathbb{R}$ . Da  $B$  orthonormal ist gilt für  $v = \Phi_B(x)$  und  $w = \Phi_B(y)$ , dass

$$\langle v, w \rangle = \langle A_{\Phi_B}^{E,B}, A_{\Phi_B}^{E,B} \rangle = x^T \underbrace{\left( A_{\Phi_B}^{E,B} \right)^T A_{\Phi_B}^{E,B}}_I y = x^T y = \langle x, y \rangle_{\mathbb{R}^n} \quad \forall v, w \in V$$

Dann gilt für  $A_f^{B,B}$

$$\langle A_f^{B,B} x, y \rangle_{\mathbb{R}^n} = \left( A_f^{B,B} x \right)^T y = x^T \left( A_f^{B,B} \right)^T y = \langle x, \left( A_f^{B,B} \right)^T y \rangle_{\mathbb{R}^n}$$

Damit ist wegen der Definition des Skalarproduktes eindeutig eine lineare Abbildung mit der Darstellungsmatrix  $\left( A_f^{B,B} \right)^T$  gegeben. Diese bestimmt eindeutig den gesuchten Endomorphismus  $g$ . □

### Definition 3.22: adjungierter Endomorphismus

Die in Lemma 3.21 eindeutig definierte Abbildung  $g \in L(V, V)$  nennt man den zu  $f \in L(V, V)$  **adjungierten Endomorphismus**. Er wird mit  $f^{\text{ad}}$  bezeichnet.

### Definition 3.23: selbstadjungiert

Sei  $V$  ein euklidischer (unitärer) Vektorraum und  $f \in L(V, V)$ . Der Endomorphismus  $f$  heißt **selbstadjungiert**, wenn

$$\langle f(v), w \rangle = \langle v, f(w) \rangle \quad \forall v, w \in V$$

gilt. D.h.  $f^{\text{ad}} = f$ .

**Bemerkungen:** Es folgt unmittelbar

- Ist  $f \in L(V, V)$  und  $B$  eine ONB, so gilt

$$f \text{ selbstadjungiert} \iff A_f^{B,B} \text{ ist symmetrisch bzw. hermitesch, d.h. } A = A^H$$

- Ist  $f$  orthogonal bzw. unitär, so ist  $f^{\text{ad}} = f^{-1}$ , denn für  $u, v \in V$  mit  $w = f(u)$  d.h.  $u = f^{-1}(w)$  gilt

$$\langle f(v), w \rangle = \langle f(v), f(u) \rangle = \langle v, u \rangle = \langle v, f^{-1}(w) \rangle$$

**Lemma 3.24:** Sei  $V$  ein euklidischer (unitärer) Vektorraum und  $f \in L(V, V)$  selbstadjungiert. Dann sind alle Eigenwerte von  $f$  reell und das charakteristische Polynom zerfällt in Linearfaktoren.

*Beweis:* Sei zunächst  $K = \mathbb{C}$ . Sei  $\lambda$  ein Eigenwert von  $f$  mit zugehörigen Eigenvektor  $v \neq 0$ . Dann gilt

$$\lambda \underbrace{\langle v, v \rangle}_{>0} = \langle \lambda v, v \rangle = \langle f(v), v \rangle = \langle v, f(v) \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \langle v, v \rangle$$

$$\Rightarrow \lambda = \bar{\lambda} \in \mathbb{R}$$

Fundamentalsatz der Algebra  $\Rightarrow p_f(\cdot) = p_A(\cdot)$  zerfällt über  $\mathbb{C}$  in Linearfaktoren.

Sei nun  $K = \mathbb{R}$ .  $B$  ONB  $\Rightarrow A := A_f^{B,B} = (A_f^{B,B})^T$  ist eine spezielle komplexe Matrix  $\Rightarrow$  wie oben folgt für  $p_A(\cdot)$  betrachtet über  $\mathbb{C}$ , dass  $p_A(\cdot)$  in Linearfaktoren zerfällt

$$(\lambda - \lambda_i) \quad (\lambda_i \text{ ist EW} \in \mathbb{R})$$

$\Rightarrow p_A(\cdot)$  zerfällt auch über  $\mathbb{R}$  in Linearfaktoren.

□

**Satz 3.25:** Sei  $V$  ein euklidischer (unitärer) Vektorraum und  $f \in L(V, V)$  selbstadjungiert. Dann gibt es eine Orthonormalbasis von  $V$  die aus Eigenvektoren zu den reellen Eigenwerten von  $f$  besteht.

