Lineare Algebra und Analytische Geometrie II
Prof. Dr. Steffen König

David Holzmüller

18. Oktober 2016

Inhaltsverzeichnis

Kapitel 1

Quotientenräume und Dualräume

Sei K ein Körper und V ein K-Vektorraum:

- $(V,+)$ ist eine abelsche Gruppe mit $+ : V \times V \to V, (v_1, v_2) \mapsto v_1 + v_2$.
- Skalarmultiplikation $\cdot : K \times V \to V, (\lambda, v_1) \mapsto \lambda \cdot v_1$.

Beispiel 1.1.

$$V = K^n, n \in \mathbb{N}, V \ni \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, x_1, \dots, x_n \in K \ .$$

$$U = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ 0 \end{pmatrix} \middle| \, x_1, \dots, x_{n-1} \in K \right\}$$

ist abgeschlossen unter + und $\cdot$ und es ist $0 \in K$, also ist U ein Untervektorraum von V.

Für K-Vektorräume V, W ist $\varphi : V \to W$ eine lineare Abbildung, falls gilt:

$$\forall v_1, v_2 \in V : \varphi(v_1 \underset{V}{+} v_2) = \varphi(v_1) \underset{W}{+} \varphi(v_2)$$
$$\forall \lambda \in K, \forall v \in V : \varphi(\lambda \underset{V}{\cdot} v) = \lambda \underset{W}{\cdot} \varphi(v) \ .$$

Wir definieren

$$\operatorname{Kern}(\varphi) := \{v \in V \mid \varphi(v) = 0_W\} < V \qquad \text{(UVR von } V\text{)}$$
$$\operatorname{Bild}(\varphi) := \operatorname{Im}(\varphi) := \{w \in W \mid \exists v \in V : w = \varphi(v)\} < W \; .$$

Gegeben einen Vektorraum V, dann ist mit $\varphi = \mathrm{id} : V \to V, v_1 \mapsto v_1$ auch $V = \operatorname{Im}(\varphi)$.

Frage: Gegeben $U < V$. Gibt es ein W und ein lineares $\varphi : V \to W$, sodass $U = \operatorname{Kern}(\varphi)$? Können wir verlangen, dass φ surjektiv ist? (Die Lösung wird eine neue Konstruktion von Vektorräumen benutzen.)

Falls $\dim V < \infty$ und $\dim W < \infty$: $\exists$ Basis $v_1, \ldots, v_n$ von V und $\exists$ Basis $w_1, \ldots, w_l$ von W.

$\varphi : V \to W$ linear ist gegeben durch eine darstellende Matrix A (bezüglich der fest gewählten Basen).

$$\begin{aligned} \varphi(v) &= Av, \\ \operatorname{Kern}(\varphi) &= \{v \in V \mid Av = 0_W\} \\ &= \text{ Lösungsmenge des homogenen LGS } Ax = 0 \; . \end{aligned}$$

Gauß-Algorithmus: Gegeben LGS $Ax = 0 \rightsquigarrow$ Lösungsmenge $U < V$. Unser Problem: Gegeben U, finde A (inverses Problem).

Zunächst: Gruppen.

Beispiel 1.2. $G = \mathbb{Z}, * = +$, abelsch. $U = 2\mathbb{Z} = \{2n \mid n \in \mathbb{Z}\} = \{$ gerade Zahlen $\}$. $U < G : U \ni 0$ und für alle $a, b \in U : a + b \in U, -a \in U$. Es ist

$$\mathbb{Z} = G = U \dot\cup \underbrace{(G \setminus U)}_{1+2\mathbb{Z}}$$

eine Partition von $\mathbb{Z}$, diese entspricht einer Äquivalenzrelation auf $\mathbb{Z}$.

Allgemein: $n \in \mathbb{N}, a, b, \in \mathbb{Z} : a \sim_n b :\Leftrightarrow n \mid (b - a)$ definiert eine Äquivalenzrelation auf $\mathbb{Z}$:

- reflexiv: $a \sim_n a : n \mid (a - a)$.
- symmetrisch: $a \sim_n b : n \mid (b - a) \Rightarrow n \mid (a - b) \Rightarrow b \sim_n a$.

- transitiv: $a \sim_n b \ \wedge \ b \sim_n c \Rightarrow n \mid (b-a) \ \wedge \ n \mid (c-b) \Rightarrow c-a = (c-b)+(b-a)$ Vielfaches von $n \Rightarrow a \sim_n c$.

Konkret: a durch n mit Rest teilen: $a = ns_1 + r_1$ wobei $0 \leq r_1 \leq n-1$ und $b = ns_2 + r_2$ mit $0 \leq r_2 \leq n-1$. $a \sim_n b \Leftrightarrow r_1 = r_2$.
Die Zahlen $0, 1, \ldots, n-1$ sind Repräsentatnen der n verschiedenen Äquivalenzklassen (= Restklassen modulo n).
Sei $r \in \{0, \ldots, n-1\}$ und $\overline{r} = \{a \in \mathbb{Z} \mid a \sim_n r\} = r + n\mathbb{Z} = \{a \in \mathbb{Z} \mid a \equiv r \pmod{n}\}$ („kongruent modulo n“).
Es gilt $\overline{r} = \overline{r+n} = \overline{r-5n}$, $\overline{0} = U$ ist eine Untergruppe von $\mathbb{Z}$ aber $\overline{1} \ldots, \overline{n-1}$ sind keine Untergruppen von $\mathbb{Z}$.

$\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$ ist selbst eine Gruppe mit der Operation $* = + : \overline{a} + \overline{b} := \overline{a+b}$.
Zu zeigen: $+$ ist wohldefiniert, d. h. nicht von der Wahl der Repräsentanten abhängig. Sei $\overline{c} = \overline{a}$ und $\overline{d} = \overline{b}$, dann müssen wir zeigen, dass $\overline{a+b} = \overline{c+d}$. Wir wissen jedoch $n \mid (a-c)$ und $n \mid (b-d)$, also gilt auch $n \mid (a-c+b-d) = (a+b)-(c+d)$.

- neutrales Element: $\overline{0}$, denn $\overline{a} + \overline{0} = \overline{a+0} = \overline{a} = \overline{0} + \overline{a}$.
- inverses Element zu $\overline{a}$: $\overline{-a} + \overline{a} = \overline{-a+a} = \overline{0} = \overline{a} + \overline{-a}$.
- Assoziativität: $(\overline{a} + \overline{b}) + \overline{c} = \overline{a+b} + \overline{c} = \overline{(a+b)+c} = \overline{a+(b+c)} = \overline{a} + (\overline{b} + \overline{c})$.

Proposition 1.1. *Sei $n \in \mathbb{N}$. Die Menge $\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$ der Restklassen ganzer Zahlen modulo n bildet eine Gruppe mit Addition $\overline{a} + \overline{b} := \overline{a+b}$. Das neutrale Element ist $\overline{0}$. Inverses zu $\overline{a}$ ist $\overline{-a}$.*

Konstruktion von $\mathbb{Z}/n\mathbb{Z}$: Input Gruppe $G = \mathbb{Z}$, Untergruppe $U = n\mathbb{Z} < \mathbb{Z} = G$. Die Gruppenoperation von $\mathbb{Z}/n\mathbb{Z}$ kommt von der Operation auf $\mathbb{Z}$.
Frage: Sei G irgendeine Gruppe und $U < G$ eine Untergruppe. Gibt es eine Gruppe G/U?

Im Folgenden wird die Gruppenoperation $*$ als Multiplikation $\cdot$ geschrieben.

Definition 1.2. Sei G eine Gruppe, U eine Untergruppe. Die Relation $a \sim_U b$ (oder $a \equiv b \pmod{U}$) ist definiert durch $a \sim_U b :\Leftrightarrow ab^{-1} \in U$.

Beispiel 1.3. $U = n\mathbb{Z} < G = \mathbb{Z}$: $ab^{-1} = a - b$, d. h. $a - b \in n\mathbb{Z} = U$ ist die Bedingung.

Ist $\sim_U$ eine Äquivalenzrelation?

- Reflexivität: $a \sim_U a$: $aa^{-1} = e$ (neutrales Element) $\in U$ (da U eine Untergruppe ist).
- Symmetrie: $a \sim_U b$: $ab^{-1} \in U \Rightarrow ba^{-1} = (b^{-1})^{-1}a^{-1} = (ab^{-1})^{-1} \in U$, also $b \sim_U a$.
- Transitivität: $a \sim_U b : ab^{-1} \in U$ und $b \sim_U c : bc^{-1} \in U \Rightarrow ac^{-1} = (ab^{-1})(bc^{-1})$, also $a \sim_U c$.

Somit ist $\sim_U$ eine Äquivalenzrelation auf G.

Definition 1.3. Die Äquivalenzklassen unter $\sim_U$ heißen *Rechtsnebenklassen* (oder Rechtsrestklassen) von U in G. Bezeichnung: $Ua := \{b \in G \mid a \sim_U b\}, a \in G$.

Es gilt: $Ua = \{ua \mid u \in U\} \ni a = e \cdot a$, wobei e das neutrale Element in U ist. Denn ist $a \sim_U b$ oder $b \sim_U a$, dann folgt mit $u := ba^{-1} \in U$, dass $ua = (ba^{-1})a = b$. D. h. jedes $b \in G$ mit $b \sim_U a$ hat immer die Form ua für ein $u \in U$. Umgekehrt sei $b = ua$ für ein $u \in U$. Dann folgt $ba^{-1} = (ua)a^{-1} = u \in U$, also $b \sim_U a$. $U = Ue$ ist die Rechtsnebenklasse von e.

Sei Ua irgendeine Rechtsnebenklasse (RNK). Wir können eine Abbildung $U \to Ua$ angeben durch $u \mapsto ua$. Umgekehrt gibt es auch eine Abbildung $Ua \to U$, nämlich $ua \mapsto (ua)a^{-1} = u$. Diese sind zueinander inverse Bijektionen. Somit sind U und Ua gleichmächtig.

U	Ua	Ub	$\dots$

G

Proposition 1.4. *Sei G eine endliche Gruppe, U eine Untergruppe. Mit $|G|$ und $|U|$ wird die Ordnung (= Anzahl der Elemente) von G bzw. U bezeichnet.*

(a) Sei Ua eine RNK. Dann gilt $|Ua| = |U|$.

(b) *Sei m die Anzahl der RNK. Dann gilt $|G| = |U| \cdot m$.*

(c) *Insbesondere gilt (Satz von Lagrange): $|U|$ teilt $|G|$.*

Weitere Notation: $\operatorname{ord}(G) \coloneqq |G|$. Für $g \in G$ definieren wir $\operatorname{ord}(g) \coloneqq \min\{n \in \mathbb{N} \mid g^n = e\}$, wobei $\operatorname{ord}(g) \coloneqq \infty$, falls kein solches n existiert.

Die Menge $\{g^n \mid n \in \mathbb{Z}\}$ ist eine Untergruppe von G. Falls $\operatorname{ord}(g) < \infty$ ist $\{g, g^2, \dots, g^{\operatorname{ord}(g)-1} = g^{-1}, g^{\operatorname{ord}(g)} = e\} = \{g^n \mid n \in \mathbb{Z}\}$, also eine Untergruppe.

Was sind Linksnebenklassen? RNK von U in G haben die Form $Ug, g \in G$. Linksnebenklassen haben die Form $aU, a \in G$, wobei $aU \coloneqq \{au \mid u \in U\}$. Ist $aU = bU$, dann folgt $b^{-1}aU = b^{-1}bU = eU = U \Rightarrow b^{-1}ae \in U \Rightarrow b^{-1}a \in U$ (vgl. für RNK: $ab^{-1} \in U$). Durch $a \,{}_U\!\sim b = b^{-1}a \in U$ können wir eine Äquivalenzrelation definieren, deren Äquivalenzklassen die Linksnebenklassen sind.

Beispiel 1.4. Sei $G \coloneqq \mathbb{Z},\ U \coloneqq n\mathbb{Z}$. Die Rechtsnebenklassen sind:

$$U + a = n\mathbb{Z} + a = \{\text{ganze Zahlen kongruent zu a modulo n}\}.$$

Die Linksnebenklasse $a + U$ ist das selbe wie $U + a$, d. h. die LNK von a ist gleich der RNK von a. Der Grund dafür ist, dass die Verknüpfung in $\mathbb{Z}$ kommutativ ist. Allgemein: G abelsch (kommutativ), $U < G,\ a \in G$, dann LNK von a = RNK von a.

Warnung: Für eine beliebige Gruppe G (nicht kommutativ) gilt im Allgemeinen: LNK von $a \neq$ RNK von a. Ein Beispiel dafür ist die symmetrische Gruppe $\Sigma_n = S_n = \{$ Permutationen von n Elementen $\}$, Verknüpfung = Komposition (Hintereinanderausführung von Abbildungen).

$$S_2 \ni \begin{matrix} 1 & 2 \\ \downarrow & \downarrow \\ 1 & 2 \end{matrix} = \mathrm{id} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$$

$$S_2 \ni \begin{matrix} 1 & 2 \\ \downarrow & \downarrow \\ 2 & 1 \end{matrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = (12) \qquad \text{(Zyklenschreibweise)}\ .$$

Es gilt $|S_2| = 2!$. Außerdem gilt $|S_3| = 3 \cdot 2 \cdot 1 = 3!$:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = \mathrm{id}, \mathrm{ord}(\mathrm{id}) = 1$$
$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = (12), \mathrm{ord}((12)) = 2$$
$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = (23)$$
$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = (13)$$
$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = (123), \mathrm{ord}((123)) = 3$$
$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = (132), \mathrm{ord}((132)) = 3\ .$$

Wähle die Untergruppe $U = \{\mathrm{id}, (12)\}$. Wegen $(12)^2 = \mathrm{id}$ ist $(12)^{-1} = (12)$ und somit ist U eine Untergruppe von S_3. Berechne die Rechtsnebenklassen von U in G:

- $U \cdot \mathrm{id} = U = \{\mathrm{id}, (12)\}$.
- $U \cdot (13) = \{\mathrm{id} \cdot (13), (12) \cdot (13)\} = \{(13), (132)\}$.
- $U \cdot (23) = \{(23), (123)\}$.

Linksnebenklassen von U in G:

- $\mathrm{id} \cdot U = U = \{\mathrm{id}, (12)\}$.
- $(13) \cdot U = \{(13), (123)\} \neq U \cdot (13)$.
- $(23) \cdot U = \{(23), (132)\} \neq U \cdot (23)$.

Daran sieht man, dass im Allgemeinen LNK $\neq$ RNK gilt. Es gilt aber immer LNK $\cap$ RNK $\neq \emptyset$, denn $eU = Ue$.

Versuch, RNK zu multiplizieren: $(U \cdot (13)) \cdot (U \cdot (23)) \stackrel{?}{=}$ genau eine RNK?

$$\begin{aligned}(13)\cdot(23) &= (132) \in U\cdot(13)\\ (13)\cdot(123) &= (12) \in U\cdot \mathrm{id}\\ (132)\cdot(23) &= (13) \in U\cdot(13)\\ (132)\cdot(123) &= \mathrm{id} \in U\cdot \mathrm{id}\ \ .\end{aligned}$$

Somit ist die Multiplikation von Rechtsnebenklassen durch die Multiplikation von Repräsentanten nicht wohldefiniert.

Frage: Wann funktioniert die Multiplikation von Rechtsnebenklassen? Die Antwort ist abhängig von U und G.

Proposition 1.5. *Sei G eine Gruppe und N eine Untergruppe. Dann sind die folgenden Bedingungen äquivalent:*

(a) $\forall a,b \in G : a \sim_N b \Leftrightarrow a\ {}_N{\sim}\ b$.

(b) $\forall a \in G : Na = aN$, *d. h. RNK = LNK.*

(c) $\forall a \in G : a^{-1}Na = N$.

(d) $\forall a \in G : a^{-1}Na \subseteq N$.

(e) $\forall a,b \in G : (aN)(bN) = abN$.

(f) $\forall a,b \in G : (Na)(Nb) = Nab$.

Wenn N diese Bedingungen erfüllt, heißt N Normalteiler *oder* normale Untergruppe *von G. Man schreibt $N \lhd G$. In einer abelschen Gruppe ist jede Untergruppe normal. Die Menge der Links- oder Rechtsnebenklassen wird mit G/N bezeichnet. G/N heißt* Faktorgruppe *oder* Quotientengruppe. *G/N ist eine Gruppe mit der Verknüpfung $(aN) * (bN) \coloneqq abN$. Für endliche Gruppen G gilt*

$$\begin{aligned}|G/N| &= \frac{|G|}{|N|} = \#\textit{Nebenklassen in } G\\ &=: [G:N]\ (\textit{Index von } N \textit{ in } G).\end{aligned}$$

Beweis.

(a) $\Leftrightarrow$ (b): $x \in aN \Leftrightarrow \exists n \in N : x = an \Leftrightarrow \exists n \in N : a^{-1}x = n \Leftrightarrow x \,{}_N\!\sim a$ und $x \in Na \Leftrightarrow xa^{-1} \in N \Leftrightarrow x \sim_N a$.

(b) $\Leftrightarrow$ (c): Multiplikation mit $a \in G$ liefert eine bijektive Abbildung (invers zu Multiplikation mit a^{-1}). Sei $aN = Na \Rightarrow N = eN = a^{-1}aN = a^{-1}Na$. Ist umgekehrt $a^{-1}Na = N$, dann folgt $Na = eNa = aa^{-1}Na = a(a^{-1}Na) = aN$.

(c) $\Rightarrow$ (d): ist klar.

(c) $\Rightarrow$ (e): $aNbN = abb^{-1}NbN = abNN = abN$, denn nach (c) gilt $b^{-1}Nb = N$. Außerdem gilt $NN \subseteq N$, da N eine Untergruppe ist und $N = Ne \subseteq NN$, also $N = NN$.

(d) $\Rightarrow$ (f): analog.

(e) $\Rightarrow$ (b): Mit $a = e$ folgt $NbN = eNbN = ebN = bN$. Also folgt $Nb \subseteq NbN = bN$. Ebenso folgt $bN \subseteq Nb$, also $Nb = bN$ für alle $b \in G$.

(f) $\Rightarrow$ (b): analog.

(d) $\Rightarrow$ (c): Nach (d) gilt $aNa^{-1} = (a^{-1})^{-1}Na^{-1} \subseteq N$. Also gilt auch

$$N = a^{-1}aNa^{-1}a \subseteq a^{-1}Na \ . \qquad \square$$

Ist $N \lhd G$, dann ist G/N, also die Menge der Nebenklassen, eine Gruppe. Die Multiplikation ist definiert durch $(aN)(bN) := abN$. Das neutrale Element ist $eN = N$, denn $eN \cdot aN = eaN = aN$. Das Inverse zu aN ist $a^{-1}N$, denn $a^{-1}N \cdot aN = a^{-1}aN = eN$. Durch $\pi : G \to G/N, g \mapsto gN$ ist ein Gruppenhomomorphismus definiert, denn

$$\begin{aligned}
\pi(ab) &= abN = (aN)(bN) = \pi(a)\pi(b)\\
\pi(e) &= eN\\
\pi(a^{-1}) &= a^{-1}N = \pi(a)^{-1} \ .
\end{aligned}$$

π ist surjektiv, denn $gN = \pi(g)$ für alle $gN \in G/N$. Was ist $\mathrm{Kern}(\pi) = \{g \in G \mid \pi(g) = eN\}$? Für $g \in G$ gilt $eN = \pi(g) = gN \Leftrightarrow ge^{-1} \in N \Leftrightarrow g \in N$.

Also ist $\text{Kern}(\pi) = N$ und $\text{Bild}(\pi) = G/N$. Ist G endlich, dann gilt $|G| = [G : N] \cdot |N|$.

Ist V ein Vektorraum über K, dann ist V eine abelsche Gruppe. Ist also U ein Unterraum von V, dann ist U eine abelsche Untergruppe von V. Somit ist U ein Normalteiler von V und damit ist die Gruppe V/U definiert (als abelsche Gruppe). Es fehlt noch die Multiplikation mit Skalaren. Sei $v_1 + U \in V/U$ und $\lambda \in K$. Dann wollen wir $\lambda \cdot (v_1 + U)$ so definieren, dass $\pi : V \to V/U$ eine lineare Abbildung, also ein Vektorraumhomomorphismus ist. π soll also additiv sein:

$$\pi(v_1) + \pi(v_2) = (v_1 + U) + (v_2 + U) = (v_1 + v_2 + U) = \pi(v_1 + v_2) \ .$$

Damit π linear ist, muss gelten:

$$\lambda(v_1 + U) = \lambda\pi(v_1) \overset{!}{=} \pi(\lambda v_1) = \lambda v_1 + U \ .$$

Also müssen wir definieren: $\lambda(v_1 + U) = \lambda v_1 + U$. Wir müssen allerdings noch die Wohldefiniertheit zeigen: Sei $v_1 + U = v_2 + U$, d. h. $v_1 - v_2 \in U$. Daraus folgt jedoch auch $\lambda v_1 - \lambda v_2 = \lambda(v_1 - v_2) \in U$ und somit

$$\lambda(v_1 + U) = \lambda v_1 + U = \lambda v_2 + U = \lambda(v_2 + U) \ .$$

Proposition 1.6. *Sei V ein K-VR, U ein Unterraum. Dann ist die Restklassenmenge V/U ein K-VR durch $(v_1 + U) + (v_2 + U) := (v_1 + v_2) + U$ und $\lambda(v_1 + U) := \lambda v_1 + U$. Die Abbildung $\pi : V \to V/U, \pi(v_1) := v_1 + U$ ist K-linear. Es gibt nur eine K-VR-Struktur auf V/U, sodass π linear wird.*

Weiter gilt: $\text{Kern}(\pi) = U, \text{Bild}(\pi) = V/U$, *d. h. π ist surjektiv. V/U heißt* Quotientenraum.

Falls $\dim V < \infty$: $\dim V = \dim(\text{Kern}(\pi)) + \dim(\text{Bild}(\pi))$, d. h. $\dim V = \dim U + \dim V/U$. Falls K endlich ist, gilt $|V| = |K|^{\dim V}, |U| = |K|^{\dim U}$ und

$$\begin{aligned} |V/U| = |K|^{\dim V/U} &= |K|^{\dim V - \dim U} \\ &= \frac{|K|^{\dim V}}{|K|^{\dim U}} \ . \end{aligned}$$

(Vergleiche dazu $|G/N| = |G|/|N|$.)

Insgesamt haben wir das erste Problem gelöst: Gegeben einen K-VR V und einen Unterraum U. Gesucht war eine lineare Abbildung $V \xrightarrow{\pi} W$ (K-VR) mit $\operatorname{Kern}(\pi) = U$. Lösung: $V \xrightarrow{\pi} W := V/U$ (Quotienten-Vektorraum) mit $\operatorname{Kern}(\pi) = U$.

Zweites Problem: Wir wollen einen Unterraum von $V = K^n$ (oder V endlich-dimensional) als Lösungsmenge eines homogenen LGS darstellen:

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1q}x_q &= 0 \\ &\vdots \\ a_{p1}x_1 + \ldots + a_{pq}x_q &= 0 \end{aligned}$$

Gesucht ist eine Matrix (a_{ij}), sodass $U < V$ die Lösungsmenge des LGS ist. Allgemeine Konstruktion: Seien V, W K-Vektorräume und $\varphi : V \to W$ eine lineare Abbildung. Definiere $\operatorname{Hom}_K(V, W) := \{$ lineare Abb. $\varphi : V \to W\}$, die Menge aller Vektorraum-Homomorphismen. $\operatorname{Hom}_K(V, W)$ ist selbst ein K-Vektorraum: Sind $\varphi, \psi \in \operatorname{Hom}_K(V, W)$, dann können wir für alle $v \in V, \lambda \in K$ definieren:

$$\begin{aligned} (\varphi + \psi)(v) &:= \varphi(v) + \psi(v) \\ (\lambda \cdot \varphi)(v) &:= \lambda \cdot \varphi(v) \ . \end{aligned}$$

Das neutrale Element von $\operatorname{Hom}_K(V, W)$ ist die Nullabbidung $0 : V \to W, v \mapsto 0_W$. Interpretiert man φ, ψ als Matrizen, dann ergibt sich die übliche Addition von Matrizen und die eintragsweise Multiplikation mit λ. Ein Spezialfall ist $W = K$.