*Beweis:* Sei  $n = \dim(V) < \infty$ .

Für  $n = 1$ : klar ✓

$n - 1 \rightarrow n$ :

Wegen Lemma 3.24 gilt

$$p_f(\lambda) = \pm(\lambda - \lambda_1) \cdot \dots \cdot (\lambda - \lambda_n)$$

mit  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ . Zu  $\lambda_1$  existiert ein Eigenvektor  $v_1$  mit  $\|v_1\| = 1$ . Dann gilt für

$$u \in U := \{u \in V \mid \langle v_1, u \rangle = 0\},$$

dass

$$\langle v_1, f(u) \rangle = \langle f(v_1), u \rangle = \lambda_1 \underbrace{\langle v_1, u \rangle}_{=0} = 0$$

d.h.  $f(U) \subseteq U$ . Also ist  $U$  invariant unter  $f$ . Die Einschränkung  $f|_U : U \rightarrow U$  ist selbstadjungiert mit

$$p_{f|_U} = \pm(\lambda - \lambda_2) \cdot \dots \cdot (\lambda - \lambda_k)$$

Nach Induktionsvoraussetzung ex. ONB für  $U$ . Die Vereinigung dieses ONB mit  $v_1$  ist ONB für  $V$ .

□

Für die Matrixform erhalten wir analog:

**Lemma 3.26:** Sei  $A \in K^{n,n}$  symmetrisch (hermitesch). Dann gibt es ein  $T \in \text{GL}_n(K)$  und  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  so dass gilt

$$TAT^{-1} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

*Beweis:* Über Darstellungsmatrix eines selbstadjungierten Endomorphismus.

□

### Definition 3.27: positiv definite Matrix

Eine symmetrische (hermitesche) Matrix  $A \in K^{n,n}$  heißt **positiv definit**, wenn

$$v^T Av > 0 \text{ bzw. } v^H Av > 0 \quad \forall v \in V \setminus \{0\}$$

Lemma 3.26  $\implies A$  symmetrisch (hermitesch)  $\implies A$  diagonalisierbar

Des Weiteren gilt:

**Satz 3.28:** Sei  $A \in K^{n,n}$  symmetrisch (hermitesch). Dann sind folgende Aussagen äquivalent:

1.  $A$  ist positiv definit
2. Alle Eigenwerte  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  von  $A$  sind positiv.

*Beweis:* Hier nur  $K = \mathbb{C}$ ,  $K = \mathbb{R}$  folgt analog.

“1  $\implies$  2”: Sei  $\lambda$  Eigenwert von  $A$ . Lemma 3.26  $\implies \lambda \in \mathbb{R}$ . Sei  $v$  ein Eigenvektor zu  $\lambda$ , dann gilt:

$$0 < v^H Av = v^H A^H v = (Av)^H v = (\lambda v)^H v = \lambda v^H v = \lambda \underbrace{\|v\|^2}_{>0} \implies \lambda > 0$$

“2  $\implies$  1”: Satz 3.25: Es existiert ONB  $\{v_1, \dots, v_n\}$  bestehend aus Eigenvektoren zu Eigenwerten von  $A$ .

$$v_i^H Av_j = \lambda_i v_i^H v_j = \lambda_i \delta_{ij}$$

Jedes  $v \in V$  besitzt eine Darstellung

$$\begin{aligned}
v &= \sum_{i=1}^n \mu_i v_i, \quad \mu_i \in \mathbb{C}, 1 \leq i \leq n \\
v^H A v &= \left\langle \sum_{i=1}^n \mu_i v_i, \sum_{j=1}^n \mu_j A v_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \mu_i \bar{\mu}_j \lambda_j \underbrace{\langle v_i, v_j \rangle}_{\delta_{ij}} \\
&= \sum_{i=1}^n \mu_i \bar{\mu}_i \underbrace{\lambda_i}_{>0} = \sum_{i=1}^n \underbrace{|\mu_i|^2}_{\geq 0} \lambda_i > 0 \quad \text{für } v \neq 0
\end{aligned}$$

□

Zur Berechnung einer solchen ONB:

**Algorithmus 3.29:** Gegeben:  $A \in K^{n,n}$  bzw.  $f \in L(V, V)$  mit  $A_f^{B,B} = A$ .

- Bestimme

$$p_A(\lambda) = \pm(\lambda - \lambda_1)^{k_1} \cdot \dots \cdot (\lambda - \lambda_m)^{k_m}$$

mit paarweise verschiedenen  $\lambda_i, 1 \leq i \leq m$ . Ist dies nicht möglich: STOP

- Für jeden Eigenwert  $\lambda_i$  der algebraischen Vielfachheit  $k_i$  bestimme eine Basis des dazugehörigen Eigenraums  $\text{Eig}(A, \lambda_i)$ . Stimmen geometrische und algebraische Vielfachheit nicht überein: STOP
- Orthonormalisiere die Vereinigung der jeweiligen Basen mit dem Gram-Schmidt-Verfahren.

**Beispiel 3.30:** Fortsetzung von Beispiel 2.32

Wir betrachten wieder

$$A = \begin{pmatrix} 3 & -1 & 4 & -3 & -1 \\ 1 & 1 & -1 & 1 & 0 \\ -1 & 0 & 2 & 0 & 0 \\ 4 & 1 & -4 & 5 & 1 \\ -2 & 0 & 2 & -2 & 1 \end{pmatrix}$$

$$p_A(\lambda) = (\lambda - 1)^4(\lambda - 2) \quad \lambda_1 = 1 > 0, \lambda_2 = 2 > 0$$

Es gilt

$$\begin{aligned}
\text{Eig}(A, 1) &= \text{Span} \left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\} = \{v_1, v_2, v_3, v_4\} \\
\text{Eig}(A, 2) &= \text{Span} \left\{ (0 \ 1 \ 2 \ 3 \ -2)^T \right\} = \{v_5\}
\end{aligned}$$

GS-Verfahren



$$w_1 = \|v_1\|^{-1} v_1 = \frac{1}{2} (1 \ -1 \ 1 \ 0 \ 1)^T$$

$j = 1$

$$\tilde{w}_2 = v_2 - \sum_{k=1}^1 \langle v_2, w_k \rangle w_k = (1 \ 1 \ 1 \ 0 \ -1)^T - 0, \quad w_2 = \frac{1}{2} (1 \ 1 \ 1 \ 0 \ -1)^T$$

$j = 2$  :

$$\begin{aligned} \tilde{w}_3 &= v_3 - \sum_{k=1}^2 \langle v_3, w_k \rangle w_k = (0 \ 0 \ 1 \ 1 \ 0)^T - \frac{1}{2} \cdot \frac{1}{2} \cdot (1, -1, 1, 0, 1)^T - \frac{1}{2} \cdot \frac{1}{2} \cdot (1 \ 1 \ 1 \ 0 \ -1)^T \\ &= \left(-\frac{1}{2} \ 0 \ \frac{1}{2} \ 1 \ 0\right)^T, \quad w_3 = \frac{1}{\sqrt{\frac{3}{2}}} \left(-\frac{1}{2} \ 0 \ \frac{1}{2} \ 1 \ 0\right)^T \end{aligned}$$

$j = 3: \dots j = 4: \dots$

## 4. Affine Geometrie

Bisher als Struktur:

Gruppen  $\Rightarrow$  Körper  $\Rightarrow$  Vektorraum über Körper  $\Rightarrow$  Unterraum  
affine Unterräume als weitere Struktur

Zur Motivation/Startpunkt,  $\mathbb{R}^3$

Gerade:

$$G := \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + a \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \mid a \in \mathbb{R} \right\}$$

Ebene:

$$E := \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + a \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + b \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$$

Def. 6.5 aus LinA:  $G$  und  $E$  sind affine Unterräume des  $\mathbb{R}^3$ . Damit ist  $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}^T$  ein Punkt im Raum zu dem ein Vektor als Repräsentant einer Äquivalenzklasse addiert wird. D.h.

$$G : P \underset{=}{+} \vec{v}$$

Wir hatten schon:  $U = v \underset{=}{+} W$  mit  $W$  UVR

### 4.1. Operation einer Gruppe auf einer Menge

#### Definition 4.1: Wirkung einer Gruppe

Es sei  $G$  eine Gruppe mit der Verknüpfung  $\circ$  und dem neutralen Element  $e$  sowie eine Menge  $M$ . Eine Abbildung der Form

$$G \times M \rightarrow M, \quad (g, m) \mapsto g \bullet m$$

nennt man **Wirkung** oder **Operation** der Gruppe  $G$  auf der Menge  $M$ , falls gilt

1.  $(g_1 \circ g_2) \bullet m = g_1 \bullet (g_2 \bullet m) \quad \forall g_1, g_2 \in G, \forall m \in M$
2.  $e \bullet m = m \quad \forall m \in M$

#### Beispiel 4.2:

- Passend zum obigen Beispiel der Gerade/Ebene:

Sei  $G = V$  ein  $K$ -Vektorraum (vgl. Def. 2.26 LinA I),  $M = V$ . Dann ist durch

$$V \times V \rightarrow V, (v, x) \mapsto v + x$$

eine Operation von  $V$  auf sich selbst definiert.

- Für die Gruppe  $G := (\mathbb{R}, +)$  und  $M = S^1$  als Einheitskreis im  $\mathbb{R}^2$ , d.h.

$$S^1 = \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}$$

wird durch

$$(a, x) \mapsto e^{ia} \cdot x \quad \forall a \in G, \forall x \in M$$

eine Operation von  $(\mathbb{R}, +)$  auf  $S^1$  gegeben.

- Sei  $G = \text{GL}_n(\mathbb{R})$  und  $M = \mathbb{R}^n$ . Dann definiert

$$(A, x) \mapsto A \cdot x \in M \quad \forall A \in G, \forall x \in M$$

eine Operation von  $\text{GL}_n(\mathbb{R})$  auf  $M$ .

#### Definition 4.3: Bahn von $m$

Eine Gruppe  $G$  wirke auf die Menge  $M$ . Für  $m \in M$  wird die Teilmenge

$$G \bullet m := \{g \bullet m \mid g \in G\} \subseteq M$$

die **Bahn von  $m$  unter  $G$**  genannt.

**Beobachtung:** In Beispiel 4.2:

In 1 und 2) entspricht die Bahn eines einzigen Elements der ganzen Menge.

In 3)

#### Definition 4.4: transitive Wirkung

Eine Gruppe  $G$  wirke auf der Menge  $M$ . Die Wirkung nennt man **transitiv**, wenn für alle Paare  $m, \tilde{m} \in M$  ein  $g \in G$  existiert, so dass

$$m = g \bullet \tilde{m}$$

Man nennt die Wirkung **einfach transitiv**, falls das Gruppenelement  $g$  eindeutig bestimmt ist.

**Lemma 4.5:** Eine Gruppe  $G$  wirke auf die Menge  $M$ . Dann gilt:

1. Ist die Wirkung transitiv, so gilt für jedes  $m \in M$  die Gleichheit  $G \bullet m = M$
2. Ist die Wirkung einfach transitiv, so existiert eine Bijektion zwischen  $M$  und  $G$ .

*Beweis:*

zu 1) Sei  $m \in M$  bel. gewählt. Dann existiert wegen der transitiven Wirkung zu  $\tilde{m} \in M$  ein  $g \in G$  mit  $\tilde{m} = g \bullet m \implies M = G \bullet m$

zu 2) Für ein fest gewähltes  $m \in M$ , definiert man

$$\psi_m : G \rightarrow M, \quad g \mapsto g \bullet m$$

Wegen der transitiven Wirkung ist  $\varphi_m$  surjektiv. Da die Wirkung einfach transitiv ist, ist  $\varphi_m$  auch injektiv  $\implies$  Bijektivität

□

## 4.2. Affine Räume

### Definition 4.6: affiner Raum

Sei  $V$  ein  $K$ -Vektorraum. Eine nichtleere Menge  $M$  heißt **affiner Raum** über dem Vektorraum  $V$ , wenn  $V$  einfach transitiv auf  $M$  wirkt. Die Elemente von  $M$  werden als **Punkte** bezeichnet. Ist  $M = \emptyset$ , so wird  $M$  ebenfalls als affiner Raum aufgefasst.

Wie passt das zu 6.5 aus LinA I?