Definition 1.7. Der K-Vektorraum $V^* := \operatorname{Hom}_K(V, K)$ heißt der *zu V duale Vektorraum* (oder der *Dualraum* von V). Die Elemente $\varphi \in V^*$, d. h. $\varphi : V \to K$, heißen *Linearformen* oder Funktionale.

Beispiel 1.5. Sei V der $\mathbb{R}$-Vektorraum der auf $[a, b]$ (stetigen oder) Riemann-integrierbaren Funktionen $f : [a, b] \to \mathbb{R}$. Definiere $\varphi : V \to K = \mathbb{R}, f \mapsto \int_a^b f(x)\,\mathrm{d}x$, dann ist φ linear: $f + g \mapsto \int_a^b (f+g)(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x$ und $\int_a^b (\lambda f)(x)\,\mathrm{d}x = \lambda \int_a^b f(x)\,\mathrm{d}x$. Sei außerdem $\psi : V \to \mathbb{R}, \psi(f) := f(c)$ für ein festes $c \in [a,b]$, dann ist ψ ebenfalls linear wegen $\psi(f+g) = (f+g)(c) = f(c) + g(c) = \psi(f) + \psi(g)$ und $\psi(\lambda f) = (\lambda f)(c) = \lambda f(c) = \lambda \psi(f)$.

Sei $V = K^n$ für ein $n \in \mathbb{N}$, $a_1, a_2 \cdots, a_n \in K$ und $\varphi : (x_1, \ldots, x_n)^T \mapsto a_1x_1 + a_2x_2 + \ldots + a_nx_n \in K$. Ist $x = (x_1, \ldots, x_n)^T$ und $y = (y_1, \ldots, y_n)^T$, dann folgt

$$\begin{aligned}\varphi(x+y) &= a_1(x_1+y_1) + \ldots + a_n(x_n+y_n) \\ &= a_1x_1 + \ldots + a_nx_n + a_1y_1 + \ldots + a_ny_n \\ &= \varphi(x) + \varphi(y)\ .\end{aligned}$$

Also ist φ additiv. Analog zeigt man $\lambda\varphi(x) = \varphi(\lambda x)$ für alle $\lambda \in K$. Somit ist φ eine Linearform $V = K^n \to K$. Die Standardbasis ist $e_1 = (1,0,\ldots,0)^T, \ldots, e_n = (0,\ldots,0,1)^T$. $\{1\}$ ist eine Basis von K. Da $\varphi(e_1) = a_1, \ldots, \varphi(e_n) = a_n$, ist die darstellende Matrix von φ bezüglich dieser Basen $A \coloneqq \begin{pmatrix} a_1 & a_2 & \ldots & a_n \end{pmatrix}$ und es gilt $\varphi(x) = Ax$.

Ist V ein K-Vektorraum und $\dim V = n$, dann existiert eine Basis $v_1, \ldots, v_n$ von V. Definiere dann die Abbildungen $\varphi_i \in V^*, i \in \{1, \ldots, n\}$ durch

$$\varphi_i(v_j) = \delta_{ij} = \begin{cases} 1 & , j = i \\ 0 & , j \neq i\ . \end{cases}$$

Ist $v \in V$ mit $v = \lambda_1v_1 + \ldots + \lambda_nv_n$, dann gilt

$$\begin{aligned}\varphi_j(v) &= \varphi_j(\lambda_1v_1 + \ldots + \lambda_nv_n) \\ &= \varphi_j(\lambda_1v_1) + \ldots + \varphi_j(\lambda_nv_n) \\ &= \lambda_1 \underbrace{\varphi(v_1)}_{0} + \ldots + \lambda_j \underbrace{\varphi_j(v_j)}_{1} + \ldots + \lambda_n \underbrace{\varphi_j(v_n)}_{0} \\ &= \lambda_j\ .\end{aligned}$$

Die darstellende Matrix von φ_j ist also $\begin{pmatrix} 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \end{pmatrix}$, wobei die 1 im j-ten Eintrag steht.

Behauptung: $\varphi_1, \ldots, \varphi_n \in V^*$ sind linear unabhängig: Die Null in V^* ist die Abbildung $0_{V^*} : V \to K, v \mapsto 0_K$. Sei jetzt $0 = \lambda_1\varphi_1 + \ldots + \lambda_n\varphi_n$ für $\lambda_1, \ldots, \lambda_n \in K$. Dann gilt für alle $v \in V$:

$$\begin{aligned}0 = 0(v) &= (\lambda_1\varphi_1 + \ldots + \lambda_n\varphi_n)(v) \\ &= \lambda_1\varphi_1(v) + \ldots + \lambda_n\varphi_n(v)\ .\end{aligned}$$

Speziell ist also für $v = v_j$ auch

$$\begin{aligned} 0 &= \lambda_1 \underbrace{\varphi_1(v_j)}_{0} + \ldots + \lambda_j \underbrace{\varphi_j(v_j)}_{1} + \ldots + \lambda_n \underbrace{\varphi_n(v_j)}_{0} \\ &= \lambda_j \ . \end{aligned}$$

Also sind $\varphi_1, \ldots, \varphi_n$ linear unabhängig in V^*. Bilden $\varphi_1, \ldots, \varphi_n$ auch ein Erzeugendensystem von V^*, insgesamt also eine Basis? Sei $\varphi \in V^*$ und $\varphi(v_1) = a_1 \in K, \varphi(v_2) = a_2 \in K, \ldots, \varphi(v_n) = a_n \in K$. Die darstellende Matrix von φ ist dann $\begin{pmatrix} a_1 & \ldots & a_n \end{pmatrix}$. Folglich ist $\varphi = a_1\varphi_1 + \ldots + a_n\varphi_n$ (denn: $v = \lambda_1 v_1 + \ldots + \lambda_n v_n \Rightarrow \varphi(v) = \lambda_1\varphi(v_1) + \ldots + \lambda_n\varphi(v_n) = \lambda_1 a_1 + \ldots + \lambda_n a_n = a_1\varphi_1(v) + \ldots + a_n\varphi_n(v)$).

Koordinatenvektor von v ist $(\lambda_1, \ldots, \lambda_n)^T$. Es gilt demnach

$$\varphi(v) = \begin{pmatrix} a_1 & \ldots & a_n \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = a_1\lambda_1 + \ldots + a_n\lambda_n \ .$$

Proposition 1.8. *Sei V ein n-dimensionaler K-Vektorraum mit der Basis $v_1, \ldots, v_n$. Sei $\varphi_j : V \to K$ definiert durch*

$$\varphi_j(v_l) = \begin{cases} 1 & , j = l \\ 0 & , j \neq l \ . \end{cases}$$

Dann bilden $\varphi_1, \ldots, \varphi_n$ eine Basis von V^*. Also ist $\dim V^* = n = \dim V$. Bezeichnung: $v_j^* \coloneqq \varphi_j$, die zu $v_1, \ldots, v_n$ duale Basis.

Achtung: $v_1^*, \ldots, v_n^*$ hängen von der Wahl der gesamten Basis $v_1, \ldots, v_n$ ab! v_j^* hängt nicht nur von v_j ab, sondern von $v_1, \ldots, v_n$! Wegen $\dim V = \dim V^*$ gilt $V \simeq V^*$ (V ist isomorph zu V^*). Einen Isomorphismus $\alpha : V \to V^*$ bekommt man durch die Festlegung $\alpha(v_j) = v_j^*$. Dieser ist surjektiv, also wegen $\dim V = \dim V^* < \infty$ auch bijektiv.

Es ist $V \simeq K^n$. Bezüglich der Basis $v_1, \ldots, v_n$ können wir die Koordinatenvektoren bilden:

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, v_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \ .$$

Wir können K^n mit dem Raum $\mathrm{Mat}(n \times 1, K)$ der $n \times 1$-Matrizen über K identifizieren. In V^* sind die Elemente gegeben durch die darstellenden Matrizen:

$$v_1^* = \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix}, \dots, v_n^* = \begin{pmatrix} 0 & \dots & 0 & 1 \end{pmatrix} .$$

Somit können wir V^* mit dem Raum $\mathrm{Mat}(1 \times n, K)$ identifizieren. Somit können wir unseren Isomorphismus α über den Matrizenräumen schreiben als $\alpha(v) = v^T$.

Nächster Schritt: Wir gehen von $V^* = \mathrm{Hom}_K(V, K)$ zu $V^{**} = (V^*)^* = \mathrm{Hom}_K(V^*, K)$. $x \in V^{**}$ ist eine Abbildung x, die eine Abbildung $\varphi : V \to K$ auf ein Element $x(\varphi) \in K$ abbildet. Was für Beispiele für Elemente $x \in V^{**}$ gibt es? Ein Beispiel ist für ein beliebiges $w \in V$ die Abbildung $x(w)$ mit $x(w)(\varphi) = \varphi(w) \in K$, die φ in w auswertet. Ist $x(w) \in V^{**}$, d. h. ist $x(w)$ eine lineare Abbildung $V^* \to K$? VR-Struktur auf $V^* : \varphi, \psi \in V^*, \lambda \in K :$ $\varphi + \psi : v \mapsto \varphi(w) + \psi(w), (\lambda\varphi)(w) = \lambda \cdot \varphi(w)$. Ist $x(w)$ linear?

$$\begin{aligned} x(w)(\varphi + \psi) &= (\varphi + \psi)(w) = \varphi(w) + \psi(w) = x(w)(\varphi) + x(w)(\psi) \\ x(w)(\lambda\varphi) &= (\lambda\varphi)(w) = \lambda\varphi(w) = \lambda x(w)(\varphi) . \end{aligned}$$

Also ist $x(w)$ eine lineare Abbildung $V^* \to K$ und somit $x(w) \in V^{**}$. V^{**} heißt der *Bidualraum* von V. $w \mapsto x(w)$ definiert eine Abbildung $V \to V^{**}$.

Theorem 1.9. *Sei V ein K-VR und V^{**} der Bidualraum von V. Die Abbildung $\iota : V \to V^{**}, w \mapsto \iota(w)$ mit $\iota(w)(\varphi) = \varphi(w)$ ist K-linear und injektiv. ι hängt nicht von der Wahl einer Basis ab. Falls $\dim V < \infty$ ist ι ein Isomorphismus. Achtung: Falls V ∞-dimensional ist, ist ι kein Isomorphismus (und es gibt keinen).*

Beweis. Gezeigt: Für $w \in V$ ist $x(w) = \iota(w)$ eine lineare Abbildung $V^* \to K$, d. h. $\iota(w) \in V^{**}$. Es bleibt noch zu zeigen, dass ι linear und injektiv ist. Seien $w_1, w_2 \in V$ und $\lambda \in V$. Zu zeigen: $\iota(w_1 + w_2) = \iota(w_1) + \iota(w_2), \iota(\lambda w_1) = \lambda\iota(w_1)$.

Es gilt für alle $w_1, w_2 \in V, \varphi \in V^*, \lambda \in K$:

$$\begin{aligned} \iota(w_1 + w_2)(\varphi) &= \varphi(w_1 + w_2) = \varphi(w_1) + \varphi(w_2) = \iota(w_1)(\varphi) + \iota(w_2)(\varphi) . \\ \iota(\lambda w_1)(\varphi) &= \varphi(\lambda w_1) = \lambda\varphi(w_1) = \lambda\iota(w_1)(\varphi) . \end{aligned}$$

Damit ist ι linear.

Es bleibt zu zeigen, dass ι injektiv ist, also $\mathrm{Kern}(\iota) = \{0_V\}$. Sei $w \in \mathrm{Kern}(\iota)$, d. h. $\iota(w) = 0$ (in V^{**}), d. h. $\forall \varphi \in V^*$: $\iota(w)(\varphi) = 0$. Falls $w \neq 0$, ergänze w zu einer Basis $B \coloneqq \{b_1 = w, b_2, \ldots\}$ von V. Definiere $\varphi \in V^*$ durch $\varphi(w) = 1$ und $\varphi(b) = 0$ für alle $b \in B \setminus \{w\}$. Somit ist aber $\iota(w)(\varphi) = 1$. Also muss $w = 0$ gewesen sein. Falls $\dim V = n < \infty$, gilt $\dim V^{**} = \dim V^* = \dim V = n$, also ist ι auch bijektiv als injektive lineare Abbildung zwischen Vektorräumen gleicher Dimension. □

ι wird auch als Auswertungsabbildung $\iota = \mathrm{ev} : V \to V^{**}$ bezeichnet. Manchmal wird ι weggelassen: Dann ist die Formel: $w(\varphi) = \varphi(w)$.

Zurück zu Unterräumen: Sei $\{0\} \neq U < V$, $U \neq V$ und $\dim V < \infty$. Sei $u_1, \ldots, u_r$ Basis von U, ergänzt zu einer Basis $u_1, \ldots, u_r, v_1, \ldots, v_s$ von V. Dann ist $\dim V = r + s$ und V^* hat eine duale Basis $u_1^*, \ldots, u_r^*, v_1^*, \ldots, v_s^*$. Frage: Was ist $\langle v_1^*, \ldots, v_s^* \rangle = \mathrm{Span}\{v_1^*, \ldots, v_s^*\} < V^*$?

Definition 1.10. Sei V ein K-Vektorraum und $U < V$ ein Unterraum. Die Menge $U^0 \coloneqq \{\varphi \in V^* \mid \forall u \in U : \varphi(u) = 0\}$ heißt der zu U *orthogonale Raum.*

U^0 ist ein Unterraum von V^*: Für alle $u \in U$ gilt $0_{V^*}(u) = 0_K$. Für alle $\varphi, \psi \in U^0$ und alle $\lambda \in K$ gilt $(\varphi + \psi)(u) = \varphi(u) + \psi(u) = 0 + 0 = 0 \Rightarrow \varphi + \psi \in U^0$ und $(\lambda\varphi)(u) = \lambda\varphi(u) = \lambda \cdot 0 = 0 \Rightarrow \lambda\varphi \in U$.

Beispiel 1.6. Sei $V \coloneqq \mathbb{R}^2$ und $(a, b)^T \in \mathbb{R}^2$ mit $(a, b)^T \neq (0, 0)$. Definiere $U \coloneqq \{\lambda(a, b)^T \mid \lambda \in \mathbb{R}\}, (a, b)^T \neq (0, 0)^T$. U ist also eine Gerade in $\mathbb{R}^2$. Dann ist $U^0 = \{\varphi \in V^* \mid \forall \lambda \in \mathbb{R} : \varphi(\lambda(a, b)^T) = 0\}$. Sei $\begin{pmatrix} c & d \end{pmatrix}$ die Darstellung von $\varphi : \mathbb{R}^2 \to \mathbb{R}$ als Zeilenvektor. Dann muss dieser die Bedingung

$$\varphi\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) = \begin{pmatrix} c & d \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = ca + db \stackrel{!}{=} 0$$

erfüllen. U^0 ist also eine Gerade im $(\mathbb{R}^2)^*$, die orthogonal zu U ist.

Beispiel 1.7. Wählen wir $U = V$, dann ist $U^0 = \{\varphi : V \to K \mid \forall v \in V : \varphi(v) = 0\} = \{0_{V^*}\}$. Für $U = \{0_V\}$ ist hingegen $U^0 = \{\varphi : V \to K \mid \varphi(0) = 0\} = V^*$.

Theorem 1.11. *Sei $U < V$, $\dim V < \infty$, $(u_1, \dots, u_r)$ eine Basis von U und $(u_1, \dots, u_r, v_1, \dots, v_s)$ eine Basis von V. Dann gilt:*

(a) $v_1^*, \dots, v_s^*$ *ist eine Basis von* U^0.

(b) $\dim U + \dim U^0 = \dim V$.

(c) $\left(U^0\right)^0 = \iota(U)$ *mit der Auswertungsabbildung* $\iota : V \to V^{**}, \iota(v)(\varphi) = \varphi(v)$.

Beweis. (a) Da $(u_1^*, \dots, u_r^*, v_1^*, \dots, v_s^*)$ eine duale Basis von V^* ist, sind $v_1^*, \dots, v_s^*$ linear unabhängig. Nach Definition gilt $v_i^*(v_i) = 1$, $v_i^*(v_j) = 0$ für $i \neq j$ und $v_i^*(u_l) = 0$ für alle l. Da also für alle $i \in \{1, \dots, s\}$ und alle $u \in U$ gilt, dass $v_i^*(u) = 0$, folgt $v_1^*, \dots, v_s^* \in U^0$. Es bleibt zu zeigen, dass diese Vektoren auch U^0 erzeugen. Sei $\varphi \in U^0 < V^*$, also auch $\varphi \in V^*$. φ kann also geschrieben werden als $\varphi = \lambda_1 u_1^* + \ldots + \lambda_r u_r^* + \mu_1 v_1^* + \ldots + \mu_s v_s^*$. Aus $\varphi \in U^0$ folgt $\varphi(u_1) = \ldots = \varphi(u_r) = 0$ und somit

$$\begin{aligned} 0 = \varphi(u_1) &= \lambda_1 \underbrace{u_1^*(u_1)}_{1} + \lambda_2 \underbrace{u_2^*(u_1)}_{0} + \ldots + \lambda_r \underbrace{u_r^*(u_1)}_{0} \\ &\quad + \mu_1 \underbrace{v_1^*(u_1)}_{0} + \ldots + \mu_s \underbrace{v_s^*(u_1)}_{0} = \lambda_1 \\ 0 = \varphi(u_2) &= \lambda_2 \\ &\vdots \\ 0 = \varphi(u_r) &= \lambda_r \ , \end{aligned}$$

also $\varphi = \mu_1 v_1^* + \ldots + \mu_s v_s^*$. Da φ eine Linearkombination der $v_1^*, \dots, v_s^*$ ist, erzeugen die $v_1^*, \dots, v_s^*$ ganz U^0 und bilden folglich eine Basis von U^0.

(b) Die Behauptung folgt unmittelbar aus $\dim U^0 = s, \dim U = r, \dim V = r + s$.

(c) Sei $\varphi \in U^0, x \in U \Rightarrow \varphi(x) = 0$. Es ist aber $\varphi(x) = \iota(x)(\varphi) \Rightarrow \iota(x) \in \left(U^0\right)^0 \Rightarrow \iota(U) \subseteq \left(U^0\right)^0$. Für Gleichheit prüfe Gleichheit der Dimensionen: Es gilt $\dim U = r$. Da ι ein Isomorphismus ist, gilt $\dim \iota(U) = r$. Außerdem gilt $\dim V^{**} = \dim V^* = \dim V = r + s$ und $\dim U^0 = s$. Aus (b) wissen wir, dass $s + \dim \left(U^0\right)^0 = \dim U^0 + \dim \left(U^0\right)^0 = \dim V^* = r + s$. Folglich gilt $\dim \left(U^0\right)^0 = (r + s) - s = r = \dim U$. □

Gegeben endlich-dimensionale Vektorräume V, W, $\alpha : V \to W$ linear mit Matrix A. Wie finden wir eine lineare Abbildung zwischen W^* und V^*?

$$\begin{tikzcd} V \arrow[r, "\alpha"] \arrow[dr, "\varphi \circ \alpha"'] & W \arrow[d, "\varphi \in W^*"] \\ & K \end{tikzcd}$$

Die Komposition mit α definiert eine Abbildung $\alpha^* : W^* \to V^*, \varphi \mapsto \varphi \circ \alpha$. Sind α und φ linear, dann ist $\varphi \circ \alpha$ linear, somit ist $\alpha^* : W^* \to V^*$ definiert. Es bleibt zu zeigen, dass α^* linear ist. Seien dazu $\varphi_1, \varphi_2 \in W^*$, dann ist

$$\alpha^*(\varphi_1 + \varphi_2) = (\varphi_1 + \varphi_2) \circ \alpha = \varphi_1 \circ \alpha + \varphi_2 \circ \alpha = \alpha^*(\varphi_1) + \alpha^*(\varphi_2) \ .$$

Für beliebige $\varphi \in W^*, \lambda \in K$ ist außerdem

$$\alpha^*(\lambda\varphi) = (\lambda\varphi) \circ \alpha = \lambda(\varphi \circ \alpha) = \lambda\alpha^*(\varphi) \ ,$$

also ist α^* linear. Was ist die darstellende Matrix von α^*?

Sei $(v_1, \ldots, v_n)$ eine Basis von V und $(w_1, \ldots, w_l)$ eine Basis von W. Dann ist $(v_1^*, \ldots, v_n^*)$ die duale Basis von V^* und $(w_1^*, \ldots, w_l^*)$ die duale Basis von W^*. Für die darstellende Matrix A von α gilt $Av_j = j$-te Spalte von A, also

$$\alpha(v_j) = Av_j = a_{1j}w_1 + \ldots + a_{lj}w_l \ .$$

Sei B die darstellende Matrix von α^*. Dann gilt

$$\alpha^*(w_i^*) = Bw_i^* = b_{1i}v_1^* + \ldots + b_{ni}v_n^* \ . \tag{1.0.1}$$

Nach Definition der w_i^* gilt

$$w_i^*(w_k) = \begin{cases} 1 & , i = k \\ 0 & , k \neq l \ . \end{cases}$$

Es folgt

$$w_i^*(\alpha(v_j)) = a_{ij}w_i^*(w_i) = a_{ij} \ .$$

Nach Definition gilt $\alpha^*(w_i^*) = w_i^* \circ \alpha$ und somit

$$\begin{aligned} \alpha^*(w_i^*)(v_j) &\overset{1.0.1}{=} b_{ji}v_j^*(v_j) = b_{ji} \\ \alpha^*(w_i^*)(v_j) &= (w_i^* \circ \alpha)(v_j) = w_i^*(\alpha(v_j)) = a_{ij} \ , \end{aligned}$$

also $b_{ji} = a_{ij}$ für alle $i, j \Rightarrow B = A^T$.

Proposition 1.12. *Seien V und W endlich-dimensionale K-Vektorräume und die lineare Abbildung $\alpha : V \to W$ durch eine Matrix A gegeben. Dann ist die duale Abbildung $\alpha^* : W^* \to V^*$ bezüglich der dualen Basen durch die transponierte Matrix A^T gegeben.*

Die Abbildung $\alpha \mapsto \alpha^*$ entspricht also der Abbildung $A \mapsto A^T$ und es gelten die Linearitätseigenschaften $(A+B)^T = A^T + B^T$, $(\lambda A)^T = \lambda A^T$.

Korollar 1.13. *Die Abbildung $\alpha \mapsto \alpha^*$ ist ein Isomorphismus zwischen* $\mathrm{Hom}_K(V,W)$ *und* $\mathrm{Hom}_K(W^*,V^*)$.

Proposition 1.14. *Sei $\alpha : V \to W$ linear, V und W endlich-dimensionale K-Vektorräume. Dann ist* $\mathrm{Bild}(\alpha^*) = (\mathrm{Kern}\,\alpha)^0$.

Beweis.

- 1. Schritt: $\mathrm{Bild}(\alpha^*) \subseteq (\mathrm{Kern}\,\alpha)^0$.

  Sei $\varphi \in \mathrm{Bild}(\alpha^*)$. Dann gibt es ein $\psi \in W^*$ mit $\varphi = \alpha^*(\psi) = \psi \circ \alpha$. Wir wollen zeigen, dass $\varphi \in (\mathrm{Kern}\,\alpha)^0$. Sei dazu $x \in \mathrm{Kern}\,\alpha$, dann gilt $\varphi(x) = \psi(\alpha(x)) = \psi(0) = 0$ und folglich $\varphi \in (\mathrm{Kern}\,\alpha)^0$.

- 2. Schritt: $\mathrm{Bild}(\alpha^*) \supseteq (\mathrm{Kern}\,\alpha)^0$.

  Sei $\varphi \in (\mathrm{Kern}\,\alpha)^0$, d. h. $\varphi : V \to K$ mit $\varphi(x) = 0$ für alle $x \in \mathrm{Kern}(\alpha)$. Gesucht: $\psi \in W^*$ mit $\varphi = \alpha^*(\psi) = \psi \circ \alpha$. Für $x \in \mathrm{Kern}\,\alpha$ folgt $\varphi(x) = (\psi \circ \alpha)(x) = \psi(\alpha(x)) = \psi(0) = 0$, was keine Bedingung an ψ stellt. Wähle jetzt eine Basis $a_1, \dots, a_n$ von $\mathrm{Kern}(\alpha)$ und ergänze diese zu einer Basis $a_1, \dots, a_n, b_1, \dots, b_s$ von V. Wähle außerdem eine Basis $c_1, \dots, c_r$ von $\mathrm{Bild}(\alpha)$ und ergänze diese zu einer Basis $c_1, \dots, c_r, d_1, \dots, d_l$ von W. Nach der Dimensionsformel gilt

  $$n + s = \dim V = \dim \mathrm{Kern}(\alpha) + \dim \mathrm{Bild}(\alpha) = n + r \Rightarrow r = s \ .$$

  Wir können demnach $c_i = \alpha(b_i)$ wählen für $i \in \{1, \dots, s = r\}$ (die $\alpha(b_1), \dots, \alpha(b_s)$ sind also eine Basis von $\mathrm{Bild}(\alpha)$). Wir wollen $\varphi = \psi \circ \alpha$ auf den Basisvektoren:

  $$\begin{aligned} \varphi(a_i) &= (\psi \circ \alpha)(a_i) = 0 \\ \varphi(b_j) &= \psi(\alpha(b_j)) = \psi(c_j) \ , \end{aligned}$$

also definieren wir $\psi(c_j) := \varphi(b_j)$ und $\psi(d_i) = 0$. Dies definiert eindeutig ein $\psi \in W^*$ mit $\varphi = \psi \circ \alpha$, also ist $\varphi = \alpha^*(\psi) \in \operatorname{Bild}(\alpha^*)$. □

Anwendung von 1.14: α hat die darstellende Matrix A. Es gilt: Spaltenrang von A = Dimension des von den Spalten von A erzeugten Unterraums von W $= \dim(\operatorname{Bild}(\alpha)) = \dim V - \dim \operatorname{Kern}(\alpha) \overset{1.11}{=} \dim(\operatorname{Kern}(\alpha)^0) \overset{1.14}{=}$ Spaltenrang der darstellenden Matrix von α^* = Spaltenrang von A^T = Zeilenrang von A. Dies ist die Aussage des Rangsatzes 7.26 aus LAAG 1.

Korollar 1.15. *Der Spaltenrang einer Matrix ist gleich dem Zeilenrang.*

Um ein homogenes LGS $Ax = 0$ mit einer Matrix A zu lösen[1], haben wir in LAAG 1 den Gauß-Algorithmus kennen gelernt. Beim Gauß-Algorithmus wird die Matrix A in eine Matrix B in Zeilenstufenform umgeformt. Diese Umformungen erfolgen durch Linksmultiplikationen mit (invertierbaren) Elementarmatrizen. Daraus folgt, dass die Lösungsmenge von $Ax = 0$ gleich der Lösungsmenge von $Bx = 0$ ist. Hat B r Zeilen, die nicht Null sind, und n Spalten, dann gibt es n Variablen $x_1, \dots, x_n$, von denen r bestimmt und $n - r$ frei wählbar sind. Die Dimension des Lösungsraums ist also $n - r$. In der Zeilenstufenform B sind r Zeilen ungleich Null, diese sind wegen der Stufenform linear unabhängig $\Rightarrow$ Zeilenrang von B ist r. Ist A die darstellende Matrix einer linearen Abbildung α, dann gilt: Lösungsraum des LGS $Ax = 0$ ist $\operatorname{Kern}(\alpha)$ mit $\dim \operatorname{Kern}(\alpha) = \dim V - \dim \operatorname{Bild}(\alpha) = n-$ Spaltenrang von $A \overset{1.15}{=} n-$ Zeilenrang von $A = n-$ Zeilenrang von $B = n - r$.

Zurück zur Ausgangsfrage: Sei V endlich-dimensional und $U < V$ ein Unterraum. Können wir U als Lösungsmenge eines homogenen LGS schreiben? Wir schreiben ein solches LGS als

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n &= 0 \\ &\vdots \\ a_{l1}x_1 + \ldots + a_{ln}x_n &= 0\ . \end{aligned}$$

Dies ist ein LGS mit l Zeilen. Der i-ten Zeile können wir eine Linearform $\varphi_i : (x_1, \dots, x_n)^T \mapsto a_{i1}x_1 + \ldots + a_{in}x_n$ zuordnen. φ_i hat die darstellende

[1] A muss nicht quadratisch sein.

Matrix $\begin{pmatrix} a_{i1} & \dots & a_{in} \end{pmatrix}$. $(x_1, \dots, x_n)^T$ ist genau dann eine Lösung des LGS, wenn es in Kern(φ_i) für alle $i \in \{1, \dots, l\}$ liegt. Die Matrix $A = (a_{ij})$ definiert eine lineare Abbildung $\alpha : V \to W, x \mapsto Ax$. Kern$(\alpha)$ ist die Lösungsmenge des LGS. Wir definieren

$$\begin{aligned} U &\coloneqq \mathrm{Kern}(\alpha) = \{x \in V \mid \varphi_1(x) = \ldots = \varphi_l(x) = 0\} \\ X &\coloneqq \langle \varphi_1, \dots, \varphi_l \rangle \subseteq V^* \text{ (der von } \varphi_1, \dots, \varphi_l \text{ erzeugte Unterraum).} \end{aligned}$$

Die Ausgangsfrage bedeutet: Gegeben ist $U \subseteq V$, gesucht ist $X \subseteq V^*$. Ein Erzeugendensystem von X liefert Gleichungen, die als Zeilen des LGS mit U als Lösung verwendet werden können.

Was ist die Beziehung zwischen $U \subseteq V$ und $X \subseteq V^*$? Ist $X = U^0$? Unter dieser Annahme wäre $X^0 = (U^0)^0 = \iota(U)$, also auch $U = \iota^{-1}(X^0)$. Dann würden U und X einander gegenseitig bestimmen.

Ist $\{\varphi_1, \dots, \varphi_l\}$ ein Erzeugendensystem von X, dann gilt:

$$\begin{aligned} \{\iota(v) \mid v \in V\} = V^{**} \supseteq X^0 &= \{w \in V^{**} \mid \forall \varphi \in X : w(\varphi) = 0\} \\ &= \{w \in V^{**} \mid w(\varphi_1) = \ldots = w(\varphi_l) = 0\} \\ &= \{\iota(v) \mid v \in V, \underbrace{\iota(v)(\varphi_1)}_{=\varphi_1(v)} = \ldots = \underbrace{\iota(v)(\varphi_l)}_{=\varphi_l(v)} = 0\} \\ &= \{\iota(v) \mid v \in V, \varphi_1(v) = \ldots = \varphi_l(v) = 0\} \\ &= \iota(U) \ . \end{aligned}$$

$\varphi_1, \dots, \varphi_l$ bilden ein $x \in U$ auf 0 ab, also $X \subseteq U^0$. Wir wollen jetzt noch zeigen, dass $\dim X = \dim U^0$, dann folgt $X = U^0$. Es gilt $\dim V = \dim V^* = \dim V^{**} = n$ und $\dim X + \dim X^0 = n \Rightarrow \dim X = n - \dim X^0$. Da $X^0 = \iota(U)$ und ι ein Isomorphismus ist, gilt $\dim(X^0) = \dim U$. Mit $\dim U + \dim U^0 = n$ folgt $\dim X^0 = \dim U = n - \dim U^0$. Daraus folgt $\dim X = n - \dim X^0 = n - (n - \dim U^0) = \dim U^0$. Also gilt $X = U^0$.

Ergebnis: Jeder Unterraum $U < V$ ist Lösungsmenge eines homogenen LGS $Ax = 0$. Die Zeilen entsprechen Linearformen $\varphi_1, \dots, \varphi_l \in V^*$. Der Vektorraum $X = \langle \varphi_1, \dots, \varphi_l \rangle \subseteq V^*$ ist eindeutig bestimmt durch $X = U^0$.

Also existiert quasi eine „Bijektion“ zwischen homogenen LGS, die Unterräumen von V^* entsprechen und Lösungsmengen von homogenen LGS, die Unterräume von V sind ($U \mapsto U^0$). Außerdem: dim Lösungsraum + minimale

Zahl der benötigten Gleichungen $= \dim U + \dim U^0 = \dim V =$ Anzahl der Unbekannten.

Konkretes Problem: Gegeben U, wie berechnet man U^0? Sei $\dim V = n, \dim U = l \leq n$. Falls $l = 0$, dann ist $U^0 = V^*$. Falls $l = n$, dann ist $U^0 = \{0_{V^*}\}$. Sei also nun $0 < l < n$ und $u_1, \dots, u_l$ eine Basis von U. Es gilt $U^0 = \{\varphi \in V^* \mid \varphi(u_1) = \ldots = \varphi(u_l) = 0\}$. Schreibe die Koordinaten von $u_1, \dots, u_l$ als Spalten einer Matrix $B = (b_1, \dots, b_l)$ (bezüglich irgendeiner festen Basis von V). u_1 hat also den Koordinatenvektor $(b_{11}, \dots, b_{1n})^T$. $\varphi \in V^*$ hat die darstellende Matrix $\begin{pmatrix} a_1 & \dots & a_n \end{pmatrix}$. Wir wollen $\varphi(u_1) = \varphi(u_2) = \ldots = \varphi(u_l) = 0$. Das bedeutet $\varphi(u_j) = 0 \Leftrightarrow a_1 b_{j1} + \ldots + a_n b_{jn} = 0$. Also

$$\begin{pmatrix} a_1 & \dots & a_n \end{pmatrix} \cdot B = \begin{pmatrix} 0 & \dots & 0 \end{pmatrix} \Leftrightarrow B^T \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} .$$

Also erfüllt die darstellende Matrix von φ genau dann dieses Gleichungssystem, wenn $\varphi \in U^0$. Somit ist U^0 die Lösungsmenge eines homogenen LGS.

Die Lösungsmenge hat Dimension $n - l =: r$. Folglich hat U^0 eine Basis $\varphi_1, \dots, \varphi_r$. Ist $\begin{pmatrix} a_{j1} & \dots & a_{jn} \end{pmatrix}$ die darstellende Matrix von φ_j, dann können wir die Matrix $A := (a_{ij})$ bilden, deren Zeilen aus den darstellenden Matrizen der φ_j bestehen. Dann gilt $A \cdot B = 0$ und $\operatorname{Rang}(A) = r$. Die Zeilen entsprechen den Basisvektoren von U^0.

Beispiel 1.8. Geometrisches Beispiel: Wir betrachten den $\mathbb{R}$-Vektorraum $V := \mathbb{R}^3$. Gegeben seien zwei Ebenen (2-dimensionale Untervektorräume) $E_1 : ax + by + cz = 0, E_2 : dx + ey + fz = 0$ mit $(a, b, c) \neq (0, 0, 0) \neq (d, e, f)$. Was für ein geometrisches Objekt kann $E_1 \cap E_2$ sein?

E_1 ist gegeben durch die Linearform $\varphi = \begin{pmatrix} a & b & c \end{pmatrix}$, E_2 ist gegeben durch $\psi = \begin{pmatrix} d & e & f \end{pmatrix}$. $E_1 \cap E_2$ wird gegeben durch $\operatorname{Span}\{\varphi, \psi\} \subseteq V^*$. Es gilt $\dim E_1 \cap E_2 = \dim V - \dim \operatorname{Span}\{\varphi, \psi\}$. Aus $\varphi \neq 0 \neq \psi$ und weil zwei Vektoren gegeben sind, folgt $1 \leq \dim \operatorname{Span}\{\varphi, \psi\} \leq 2$. Somit ist $\dim(E_1 \cap E_2) \in \{3 - 1, 3 - 2\} = \{1, 2\}$.

Es gilt $\dim \operatorname{Span}\{\varphi, \psi\} = 1$ genau dann, wenn φ und ψ linear abhängig sind. Dies ist äquivalent dazu, dass $E_1 = E_2$ gilt; in diesem Fall ist $E_1 \cap E_2 = E_1 = E_2$ eine Ebene. Umgekehrt ist $\dim \operatorname{Span}\{\varphi, \psi\} = 2$ äquivalent dazu, dass φ und ψ linear unabhängig sind. Dann ist $\dim(E_1 \cap E_2) = 1$, also ist $E_1 \cap E_2$ eine Gerade.

Kapitel 2

Normalformen von Endomorphismen

Siehe dazu auch das Kapitel 10 in LAAG 1.

Problem 2.1. Seien V, W endlich-dimensionale K-Vektorräume, $\varphi : V \to W$ linear. Die Wahl von Basen von V und W führt auf eine darstellende Matrix A. Durch einen Basiswechsel soll A vereinfacht werden.

$$\begin{tikzcd} V \arrow[r, "A"] \arrow[d, "T"'] & W \arrow[d, "S"] \\ V \arrow[r] & W \end{tikzcd}$$

Die neue entstehende Matrix SAT^{-1} entsteht durch die Basiswechselmatrizen S und T. Die neue Matrix ist zur alten äquivalent: $A \sim SAT^{-1}$.

Mit dem Gauß-Algorithmus folgt

$$A \sim \begin{pmatrix} \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix}.$$

Die Anzahl der Einsen in dieser Normalform ist genau der Rang von A.

Problem 2.2. Sei φ ein Endomorphismus, d. h. $\varphi : V \to V$. Es darf eine Basis von V gewählt werden.

$$\begin{tikzcd} V \arrow[r, "A"] \arrow[d, "T"'] & V \arrow[d, "T"] \\ V \arrow[r, "TAT^{-1}"] & V \end{tikzcd}$$

Hier sind Zeilen- und Spaltenumformungen gekoppelt. Die durch $A \sim TAT^{-1}$ gegebene Äquivalenzrelation heißt Ähnlichkeit. Es gilt

$$A^n \sim (TAT^{-1}) \cdots (TAT^{-1}) = TA^nT^{-1} ,$$

d. h. diese Normalform ist kompatibel mit der Multiplikation.

Wir haben bereits definiert: $0 \neq v_1 \in V$ ist Eigenvektor zum Eigenwert $\lambda \Leftrightarrow \varphi(v) = \lambda v$. Falls eine Basis von V aus Eigenvektoren $v_1, \ldots, v_n$ zu Eigenwerten $\lambda_1, \ldots, \lambda_n$ existiert, dann hat φ bezüglich dieser Basis die darstellende Matrix

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

und φ ist diagonalisierbar.

Verfahren:

- Gegeben φ und eine darstellende Matrix A, bestimme das charakteristische Polynom $\chi_\varphi = \chi_A = \det(A - tE_n)$.
- Bestimme die Nullstellen $\lambda_1, \ldots, \lambda_l$ $(l \leq n)$ von χ_A. Diese sind die Eigenwerte von φ. (*hier kann man scheitern*)
- Bestimme die Eigenvektoren zu den Eigenwerten λ_j durch die Bestimmung des Eigenraums $\{x \in V \mid \varphi(x) = \lambda_j x\} = \mathrm{Kern}(A - \lambda_j E_n)$. (*Problem: vielleicht zu wenige EV*)

Aber: *Nicht alle φ sind diagonalisierbar!*

Beispiele für φ nicht diagonalisierbar:

- Sei $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ die Drehung um $90°$ gegen den Uhrzeigersinn. Dabei wird $(1,0)^T$ auf $(0,1)^T$ und $(0,1)^T$ auf $(-1,0)^T$ abgebildet. Die darstellende Matrix bezüglich der Einheitsbasis ist

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

also hat das charakteristische Polynom $\chi_A(t) = t^2 + 1$ keine Nullstellen in $\mathbb{R}$ und somit hat A keine Eigenwerte (geometrisch ist es ohnehin klar, dass es keine Eigenvektoren geben kann).

- Sei jetzt

$$A := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} .$$

$\chi_A = (t-1)^2$ hat $\lambda = 1$ als doppelte Nullstelle, also ist λ der einzige Eigenwert mit algebraischer Vielfachheit 2. Aber der Eigenraum $\mathrm{Eig}(A,1)$ hat Dimension 1. (Aus $\dim \mathrm{Eig}(A,1) = 2$ würde $\mathrm{Eig}(A,1) = \mathbb{R}^2$ folgen, aber $(0,1)^T$ ist kein Eigenvektor von A – Widerspruch. Alternativer Beweis: Wäre A ähnlich zu E_2, dann wäre $A = TE_2T^{-1} = E_2$.)

Hier ist die algebraische Vielfachheit von λ größer als die geometrische Vielfachheit.

Es gilt: A diagonalisierbar $\Leftrightarrow$ $\chi_A(t)$ ist Produkt von Linearfaktoren ($\chi_A(t) = \prod_{j=1}^{l}(t-\lambda_j)^{n_j}$) und für alle Eigenwerte λ_j gilt, dass die algebraische Vielfachheit n_j gleich der geometrischen Vielfachheit, $\dim \mathrm{Eig}(A, \lambda_j)$, ist. Dann gilt

$$A \sim \begin{pmatrix} \lambda_1 &&&&&&&& \\ & \ddots &&&&&&& \\ && \lambda_1 &&&&&& \\ &&& \lambda_2 &&&&& \\ &&&& \ddots &&&& \\ &&&&& \lambda_2 &&& \\ &&&&&& \ddots && \\ &&&&&&& \lambda_l & \\ &&&&&&&& \ddots & \\ &&&&&&&&& \lambda_l \end{pmatrix}$$

und $V = \mathrm{Eig}(A, \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(A, \lambda_l)$. $\mathrm{Eig}(A, \lambda_j)$ ist invariant: $\varphi(\mathrm{Eig}(A, \lambda_j)) \subseteq \mathrm{Eig}(A, \lambda_j)$.

Wir wollen eine allgemeine Normalform (unter Ähnlichkeit), auch für nicht-diagonalisierbare Matrizen (zumindest über $\mathbb{C}, \mathbb{R}$), finden. Hilfsmittel: Polynome und invariante Unterräume, d. h. Unterräume $U < V$ mit $\varphi(U) \subseteq U$. Zuerst beschäftigen wir uns mit der Theorie der Polynome. Für einen Körper K bezeichnen wir mit $K[t]$ die Menge der Polynome $f(t) = a_0 + a_1 t + \ldots + a_n t^n$ mit $a_0, \ldots a_n \in K$, bestimmt durch $a_0, \ldots, a_n$. (Ein Polynom wird nicht als Funktion betrachtet: Sei z. B. $K := \{0, 1\}$ und $f = t^2 + t$, dann ist $f(0) = 0, f(1) = 1 + 1 = 0$, aber $f \neq 0$).

Auf $K[t]$ gibt es eine Skalarmultiplikation: Für $\lambda \in K, f = (a_0 + \ldots + a_n t^n) \in K[t]$ definieren wir $\lambda f \in K[t]$ durch

$$\lambda(a_0 + \ldots + a_n t^n) = \lambda a_0 + \ldots + \lambda a_n t^n \ .$$

Für $f = a_0 + \ldots + a_n t^n, g = b_0 + \ldots + b_l t^l \in K[t]$ definieren wir $f + g \in K[t]$ durch

$$f + g = (a_0 + b_0) + (a_1 + b_1)t + \ldots \ .$$

Wir definieren außerdem $f \cdot g \in K[t]$ durch

$$f \cdot g = a_0 b_0 + (a_1 b_0 + a_0 b_1)t + (a_2 b_0 + a_1 b_1 + a_0 b_2)t^2 + \ldots \ .$$

$f = 0$ ist neutral bezüglich $+$ und $f = 1$ ist neutral bezüglich $\cdot$. $K[t]$ ist ein Ring[1] und ein K-Vektorraum. Außerdem ist die Multiplikation in $K[t]$ kommutativ.

Definition 2.1. Sei R ein Ring. Eine Teilmenge $I \subseteq R$ heißt *Ideal* in R, falls $I \neq \emptyset$, für alle $a, b \in I$ auch $a + b \in I$ und $\forall a \in I, x \in R : ax, xa \in I$.

Ist I ein Ideal und $a \in I$, dann ist $0 = 0a \in I$ und $(-a) = (-1)a \in I$. I ist bezüglich der Addition ein Normalteiler in R. Außerdem folgt $I = R \Leftrightarrow 1 \in I$, denn $1 \in I \Rightarrow \forall x \in R : x = x \cdot 1 \in I$. Für $I \neq R$ ist R/I selbst ein Ring mit

$$(a + I) + (b + I) := (a + b) + I$$
$$(a + I) \cdot (b + I) := ab + I \ .$$

[1] In Zukunft ist hiermit immer ein Ring mit Einselement gemeint.

Beispiel: Für $R = \mathbb{Z}$ ist $I = n\mathbb{Z} = \{na \mid a \in \mathbb{Z}\}$ ein Ideal.

Sei $I \subseteq \mathbb{Z}$ ein Ideal. Falls $1 \in I$, folgt $I = \mathbb{Z}$. Ein anderes Ideal ist $I = \{0\}$. Sei nun $I \neq \mathbb{Z}$ und $I \neq \{0\}$. Es existiert dann ein $a \in I \setminus \{0, 1\}$, also folgt auch $-a \in I$. Somit enthält I positive ganze Zahlen $\Rightarrow \exists n \in I$: n ist kleinstes positives Element in $I \Rightarrow n\mathbb{Z} \subseteq I$. Behauptung: $I = n\mathbb{Z}$. Beweis: Sei $0 < b \in I$, dann ist $b = nc + r$ (Division mit Rest), $0 \leq r < n$. Dann ist $r = b - nc \in I$. Da $r \in I$ und $r < n$ folgt $r = 0$ und somit $b = nc \in n\mathbb{Z}$. Insgesamt folgt also $I = n\mathbb{Z}$.

Proposition 2.2. *Die Ideale in $R = \mathbb{Z}$ sind genau die Teilmengen $n\mathbb{Z}, n \in \mathbb{N}_0$. Insbesondere ist jedes Ideal ein* Hauptideal*, d. h. von einem Element erzeugt ($I = Ra$ für ein $a \in R$).*

Proposition 2.3. *Alle Ideale im Polynomring $K[t]$ über einem Körper K sind Hauptideale.*

Beweis. Sei $I \subseteq K[t]$ ein Ideal. Gesucht ist ein Polynom $f \in K[t]$ mit $I = K[t] \cdot f = \{gf \mid g \in K[t]\}$. Im Fall $I = \{0\}$ ist $I = K[t] \cdot 0$. Sei nun $I \neq \{0\}$. Sei nun $I \neq \{0\}$. Falls $\exists c \in K \setminus \{0\} : c \in I$, dann existiert $c^{-1} \in K[t]$ und es folgt $1 = c^{-1}c \in I$, also $I = K[t] \cdot 1 = K[t]$.

Sei jetzt $I \neq \{0\}, \forall c \in K : c \in I \Rightarrow c = 0$. Dann existiert $f \in K[t]$ mit $f(t) \in I, \mathrm{Grad}(f(t)) \geq 1$. Wähle ein Polynom $f \in I$ mit minimalem positivem Grad. Dann gilt $K[t] \cdot f \subseteq I$. Wir wollen zeigen, dass $K[t] \cdot f = I$ gilt. Sei $0 \neq g \in I$. Dann ist g nicht konstant. Zu zeigen: g ist Vielfaches von f. Division mit Rest liefert $g = fh + r$ mit $0 \leq \mathrm{Grad}\, r < \mathrm{Grad}\, f$ oder $r = 0$. Es gilt $r = g - fh \in I$. Da $\mathrm{Grad}(f)$ minimal war unter allen Polynomen ungleich 0 in I, folgt $r = 0$ und somit $g = fh \in K[t] \cdot f$. □

Sei V ein endlich-dimensionaler K-Vektorraum und $\alpha \in \mathrm{End}_K(V)$, d. h.