**Beispiel 4.7:** Sei  $\mathcal{L}(A, b)$  die Lösungsmenge des LGS  $Ax = b$  mit  $A \in \mathbb{R}^{m,n}$ ,  $x \in \mathbb{R}^n$ ,  $b \in \mathbb{R}^m$ . Im Satz 6.3, LinA I, haben wir gezeigt, dass  $\mathcal{L}(A, 0)$  ist ein Unterraum des  $\mathbb{R}^n$  ( $\implies$  Gruppe). Ist  $\mathcal{L}(A, b) \neq \emptyset$ , gilt nach Satz 6.4, LinA I, dass

$$\mathcal{L}(A, b) = x_* + \mathcal{L}(A, 0)$$

für ein beliebiges  $x_* \in \mathcal{L}(A, b)$ . Dann ist  $M = \mathcal{L}(A, b)$  ein affiner Raum über dem Vektorraum  $\mathcal{L}(A, 0)$ , denn es gilt

$$+ : \mathcal{L}(A, 0) \times \mathcal{L}(A, b) \rightarrow \mathcal{L}(A, b), \quad (y, x) \mapsto y + x$$

dass  $\forall y \in G, \forall x \in M$ :

$$A(y + x) = Ay + Ax = 0 + b = b$$

$$\implies y + x \in M$$

$\forall y, \tilde{y} \in G, \forall x \in M$  gilt

1.  $(y +_G \tilde{y}) +_W x = y +_W (\tilde{y} +_W x) = y +_{\mathbb{R}^n} \tilde{y} +_{\mathbb{R}^n} x$
2.  $0 +_W x = 0 +_{\mathbb{R}^n} x = x$

$\implies +$  ist eine Wirkung der Gruppe  $G$  auf die Menge  $M$ . Sind  $x, \tilde{x} \in \mathcal{L}(A, b)$  ist  $Ax = A\tilde{x} = b \implies$

$$A(x - \tilde{x}) = b - b = 0$$

$y := x - \tilde{x} \in \mathcal{L}(A, 0) \implies x = (x - \tilde{x}) + \tilde{x} = y + \tilde{x} \implies$  Wirkung ist transitiv

Sei  $\tilde{y} \in \mathcal{L}(A, 0) = G$  so gewählt, dass auch

$$x = \tilde{y} + \tilde{x}$$

gilt.

$$x = y + \tilde{x}$$

$$x = \tilde{y} + \tilde{x}$$

$$\implies 0 = y - \tilde{y} \implies y = \tilde{y} \implies \text{einfach transitiv}$$

□

**Korollar 4.8:** Sind  $M$  und  $\tilde{M}$  zwei affine Räume, so existiert eine Bijektion zwischen  $M$  und  $\tilde{M}$ .

*Beweis:* Folgt aus Lemma 4.5 und Komposition zweier bijektiver Abbildungen.

□

**Folgerung:** Ein affiner Raum ist bis auf eine Bijektion eindeutig bestimmt. Damit ist folgendes sinnvoll:

**Definition 4.9:**

Wir bezeichnen den affinen Raum über den zugehörigen  $K$ -Vektorraum  $V$  mit  $A(v)$  bzw.  $A$ , wenn der Kontext klar ist. Die einfach transitive Wirkung  $\bullet$  von  $V$  auf  $A(v)$  wird mit  $+$  bezeichnet, d.h.

$$x \bullet P := P + x, \quad x \in V, P \in A(v)$$

**Lemma 4.10:** Sei  $V$  ein  $K$ -Vektorraum und  $A$  ein affiner Raum über  $V$ . Sei  $P, Q, R, S \in A$  und  $v, w \in V$ . Dann gelten folgende Aussagen:

1.  $P + v = P + w \implies v = w$   
D.h. für  $Q = P + x \in A$  ist der Vektor  $x \in V$  eindeutig bestimmt.
2.  $P + v = Q + v \implies P = Q$
3.  $P + v = Q \implies P = Q + (-v)$
4. Für  $Q = P + v \in A$  wird  $v$  als Verbindungsvektor von  $P$  nach  $Q$  bezeichnet und man schreibt

$$v = \overrightarrow{PQ}$$

Es gilt

$$\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$$

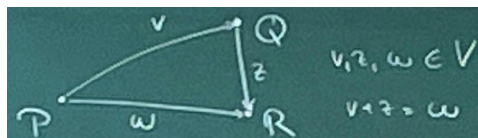
5. Für  $n \in \mathbb{N}$  Punkte,  $n > 1$ ,  $P_1, \dots, P_n \in A$  gilt

$$\overrightarrow{P_1 P_2} + \overrightarrow{P_2 P_3} + \dots + \overrightarrow{P_{n-1} P_n} = \sum_{i=1}^{n-1} \overrightarrow{P_i P_{i+1}} = \overrightarrow{P_1 P_n}$$

6.  $\overrightarrow{PQ} + \overrightarrow{QP} = 0 \in V$ , d.h.  $\overrightarrow{PQ} = -\overrightarrow{QP} \in V$
7.  $\overrightarrow{PQ} = \overrightarrow{RS} \implies \overrightarrow{PR} = \overrightarrow{QS}$

*Beweis:* Hier nur der Beweis von einigen Aussagen

zu 1: Wegen der einfachen Transitivität existiert genau ein Vektor  $v \in V$  mit  $Q = P + v = P + w$



Formal:  $\overrightarrow{PQ}, \overrightarrow{QR}, \overrightarrow{PR}$  sind definitionsgemäß die eindeutig bestimmten Vektoren, für die gilt

$$Q = P + \overrightarrow{PQ}, R = Q + \overrightarrow{QR}, R = P + \overrightarrow{PR}$$

Damit folgt

$$R = \left( P + \overrightarrow{PQ} \right) + \overrightarrow{QR} = P + \left( \overrightarrow{PQ} + \overrightarrow{QR} \right)$$

zu 7: Sei  $\overrightarrow{PQ} = \overrightarrow{RS}$ . Dann folgt mit 4):

$$\overrightarrow{PR} = \overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{RS} + \overrightarrow{QR} = \overrightarrow{QR} + \overrightarrow{RS} \stackrel{4)}{=} \overrightarrow{QS}$$

□

**Definition 4.11:**

Sei  $V$  ein  $K$ -Vektorraum und  $A$  ein affiner Raum über  $V$ . Für einen Punkt  $O \in A$  definiert man:

$$\psi_O : V \rightarrow A, \quad x \mapsto P := O + x$$

Aus Lemma 4.10 folgt unmittelbar, dass  $\psi_O$  eine Bijektion ist. D.h. für alle  $P \in A$  ist der Vektor  $v_O(P)$  das eindeutig bestimmte Element in  $V$  mit

$$P = O + v_O(P)$$

**Definition 4.12: Dimension affiner Räume**

Sei  $V$  ein  $K$ -Vektorraum und  $A$  ein affiner Raum über  $V$ . Dann ist

$$\dim A := \dim V$$

die **Dimension** von  $A$ . Ist  $A = \emptyset$ , so definiert man  $\dim A = -1$ .

Als Verallgemeinerung von Def. 6.5, LinA I:

**Definition 4.13: affiner Unterraum**

Sei  $V$  ein  $K$ -Vektorraum,  $A$  ein affiner Raum über  $V$  mit der Verknüpfung  $+: V \times A \rightarrow A$  und  $P \in A$ . Ist  $U \subseteq V$  ein Untervektorraum von  $V$ , so nennt man die Menge

$$B := P + U := \{Q \in A \mid \exists u \in U : Q = u + P\}$$

einen **affinen Unterraum** von  $A$ .

**Lemma 4.14:** Sei  $B$  ein affiner Unterraum des affinen Raums  $A(V)$ , d.h.  $B \subseteq A(V)$ . Damit ist  $B$  selbst ein affiner Raum über einen Vektorraum  $U \subseteq V$ .

*Beweis:* Nach Definition existiert zu  $B$  ein  $P \in A(V)$  und ein Unterraum  $U \subseteq V$ :

$$B = \{Q \in A \mid \exists u \in U : Q = P + u\}$$

$A$  affiner Raum  $\implies \exists + : V \times A \rightarrow A$

Einschränkung auf  $B$  liefert:

$$+ : V \times B \rightarrow B$$

$$B = \{Q \in A \mid \exists v \in U : Q = P + v\}, \quad + : U \times B \rightarrow B \text{ wohldefiniert?}$$

$\forall Q \in A : \forall v \in V$  ist  $Q = P + v$  definiert.

$\implies \forall Q \in B \subseteq A, \forall u \in U \subseteq V$  ist  $Q = P + u$  wohldefiniert. Des Weiteren gilt: für  $Q \in B$ ,  $u \in U$  sowie  $v \in U$  erhält man

$$Q + u = (P + v) + u = \underbrace{\vec{U} + (v + u)}_{\in U} \in B$$

Auch bei der Einschränkung auf  $B$  bzw.  $U$  bleibt die einfache Transitivität erhalten.