$\alpha : V \to V$ ist linear. Definiere eine Abbildung $K[t] \to \mathrm{End}_K(v)$ durch

$$\begin{aligned} 1 &\mapsto \mathrm{id} \\ t &\mapsto \alpha \\ t^2 &\mapsto \alpha^2 = \alpha \circ \alpha \\ &\vdots \\ t^n &\mapsto \alpha^n = \underbrace{\alpha \circ \ldots \circ \alpha}_{n \text{ mal}} \\ f = a_0 + a_1 t + \ldots + a_n t^n &\mapsto a_0 \,\mathrm{id} + a_1\alpha + a_2\alpha^2 + \ldots + a_n\alpha^n \ . \end{aligned}$$

Diese Abbildung ist additiv, K-linear und multiplikativ, denn $t^{a+b} = t^a \cdot t^b \mapsto \alpha^a \circ \alpha^b = \alpha^{a+b}$. Sie ist daher ein K-linearer Ringhomomorphismus.

Der Ringhomomorphismus $K[t] \to \mathrm{End}_K(V)$ mit $t \mapsto \alpha$ heißt Auswertungshomomorphismus. Bezeichnung: $\mathrm{ev} = \mathrm{ev}_\alpha$ (hängt von α ab).

Proposition 2.4. *Für $\alpha \in \mathrm{End}_K(V)$ ist der Auswertungshomomorphismus $\mathrm{ev} = \mathrm{ev}_\alpha : K[t] \to \mathrm{End}_K(V)$ ein K-linearer Ringhomomorphismus, d. h. additiv, multiplikativ, K-linear, $0 \mapsto 0$ und $1 \mapsto \mathrm{id}$.*

Beispiel 2.3.

- Sei $\alpha = 0$. Dann gilt

$$\mathrm{ev}_\alpha(a_0 + a_1 t + \ldots + a_n t^n) = a_0 \,\mathrm{id} + a_1 \cdot 0 + \ldots + a_n \cdot 0 = a_0 \,\mathrm{id} \ .$$

- Sei $\alpha = \mathrm{id}$. Dann gilt

$$\begin{aligned} \mathrm{ev}_\alpha(a_0 + a_1 t + \ldots + a_n t^n) &= a_0 \,\mathrm{id} + a_1 \,\mathrm{id} + \ldots + a_n \,\mathrm{id} \\ &= (a_0 + \ldots + a_n)\,\mathrm{id} \ . \end{aligned}$$

- Sei $\dim V = l$ und $(v_1, \ldots, v_l)$ eine Basis von V. Sei α die Abbildung mit $\alpha(v_1) = v_2, \ldots, \alpha(v_{l-1}) = v_l$ und $\alpha(v_l) = 0$. Die darstellende Matrix von α bezüglich dieser Basis ist

$$A = \begin{pmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix} \ .$$

Es gilt $\alpha^2(v_1) = v_3, \ldots, \alpha^2(v_{l-2}) = v_l$ und $\alpha^2(v_{l-1}) = \alpha^2(v_l) = 0$. α^2 hat also die darstellende Matrix

$$A^2 = \begin{pmatrix} 0 & & & & \\ 0 & \ddots & & & \\ 1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 0 & 0 \end{pmatrix}.$$

Ensprechend hat α^{l-1} die Matrix

$$A^{l-1} = \begin{pmatrix} 0 & & & \\ \vdots & \ddots & & \\ 0 & & \ddots & \\ 1 & 0 & \ldots & 0 \end{pmatrix}.$$

Es gilt $\alpha^l = 0$ und $A^l = 0$. Somit folgt

$$\mathrm{ev}_\alpha(a_0 + a_1 t + \ldots + a_n t^n) = a_0 \,\mathrm{id} + a_1\alpha + \ldots + a_{l-1}\alpha^{l-1} + 0 + \ldots + 0$$

mit der darstellenden Matrix

$$\begin{pmatrix} a_0 & & & & \\ a_1 & \ddots & & & \\ a_2 & \ddots & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & \\ a_{l-1} & \ldots & a_2 & a_1 & a_0 \end{pmatrix}.$$

- Sei α eine Abbildung mit der darstellenden Matrix $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Es gilt $A^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$. Also hat $\mathrm{ev}_\alpha(a_0 + \ldots + a_n t^n)$ die darstellende Matrix $\begin{pmatrix} x & y \\ 0 & x \end{pmatrix}$ mit $x = a_0 + a_1 + \ldots + a_n$ und $y = a_1 + 2a_2 + 3a_3 + \ldots + na_n$.

Grundidee unseres Vorgehens: Verwende unsere Kenntnisse über Polynome (z. B. Euklidischer Algorithmus), um Informationen über $\alpha \in \operatorname{End}_K(V)$ zu gewinnen. Ziel: Normalform für die darstellende Matrix von A.

Für $f \in K[t]$ ist $f(\alpha) = \operatorname{ev}_\alpha(f)$ eine lineare Abbildung $a_0 \operatorname{id} + a_1\alpha + \ldots + a_n\alpha^n : V \to V$. Wir führen eine neue Bezeichnung ein: $K(f) \coloneqq \operatorname{Kern}(f(\alpha)) < V$ für festes α.

Lemma 2.5. *$\alpha(K(f)) \subseteq K(f)$, d. h. die Einschränkung von α auf $K(f)$ bildet $K(f)$ in sich selbst ab. $K(f)$ ist demnach ein* invarianter Unterraum *unter α. Das ist interessant für eine Blockzerlegung der darstellende Matrix: Ergänzen wir eine Basis $u_1, \ldots, u_l$ von diesem Unterraum zu einer Basis $u_1, \ldots, u_l, v_1, \ldots, v_p$ von V, dann ist die darstellende Matrix von α bezüglich dieser Basis von der Form*

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

mit einer $l \times l$-Matrix A.

Beweis. Sei $x \in K(f) = \operatorname{Kern}(f(\alpha))$, d. h. $f(\alpha)(x) = 0$. Zu zeigen: $\alpha(x) \in K(f)$, d. h. $f(\alpha)(\alpha(x)) = 0$.

Sei $f = a_0 + a_1 t + \ldots + a_n t^n$, also $f(\alpha) = a_0 \operatorname{id} + a_1\alpha + \ldots + a_n\alpha^n$. Nach Voraussetzung gilt

$$0 = f(\alpha)(x) = a_0 \operatorname{id}(x) + a_1\alpha(x) + \ldots + a_n\alpha^n(x) ,$$

also

$$\begin{aligned} 0 = \alpha(0) &= \alpha(a_0 \operatorname{id}(x) + \ldots + a_n\alpha^n(x)) \\ &= a_0\alpha(x) + a_1\alpha^2(x) + \ldots + a_n\alpha^{n+1}(x) \\ &= a_0 \operatorname{id}(\alpha(x)) + a_1\alpha(\alpha(x)) + \ldots + a_n\alpha^n(\alpha(x)) \\ &= (a_0 \operatorname{id} + \ldots + a_n\alpha^n)(\alpha(x)) \\ &= f(\alpha)(\alpha(x)) . \end{aligned}$$

Damit gilt $f(\alpha)(\alpha(x)) = 0$, was zu zeigen war. □

Übertrage Information von Polynomen auf $K(t)$: $f + g$, $f \cdot g$, $\operatorname{ggT}(f, g)$, $\operatorname{kgV}(f, g) = fg / \operatorname{ggT}(f, g)$.

Proposition 2.6. *Seien $\alpha \in \operatorname{End}_K(V)$ und $f, g \in K[t]$. Dann gilt:*

(a) $K(f) \cap K(g) \subseteq K(f+g)$.

(b) $K(f) + K(g) \subseteq K(fg)$.

(c) $g \mid f \Rightarrow K(g) \subseteq K(f)$.

(d) Ist $d = \operatorname{ggT}(f, g)$*, dann gilt* $K(d) = K(f) \cap K(g)$*. Sind insbesondere* f, g *teilerfremd, d. h.* $\operatorname{ggT}(f, g) = 1$*, dann folgt* $K(f) \cap K(g) = \{0\}$.

(e) Ist $m = \operatorname{kgV}(f, g)$*, dann gilt* $K(m) = K(f) + K(g)$.

(f) Sind f, g *teilerfremd, dann gilt* $K(fg) = K(f) \oplus K(g)$.

Beweis.

(a) Sei $x \in K(f) \cap K(g)$, dann müssen wir zeigen, dass $x \in K(f+g)$. $x \in K(f)$ heißt $f(\alpha)(x) = 0$, entsprechend folgt aus der Voraussetzung auch $g(\alpha)(x) = 0$. Dann folgt $(f+g)(\alpha)(x) = (f(\alpha) + g(\alpha))(x) = f(\alpha)(x) + g(\alpha)(x) = 0 + 0 = 0$, also $x \in K(f+g)$.

(b) Sei $x \in K(f) + K(g)$, also $x = x_1 + x_2$ mit $x_1 \in K(f)$ und $x_2 \in K(g)$. Das bedeutet $f(\alpha)(x_1) = 0$ und $g(\alpha)(x_2) = 0$. Zu zeigen: $x \in K(fg)$, d. h. $(f(\alpha)g(\alpha))(x) = 0$. Da $fg = gf$ gilt, folgt $f(\alpha)g(\alpha) = g(\alpha)f(\alpha)$. Folglich gilt

$$\begin{aligned}
(f(\alpha)g(\alpha))(x_1 + x_2) &= f(\alpha)g(\alpha)(x_1) + f(\alpha)g(\alpha)(x_2) \\
&= g(\alpha)f(\alpha)(x_1) + f(\alpha)g(\alpha)(x_2) \\
&= 0 + 0 = 0\ .
\end{aligned}$$

(c) $g \mid f$ bedeutet $f = gh = hg$. Sei $x \in K(g)$, d. h. $g(\alpha)(x) = 0$. Dann folgt $f(\alpha)(x) = hg(\alpha)(x) = h(\alpha) \cdot g(\alpha)(x) = 0$. Das folgt auch aus (b): $K(f) = K(gh) \supseteq K(g) + K(h) \supseteq K(g)$.

(d) Sei $d = \operatorname{ggT}(f, g)$, also gilt nach (c): $K(d) \subseteq K(f)$ und $K(d) \subseteq K(g) \Rightarrow K(d) \subseteq K(f) \cap K(g)$. Es bleibt zu zeigen, dass $K(f) \cap K(g) \subseteq K(d)$ gilt.

Mit dem euklidischen Algorithmus (LAAG 1: Lemma von Bézout, 11.13) folgt $\exists h, j : d = fh + gj$. Dann folgt

$$K(d) = K(fh + gj) \overset{(a)}{\supseteq} K(fh) \cap K(gj) \overset{(c)}{\supseteq} K(f) \cap K(g) \ .$$

Spezialfall: Ist $d = 1$, dann ist $K(f) \cap K(g) = K(1) = \mathrm{Kern}(\mathrm{id}) = \{0\}$.

(e) Sei $m \coloneqq \mathrm{kgV}(f, g)$. Dann gilt $f \mid m$ und $g \mid m$, also $K(f) \subseteq K(m)$ und $K(g) \subseteq K(m) \Rightarrow K(f) + K(g) \subseteq K(m)$, da alles Unterräume sind.

Da m kleinstes gemeinsames Vielfaches von f und g ist, gilt $m = fh = gj$ mit $\mathrm{ggT}(h, j) = 1$. Mit dem Lemma von Bézout folgt, dass Polynome p, q existieren mit $1 = hp+jq$. Somit ist $1(\alpha) = \mathrm{id} = h(\alpha)\cdot p(\alpha)+j(\alpha)\cdot q(\alpha)$. Also gilt für alle $x \in K(m)$:

$$x = \mathrm{id}(x) = \underbrace{(h(\alpha) \cdot p(\alpha))(x)}_{=:y} + \underbrace{(j(\alpha) \cdot q(\alpha))(x)}_{=:z} = y + z \ .$$

Wir wollen zeigen, dass $y \in K(f)$ und $z \in K(g)$, dann haben wir $x \in K(f) + K(g)$ gezeigt. $y \in K(f)$ bedeutet $f(\alpha)(y) = 0$. Das gilt, denn:

$$\begin{aligned} f(\alpha)(y) &= (f(\alpha) \cdot h(\alpha) \cdot p(\alpha))(x) = (m(\alpha) \cdot p(\alpha))(x) \\ &= p(\alpha)(m(\alpha) \underbrace{(x)}_{\in K(m)}) = 0 \Rightarrow y \in K(f) \ . \end{aligned}$$

Ebenso folgt $z \in K(g)$, denn:

$$\begin{aligned} g(\alpha)(z) &= (g(\alpha) \cdot j(\alpha) \cdot q(\alpha))(x) = (m(\alpha) \cdot q(\alpha))(x) \\ &= q(\alpha)(m(\alpha)(x)) = 0 \Rightarrow z \in K(g) \ . \end{aligned}$$

(f) Seien f, g teilerfremd, also $\mathrm{ggT}(f, g) = 1$ und $\mathrm{kgV}(f, g) = fg/\mathrm{ggT}(f, g) = fg$. Aus (e) folgt $K(fg) = K(f)+K(g)$ und aus (d) folgt $K(f)\cap K(g) = \{0\}$, insgesamt also $K(fg) = K(f) \oplus K(g)$. □

Daraus folgt mit Induktion: Sind $f_1, \ldots, f_r$ paarweise teilerfremd, dann gilt $K(f_1 \cdots f_r) = K(f_1) \oplus \ldots \oplus K(f_r)$.

Betrachte $\mathrm{ev}_\alpha : K[t] \to \mathrm{End}_K(V), t \mapsto \alpha$. Dies ist ein Ringhomomorphismus, der für $\dim V > 1$ nicht surjektiv ist, weil $K[t]$ kommutativ ist. Er ist auch nicht injektiv, denn $K[t]$ ist ein unendlich-dimensionaler K-Vektorraum, aber $\mathrm{End}_K(V)$ hat die Dimension n^2, falls $\dim V = n$ ist. Es folgt $\mathrm{Kern}(\mathrm{ev}_\alpha) \neq \{0\}$. $\mathrm{Kern}(\mathrm{ev}_\alpha)$ ist ein Ideal in $K[t]$. Die Ideale in $K[t]$ sind Hauptideale (Proposition 2.3), erzeugt von Polynomen kleinsten Grades im Ideal. Somit existiert ein Polynom m_α mit $\mathrm{Kern}(\mathrm{ev}_\alpha) = K[t] \cdot m_\alpha$. Insbesondere: $m_\alpha(\alpha) = 0$. Wir verlangen: $\mathrm{Grad}(m_\alpha) \geq 1$ und m_α ist normiert, d. h. der Höchstkoeffizient ist 1 (also $m_\alpha = a_0 + \ldots + a_{n-1}t^{n-1} + 1 \cdot t^n$).

Definition 2.7. Sei $V \neq \{0\}$ und $\alpha \in \mathrm{End}_K(V)$. Das eindeutig bestimmte normierte Polynom m_α minimalen Grades mit $m_\alpha(\alpha) = 0$ heißt *Minimalpolynom* von α. Wenn α durch eine darstellende Matrix A gegeben ist, gilt $m_\alpha(A) = 0$. m_α heißt dann auch Minimalpolynom von A.

Strategie: $m_\alpha(\alpha) = 0 \Rightarrow K(m_\alpha) = V$. Eine Zerlegung $m_\alpha = f_1 \ldots f_r$ in teilerfremde Faktoren liefert eine Zerlegung $V = K(f_1) \oplus \ldots \oplus K(f_r)$ in α-invariante Unterräume. Dies führt auf eine Blockzerlegung der Matrix A zu α:

$$\begin{array}{cc} \begin{array}{cccc} K(f_1) & K(f_2) & \ldots & K(f_r) \end{array} & \\ \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_r \end{pmatrix} & \begin{array}{c} K(f_1) \\ K(f_2) \\ \vdots \\ K(f_r) \end{array} \end{array}$$

Beispiel 2.4.

- $\alpha = 0, m_\alpha = t : \mathrm{ev}_\alpha : t \mapsto \alpha = 0$.
- $\alpha = \mathrm{id}, m_\alpha(t) = t - 1 : m_\alpha(\mathrm{id}) = \mathrm{id} - 1 \cdot \mathrm{id} = 0$.
- Sei $\alpha(v_1) = v_2, \ldots, \alpha(v_{n-1}) = v_n, \alpha(v_n) = 0$, also ist die darstellende Matrix A von α bezüglich der Basis $(v_1, \ldots, v_n)$ gegeben durch

$$A = \begin{pmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}.$$

Es folgt $A^n = 0$ und somit $m_\alpha \mid t^n \in \operatorname{Kern}(\operatorname{ev}_\alpha)$. Somit ist $m_\alpha = t^l$ für ein $l \leq n$. Es muss gelten, dass $m_\alpha(\alpha) = \alpha^l = 0$. Da $A^{n-1} \neq 0$ gilt, folgt $l = n$ und $m_\alpha(t) = t^n$.

Vorgehen hier: Errate ein Polynom f mit $f(\alpha) = 0$. m_α muss ein Teiler von f sein.

Allgemeines Vorgehen: Wähle die darstellende Matrix A zu α und bilde $A, A^2, A^3, \ldots$. Es gilt $\dim(\operatorname{End}_K(V)) = (\dim V)^2$. Folglich sind $\operatorname{id}, A, \ldots, A^l$ linear abhängig für $l \geq n^2$, wobei $n := \dim V$. Somit existieren $a_0, \ldots, a_l \in K$ mit $a_0 \operatorname{id} + a_1 A + \ldots a_l A^l = 0$ und nicht alle a_i sind gleich 0. Sei $f = a_0 + a_1 t + \ldots + a_l t^l$ (z. B. $a_l = 1$). Dann folgt $f(A) = 0 \Rightarrow m_A \mid f$. m_A ist das normierte Polynom kleinsten Grades, das durch dieses Verfahren erzeugt wird.

Als nächstes stellen wir die Eigenschaften des Minimalpolynoms m_α zusammen.

- Sei $\alpha \in \operatorname{End}_K(V)$ mit der darstellenden Matrix A. Dann gehört A^T zur dualen Abbildung $\alpha^* : V^* \to V^*$. Ist $f = a_0 + \ldots + a_n t^n$ ein Polynom mit $f(A) = 0$, dann ist auch $f(A^T) = a_0 E_n^T + \ldots + a_n (A^T)^n = (a_0 E_n + \ldots + a_n A^n)^T = 0^T = 0$. Deshalb gilt $m_\alpha = m_{\alpha^*}$, d. h. α und α^* haben das selbe Minimalpolynom. Alternativ: Wähle eine Basis $v_1, \ldots, v_n$ von V. Dann folgt $\alpha^*(v_j^*)(x) = v_j^*(\alpha(x))$, also $f \in K[t] \Rightarrow f(\alpha^*)(v_j^*)(x) = v_j^*(f(\alpha)(x)) \Rightarrow (f(\alpha^*) = 0 \Leftrightarrow f(\alpha) = 0)$.

- Ist m_α das Minimalpolynom von α, dann ist $m_\alpha(\alpha) = 0$ (und m_α ist von kleinstem Grad). Daher ist $m_\alpha(\alpha)(x) = 0$ für alle $x \in V$, also $K(m_\alpha) = V$. Für $f = m_\alpha \cdot g$ gilt $K(f) = V$. Ist hingegen $f \mid m_\alpha$ ein echter Teiler, dann ist $f(\alpha) \neq 0$, d. h. f hat einen echt kleineren Kern.

Lemma 2.8. *Sei f ein Teiler von m_α und g ein echter Teiler von f. Dann ist $K(g) \subsetneq K(f)$.*

Beweis. Nach Voraussetzung gilt $\operatorname{Grad}(g) < \operatorname{Grad}(f)$. Sei $m_\alpha = fh$ und $f = gj$, wobei j nicht konstant ist, also einen Grad > 0 hat. Sei $p := gh \mid fh = m_\alpha$. Dann ist $\operatorname{Grad}(p) = \operatorname{Grad}(g) + \operatorname{Grad}(h) < \operatorname{Grad}(f) + \operatorname{Grad}(h) = \operatorname{Grad}(m_\alpha)$ und somit ist p ein echter Teiler von m_α. Es folgt $p(\alpha) \neq 0$, also existiert $x \in V$ mit $p(\alpha)(x) \neq 0$. Sei $y := h(\alpha)(x) \Rightarrow f(\alpha)(y) = (f(\alpha) \cdot h(\alpha))(x) = m_\alpha(\alpha)(x) =$

$0(x) = 0$. Daher gilt $y \in K(f)$. Es gilt aber $g(\alpha)(y) = (g(\alpha) \cdot h(\alpha))(x) = p(\alpha)(x) \neq 0$ nach Voraussetzung, also gilt $y \in K(f) \setminus K(g)$. Nach 2.6 gilt $K(g) \subseteq K(f)$, also ist die Behauptung gezeigt. □

Beispiel 2.5. Sei α gegeben durch

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

bezüglich der Basis v_1, v_2, v_3, v_4. Dann ist $A^4 = 0$ und $m_\alpha(t) = t^4$. Sei z. B. $f(t) = t^3, g(t) = t^2$. Dann ist

$$f(A) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, g(A) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \ ,$$

also $K(f) = \mathrm{Kern}(A^3) = \langle v_2, v_3, v_4 \rangle$ und $K(g) = \langle v_3, v_4 \rangle$, also $K(g) \subsetneq K(f)$.

Lemma 2.9. *Sei $m_\alpha = f_1 \cdot \ldots \cdot f_r$ eine Faktorisierung von m_α in ein Produkt normierter Polynome $f_1, \ldots, f_r$ mit $\mathrm{ggT}(f_p, f_q) = 1$ für $p \neq q$. Dann ist $V = K(f_1) \oplus K(f_2) \oplus \ldots \oplus K(f_r)$, wobei $\forall j \in \{1, \ldots, r\} : \alpha(K(f_j)) \subseteq K(f_j)$. Sei $\alpha_j : K(f_j) \to K(f_j)$ die Einschränkung von α auf $K(f_j)$. Dann ist f_j das Minimalpolynom von α_j.*

Beweis. Die Zerlegung von V folgt aus 2.6 (f) mit Induktion. Aus 2.5 folgt, dass $K(f_j)$ α-invariant ist. Es gilt

$$\begin{aligned} K(f_j) &= \mathrm{Kern}(f_j(\alpha)) \\ \Rightarrow \forall x \in K(f_j) \ &: \ f_j(\alpha)(x) = 0 \\ \Rightarrow m_{\alpha_j} &\mid f_j \ . \end{aligned}$$

Angenommen, m_{α_j} ist ein echter Teiler von f_j. Dann gilt $\mathrm{Kern}(m_{\alpha_j}(\alpha)) \not\subseteq K(f_j)$ nach Lemma 2.8. Damit folgt ein Widerspruch:

$$\mathrm{Kern}(m_{\alpha_j}(\alpha)) \not\subseteq K(f_j) = \mathrm{Kern}(m_{\alpha_j}(\alpha_j)) \subseteq \mathrm{Kern}(m_{\alpha_j}(\alpha)).$$

Also ist m_{α_j} kein echter Teiler von f_j. Da beide Polynome normiert sind, müssen sie gleich sein. □

Wählen wir Basen von $K(f_1), \ldots, K(f_r)$, dann bekommen wir zusammmen eine Basis von V. Bezüglich dieser Basis hat α eine Matrix der folgenden Form.

$$\begin{array}{cc} \begin{array}{cccc} K(f_1) & K(f_2) & \ldots & K(f_r) \end{array} & \\ \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_r \end{pmatrix} & \begin{array}{c} K(f_1) \\ K(f_2) \\ \vdots \\ K(f_r) \end{array} \end{array}$$

Das war der erste Schritt zur Blockzerlegung. Fragen: Ist diese Zerlegung optimal (Blöcke so klein wie möglich)? Einfache Form für die Blöcke?

Lemma 2.10.

(a) Sei λ eine Nullstelle von m_α. Dann ist λ ein Eigenwert von α.

(b) Sei λ ein Eigenwert von α. Dann ist $m_\alpha(\lambda) = 0$.

Beweis.

(a) Sei $m_\alpha(\lambda) = 0$. Zu zeigen: λ ist Eigenwert von α, d. h. $\exists x \neq 0 : \alpha(x) = \lambda x$, d. h. $x \in \operatorname{Kern}(\alpha - \lambda \operatorname{id})$. Aus $m_\alpha(\lambda) = 0$ folgt $(t - \lambda) \mid m_\alpha$. Wende jetzt 2.8 an auf $(t-\lambda)$: Mit $g(t) := 1$ folgt, dass $\{0\} = K(g) \subsetneq K(t-\lambda)$. Somit ist $\operatorname{Kern}(\alpha - \lambda \operatorname{id}) \neq 0$, also ist λ ein Eigenwert von α.

(b) Sei λ ein EW von α mit Eigenvektor $x \neq 0$. Sei wieder $f = (t - \lambda)$, dann folgt $x \in \operatorname{Kern}(f(\alpha)) = K(f)$, also insbesondere $K(f) \neq \{0\}$. Es gilt $K(m_\alpha) = V$, also $K(f) \cap K(m_\alpha) \neq \{0\}$. Nach 2.6 (d) gilt $K(f) \cap K(m_\alpha) = K(\operatorname{ggT}(f, m_\alpha))$. Wegen $K(1) = \{0\}$ folgt $\operatorname{ggT}(f, m_\alpha) \neq 1 \Rightarrow \operatorname{ggT}(f, m_\alpha) = f = t - \lambda$, also $t - \lambda \mid m_\alpha$ und folglich ist λ Nullstelle von m_α. □

Also: $\{\text{Nullstellen von } m_\alpha\} = \{\text{EW von } \alpha\} = \{\text{Nullstellen von } \chi_\alpha\}$. Aber: Im Allgemeinen ist $m_\alpha \neq \chi_\alpha$. Beispiel: Sei

$$A := \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} .$$

Dann ist $A^2 = 0$ und $A \neq 0$, also $m_A = t^2$, aber $\chi_A = t^4$ (Später: $m_\alpha \mid \chi_\alpha$).

Definition 2.11. Ein Unterraum $U < V$ heißt (bezüglich α) *invarianter Unterraum* $:\Leftrightarrow \alpha(U) \subseteq U$, d. h. die Einschränkung von α ist ein lineare Abbildung $\alpha : U \to U$. Ein invarianter Unterraum heißt unzerlegbar $:\Leftrightarrow$ Für alle $U_1, U_2 < U$ mit $U = U_1 \oplus U_2$ und U_1, U_2 α-invariant folgt $U_1 = \{0\}$ oder $U_2 = \{0\}$.

Unzerlegbar heißt: Es gibt keine Zerlegung in echt kleinere α-invariante Unterräume. D. h. der Block in der Matrix kann nicht in kleinere Blöcke aufgeteilt werden. Eine Zerlegung $V = U_1 \oplus \ldots \oplus U_r$ mit invarianten Unterräumen $U_1, \ldots, U_r$ heißt Blockzerlegung. Hierbei müssen die U_i nicht unzerlegbar sein. Wir suchen die bestmögliche Blockzerlegung mit U_i unzerlegbar für alle i.

Beispiel: Ist $V = U_1 \oplus \ldots \oplus U_r$ mit α-invarianten U_i und $\dim U_i = 1$ für alle i. Dies liefert eine Blockzerlegung mit 1×1-Blöcken. Wähle eine Basis von V: $x_1 \in U_1, \ldots, x_r \in U_r$ mit $x_i \neq 0$ für alle i. Da für alle i, U_i α-invariant ist, folgt $\alpha(x_i) \in U_i = \langle x_i \rangle$, also existiert ein $\lambda_i \in K$ mit $\alpha(x_i) = \lambda x_i$ und x_i ist dann ein Eigenvektor zum Eigenwert λ_i. Somit hat V eine Basis aus Eigenvektoren und α ist diagonalisierbar. Die Existenz einer solchen Zerlegung ist somit äquivalent zur Diagonalisierbarkeit.

Was machen wir, wenn α nicht diagonalisierbar ist? Weiteres Vorgehen:

- 1. Schritt: Betrachte eine spezielle Klasse von linearen Abbildungen oder Matrizen, die im Allgemeinen nicht diagonalisierbar sind, aber bekanntes m_α haben.
- 2. Schritt: Setze den allgemeinen Fall zusammen aus diagonalisierbar und den Matrizen aus dem ersten Schritt.

Zum 1. Schritt: Betrachte z. B. die Matrix

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} .$$

Dann gilt $A^5 = 0, m_A(t) = t^5$. Diese Matrix hat also nur den Eigenwert 0 und ist daher nicht diagonalisierbar (folgt z. B. aus $\operatorname{Rang} A = 4$).

Definition 2.12. Ein Endomorphismus $\alpha \in \mathrm{End}_K(V)$ heißt *nilpotent* $:\Leftrightarrow$ $\exists l > 0 : m_\alpha(t) = t^l$, d. h. $\alpha^l = 0$. $\mathrm{Grad}\, m_\alpha = l$ heißt die *Ordnung* von α.

Die Matrix A von oben ist also nilpotent mit Ordnung 5. Die $n \times n$-Matrix $(n \geq 2)$ mit einer 1 in der linken unteren Ecke und sonst 0, ist nilpotent von Ordnung 2.

Im 1. Schritt betrachten wir nilpotente Matrizen.

Lemma 2.13. *Sei* $\dim V = n \neq 0$ *und* $\alpha \in \mathrm{End}_K(V)$ *nilpotent von Ordnung* l*. Dann gilt* $l \leq n$*. (Das vorherige Beispiel zeigt, dass* $l = n$ *möglich ist.)*

Beweis. Sei α von Ordnung l. Also $\alpha^{l-1} \neq 0 \Rightarrow \exists x \in V : \alpha^{l-1}(x) \neq 0$. Betrachte $x, \alpha x, \ldots, \alpha^{l-1}x$, das sind l Vektoren. Wir zeigen: diese Vektoren sind linear unabhängig. Dann folgt: $l \leq n = \dim V$.

Seien $\lambda_0, \ldots, \lambda_{l-1} \in K$ mit

$$\lambda_0 x + \lambda_1 \alpha x + \ldots + \lambda_{l-1}\alpha^{l-1}x = 0 \ . \tag{$*$}$$

Zu zeigen ist $\lambda_0 = \ldots = \lambda_{l-1} = 0$. Wende α^{l-1} auf $(*)$ an:

$$\begin{aligned} 0 &= \alpha^{l-1}(0) \\ &= \alpha^{l-1}(\lambda_0 x + \lambda_1 \alpha x + \ldots + \lambda_{l-1}\alpha^{l-1}x) \\ &= \lambda_0 \alpha^{l-1}x + \lambda_1 \underbrace{\alpha^l x}_{=0} + \ldots + \lambda_{l-1} \underbrace{\alpha^{2l-2}x}_{=0} \\ &= \lambda_0 \underbrace{\alpha^{l-1}x}_{\neq 0} \ , \end{aligned}$$

also folgt $\lambda_0 = 0$. Jetzt können wir α^{l-2} anwenden und es folgt $\lambda_1 = 0$ usw. Per Induktion folgt also $\lambda_0 = \ldots = \lambda_{l-1} = 0$, somit sind die Vektoren $x, \alpha x, \ldots, \alpha^{l-1}x$ linear unabhängig. □

Idee für Normalform: Setze $U := \langle x, \alpha x, \ldots, \alpha^{l-1}x \rangle$ wie im vorherigen Beweis. U ist α-invariant:

$$\alpha(\alpha^j x) = \begin{cases} \alpha^{j+1}x & , j < l-1 \\ 0 & , j = l-1 \ . \end{cases}$$

Die Matrix von α auf U ist also

$$\begin{pmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}.$$

Notation: Sei $U < V$ ein Unterraum. Wähle eine Basis $u_1, \dots, u_l$ und setze diese zu einer Basis $u_1, \dots, u_l, v_1, \dots, v_r$ fort. Es gilt $l + r = \dim V =: n$ und $U^0 = \langle v_1^*, \dots, v_r^* \rangle < V^*$. Sei $W =: \langle u_1^*, \dots, u_l^* \rangle < V^*$. Es gilt $W \cap U^0 = \{0\}$, denn $u_1^*, \dots, u_l^*, v_1^*, \dots, v_r^*$ ist eine Basis von V^*. Dann folgt $W^0 \cap \iota(U) = \{0\}$ für den Auswertungsisomorphismus $\iota : V \to V^{**}$. $u_1^*, \dots, u_l^*$ hängen von der gesamten Basis $u_1, \dots, u_l, v_1, \dots, v_r$ ab und nicht nur von $u_1, \dots, u_l$, also hängt auch W von dieser Basiswahl ab, während U^0 durch U eindeutig bestimmt ist.

Beispiel $V = \mathbb{R}^2$, Basis $e_1 = (1,0)^T, e_2 = (0,1)^T, U = \langle e_1 \rangle, e_1 = u_1$. Wähle $v_1 = e_2$. Dann ist $u_1^*(u_1) = 1$ und $u_1^*(v_1) = 0$. Wählt man stattdessen $u_1' = u_1$ und $v_1' = (1,1)^T$, dann ist $(u_1')^*(e_2) = (u_1')^*(v_1' - u_1') = -1$.

Definition 2.14. Unterräume $U < V$ und $W < V^*$ heißen *dual* zueinander $:\Leftrightarrow U^0 \cap W = \{0\}$ und $W^0 \cap \iota(U) = \{0\}$.

Daraus folgt: $\dim W = \dim U$. Denn: $U^0 \cap W = \{0\} \Rightarrow \dim U_0 + \dim W \leq \dim V^*$ also $(n-l) + \dim W \leq n$, also $\dim W \leq l$. Andererseits gilt $\dim W^0 + \dim \iota(U) \leq dim V^{**}$. Mit $\dim \iota(U) = \dim U = l$ und $\dim V^{**} = n$ folgt $\dim W = n - \dim W^0 \geq l$.

Proposition 2.15.

(a) $\alpha \in \operatorname{End}_K(V)$ *und* α^* *haben dasselbe Minimalpolynom.*

(b) Ist $V_1 \subseteq V$ *ein invarianter Unterraum bezüglich* α*, dann ist* $V_1^0 \subseteq V^*$ *ein invarianter Unterraum bezüglich* α^**.*

(c) Ist $V = V_1 \oplus V_2$ *eine Blockzerlegung von* V*, dann ist* $V^* = V_1^0 \oplus V_2^0$ *eine Blockzerlegung von* V^**.*

(d) *Sind $V_1 \subseteq V$ und $W_1 \subseteq V^*$ dual zueinander und invariant bezüglich α bzw. α^*, dann sind V_1^0, W_1^0 invariant unter α^* bzw. $\iota(\alpha)$ und $V = V_1 \oplus \iota^{-1}(W_1^0)$ sowie $V^* = W_1 \oplus V_1^0$. Hierbei ist $\iota(\alpha) = \iota \circ \alpha \circ \iota^{-1}$, insbesondere haben α und $\iota(\alpha)$ die gleiche darstellende Matrix.*

Beweis.

(a) Diese Aussage ist schon bewiesen (α^* hat die Matrix A^T).

(b) Per Definition ist $V_1^0 = \{\varphi \in V^* \mid \forall x \in V_1 : \varphi(x) = 0\}$. Zu zeigen ist, dass V_1^0 invariant ist unter α^*. Sei $\varphi \in V_1^0$, zu zeigen: $\alpha^*(\varphi) \in V_1^0 \Leftrightarrow \forall x \in V_1 : \alpha^*(\varphi)(x) = 0$. Sei also $x \in V_1$. Dann ist $\alpha(x) \in V_1$ nach Voraussetzung und somit

$$\alpha^*(\varphi)(x) = \varphi(\alpha(x)) = 0 \ .$$

(c) Nach 1.11 ist V_1^0 dual zu V_2 und umgekehrt, also folgt $V^* = V_1^0 \oplus V_2^0$. Nach (b) sind V_1^0 und V_2^0 α-invariant.

(d) Wir haben bereits gezeigt, dass V_1^0 und W_1^0 invariant sind. Es bleibt zu zeigen, dass $V = V_1 \oplus \iota^{-1}(W_1^0)$ (und analog $V^* = W_1 \oplus V_1^0$). Zuerst zeigen wir, dass $V_1 \cap \iota^{-1}(W_1^0) = \{0\}$ oder äquivalent dazu (weil ι ein Isomorphismus ist) $\iota(V_1) \cap W_1^0 = \{0\}$. Wir wissen schon, dass V_1 und W_1 dual sind, also folgt per Definition $\iota(V_1) \cap W_1^0 = \{0\}$.

Es bleibt zu zeigen, dass $n = \dim V = \dim V_1 + \dim \iota^{-1}(W_1^0)$, also dass $\dim \iota^{-1}(W_1^0) = n - l$. Da V_1 und W_1 dual zueinander sind, folgt $\dim W_1 = \dim V_1 = l$, also folgt $\dim W_1^0 = \dim V - \dim W_1 = n - l$. □

Der Nutzen ist: $V = V_1 \oplus \iota^{-1}(W_1^0)$ ist eine Blockzerlegung in invariante Unterräume. Wenn V_1 bekannt ist, können wir $\iota^{-1}(W_1^0)$ induktiv weiter zerlegen. $\iota^{-1}(W_1^0)$ ist invariantes Komplement zu V_1 bezüglich α.

Bemerkung 2.16. Konstruktion einer Blockzerlegung für α nilpotent von Ordnung k, d. h. $\alpha^k = 0$ mit k minimal.

Sei $x \in V$ mit $x, \alpha(x), \ldots, \alpha^{k-1}(x)$ linear unabhängig. Sei $U := \langle x, \alpha(x), \ldots, \alpha^{k-1}(x)\rangle < V$, U ist α-invarianter Unterraum und V hat eine Blockzerlegung $V =$

$U \oplus \iota^{-1}(W^0)$ mit W dual zu U. Bezüglich der Basis $x, \alpha(x), \ldots, \alpha^{k-1}(x)$ von U hat α auf U die Matrix

$$\begin{pmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}.$$

Auf $\iota^{-1}(W^0)$ ist α nilpotent von Ordnung $\leq k$. Weiter mit Induktion.

In 2.17 werden wir zeigen, dass U unzerlegbar ist, d. h. der zugehörige Block nicht weiter verkleinert werden kann. Noch zu zeigen ist, dass die Konstruktion funktioniert. Wir müssen auch noch eine konkrete Berechnungsvorschrift für die Normalform finden:

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

Es kann 1×1-Blöcke geben, nämlich die von der Form (0) zum Eigenraum des Eigenwerts 0.

Begründung, dass die Konstruktion funktioniert: Ist $k = 1$, dann ist $\alpha = 0$, also ist das Ergebnis die Nullmatrix. Für $k > 1$ ist $\alpha^{k-1} \neq 0$, also existiert $x \neq 0$ mit $\alpha^{k-1}(x) \neq 0$ (wähle zur Berechnung z. B. einen Basisvektor, zu dem in A^{k-1} nicht eine Nullspalte gehört). Sei $U = \langle x, \alpha(x), \ldots, \alpha^{k-1}(x) \rangle$ mit linear unabhängigen Vektoren (folgt wie im Beweis von Lemma 2.13). Dann ist $\dim U = k$ und die Vektoren $x, \ldots, \alpha^{k-1}(x)$ bilden eine Basis von U, bezüglich der die darstellende Matrix von $\alpha|_U$ die angegebene Form annimmt.

Nächster Schritt: Suche α-invariantes Komplement zu U. Wir haben $\alpha^{k-1}(x) \neq 0$ gefordert und daher existiert ein $\varphi \in V^*$ mit $\varphi(\alpha^{k-1}(x)) \neq 0$. Es folgt $0 \neq \varphi(\alpha^{k-1}(x)) = (\alpha^*)^{k-1}(\varphi)(x)$, also $\varphi \neq 0, \alpha^*(\varphi) \neq 0, \ldots, (\alpha^*)^{k-1}(\varphi) \neq 0$. Wie im Beweis von Lemma 2.13 folgt, dass $\varphi, \ldots, (\alpha^*)^{k-1}(\varphi)$ linear unabhängig sind.

Setze $W \coloneqq \langle \varphi, \ldots, (\alpha^*)^{k-1}(\varphi) \rangle < V^*$. W hat analoge Eigenschaften wie U: W ist α^*-invariant, $\dim(W) = k$ und bezüglich der angegebenen Basis hat α^* die altbekannte Normalform. Wir müssen noch zeigen, dass U und W dual zueinander sind. Dann können wir 2.15 (d) anwenden und wir erhalten eine Blockzerlegung $V = U \oplus \iota^{-1}(W^0)$, sodass wir mit Induktion auf $\iota^{-1}(W^0)$ weitermachen können. Es ist also zu zeigen, dass $U^0 \cap W = \{0\}$ und $W^0 \cap \iota(U) = \{0\}$. Sei $\psi \in U^0 \cap W$. Wegen $\psi \in W$ gilt $\psi = \lambda_0 \varphi + \ldots + \lambda_{k-1} (\alpha^*)^{k-1}(\varphi)$. $\psi \in U^0$ bedeutet $\psi(u) = 0$ für alle $u \in U$. Setze $u = \alpha^{k-1}(x)$ ein:

$$\begin{aligned} 0 = \psi(u) &= \lambda_0 \underbrace{\varphi(\alpha^{k-1}(x))}_{\neq 0} + \lambda_1 \underbrace{\varphi(\alpha^k(x))}_{=0} + \ldots \\ &= \lambda_0 \varphi(\alpha^{k-1}(x)) \ , \end{aligned}$$

also $\lambda_0 = 0$. Mit Induktion folgt $\lambda_i = 0$ für alle i.

Proposition 2.17. *Sei* $\dim U = k$ *und* $\alpha : U \to U$ *nilpotent mit Ordnung* l. *Dann ist* U *unzerlegbar genau dann, wenn* $k = l$.

Beweis. Aus 2.13 folgt $l \leq k$. Falls $l < k$ gilt, existiert ein invariantes $W < U$ mit $\dim W = l < \dim U$. Es folgt, dass U zerlegbar ist, denn nach 2.16 existiert ein invariantes Komplement, also eine Zerlegung.

Sei $l = k$. Zu zeigen: U ist unzerlegbar. Angenommen, $U = U_1 \oplus U_2$ mit α-invarianten Unterräumen U_1, U_2. Zu zeigen ist, dass $U_1 = \{0\}$ oder $U_2 = \{0\}$. Das Minimalpolynom von α auf U ist $m_\alpha(t) = t^l = t^k$. Das Minimalpolynom von α auf U_1 ist $m_1(t) = t^{k_1}$ mit $k_1 \leq k$. Entsprechend ist das Minimalpolynom von α auf U_2 das Polynom $m_2(t) = t^{k_2}, k_2 \leq k$. Da α die Ordnung k hat, gibt es ein x, sodass $x, \alpha x, \ldots, \alpha^{k-1}(x) \neq 0$ gilt. Da $U = U_1 \oplus U_2$ gilt, können wir $x = x_1 + x_2$ schreiben mit $x_1 \in U_1$ und $x_2 \in U_2$. Es folgt

$$\alpha^{\max\{k_1,k_2\}}(x) = \alpha^{\max\{k_1,k_2\}}(x_1) + \alpha^{\max\{k_1,k_2\}}(x_2) = 0 \ ,$$

also $\max\{k_1, k_2\} = k$. Sei o.B.d.A. $k_1 = k$. Es folgt $k = k_1 \leq \dim U_1$, also $U_1 = U$ und $U_2 = \{0\}$. □

2.17 sagt: unser Verfahren ist optimal, die Blöcke sind kleinstmöglich.

Theorem 2.18. *Sei $\alpha : V \to V$ nilpotent von Ordnung k. Dann existiert eine Zerlegung von V in α-invariante unzerlegbare Unterräume. Mindestens einer der Unterräume hat die Dimension k. Für jeden Unterraum kann man eine Basis wählen, bezüglich derer die darstellende Matrix von α auf dem Unterraum die Form*

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

hat. (Wahlweise ist auch die Transponierte davon möglich. Die Anordnung der Blöcke kann gewählt werden, die Basis auch. Die Anzahl der Blöcke fester Größe liegt fest.)

1×1-Block: (0) zum Basisvektor x. Es gilt $\alpha(x) = 0 \cdot x$, d. h. x ist Eigenvektor zum EW 0, $x \in \operatorname{Kern}(\alpha)$. Jeder Block hat einen Basisvektor x mit $x \in \operatorname{Kern}(\alpha)$ (den, der zur letzten Spalte in der Normalform gehört). Es gilt also $\dim V - r(\alpha) = \dim \operatorname{Kern}(\alpha) =$ Anzahl der Blöcke.

Zum Beispiel 2×2-Block:

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} .$$

Basisvektoren $x, \alpha(x)$ mit $\alpha^2(x) = 0$. Das Minimalpolynom eines unzerlegbaren $l \times l$-Blocks ist $m(t) = t^l$, sein Rang ist $l - 1$.

Verfahren zur Bestimmung der Normalform (und der zugehörigen Basis): Wähle irgendeine Matrix A zu α. Berechne $\dim V - r(A) = \dim \operatorname{Kern}(A) =$ Anzahl der Blöcke (mit Gauß-Algorithmus). Bestimme Blöcke der Größe nach, beginnend mit den Größten: Bestimme m_α (Ausprobieren: $A, A^2, A^{k-1} \neq 0, A^k = 0 \Rightarrow m_\alpha(t) = t^k$). Ist $m_\alpha = t^k$, dann existiert ein $k \times k$-Block, aber kein größerer. Gesucht: x mit $\alpha^{k-1}(x) \neq 0$ (dann ist $\langle x, \dots, \alpha^{k-1}(x) \rangle$ ein Block): Löse $A^{k-1}x = 0$ mit dem Gauß-Algorithmus und wähle irgendein x, das nicht in der Lösungsmenge liegt. Die Lösungsmenge von $A^{k-1}x = 0$ ist ein echter Unterraum $W < V,\ W \neq V$. Wähle irgendeine

Basis von W und ergänze sie durch $v_1, v_2, \ldots$ zu einer Basis von V. Es gilt: $A^{k-1}v_1 \neq 0, A^{k-1}v_2 \neq 0, \ldots$. $v_1, v_2, \ldots$ sind linear unabhängig nicht nur in V, sondern auch in V/W. Das bedeutet: Ist $\lambda_1 v_1 + \lambda_2 v_2 + \ldots \in W$, dann folgt $0 = \lambda_1 = \lambda_2 = \ldots$. Daraus folgt $\forall l < k : A^l v_1, A^l v_2, \ldots$ sind linear unabhängig. Also $\langle v_1, \ldots, \alpha^{k-1}(v_1)\rangle \cap \langle v_2, \ldots, \alpha^{k-1}(v_2)\rangle = \{0\}$ und so weiter. Durch diese Wahl ist gesichert, dass die von $v_1, v_2, \ldots$ erzeugten Blöcke der Größe $k \times k$ disjunkt sind. Insbesondere sind $\alpha^{k-1}(v_1), \alpha^{k-1}(v_2), \ldots$ linear unabhängig. Damit sind die $k \times k$-Blöcke konstruiert.

Als nächstes: $(k-1) \times (k-1)$-Blöcke. Gesucht: x mit $\alpha^{k-2}(x) \neq 0$. Zu vermeiden sind $\alpha(v_1), \alpha(v_2), \ldots$, denn diese Vektoren sind schon in $k \times k$-Blöcken. Gesucht sind Vektoren x mit $\alpha^{k-1}(x) = 0, \alpha^{k-2}(x) \neq 0, x \notin \langle \alpha(v_1), \alpha(v_2), \ldots\rangle$. Also: Löse LGS $A^{k-2}x = 0$, bestimme also die Lösungsmenge W_2. Wähle eine Basis von W_2. Wähle eine Basis von $W_2 \oplus \langle v_1, v_2, \ldots, \alpha(v_1), \alpha(v_2), \ldots\rangle$ und ergänze sie durch Vektoren $w_1, w_2, \ldots$ zu einer Basis von V. $w_1, w_2, \ldots$ erfüllen: $\alpha^{k-2}(w_1) \neq 0, \alpha^{k-2}(w_2) \neq 0, \ldots$. Diese Vektoren liefern die $(k-1) \times (k-1)$-Blöcke. Ein solcher Block ist z. B. $\langle w_1, \ldots, \alpha^{k-2}(w_1)\rangle$. Dann machen wir weiter mit $(k-2) \times (k-2)$-Blöcken usw.