□

Analogo zu Satz 6.6 aus LinA I kann man zeigen:

**Satz 4.15:** Sei  $V$  ein  $K$ -Vektorraum,  $A$  ein affiner Raum über  $V$ ,  $P, \tilde{P} \in A$  und  $U, \tilde{U} \subseteq V$  Untervektorräume von  $V$ . Dann gilt:

1. Für jedes  $Q \in P + U$  ist  $P + U = Q + U$
2. Gilt  $P + U = \tilde{P} + \tilde{U}$ , so ist  $U = \tilde{U}$  und  $\vec{P\tilde{P}} \in U = \tilde{U}$

*Beweis:* siehe LinA I

□

#### Definition 4.16: Aufpunkt und Richtung

Sei  $V$  ein  $K$ -Vektorraum,  $A$  ein affiner Raum über  $V$  und  $A(W)$  ein affiner Unterraum von  $A$ . Gilt

$$A(W) = P + W$$

so nennt man  $P$  einen **Aufpunkt** von  $A(W)$  und den Untervektorraum  $W$  die **Richtung** von  $A(W)$

### 4.3. Lagebeziehungen von affinen Unterräumen

#### Definition 4.17: (schwach) parallel

Sei  $V$  ein  $K$ -Vektorraum,  $A(V)$  ein affiner Raum und  $A(W_1), A(W_2)$  zwei affine Unterräume von  $A(V)$ .

- $A(W_1)$  und  $A(W_2)$  heißen **parallel**, wenn  $W_1 = W_2$  gilt ( $A(W_1) \parallel A(W_2)$ )
- $A(W_1)$  und  $A(W_2)$  heißen **schwach parallel**, falls  $W_1 \subset W_2$  gilt ( $A(W_1) \triangleleft A(W_2)$ )

**Satz 4.18:** Sei  $V$  ein  $K$ -Vektorraum,  $A(V)$  ein affiner Raum über  $V$  und  $A(W_1), A(W_2)$  zwei parallele affine Unterräume. Dann gilt entweder  $A(W_1) = A(W_2)$  oder  $A(W_1) \cap A(W_2) = \emptyset$

*Beweis:* Gilt  $A(W_1) \parallel A(W_2) \implies W_1 = W_2$

Annahme:  $A(W_1) \cap A(W_2) \neq \emptyset$ . Dann existiert ein  $p \in A(W_1) \cap A(W_2)$ . Satz 4.15 liefert



$$A(W_1) = P + W_1 = P + W_2 = A(W_2)$$

□

Bekannt ist:

- Ein 0-dimensionaler affine Unterraum  $\mathbb{R}^3$  heißt Punkt im  $\mathbb{R}^3$ .
- Ein 1-dimensionaler affine Unterraum des  $\mathbb{R}^3$  heißt Gerade im  $\mathbb{R}^3$ .
- Ein 2-dimensionaler affine Unterraum des  $\mathbb{R}^3$  heißt Ebene in  $\mathbb{R}^3$ .

Verallgemeinerung:

**Definition 4.19:**

Sei  $V$  ein  $K$ -Vektorraum,  $A(V)$  ein affiner Raum über  $V$  und  $A(W)$  ein affiner Unterraum von  $A(V)$ .

- Ist  $\dim(A(W)) = 0$ , so heißt  $A(W)$  **(affiner) Punkt** von  $A(V)$ .
- Ist  $\dim(A(W)) = 1$ , so heißt  $A(W)$  **(affine) Gerade** von  $A(V)$ .
- Ist  $\dim(A(W)) = 2$ , so heißt  $A(W)$  **(affine) Ebene** von  $A(V)$ .

**Bemerkung:** Geraden können maximal schwach parallel zu Ebenen sein!

Für Untervektorräume gilt:  $\dim(U_1 \cap U_2) = \dim(U_1) + \dim(U_2) - \dim(U_1 + U_2)$ , Satz 3.40, LinA I

**Lemma 4.20:** Es seien  $U_1$  und  $U_2$  zwei Untervektorräume eines  $K$ -Vektorraums  $V$  sowie  $A(U_1) =: A_1$  und  $A(U_2) =: A_2$  zwei affine Unterräume des affinen Raums  $A(V)$ . Ist  $A_1 \cap A_2 \neq \emptyset$ , so ist  $A_1 \cap A_2$  ein affiner Unterraum von  $A(V)$  mit dem zugehörigen Untervektorraum  $U_1 \cap U_2$  und es gilt

$$\dim(A_1 \cap A_2) = \dim(U_1 \cap U_2)$$

*Beweis:* Es gilt:

$$A_1 = P_1 + U_1 \quad \text{und} \quad A_2 = P_2 + U_2$$

$$A_1 \cap A_2 \neq \emptyset \implies \exists Q \in A_1 \cap A_2$$

$$A_1 \cap A_2 = \{P \in A \mid \exists u_1 \in U_1, u_2 \in U_2 : P = Q + u_1 = Q + u_2\}$$