Beispiel 2.6. Sei A eine 7×7-Matrix mit $A^3 = 0$ (also ist A nilpotent). Außerdem sei $r(A) = 4, r(A^2) = 1$. Aufgabe: Bestimme die Normalform.

Wegen $A^3 = 0, A^2 \neq 0$ gilt $m_A = t^3$, also gibt es 3×3-Blöcke, aber keine größeren Blöcke. Vielleicht gibt es kleinere Blöcke. Möglich wären erst einmal zwei Dreierblöcke und ein Einerblock oder ein Dreierblock und zwei Zweierblöcke oder ein Zweierblock mit zwei Einerblöcken oder vier Einerblöcke. Quadriert man die Normalform mit zwei Dreierblöcken und einem Einerblock, dann sieht man, dass $r(A^2) = 2$ gelten müsste. Wegen $\dim V - r(A) = 7 - 4 = 3$ muss es drei Blöcke geben, also bleibt nur die Möglichkeit mit einem 3×3-Block und zwei 2×2-Blöcken übrig:

$$\begin{pmatrix} 0 & 1 & & & & & \\ & 0 & 1 & & & & \\ & & 0 & & & & \\ & & & 0 & 1 & & \\ & & & & 0 & & \\ & & & & & 0 & 1 \\ & & & & & & 0 \end{pmatrix}$$

Die Ränge bestimmen nicht immer die Normalform, aber sie schränken die Möglichkeiten ein.

Beispiel 2.7. Sei

$$A := \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow A^2 = \begin{pmatrix} 0 & 0 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, A^3 = 0 \Rightarrow m_A = t^3 \; .$$

A hat die Ordnung 3, also folgt

$$A \sim \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \; .$$

Gesucht ist eine Basis für die Blockform. Es gilt $r(A) = 2 \Rightarrow \dim \operatorname{Kern}(A) = 3 - 2 = 1$, $\operatorname{Kern}(A) = \langle e_1 \rangle$. Suche x mit $A^2 x \neq 0$. Wähle $x = e_3$, denn $A^2 e_3 = 3e_1 \neq 0$. Basis: $\{e_3, Ae_3, A^2 e_3\} = \{e_3, (2,3,0)^T, 3e_1\}$. Bezüglich dieser Reihenfolge ist die darstellende Matrix

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \; .$$

Nachdem wir eine Normalform für nilpotente Matrizen eingeführt haben, wollen wir jetzt den allgemeinen Fall (zumindest für $K = \mathbb{C}$) betrachten. Sei also V ein endlich-dimensionaler K-Vektorraum und $\alpha \in \operatorname{End}_K(V)$. Erster Schritt: bestimme das Minimalpolynom m_α. Über $K = \mathbb{C}$ ist m_α ein Produkt von Linearfaktoren (Fundamentalsatz der Algebra). Sei also

$$m_\alpha = (t - \lambda_1)^{k_1} \cdots (t - \lambda_r)^{k_r}$$

mit paarweise verschiedenen λ_j. Nach 2.10 sind $\lambda_1, \ldots, \lambda_r$ die Eigenwerte von α. Unser Verfahren zur Bestimmung einer Normalform von α wird funktionieren, wenn m_α ein Produkt von Linearfaktoren ist (also immer für $K = \mathbb{C}$). $(t - \lambda_j)^{k_j}$ und $(t - \lambda_l)^{k_l}$ sind für $j \neq l$ teilerfremd. Nach 2.6 (f) liefert dies eine Zerlegung in α-invariante Unterräume: $V = K(f_1) \oplus \ldots \oplus K(f_r)$ mit $f_j = (t - \lambda_j)^{k_j}$. Hierbei ist $K(f_j) = \{x \in V \mid (\alpha - \lambda_j)^{k_j}(x) = 0\}$. Im Fall $k_j = 1$ ist die Bedingung $(\alpha - \lambda_j \operatorname{id})(x) = 0$, d. h. x ist EV zum EW λ_j (oder $x = 0$).

Die Zerlegung von V liefert eine Blockform.

Definition 2.19. Die Zerlegung $V = K(f_1) \oplus \ldots \oplus K(f_r)$ heißt *Hauptraumzerlegung*. Die invarianten Unterräume $K(f_j)$ heißen *Haupträume* oder *verallgemeinerte Eigenräume*, iher Elemente ($\neq 0$) *verallgemeinerte Eigenvektoren* oder *Hauptvektoren*.

Für eine Normalform genügt es, einen einzelnen Hauptraum zu betrachten, d. h. $m_\alpha = (t - \lambda)^k$. Im Fall $\lambda = 0$ ist α nilpotent von Ordnung k und die Normalform ist bekannt. Sei $\sigma \coloneqq \alpha - \lambda \operatorname{id}$. σ ist linear, d. h. $\sigma \in \operatorname{End}_K(V)$. Es folgt $0 = m_\alpha(\alpha) = (\alpha - \lambda \operatorname{id})^k = \sigma^k$ und somit ist σ nilpotent. Also ist die Normalform von σ bekannt. Wähle eine Basis, bezüglich derer σ die Normalform mit Blöcken

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

annimmt. Bezüglich dieser Basis hat id die Matrix E_k, also hat $\lambda \operatorname{id}$ die Matrix λE_k, also hat $\alpha = \sigma + \lambda \operatorname{id}$ Blöcke der Form

$$\begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}.$$

Das ist die Normalform von α.

Theorem 2.20 (Jordansche Normalform)**.** *Sei $\alpha \in \operatorname{End}_K(V)$, $\dim V < \infty$ und m_α zerfalle in Linearfaktoren (über $\mathbb{C}$ ist das immer erfüllt). Dann existiert eine Zerlegung von V in α-invariante Unterräume und dazu eine Basis, sodass die darstellende Matrix von α Blockform hat mit Blöcken der Form*

$$\begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \quad \textit{oder wahlweise} \quad \begin{pmatrix} \lambda & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{pmatrix}.$$

Für jeden solchen Block

$$\begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$$

zu einer Basis $v_1, \dots, v_l$ in der Jordan-Normalform von α ist v_1 ein Eigenvektor von α zum Eigenwert λ. Entsprechend findet man auch einen Eigenvektor, falls der Block transponiert ist. Hat α keine Eigenwerte (z. B. eine geeignete Drehung in 2D), dann kann α nicht in diese Normalform gebracht werden. Noch zu klären ist die Anzahl und Größe der Jordanblöcke zum Eignwert λ.

Beispiel 2.8. Für 2×2-Matrizen hat m_α den Grad 1 oder 2. Erster Fall: Grad 1, d. h. $m_\alpha = t - \lambda$. Also ist $\alpha = \lambda \operatorname{id}$ und α hat die Normalform λE_2. Die verallgemeinerten Eigenräume sind in diesem Fall also gleich den normalen Eigenräumen.

Zweiter Fall: Grad 2, dann ist $m_\alpha = (t - \lambda)(t - \mu)$.

- Erster Teilfall: $\lambda \neq \mu$, dann gibt es zwei Haupträume: $V = K(t - \lambda) \oplus K(t - \mu)$. Die Dimensionen beider Haupträume sind 1, also ergeben sich zwei 1×1-Blöcke. Insgesamt ist die Normalform (bis auf Permutation von λ und μ)

$$\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} .$$

- Zweiter Teilfall: $\lambda = \mu$, also $m_\alpha = (t - \lambda)^2$. Die Normalform könnte einen oder zwei Blöcke enthalten. Falls dies zwei Blöcke wären, wäre die darstellende Matrix also λE_n mit Minimalpolynom $t - \lambda$, ein Widerspruch. Also hat die darstellende Matrix nur einen Block und ist daher

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} .$$

Sind v_1, v_2 die beiden Basisvektoren bezüglich dieser Normalform, dann ist v_1 ein EW, aber v_2 nicht. Es folgt $\dim \operatorname{Eig}(\alpha, \lambda) = 1$.

Diese Fälle können wie folgt zusammengefasst werden:

Normalform	Minimalpolynom	charakteristisches Polynom
$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$	$t-\lambda$	$(t-\lambda)^2$
$\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}, \lambda \neq \mu$	$(t-\lambda)(t-\mu)$	$(t-\lambda)(t-\mu)$
$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$	$(t-\lambda)^2$	$(t-\lambda)^2$

Beispiel 2.9. Allgemeiner Jordanblock zum EW λ: Sei β gegeben durch eine $l \times l$-Matrix

$$\begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}.$$

Es gilt dann $m_\beta(t) = (t-\lambda)^l$ und $\chi_\beta(t) = (-1)^l(t-\lambda)^l$ (folgt zum Beispiel durch Laplace-Entwicklung nach den Spalten). D. h. $m_\beta(t) = \pm\chi_\beta(t)$.

Hat β einen EW λ, aber mehrere Jordan-Blöcke, dann können wir diese in der Normalform nach ihrer Größe l_i sortieren, sodass die Normalform von oben links nach unten rechts Blöcke der Größe $l_1 \times l_1, \ldots, l_r \times l_r$ hat mit $l_i \geq l_{i+1}$. Dann ist $m_\beta(t) = (t-\lambda)^{l_1}$, denn das Minimalpolynom des i-ten Blocks ist $(t-\lambda)^{l_i}$. Der größte λ-Block bestimmt also den Exponenten von $(t-\lambda)$. Wir erhalten außerdem $\chi_\beta(t) = (\lambda-t)^{l_1}\cdots(\lambda-t)^{l_r} = (\lambda-t)^l = (-1)^l(t-\lambda)^l$.

Insbesondere gilt unter unserer Voraussetzung, dass m_β in Linearfaktoren zerfällt, dass $m_\beta \mid \chi_\beta$ und somit $\chi_\beta(\beta) = 0$.

Theorem 2.21 (Cayley-Hamilton)**.** *Sei A eine Matrix, χ_A ihr charakteristisches Polynom. Dann ist $\chi_A(A) = 0$.*

Beweis. Beweisskizze: Für die Jordan-Normalform stimmt das, wie wir gerade gezeigt haben. Es gilt $\chi_A = \chi_{TAT^{-1}}$, d. h. χ_A ändert sich nicht beim Übergang zu ähnlichen Matrizen. Damit sind wir fertig über $\mathbb{C}$. Ist $A \in M_n(\mathbb{R}) \subseteq M_n(\mathbb{C})$, dann hat A über $\mathbb{C}$ eine Jordan-Normalform und $\chi_A(A) = 0$. Damit sind wir fertig über $\mathbb{R}$. Für einen allgemeinen Grundkörper: erweitere zu $\overline{K}$, sodass

A über $\overline{K}$ Jordan-Normalform hat. Dass eine solche Körpererweiterung $\overline{K}$ existiert, wird in der Vorlesung „Algebra“ bewiesen. $\square$

Verfahren zur Bestimmung einer Jordan-Normalform für $\alpha \in \mathrm{End}_K(V)$:

(1) Bestimme irgendeine darstellende Matrix A.

(2) Bestimme das charakteristische Polynom χ_A oder das Minimalpolynom m_A. Bestimme die Eigenwerte von α als Nullstellen von χ_A oder von m_A. Falls m_A kein Produkt von Linearfaktoren ist, gibt es keine Jordan-Normalform. (Möglicher Ausweg z. B. über $\mathbb{C}$ statt über $\mathbb{R}$ rechnen.) Falls m_A ein Produkt von Linearfaktoren ist, weiter mit Schritt 3.

(3) Bestimme für jeden EW λ die verallgemeinerten Eigenräume bzw. ihre Dimension: Berechne

$$\begin{aligned} d_1 &:= \dim \mathrm{Kern}(A - \lambda E) \\ d_2 &:= \dim \mathrm{Kern}((A - \lambda E)^2) \\ &\ \vdots \\ d_l &:= \dim \mathrm{Kern}((A - \lambda E)^l) \end{aligned}$$

so lange, bis sich der Kern bzw. die Dimension nicht mehr ändert. Ist $d_{l-1} < d_l = d_{l+1}$, dann ist $\mathrm{Kern}(A - \lambda E)^l$ der Hauptraum zu λ. Es gilt $d_1 = \dim \mathrm{Eig}(A, \lambda)$ und $d_j = d_{j-1} +$ Anzahl Blöcke $\geq j$. Folglich gilt

- $d_l - d_{l-1} =$ Anzahl Blöcke $\geq l =$ Anzahl Blöcke der Größe $l \times l$ (maximale Größe).
- $d_{j+1} - d_j =$ Anzahl Blöcke $\geq j + 1$ und $d_j - d_{j-1} =$ Anzahl Blöcke $\geq j$. Also ist die Anzahl der Blöcke der Größe $j \times j$ genau

$$(d_j - d_{j-1}) - (d_{j+1} - d_j) = 2d_j - d_{j-1} - d_{j+1} \ .$$

Beispiel 2.10. Sei

$$A = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 3 & 3 \\ -1 & -2 & -2 \end{pmatrix} \ .$$

Dann ist $\chi_A(t) = -t^3 + 3t^2 - 3t + 1 = -(t-1)^3$. Da $\lambda = 1$ dreifache Nullstelle ist, ist der Hauptraum zu λ ganz V. Es gilt

$$A - \lambda E_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ -1 & -2 & -3 \end{pmatrix}$$

Es gilt $r(A - \lambda E_3) = 1$, mit der Dimensionsformel folgt $d_1 = \dim \operatorname{Kern}(A - \lambda E_3) = 3 - 1 = 2$. Außerdem gilt $(A - \lambda E_3)^2 = 0$, also $d_2 = \dim \operatorname{Kern}((A - \lambda E_3)^2) = 3$ und dann $d_3 = d_2$ usw. Wegen $d_2 - d_1 = 1$ folgt, dass es einen 2×2-Block gibt. Es bleibt ein 1×1-Block übrig. Also gilt

$$A \sim \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}$$

mit $\lambda = 1$.

Allgemeiner: Gesucht ist eine Normalform und eine Basis dazu. Bestimme Basen für jeden EW λ wie folgt: Beginne mit l, wobei $d_l = d_{l+1}$ (also: größte Blöcke sind $l \times l$). Gesucht sind Vektoren in $\operatorname{Kern}(A - \lambda E)^l \setminus \operatorname{Kern}(A - \lambda E)^{l-1}$ usw., verwende also die Konstruktion der Normalform für nilpotente Matrizen (außer, dass die Matrix hier nur auf einem Hauptraum nilpotent ist). Wähle also eine Basis von $\operatorname{Kern}((A - \lambda E)^{l-1})$ und ergänze diese zu einer Basis von $\operatorname{Kern}((A - \lambda E)^l)$. Sind $v_1, w_1, \ldots$ die ergänzten Vektoren, dann ergeben sich Blöcke durch $v_1, v_2 := (A - \lambda E)v_1, \ldots, v_{l-1} = (A - \lambda E)^{l-1} v_1$ und das gleiche für w_1 usw. Weiter mit $(l-1) \times (l-1)$-Blöcken (falls existent). $\operatorname{Kern}((A - \lambda E)^{l-1})$ enthält $v_2, v_3, \ldots, w_2, w_3, \ldots$ usw. Ergänze eine Basis von $\operatorname{Kern}(A - \lambda E)^{l-2} \oplus \langle v_2, w_2, \ldots \rangle$ zu einer Basis von $\operatorname{Kern}(A - \lambda E)^{l-1}$ und fahre wie zuvor fort, wobei wie in diesem Schritt bereits verwendete Vektoren ausgeschlossen werden müssen.

Beispiel 2.11. Im Beispiel

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

haben wir oben bereits eine Normalform bestimmt. Beginne mit $l = 2$. Ergänze eine Basis von $\operatorname{Kern}(A - \lambda E)$ durch v_1 zu einer Basis von $\operatorname{Kern}((A - \lambda E)^2)$. Wähle z. B. $v_1 = e_1 \notin \operatorname{Kern}(A - \lambda E)$. Wir berechnen $v_2 = (A - \lambda E)v_1 = (1, 1, -1)^T$ und machen ggf. eine Proberechnung ($(A - \lambda E)v_2 = 0$, also $Av_2 = \lambda v_2$). Damit haben wir Basisvektoren für den 2×2-Block in der Matrix gefunden. Für den verbleibenden 1×1-Block ergänzen wir v_2 zu einer Basis von $\operatorname{Kern}(A - \lambda E)$, z. B. durch $x_1 = (0, 3, -2)^T$. Probe: $Ax_1 = x_1$, also ist x_1 ein Eigenvektor und erzeugt einen 1×1-Block. Die obere Normalformdarstellung ergibt sich aus der Basisanordnung (v_2, v_1, x_1).

Noch zu begründen: Jordanblöcke sind unzerlegbar. Für $\lambda = 0$ ist ein Jordan-Block nilpotent und dann folgt mit Proposition 2.17, dass U genau dann unzerlegbar ist, wenn die Ordnung von α gleich der Dimension von U ist, wobei die Ordnung von α genau der Grad des Minimalpolynoms ist.

Proposition 2.22. *Sei $\alpha : U \to U$ linear und m_α ein Produkt von Linearfaktoren. Dann ist U bezüglich α unzerlegbar genau dann, wenn $m_\alpha(t) = (t - \lambda)^{\dim U}$ für ein λ. (Daraus folgt: Jordanblöcke sind unzerlegbar.)*

Beweis. Falls α mindestens zwei Eigenwerte hat, d. h. $m_\alpha(t) = (t - \lambda)^a (t - \mu)^b \cdots$, dann sagt die Hauptraumzerlegung, dass $U = K((t - \lambda)^a) \oplus K((t - \mu)^b) \oplus \ldots$, also ist U zerlegbar. Es bleibt der Fall, dass genau ein Eigenwert λ existiert, d. h. $m_\alpha(t) = (t - \lambda)^a$. Sei $\beta := \alpha - \lambda \operatorname{id}$, dann ist $\beta^a = 0$, d. h. β ist nilpotent. Nach 2.17 ist U unzerlegbar bezüglich β genau dann, wenn $a = \dim U$. Eine Zerlegung bezüglich β ist auch eine Zerlegung bezüglich α und umgekehrt. Also überträgt sich die Aussage auf α. $\square$

Proposition 2.23. *Sei $\alpha : V \to V$ linear und m_α ein Produkt von Linearfaktoren. Dann sind äquivalent:*

(1) α ist diagonalisierbar.

(2) $V = \bigoplus_{j=1}^{n} U_j$ ist eine direkte Summe von eindimensionalen α-invarianten Unterräumen ($n = \dim V$).

(3) $m_\alpha(t) = (t - \lambda_1) \cdots (t - \lambda_l)$ für ein $l \leq \dim V = n$ mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_l$. (Also kann man Diagonalisierbarkeit direkt an m_α ablesen.)

Beweis.

(1) $\Rightarrow$ (2), (3): Ist α diagonalisierbar, dann existiert eine Basis $\{v_1, \ldots, v_n\}$ von V aus Eigenvektoren von α, also $\alpha(v_j) = \mu_j v_j$. Setze $U_j \coloneqq \operatorname{Span}\{v_j\}$. U_j ist α-invariant. Da die v_j eine Basis bilden, ist $V = U_1 \oplus \ldots \oplus U_n$.

Für das Minimalpolynom betrachten wir eine Hauptraumzerlegung $V = W_1 \oplus \ldots \oplus W_l$, wobei W_j eine Basis aus Eigenvektoren mit demselben Eigenwert μ_{i_j} hat. Dann ist $m_{\alpha|W_j} = (t - \mu_{i_j})$. Es folgt, dass das Minimalpolynom der Matrix das Produkt dieser einzelnen Minimalpolynome teilt, also die Form wie in (3) haben muss.

(2) $\Rightarrow$ (1), (3): U_j hat eine Basis v_j. Da U_j α-invariant ist, ist $\alpha(v_j) \in U_j$, also gibt es ein μ_j mit $\alpha(v_j) = \mu_j v_j$, also bilden die v_j eine Basis von α aus Eigenvektoren. Daraus folgt (1) und somit (3).

(3) $\Rightarrow$ (1), (2): Da m_α ein Produkt von Linearfaktoren ist, hat α eine Jordan-Normalform. Da ein Jordan-Block zum Eigenwert λ der Größe $a \times a$ das Minimalpolynom $(t - \lambda)^a$ hat, muss jeder Block in der Jordan-Normalform ein 1×1-Block sein und damit folgen (1) und (2). □

Die Jordan-Normalform existiert nur, wenn m_α ein Produkt von Linearfaktoren ist. Was tun wir, wenn das nicht erfüllt ist? Gegeben $\alpha \in \operatorname{End}_K(V)$.

- Bestimme $m_\alpha(t)$.
- Zerlege $m_\alpha(t) = f_1 \cdots f_r$, wobei die f_j paarweise teilerfremd und irreduzibel sind (die Zerlegung ist bestmöglich). Damit funktioniert eine Hauptraumzerlegung $V = W_1 \oplus \ldots \oplus W_r$ mit $W_i = K(f_i)$. Ist $r \geq 2$, dann können wir auf den einzelnen Haupträumen weitermachen. Es bleibt der Fall, dass m_α nicht in teilerfremde Faktoren faktorisierbar ist (es also nur einen „Hauptraum“ gibt). Sei $m_\alpha(t) = t^n + a_{n-1}t^{n-1} + \ldots + a_0$. Dann gilt

$$0 = m_\alpha(\alpha) = \alpha^n + a_{n-1}\alpha^{n-1} + \ldots + a_0 \operatorname{id} \ .$$

Man kann zeigen, dass es ein $x \in V$ gibt, sodass $x, \alpha(x), \ldots, \alpha^{n-1}(x)$ linear unabhängig sind. $U \coloneqq \langle x, \alpha(x), \ldots, \alpha^{n-1}(x) \rangle$ hat die Dimension n.

U ist α-invariant, denn $\alpha(\alpha^{n-1}(x)) = \alpha^n(x) = -a_{n-1}\alpha^{n-1}(x) - \ldots - \alpha_0 x$. Man kann auch zeigen, es ein zu U invariantes Komplement gibt.

Betrachte $\alpha : U \to U$. Bezüglich der Basis $x, \alpha(x), \ldots, \alpha^{n-1}(x)$ hat α die darstellende Matrix

$$\begin{pmatrix} & & & -a_0 \\ 1 & & & -a_1 \\ & \ddots & & \vdots \\ & & 1 & -a_{n-1} \end{pmatrix} .$$

Die letzte Spalte enthält die negierten Koeffizienten von m_α. Induktiv kann man weitere Blöcke bestimmen. Dies liefert die *Rationale Normalform*. Diese existiert immer.

Über $\mathbb{C}$ geht immer die Jordan-Normalform, über $\mathbb{R}$ dagegen nicht. Über $\mathbb{R}$ ist die Situation trotzdem besser: Jedes Polynom über $\mathbb{R}$ ist ein Produkt aus linearen und quadratischen Faktoren. Begründung: Über $\mathbb{C}$ existiert immer ein Linearfaktor $(t-z)$, also ist z und somit auch $\overline{z}$ eine Nullstelle des reellen Polynoms und $(t-z)(t-\overline{z})$ ist eine reelles quadratisches Polynom. D. h. nach der „Hauptraumzerlegung“ $V = K(f_1) \oplus \ldots \oplus K(f_r)$ bleiben Fälle der Form $f_j = (t^2 + at + b)^l$ mit $a^2 - 4b < 0$ (keine reelle Nullstelle).