Für jedes Paar  $(P, Q)$  von Punkten aus  $A$  genau einen Vektor  $v \in V$  mit  $P = Q + v$  (Lemma 4.10, 1)

$$\implies u_1 = u_2 \implies A_1 \cap A_2 = \left\{ P \in A \mid \exists u \in \underbrace{U_1 \cap U_2}_{\text{UVR}} : P = Q + u \right\}$$

$\implies A_1 \cap A_2$  affiner Raum.

$\dim(A_1 \cap A_2) = \dim(U_1 \cap U_2)$  nach Def.

□

**Lemma 4.21:** Es seien  $U_1$  und  $U_2$  zwei Untervektorräume des  $K$ -Vektorraums  $V$ ,  $A_1 = A(U_1)$  und  $A_2 = A(U_2)$  zwei affine Unterräume eines affinen Raums  $A(V)$  sowie  $P_1 \in A_1$  und  $P_2 \in A_2$  zwei beliebige Punkte

$$A_1 \cap A_2 \neq \emptyset \iff \overrightarrow{P_1 P_2} \in U_1 + U_2$$

*Beweis:*

“ $\implies$ ”:  $A_1 \cap A_2 \neq \emptyset \implies \exists Q \in A_1 \cap A_2$

Dann liegen die Verbindungsvektoren  $\overrightarrow{P_1 Q}$  bzw.  $\overrightarrow{P_2 Q}$  in den jeweiligen Untervektorräumen  $U_1$  bzw.  $U_2$ . Lemma 4.10, 4):

$$\overrightarrow{P_1 P_2} = \underbrace{\overrightarrow{P_1 Q}}_{\in U_1} + \underbrace{\overrightarrow{Q P_2}}_{\in U_2} \in U_1 + U_2 \quad \checkmark$$

“ $\impliedby$ ”: Sei  $\overrightarrow{P_1 P_2} \in U_1 + U_2 \implies$

$\exists u_1 \in U_1, u_2 \in U_2 : \overrightarrow{P_1 P_2} = u_1 + u_2$ . Setzt man  $Q := P_1 + u_1 \in A_1$ , so gilt

$$\begin{aligned} Q &= P_1 + u_1 = P_1 + ((u_1 + u_2) - u_2) = P_1 + \left( \overrightarrow{P_1 P_2} - u_2 \right) \\ &= \left( P_1 + \overrightarrow{P_1 P_2} \right) - u_2 = \underbrace{P_2}_{\in A_2} + \underbrace{(-u_2)}_{\in U_2} \in A_2 \end{aligned}$$

$\implies A_1 \cap A_2 \neq \emptyset$

□

### Definition 4.22: affine Hülle

Sei  $M \subset A(V)$  eine Teilmenge eines affinen Raumes  $A(V)$  über einem  $K$ -Vektorraum  $V$ . Der kleinste affine Unterraum von  $A$ , der  $M$  enthält, wird **affine Hülle** von  $M$  genannt und mit  $\langle M \rangle_{\text{aff}}$  bezeichnet.

Sind  $A(U_1)$  und  $A(U_2)$  zwei affine Unterräume eines affinen Raums  $A(V)$ , so bezeichnen wir die affine hülle  $\langle A(U_1) \cup A(U_2) \rangle_{\text{aff}}$  als Verbindungsraum von  $A_1$  und  $A_2$ .

**Lemma 4.23:** Seien  $U_1, U_2 \subseteq V$  zwei Untervektorräume des  $K$ -Vektorraums  $V$ ,  $A_1 = A(U_1)$  zwei offene Unterräume eines affinen Raums  $A(V)$ , sowie  $P_1 \in A_1$  und  $P_2 \in A_2$  d.h.

$$A_1 = P_1 + U_1 \quad \text{und} \quad A_2 = P_2 + U_2$$

Dann ist der Verbindungsraum durch

$$\langle A_1 \cup A_2 \rangle_{\text{aff}} = P_1 + \left( \text{Span} \left( \overrightarrow{P_1 P_2} \right) + U_1 + U_2 \right)$$