Spezialfall: $l = 1,\ 0 = m_\alpha(\alpha) = \alpha^2 + a\alpha + b\,\mathrm{id}$. Wähle $v_0 \neq 0,\ v_1 := \alpha(v_0)$. Dann hat $U := \langle v_0, v_1 \rangle$ die Dimension 2 (sonst wäre $v_1 = \lambda v_0$, also $(t-\lambda) \mid m_\alpha(t)$). Wie bei der rationalen Normalform folgt, dass U α-invariant ist und es entsteht ein Block der Form

$$\begin{pmatrix} 0 & -b \\ 1 & -a \end{pmatrix} .$$

Allgemeiner Fall: $l > 1$. Kopiere das Vorgehen bei nilpotenten Matrizen: Sei $\beta := \alpha^2 + a\alpha + b\,\mathrm{id}$. β ist nilpotent: $\beta^l = 0$, denn $m_\alpha(t) = (t^2 + at + b)^l$. Wähle einen Basisvektor v_0 mit $v_0, \beta(v_0), \ldots, \beta^{l-1}(v_0)$ linear unabhängig. Bei β liefert das einen Block der Form

$$\begin{pmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix} .$$

Es gilt $\mathrm{Grad}(m_\beta) = l$ und $\mathrm{Grad}(m_\alpha) = 2l$. Also sollte der größte Block bei α die Größe $(2l) \times (2l)$ haben.

Wähle jetzt v_0 so, dass $v_0,\ v_2 = \beta(v_0), \ldots, v_{2(l-1)} = \beta^{l-1}(v_0)$ linear unabhängig sind. Man kann zeigen, dass diese Vektoren nach Hinzufügen von Vektoren $v_{2j+1} \coloneqq \alpha(v_{2j}), j \in \{0, \ldots, l-1\}$ immer noch linear unabhängig sind. Es folgt für alle j: $\alpha(v_{2j}) = v_{2j+1}$ und $\alpha(v_{2j+1}) = \alpha(\alpha(\beta^j(v_0))) = \alpha^2(\beta^j(v_0)) = \beta^{j+1}(v_0) - a\alpha\beta^j(v_0) - b\beta^j(v_0) = v_{2j+2} - av_{2j+1} - bv_{2j}$. Der sich ergebende Block sieht wie folgt aus:

$$\begin{pmatrix} 0 & -b & & & \\ 1 & -a & & & \\ 0 & 1 & 0 & -b & \\ & & 1 & -a & \\ & & 0 & 1 & \ddots \\ & & & & \ddots \end{pmatrix}.$$

Diagonale: $0, -a, 0, -a, \ldots$. Darüber: $-b, 0, -b, 0, \ldots$. Darunter: $1, 1, 1, \ldots$. Diese „reelle Jordan-Normalform“ ist eine spezielle Form der rationalen Normalform. Es kann weitere Blöcke dieser Form geben (induktiv zu bestimmen).

Kapitel 3

Affine Räume

Erster Schritt der Anwendung linearer Algebra auf Geometrie. Die Algebraisierung der klassischen Geometrie (Euklid, 3. Jh. v. Chr.) wurde durch Descartes (1596 – 1650) vorangetrieben. Euklid: synthetische Geometrie durch Axiome, aus denen Aussagen abgeleitet werden (mit Intelligenz, nicht mit Rechnung). Zugänge zur Geometrie:

- Synthetische Geometrie (Euklid): Axiome, keine Hilfsmittel außer Logik.
- Analytische Geometrie (Descartes): Hilfsmittel lineare Algebra.
- Differentialgeometrie: Hilfsmittel Analysis.
- Algebraische Geometrie: Hilfsmittel Algebra, vor allem Polynome.

Beispiel homogene LGS: Ist $V = K^n$, dann ist U genau dann ein Untervektorraum von V, wenn U die Lösungsmenge eines homogenen LGS ist. Eine Gleichung $a_{j1}x_1 + \ldots + a_{jn}x_n = 0$ hat als Lösungsmenge entweder ganz K^n $(0 = 0)$ oder $W \cong K^{n-1}$ („Hyperebene"). Die Lösungsmenge eines homogenen LGS ist K^n oder ein Schnitt von Hyperebenen. Aber nicht jede Gerade, Ebene, $\ldots$, Hyperebene ist ein Untervektorraum, z. B. weil 0 nicht enthalten sein muss. Affine Räume erlauben eine Beschreibung aller Geraden, Ebenen, $\ldots$. Beispiel: Geraden in einer Ebene K^2 haben verschiedene Beschreibungen:

- Parameterdarstellung: $(x, y)^T = (a_1, a_2)^T + \lambda(r_1, r_2)^T$, $\lambda \in K$. Hierbei muss der Richtungsvektor $(r_1, r_2)^T$ ungleich dem Nullvektor sein.

- Gleichung in x, y: $x = a_1 + \lambda r_1, y = a_2 + \lambda r_2$. Nach Elimination von λ ergibt sich $(x - a_1)r_2 = (y - a_2)r_1 \Leftrightarrow r_2 x - r_1 y = a_1 r_2 - a_2 r_1$.

- Zwei-Punkte-Form: Gerade gegeben durch zwei verschiedene Punkte $(a_1, a_2)^T$ und $(b_1, b_2)^T$.

$$\rightsquigarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \lambda \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \end{pmatrix} .$$

Die Gerade g erfüllt eine homogene Gleichung $(a_1 r_2 - a_2 r_1 = 0) \Leftrightarrow (0,0)^T \in g \Leftrightarrow g$ ist ein Untervektorraum. Affine Geometrie: Erweiterung zu inhomogenen LGS, allen Geraden ...

Definition 3.1. Sei K ein Körper. Eine nichtleere Menge A heißt *affiner Raum* über K bezüglich einem K-Vektorraum $V \Leftrightarrow \exists$ Abbildung $+ : V \times A \to A, (x, p) \mapsto p + x$, sodass gilt

(A1) $\forall x, y \in V : \forall p \in A : p + (x + y) = (p + x) + y$.

(A2) $\forall p \in A : p + 0 = p$.

(A3) $\forall p \in A, x \in V : p + x = p \Rightarrow x = 0$.

(A4) $\forall p, q \in A : \exists x \in V : q = p + x$.

Die Elemente von A heißen *Punkte*. Die leere Menge ist auch ein affiner Raum.

Achtung: In einem affinen Raum gibt es keine Addition oder Multiplikation von Punkten.

Beispiel 3.1. Beispiele von affinen Räumen:

- Gegeben K und ein K-Vektorraum V, setze $A = V$. Definiere die Addition von Punkten mit Vektoren durch die normale Vektoraddition in V. Verifiziere die Axiome:

  (A1) $p + (x + y) = (p + x) + y$ gilt wegen der Assoziativität der Addition von Vektoren.

(A2) $p + 0 = p$, weil 0 das neutrale Element bzgl. Addition in V ist.

(A3) $p + x = p \Rightarrow x = -p + (p + x) = -p + p = 0$.

(A4) $q = (p+(-p))+q = p+(q-p)$, d. h. wir können $x = q-p$ wählen.

- Beispiel mit $A \neq V$: Gerade g im $\mathbb{R}^2$: Seien $a, r \in \mathbb{R}^2$ mit $r \neq 0$. Sei $A = g = \{a + \lambda r \mid \lambda \in \mathbb{R}\}$. Definiere $K = \mathbb{R}, V = \mathbb{R}$ und die Addition von Punkten mit Vektoren durch

$$+ : V \times A \to A, (x, a + \lambda r) \mapsto a + (\lambda + x)r \ ,$$

d. h. $+$ bewirkt eine Verschiebung um das $(\lambda + x)$-fache des Richtungsvektors. Verifiziere die Axiome mit $p = a + \lambda r, q = a + \mu r$:

(A1) $p+(x+y) = a+(\lambda+(x+y))r = a+((\lambda+x)+y)r = (a+(\lambda+x)r)+y = (p + x) + y$.

(A2) $p + 0 = a + (\lambda + 0)r = a + \lambda r = p$.

(A3) $p + x = p \Rightarrow a + \lambda r = a + (\lambda + x)r \Rightarrow xr = 0 \Rightarrow x = 0$, da $r \neq 0$ nach Voraussetzung.

(A4) Wähle $x = \mu - \lambda$. Dann ist $q = a + \mu r = a + (\lambda + x)r = p + x$.

Also ist jede Gerade ein affiner Raum bezüglich $V = K$.

Sei V ein K-Vektorraum, $U < V$ ein Untervektorraum und $W := V/U$ der Quotientenvektorraum, dann ist $W = \{v+U \mid v \in V\}, v+U = \{v+u \mid u \in U\}$. Ein Beispiel ist $V = \mathbb{R}^3, U = \langle e_1, e_2\rangle$, dann ist U die x-y-Ebene und die Elemente aus V/U sind zu U parallele Ebenen.

Sei $v + U$ ein Element in W. $v + U$ ist ein Vektorraum genau dann, wenn $v \in U$. $v + U$ ist immer ein affiner Raum bezüglich U: Sei $A = v + U$, dann definieren wir die Abbildung $+ : A \times U \to A, (v + u, x) \mapsto v + u + x \in v + U$, d. h. die Addition von Punkten und Vektoren wird durch die Addition in V definiert. Wir überprüfen die Axiome für $p = v + u_1, q = v + u_2$:

(A1) $p + (x + y) = v + u_1 + (x + y) = (v + u_1 + x) + y = (p + x) + y$.

(A2) $p + 0 = p$.

(A3) $p + x = p \Rightarrow v + u_1 + x = v + u_1 \Rightarrow x = 0$.

(A4) Wähle $x = q - p = v + u_2 - (v + u_1) = u_2 - u_1 \in U$. Dann ist $q = v + u_2 = v + u_1 + (u_2 - u_1) = p + x$.

Vorstellung: $p + x$ ist der Punkt p verschoben um den Vektor x. Also definiert der Vektor x eine Abbildung von A nach A.

Sei A ein affiner Raum bezüglich dem Vektorraum V. Definiere für $x \in V$ die Abbildung $t_x : A \to A$ (Translation um x) durch $t_x(p) = p + x$. Für $x = 0$ ergibt sich nach (A2) die identische Abbildung $t_0(p) = p + 0 = p = \mathrm{id}(p)$. Für $x \in V$ ist t_x immer eine bijektive Abbildung von A nach A, denn $(t_{-x} \circ t_x)(p) = p + x + (-x) = p + (x - x) = p + 0 = p$ und somit $t_{-x} \circ t_x = \mathrm{id}$. Damit folgt $t_x \circ t_{-x} = t_{-(-x)} \circ t_{-x} = \mathrm{id}$ nach dem bereits Gezeigten.

Proposition 3.2. *Sei A ein affiner Raum über K bezüglich V. Dann gilt:*

(a) Jedes $x \in V$ definiert eine bijektive Abbildung $t_x : A \to A$ mit $t_x(p) = p + x$ (Verschiebung um x).

(b) Für $x \neq y$ ist $t_x \neq t_y$ und sogar $t_x(p) \neq t_y(p)$ für alle $p \in A$, d. h. die Abbildung $V \to \mathrm{Bij}(A, A) := \{\textit{bijektive Abbildungen } A \to A\}, x \mapsto t_x$ ist injektiv. Insbesondere folgt aus $t_x(p) = p$ für ein $p \in A$, dass $x = 0$. Für alle $p, q \in A$ existiert ein $x \in V$ mit $t_x(p) = q$, dieses x ist sogar eindeutig.

Beweis.

(a) Diese Aussage wurde bereits oben bewiesen.

(b) Falls $\exists p : t_x(p) = t_y(p)$, dann folgt $p = p + x - x = t_{-x}(t_x(p)) = t_{-x}(t_y(p)) = p + (y - x)$ und somit nach (A3), dass $y - x = 0$. Wegen $t_0 = \mathrm{id}$ folgt $t_x(p) = p \Rightarrow x = 0$ aus der vorherigen Aussage. Gegeben $p, q \in A$, dann folgt mit (A4), dass $\exists x : q = p + x = t_x(p)$. □

Der Vektor x mit $p + x = q$ ist eindeutig bestimmt und heißt Verbindungsvektor zwischen p und q. Bezeichnung: $x = \overrightarrow{pq}$, also $p + \overrightarrow{pq} = q$.

Lemma 3.3. *$\forall p, q, r \in A$ gilt:*

(a) $\overrightarrow{pq} + \overrightarrow{qr} = \overrightarrow{pr}$.

(b) $\overrightarrow{pp} = 0$.

(c) $\overrightarrow{pq} = -\overrightarrow{qp}$.

Beweis.

(a) Seien $x := \overrightarrow{pq} + \overrightarrow{qr}, y := \overrightarrow{pr}$. Nach Proposition 3.2 genügt es, zu zeigen, dass $\exists a \in A : t_x(a) = t_y(a)$. Mit der Wahl $a = p$ folgt $t_y(p) = p + y = p + \overrightarrow{pr} = r$ und $t_x(p) = p + x = p + (\overrightarrow{pq} + \overrightarrow{qr}) = (p + \overrightarrow{pq}) + \overrightarrow{qr} = q + \overrightarrow{qr} = r$.

(b) Wähle $a = p, x = \overrightarrow{pp}, y = 0 \Rightarrow t_x(p) = p + \overrightarrow{pp} = p, t_y(p) = p + 0 = p$. Es folgt $x = y$.

(c) Nach (a) und (b) gilt $\overrightarrow{pq} + \overrightarrow{qp} = \overrightarrow{pp} = 0$. Dies zeigt die Behauptung. □

Theorem 3.4. *Seien $A, B \neq \emptyset$ affine Räume bezüglich demselben Vektorraum V. Dann gibt es eine Bijektion $f : A \to B$ mit $f(p + x) = f(p) + x$ für alle $p \in A, x \in V$. Spezialfall: A affiner Raum bezüglich V, $B = V$ als affiner Raum $\Rightarrow \exists$ Bijektion $f : A \to V$. D. h. Vektorräume sind „Standardmodelle“ von affinen Räumen.*

Beweis. Gegeben A, B. Zu definieren: $f : A \to B$. Wähle irgendwelche Punkte $a \in A, b \in B$. Setze $f(a) = b$. Definiere $f(a + x) := b + x$. Das ist wohldefiniert, da das x eindeutig ist. Anders gesagt: Für $a' \in A$ schreibe $a' = a + \overrightarrow{aa'} = a + x, x = \overrightarrow{aa'}$. Dann ist $f(a') = f(a) + \overrightarrow{aa'} = b + x$. Zu zeigen ist, dass f bijektiv ist und $f(p + x) = f(p) + x$ für alle $p \in A, x \in V$ (für $p = a$ folgt dies nach Definition von f). Sei $p \in A, x \in V$. Berechne $f(p)$ und $f(p + x)$:

$$\begin{aligned} p = a + \overrightarrow{ap} \Rightarrow f(p) &= f(a) + \overrightarrow{ap} = b + \overrightarrow{ap} \\ p + x = a + \overrightarrow{a(p+x)} = a + \overrightarrow{ap} + x \Rightarrow f(p + x) &= f(a) + (\overrightarrow{ap} + x) \\ &= (b + \overrightarrow{ap}) + x \\ &= f(p) + x\ . \end{aligned}$$

Um zu zeigen, dass f surjektiv ist: Sei $b' \in B$. Dann ist $b' = b + x, x := \overrightarrow{bb'} \in V$. Dann ist $f(a + x) = b + x = b'$. Folglich ist f surjektiv. Jetzt zeigen wir, dass f injektiv ist: Sei $f(p) = f(q)$ für $p, q \in A$. Es folgt $b + \overrightarrow{ap} = f(a + \overrightarrow{ap}) = f(p) = f(q) = f(a + \overrightarrow{aq}) = b + \overrightarrow{aq}$. Daraus folgt $\overrightarrow{ap} = \overrightarrow{aq}$ und somit $p = a + \overrightarrow{ap} = a + \overrightarrow{aq} = q$. Somit ist f eine Bijektion. □

Beispiel (inhomogene) LGS: $Mx = 0$ (homogen) hat als Lösungsmenge einen Untervektorraum. Allgemein: $Mx = b$ mit irgendeinem Vektor $b \neq 0$ impliziert $x \neq 0$, d. h. die Lösungsmenge L ist kein Vektorraum. $L = \emptyset$ kann vorkommen (z. B. $M = 0, b \neq 0$). Falls $L \neq \emptyset$: $\exists$ Lösung x_1: Weitere Lösung x_2: $Mx_1 = b$ und $Mx_2 = b \Rightarrow M(x_1 - x_2) = 0 \rightarrow x_1 - x_2$ ist Lösung des homogenen Systems. Sei U die Lösungsmenge des homogenen Systems. Ist $y \in U$, dann ist $x_1 + y \in L$, dann $Mx_1 = b, My = 0 \Rightarrow M(x_1 + y) = b$. Es folgt $L = x_1 + U = x_2 + U$ (spezielle inhomogene Lösung + allgemeine homogene Lösung). Bezüglich dem Vektorraum U ist L ein affiner Raum. Ist $A_1 = x_1 + U, A_2 = x_2 + U$, dann können wir f aus Theorem 3.4 wählen als $f(x_1 + u) = x_2 + u$ mit $u \in U$. Für $x_1 \neq x_2$ ist $f \neq \mathrm{id}$.

Wie unterscheidet man $L = \emptyset$ und $L \neq \emptyset$, wenn man M und b gegeben hat ($\exists x : Mx = b$)? Dies ist äquivalent dazu, ob b im Bild der linearen Abbildung $x \mapsto Mx$ liegt. Das Bild von M ist erzeugt von den Bildern der Basisvektoren, d. h. von den Spalten von M. Es gilt also $L = \emptyset$ genau dann, wenn b keine Linearkombination der Spalten von M ist. In diesem Fall ist $\mathrm{Rang}(M|b) > \mathrm{Rang}(M)$, ansonsten ist $\mathrm{Rang}(M|b) = \mathrm{Rang}(M)$.

Theorem 3.5. *Sei $Mx = b$ ein inhomogenes LGS und $\varphi(x) := Mx$. Dann gilt entweder (1) $L = \emptyset$ und $b \notin \mathrm{Bild}(\varphi)$ oder (2) $L = x_0 + \mathrm{Kern}(\varphi)$ für ein x_0 und $b \in \mathrm{Bild}(\varphi)$. Sei $(M|b)$ die aus den Spalten von M sowie b gebildete* erweiterte Matrix *des LGS. Dann ist (1) $\Leftrightarrow \mathrm{Rang}(M) < \mathrm{Rang}(M|b)$ und (2) $\Leftrightarrow \mathrm{Rang}(M) = \mathrm{Rang}(M|b)$. In beiden Fällen ist L ein affiner Raum. Im Fall (2) ist der zugehörige Vektorraum $\mathrm{Kern}(\varphi) =$ Lösungsmenge des homogenen Systems.*

Definition 3.6. Sei A ein affiner Raum über K bezüglich V und $\emptyset \neq A_0 \subseteq A$ eine nichtleere Teilmenge. A_0 heißt *affiner Teilraum* $:\Leftrightarrow \exists p \in A_0$ und einen Untervektorraum $U < V$, sodass $A_0 = p + U = \{p + u \mid u \in U\}$. $A_0 = \emptyset$ ist auch ein affiner Teilraum.

Fragen: Ist A_0 ein affiner Raum? Ist p durch A_0 eindeutig bestimmt? Ist U durch A_0 eindeutig bestimmt? Wann sind zwei affine Teilräume A_1 und A_2 gleich?

Beispiel: Sei $f : V \to W$ eine lineare Abbildung zwischen K-Vektorräumen. Sei $w \in W$. Die Menge $A_0 := f^{-1}(w) := \{v \in V : f(v) = w\}$ der Urbilder von w ist eine Teilmenge von $A = V$. $A_0 = \emptyset$ kann vorkommen. Falls $A_0 \neq \emptyset : \exists p \in A_0$ mit $f(p) = w$. Falls $q \in A_0 : f(p) = f(q) \Rightarrow f(p-q) = 0 \Rightarrow p - q \in \text{Kern}(f)$. Falls $x \in \text{Kern}(f) : f(x) = 0 \Rightarrow f(p+x) = f(p) + f(x) = w + 0 \Rightarrow p + x \in A_0$. Somit ist $A_0 = p + \text{Kern}(f)$. Mit dem Untervektorraum $U \coloneqq \text{Kern}(f)$ folgt, dass A_0 ein affiner Teilraum von A ist.

Falls $f(x) = Mx$ mit einer Matrix M, d. h. $V = K^n$ und $W = K^m$ sind endlich-dimensional, ist das genau das Beispiel der inhomogenen LGS $Mx = w$.

Lemma 3.7. *Sei A ein affiner Raum über K bezüglich V und A_0 ein Teilraum von A.*

(a) Sei $A_0 = p + U$ und $q \in A$. Dann ist $A_0 = q + U \Leftrightarrow q \in p + U = A_0$. (D. h. man kann U an jedem Punkt in A_0 „befestigen“, um A_0 zu beschreiben.)

(b) Für jedes $p \in A_0$ gilt: $U = \{\overrightarrow{pq} : q \in A_0\}$. U ist also durch A_0 eindeutig bestimmt als Vektorraum der Verbindungsvektoren von einem festen Punkt zu allen Punkten.

(c) Aus $x \in U, q \in A, q + x \in A_0$ folgt $q \in A_0$.

(d) Für zwei affine Teilräume $A_1 = p + U_1, A_2 = q + U_2$ von A gilt: $A_1 = A_2 \Leftrightarrow U_1 = U_2$ und $p \in A_2$.

(e) Der affine Teilraum $A_0 = p + U$ ist ein affiner Raum über K bezüglich U.

Beweis.

(a) $q \in A_0$, z. Z. $A_0 = q + U$. $q \in A_0 \Rightarrow q = p + x, x \in U, x = \overrightarrow{pq}$. $q + U = \{q + y \mid y \in U\} = \{p + x + y \mid y \in U\} = \{p + z \mid z \in U\} = p + U$, denn $x + U = U$ wegen $x \in U$. Umgekehrt gilt $A_0 = q + U \Rightarrow q = q + 0 \in A_0$.

(b) Sei $p \in A_0, x \in U$, dann ist zu zeigen, dass $x = \overrightarrow{pq}$ für ein $q \in A_0$. $x \in U$ bedeutet $p + x \in A_0$, also $q = p + x \Rightarrow x = \overrightarrow{pq}$.

(c) Sei $x \in U, q \in A, q + x \in A_0$, zz $q \in A_0$. $x \in U \Rightarrow -x \in U \Rightarrow q = (q + x) - x \in A_0$.

(d) Sei $A_1 = A_2$. Nach (b) folgt $U_1 = U_2$. Außerdem gilt dann $p \in A_1 = A_2$. Umgekehrt: Sei $U_1 = U_2$ und $p \in A_2$, dann folgt $A_1 = A_2$ aus (a).

(e) Zu zeigen ist, dass A_0 ein affiner Raum ist. Nach Voraussetzung ist A ein affiner Raum bezüglich $V > U$ mit der Abbildung $+ : V \times A \to A$. Falls $x \in U, p \in A_0 \Rightarrow p + x \in A_0$. Also können wir die Abbildung $+$ einschränken und erhalten eine Abbildung $+ : U \times A_0 \to A_0$. Die Axiome übertragen sich von A auf A_0. □

Definition 3.8. Sei $A \neq \emptyset$ ein affiner Raum über K bezüglich V. Die *Dimension* von A ist $\dim(A) \coloneqq \dim(V)$. Bezeichnungen:

- A heißt *Punkt*, falls $\dim A = 0$.
- A heißt *Gerade*, falls $\dim A = 1$.
- A heißt *Ebene*, falls $\dim A = 2$:

Ist $A_0 \subseteq A$ ein affiner Teilraum von A mit $\dim A_0 = \dim A - 1$, dann heißt A_0 *Hyperebene*. Wir definieren $\dim \emptyset \coloneqq -1$.

Beispiel 3.2. Geraden in einer Ebene: $\dim A = 2, g = A_1, h = A_2, \dim A_1 = \dim A_2 = 1$. Wir können uns die folgenden Fälle vorstellen:

- g und h schneiden sich in einem Punkt, also $g \cap h$ ist ein Punkt (Dimension 0) und $U_1 \cap U_2 = \{0\}$.
- $g = h$, also $\dim(g \cap h) = 1$ und es gilt $U_1 = U_2$.
- g und h sind parallel und nicht identisch, also $g \cap h = \emptyset$, $\dim(g \cap h) = -1$ und $U_1 = U_2$.

Sei A ein affiner Raum über K bezüglich V. Seien A_j $(j \in J)$ affine Teilräume von A. $D \coloneqq \cap_{j \in J} A_j$ ist ein affiner Teilraum von A: Falls $D = \emptyset$ ist die Aussage klar. Ansonsten sei $p \in D$, also $\forall j : p \in A_j \Rightarrow A_j = p + U_j, U_j < V$. Somit ist $D = p + \bigcap_{j \in J} U_j$ und da $\bigcap_{j \in J} U_j$ ein Untervektorraum ist, ist D ein affiner Teilraum.

$\bigcup_{j \in J} A_j$ ist im Allgemeinen kein affiner Raum, genauso wie die Vereinigung von Untervektorräumen im Allgemeinen kein Untervektorraum ist.

Die Summe von Untervektorräumen ist aber ein Untervektorraum. Dimensionsformel: $\dim(U_1 + U_2) + \dim(U_1 \cap U_2) = \dim U_1 + \dim U_2$. Frage: Gibt es ein Analogon für affine Teilräume?

Definition 3.9. Sei $M \subseteq A$ eine Teilmenge. Dann ist die *affine Hülle* von M definiert als

$$[M] \coloneqq \bigcap_{\substack{A_j \supseteq M \\ A_j \text{ affiner Teilraum von A}}} A_j \ .$$

(Vorstellung: $[M]$ ist der kleinste affine Teilraum von A, der M enthält.)

Für affine Teilräume $B_j \quad (j \in J)$:

$$[B_j \mid j \in J] \coloneqq \left[\bigcup_{j \in J} B_j\right] \ .$$

Gemäß unserer Vorstellung aus dem obigen Beispiel ist bei zwei Geraden g, h in der Ebene $[g \cup h]$ eine Ebene, falls $g \neq h$ und $[g \cup h] = g = h$, falls $g = h$.

Sei $p \neq q$. Behauptung: $[\{p, q\}] = p + \{\lambda\overrightarrow{pq} \mid \lambda \in K\}$. Denn: Sei $A_1 \supseteq \{p, q\}$ ein affiner Teilraum von A, dann ist $A_1 = p + U_1$. Wegen $q \in A_1$ ist $\overrightarrow{pq} \in U_1 \Rightarrow \operatorname{Span}\{pq\} \subseteq U_1$. Somit folgt $A_2 \coloneqq \{p + \lambda\overrightarrow{pq} \mid \lambda \in K\} \subseteq A_1$ und A_2 ist selbst ein affiner Teilraum mit $p, q \in A_2$, also folgt die Behauptung.

Theorem 3.10 (Dimensionsformel für affine Teilräume)**.** *Seien $A_1 = p_1 + U_1$ und $A_2 = p_2 + U_2$ Teilräume des affinen Raums A. Dann gilt:*

$$\dim A_1 + \dim A_2 = \begin{cases} \dim[A_1 \cup A_2] + \dim(A_1 \cap A_2), & A_1 \cap A_2 \neq \emptyset \\ \dim[A_1 \cup A_2] + \dim(U_1 \cap U_2) - 1, & A_1 \cap A_2 = \emptyset \ . \end{cases}$$

Beweis. Erster Fall: $A_1 \cap A_2 \neq \emptyset \Rightarrow \exists p \in A_1 \cap A_2 \Rightarrow A_1 = p + U_1, A_2 = p + U_2$. Dann ist $A_1 \cap A_2 = p + (U_1 \cap U_2)$ und $[A_1 \cup A_2] = p + (U_1 + U_2)$. Damit folgt

$$\begin{aligned} \dim A_1 + \dim A_2 &= \dim U_1 + \dim U_2 = \dim(U_1 + U_2) + \dim(U_1 \cap U_2) \\ &= \dim[A_1 \cup A_2] + \dim(A_1 \cap A_2) \end{aligned}$$

mit der Dimensionsformel für Untervektorräume.

Zweiter Fall: $A_1 \cap A_2 = \emptyset$. Was ist $[A_1 \cup A_2]$? Idee: A_1, A_2 parallele Geraden $\Rightarrow$ Gerade durch p_1 und p_2 muss auch in der affinen Hülle enthalten sein. Seien $p_1, p_2 \in A_1 \cup A_2 \subseteq [A_1 \cup A_2] = q + U$, dann folgt $\overrightarrow{p_1p_2} \in U$. Es folgt $\lambda\overrightarrow{p_1p_2} \in U$ für alle $\lambda \in K$ und somit

$$W \coloneqq \{p_1 + \lambda\overrightarrow{p_1p_2} \mid \lambda \in K\} \subseteq [A_1 \cup A_2] \ .$$

Es gilt dann $\dim W = 1$ und $\overrightarrow{p_1p_2} \notin U_1 \cup U_2$.

Behauptungen:

(1) $(U_1 + U_2) \cap W = \{0\}$.

(2) $[A_1 \cup A_2] = p_1 + (U_1 + U_2 + W)$.

Aus (1) und (2) folgt die Formel:

$$\begin{aligned} \dim A_1 + \dim A_2 = \dim U_1 + \dim U_2 &= \dim(U_1 + U_2) + \dim(U_1 \cap U_2) \\ &= \dim[A_1 \cup A_2] - 1 + \dim(U_1 \cap U_2) \ . \end{aligned}$$

Beweis der Behauptungen:

(1) Falls $(U_1 + U_2) \cap W \neq \{0\} \Rightarrow U_1 + U_2 \cap W = W \Rightarrow W \subseteq U_1 + U_2$ (da $\dim W = 1$). In diesem Fall folgt $\overrightarrow{p_1p_2} = x + y$ mit $x \in U_1, y \in U_2$, also $p_1 + x = p_1 + \overrightarrow{p_1p_2} - y = p_2 - y \in A_1 \cap A_2$, was ein Widerspruch ist. Also $(U_1 + U_2) \cap W = \{0\}$.

(2) Es gilt $p_1, p_2 \in A_1 \cup A_2$. Sei $B \coloneqq p_1 + (U_1 + U_2 + W)$ ein affiner Teilraum. Dann gilt $A_1 = p_1 + U_1 \subseteq B$ und wegen $\overrightarrow{p_1p_2} \in W$ gilt $A_2 = p_2 + U_2 = p_1 + \overrightarrow{p_1p_2} + U_2 \subseteq B$. Somit folgt $A_1 \cup A_2 \subseteq B$, also $[A_1 \cup A_2] \subseteq B$.

Sei C ein affiner Teilraum mit $C \supseteq A_1 \cup A_2$, z. Z.: $C \supseteq B$. Sei $C = p_1 + \tilde{V}$, wobei $\tilde{V}$ ein UVR von V ist. Dann müssen wir zeigen, dass $\tilde{V} \supseteq U_1 + U_2 + W$. Wegen $A_1 = p_1 + U_1$ folgt $U_1 \subseteq \tilde{V}$. Wegen $p_1, p_2 \in C$ gilt $\overrightarrow{p_1p_2} \in \tilde{V}$, also $W \subseteq \tilde{V}$. Außerdem gilt $p_2 \in C$, also $C = p_2 + \tilde{V} \supseteq A_2 = p_2 + U_2 \Rightarrow U_2 \subseteq \tilde{V}$. Aus $U_1, U_2, W \subseteq \tilde{V}$ folgt aber auch $U_1 + U_2 + W \subseteq \tilde{V}$, woraus die Behauptung folgt. $\square$

Definition 3.11. Seien $A_1 = p_1 + U_1, A_2 = p_2 + U_2$ nichtleere affine Teilräume. A_1 und A_2 heißen *parallel* $:\Leftrightarrow U_1 \subseteq U_2$ oder $U_2 \subseteq U_1$. Bezeichnung $A_1 \parallel A_2$.

Beispiel 3.3.

- Ein Punkt $A_1 = p_1 + \{0\}$ ist parallel zu jedem nichtleeren affinen Teilraum A_2.

- Sei A eine Ebene und A_1, A_2 Geraden. Es gibt zwei Möglichkeiten: $U_1 = U_2$, d. h. $A_1 \parallel A_2$, oder $U_1 \neq U_2$, d. h. $A_1 \nparallel A_2$. Sei $U_1 \neq U_2$. Kann $A_1 \cap A_2 = \emptyset$ bei nicht parallelen Geraden vorkommen? Annahme: $A_1 \cap A_2 = \emptyset$. Zweiter Fall der Dimensionsformel:

$$\begin{aligned} 2 = 1 + 1 = \dim A_1 + \dim A_2 &= \dim[A_1 \cup A_2] + \dim(U_1 \cap U_2) - 1 \\ &= 2 + 0 - 1 = 1 \ , \end{aligned}$$

  was ein Widerspruch ist. Ergebnis: A_1 und A_2 sind parallel oder sie schneiden sich.

- Falls $\dim A = 3$, dann existieren windschiefe Geraden (nicht parallel, kein Schnittpunkt).

- Sei $\dim A = 3$. Seien A_1, A_2 Ebenen. $A_1 = p_1 + U_1, A_2 = p_2 + U_2$. Falls $A_1 \cap A_2 \neq \emptyset$:

$$4 = 2 + 2 = \dim A_1 + \dim A_2 = \dim[A_1 \cup A_2] + \dim(A_1 \cap A_2) \ .$$

  Die beiden Möglichkeiten sind:

  - $\dim[A_1 \cup A_2] = 2$ und $[A_1 \cup A_2] = p_1 + U_0$ mit $p_1 \in A_1$. Aus $A_1 = p_1 + U_1 \subseteq [A_1 \cup A_2] = p_1 + U_0$ folgt mit $\dim U_1 = \dim U_0$, dass $U_1 = U_0$ und somit $A_1 = [A_1 \cup A_2] \Rightarrow A_1 \supseteq A_2 \Rightarrow A_1 = A_2$.
  - $[A_1 \cup A_2] \supsetneq A_1, A_2$, dann folgt $\dim[A_1 \cup A_2] = 3$ und somit $\dim(A_1 \cap A_2) = 1$. Somit sind A_1 und A_2 zwei verschiedene Ebenen, die sich in einer Geraden schneiden.

  Falls $A_1 \cap A_2 = \emptyset$:

$$\begin{aligned} 4 = 2 + 2 = \dim A_1 + \dim A_2 &= \dim[A_1 \cup A_2] + \dim(U_1 \cap U_2) - 1 \\ &= 3 + \dim(U_1 \cap U_2) - 1 \ , \end{aligned}$$

denn aus $\dim[A_1 \cup A_2] = 2$ würde $A_1 = A_2$ folgen.
Daher gilt $\dim(U_1 \cap U_2) = 2 = \dim U_1 = \dim U_2$, also $U_1 = U_2$ und folglich $A_1 \parallel A_2$.

Ergebnis: Zwei Ebenen im Raum sind parallel (eventuell gleich) oder sie schneiden sich in einer Geraden.

- $\dim A = 3, \dim A_1 = 1, \dim A_2 = 2$.
  Falls $A_1 \cap A_2 \neq \emptyset$:

  $$3 = 1 + 2 = \dim A_1 + \dim A_2 = \dim[A_1 \cup A_2] + \dim(A_1 \cap A_2)\ .$$

  Es gilt entweder $A_1 \subseteq A_2$, und somit $A_1 \parallel A_2$, oder $A_1 \cap A_2$ ist ein Punkt.
  Falls $A_1 \cap A_2 = \emptyset$:

  $$\begin{aligned} 3 = 1 + 2 = \dim A_1 + \dim A_2 &= \dim[A_1 \cup A_2] + \dim(U_1 \cap U_2) - 1 \\ &= 3 + \dim(U_1 \cap U_2) - 1\ , \end{aligned}$$

  also folgt $\dim(U_1 \cap U_2) = 1 \Rightarrow U_1 \subseteq U_2 \Rightarrow A_1 \parallel A_2$.

- $\dim A = n \geq 1$, H Hyperebene (d. h. $\dim H = n - 1$), $A_0 \neq \emptyset$ affiner Teilraum.
  Falls $A_0 \cap H \neq \emptyset$:
  Seien $H = p + U_H$ und $A_0 = p_0 + U_0$. Dann gilt:

  $$0 + (n-1) \leq \dim A_0 + \dim H = \dim[A_0 \cup H] + \dim(A_0 \cap H)\ .$$

  Falls $\dim[A_0 \cup H] = n-1$, dann folgt $H = [A_0 \cup H] \Rightarrow A_0 \subseteq H, A_0 \cap H = A_0$. Falls $\dim[A_0 \cup H] = n$, dann folgt $\dim(A_0 \cap H) = \dim A_0 - 1$. Falls $A_0 \cap H = \emptyset$:

  $$\begin{aligned} 0 + (n-1) \leq \dim A_0 + \dim H &= \dim[A_0 \cup H] + \dim(U_0 \cap U_H) - 1 \\ &= n + \dim(U_0 \cap U_H) - 1\ , \end{aligned}$$

  also $\dim U_0 = \dim A_0 = \dim(U_0 \cap U_H) \Rightarrow U_0 \subseteq U_H \Rightarrow A_0 \parallel H$.

Gesucht: geometrische Beschreibung von affinen Teilräumen. Zunächst ein Begriff:

Definition 3.12. Ein Körper K hat *Charakteristik*

$$\operatorname{char}(K) := \begin{cases} 0, & \forall n \in \mathbb{N} : n \cdot 1_K := \sum_{i=1}^{n} 1_K \neq 0_K \\ \min\{n \in \mathbb{N} \mid n \cdot 1_K = 0_K\}, & \text{sonst.} \end{cases}$$

Beispiel 3.4. Für $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$: $\operatorname{char}(K) = 0$. Es gilt $\operatorname{char}(\mathbb{Z}/p\mathbb{Z}) = p$ für jede Primzahl p.

$\operatorname{char}(K) \neq 0 \Rightarrow \operatorname{char}(K)$ ist eine Primzahl: Falls $\operatorname{char}(K) = n = a \cdot b$ mit $a, b \in \mathbb{N}$, dann ist $(a \cdot 1_K) \cdot (b \cdot 1_K) = n \cdot 1_K = 0_K$. Dann muss $a \cdot 1_K$ oder $b \cdot 1_K$ gleich 0 sein. Da n die minimale solche Zahl ist, folgt dann $a = n$ oder $b = n$.

Theorem 3.13. *Sei A ein affiner Raum über K und* $\operatorname{char}(K) \neq 2$. *Sei $X \subseteq A$. Dann sind äquivalent:*

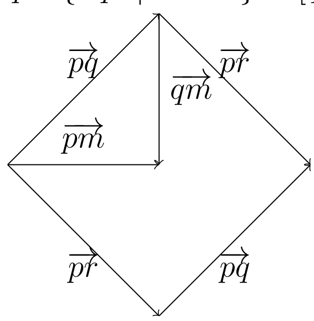
(a) *X ist ein affiner Teilraum von A.*

(b) *$\forall p, q \in X$ mit $p \neq q$ gilt $[p, q] \subseteq X$, d. h. X enthält mit zwei Punkten auch ihre Verbindungsgerade.*

Beweis.

$(a) \Rightarrow (b)$: Seien $p, q \in X$ und X ein affiner Teilraum, also $X = p + U$ mit einem Unterraum U. Dann gilt $\overrightarrow{pq} \in U$ und somit $p + \{\lambda\overrightarrow{pq} \mid \lambda \in K\} = [p, q] \subseteq X$.

$(b) \Rightarrow (a)$: Gegeben X. Sei $X \neq \emptyset$, also $\exists p \in X$. Ziel: $X = p + U$, U Untervektorraum von V. Falls $X = p + U$, dann muss $U = \{\overrightarrow{pq} \mid q \in X\}$ sein. Also definieren wir $U := \{\overrightarrow{pq} \mid q \in X\}$. Zu zeigen ist, dass U ein Untervektorraum ist. Es gilt $0 = \overrightarrow{pp} \in U$. Seien $q, r \in X$ mit $q \neq r, q \neq p, r \neq p$. Zu zeigen: $\overrightarrow{pq} + \overrightarrow{pr} \in U$. Wegen $q \neq r$ ist $q + \{\lambda\overrightarrow{qr} \mid \lambda \in K\} = [q, r] \subseteq X$.

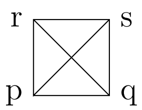


Sei $m \coloneqq q + \frac{1}{2}\overrightarrow{qr} \in X$ (Hier wird benötigt, dass 1/2 existiert, also $2 \neq 0$). Falls $m = p$, dann ist $\overrightarrow{pq} = -\overrightarrow{pr} \Rightarrow \overrightarrow{pq} + \overrightarrow{pr} = 0 \in U$. Andernfalls ist $\overrightarrow{pm} \in U$ und $p + \{\lambda\overrightarrow{pm} \mid \lambda \in K\} = [p, m] \subseteq X$. Insbesondere ist $z \coloneqq p + 2\overrightarrow{pm} \in X$, also $\overrightarrow{pz} \in U$. Wir zeigen: $\overrightarrow{pq} + \overrightarrow{pr} = \overrightarrow{pz} \in U$. Es gilt

$$\begin{aligned}\overrightarrow{pz} &= 2\overrightarrow{pm} = \overrightarrow{pm} + \overrightarrow{pm} = (\overrightarrow{pq} + \overrightarrow{qm}) + (\overrightarrow{pr} + \overrightarrow{rm}) \\ &= \overrightarrow{pq} + \overrightarrow{pr} + \underbrace{(\overrightarrow{qm} + \overrightarrow{rm})}_{=0} = \overrightarrow{pq} + \overrightarrow{qr}\ ,\end{aligned}$$

denn $m = q + \overrightarrow{qm} = q + \overrightarrow{qr} + \overrightarrow{rm} = q + 2\overrightarrow{qm} + \overrightarrow{rm} \Rightarrow \overrightarrow{qm} = -\overrightarrow{rm}$.
Sei $\overrightarrow{pq} \in U$, d. h. $q = p + \overrightarrow{pq} \in X$, z.Z.: $\lambda\overrightarrow{pq} \in U$ für $\lambda \in K$. Wegen $p, q \in X$ und $p \neq q$ folgt $p + \{\mu\overrightarrow{pq} \mid \mu \in K\} = [p, q] \subseteq X$. Wähle $\mu = \lambda \Rightarrow p + \lambda\overrightarrow{pq} \in X \Rightarrow \lambda\overrightarrow{pq} \in U$. $\square$

Für $K = \mathbb{Z}/2\mathbb{Z}$ gilt der Satz nicht. (Aber der Satz stimmt für K mit $|K| \geq 3$.) Hier enthält jede Gerade $p + \{\lambda\overrightarrow{pq} \mid \lambda \in K\}$ genau zwei Elemente: p für $\lambda = 0$ und q für $\lambda = 1$, denn $|K| = 2$. Sei $A = K^2 = \{p \coloneqq (0,0)^T, q \coloneqq (1,0)^T, r \coloneqq (0,1)^T, s \coloneqq (1,1)^T\}$ (vier Punkte). Zwei Punkte definieren eine Gerade, also gibt es $\binom{4}{2} = 6$ Geraden.



Es gilt $[p, q] \parallel [r, s]$, $[q, r] \parallel [p, s]$ mit Richtungsvektor $\overrightarrow{qr} = (0,1)^T - (1,0)^T = (1,1)^T = \overrightarrow{ps}$. Sei $X \subseteq A$. X erfüllt (b): $a, b \in X \Rightarrow [a, b] = \{a, b\} \subseteq X$. Falls der Satz hier richtig ist, muss jede Teilmenge $X \subseteq A$ ein affiner Teilraum sein. Das stimmt für $|X| \in \{0, 1, 2, 4\}$. Betrachte X mit $|X| = 3$. Falls X ein affiner Teilraum ist, ist $X = q + U$, wobei U ein Untervektorraum von K^2 ist. Also $|U| = |X| = 3$ und eine Fallunterscheidung nach $\dim U$ liefert einen Widerspruch. Konkretes Beispiel: $X = \{p, q, r\} = [p, q] \cup [p, r] \cup [q, r]$. Falls $X = p + U$: $0, \overrightarrow{pq}, \overrightarrow{pr} \in U$, aber $\overrightarrow{pq} + \overrightarrow{pr} = \overrightarrow{ps} \notin U$, denn $s \notin X$.

Definition 3.14. Sei K ein Körper, V, W K-Vektorräume und A, B affine Räume bezüglich V und W. Eine Abbildung $f : A \to B$ heißt *affin* (oder affine Abbildung) $:\Leftrightarrow \exists f_\star : V \to W$ linear, sodass $\forall p, q \in A : \overrightarrow{f(p)f(q)} = f_\star(\overrightarrow{pq})$. Eine bijektive affine Abbildung heißt *Affinität*.

Lemma 3.15.

(a) *$f : A \to B$ affin $\Leftrightarrow \exists p_0 \in A : \overrightarrow{p_0 q} \mapsto \overrightarrow{f(p_0)f(q)}$ ist eine lineare Abbildung $V \to W$.*

(b) *Sei $p_0 \in A, q_0 \in B, g : V \to W$ linear $\Rightarrow \exists f : A \to B$ affin mit $f(p_0) = q_0$ und $f_\star = g$. f ist eindeutig bestimmt. (D. h. eine affine Abbildung f ist bestimmt durch die Wahl eines Bildpunkts und von $f_\star$.)*

Beweis.

(a) „$\Rightarrow$“: klar.
„$\Leftarrow$“: $h : \overrightarrow{p_0 q} \mapsto \overrightarrow{f(p_0)f(q)}$ ist linear nach Voraussetzung. Seien $p, q \in A$ und $f_\star(\overrightarrow{pq}) = \overrightarrow{f(p)f(q)}$. Zu zeigen: $f_\star$ linear, wir zeigen $f_\star = h$. Es gilt $\overrightarrow{pq} = \overrightarrow{pp_0} + \overrightarrow{p_0 q} = -\overrightarrow{p_0 p} + \overrightarrow{p_0 q}$ und somit wegen der Linearität von h:

$$\begin{aligned} h(\overrightarrow{pq}) &= h(-\overrightarrow{p_0 p} + \overrightarrow{p_0 q}) = -h(\overrightarrow{p_0 p}) + h(\overrightarrow{p_0 q}) \\ &= -\overrightarrow{f(p_0)f(p)} + \overrightarrow{f(p_0)f(q)} = \overrightarrow{f(p)f(p_0)} + \overrightarrow{f(p_0)f(q)} \\ &= \overrightarrow{f(p)f(q)} = f_\star(\overrightarrow{pq}) \; . \end{aligned}$$

Somit ist $f_\star$ linear, weil h es ist.

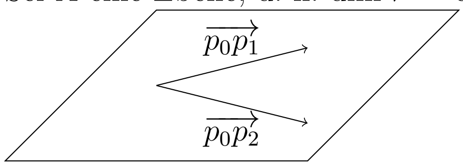
(b) Falls f existiert, muss für $p \in A$ gelten: $\overrightarrow{f(p_0)f(p)} = f_\star(\overrightarrow{p_0 p}) = g(\overrightarrow{p_0 p}) \Rightarrow f(p) = f(p_0) + g(\overrightarrow{p_0 p})$. (D. h. f ist eindeutig, falls existent.)
Definiere f durch $f(p) \coloneqq f(p_0) + g(\overrightarrow{p_0 p})$. g ist linear und $f_\star = g$ nach (a). Somit folgt $f(p_0) = q_0$, $f_\star = g$. □

Die Verknüpfung affiner Abbildungen ist affin. Die Inverse einer Affinität ist eine Affinität. Beispiele von Affinitäten: Seien A affin, V ein Vektorraum und $x \in V$. Es gilt $t_x : p \mapsto p + x$. t_x ist bijektiv (Inverse: t_{-x}), $t_0 = \mathrm{id}$. Sei $p_0 \in A$, $q_0 \coloneqq t_x(p_0) = p_0 + x$. Sei $r \in A$, dann gilt $r = p_0 + \overrightarrow{p_0 r}$. Es gilt $t_x(r) = r + x = p_0 + \overrightarrow{p_0 r} + x = p_0 + x + \overrightarrow{p_0 r} = q_0 + \overrightarrow{p_0 r}$. Setze $f_\star(\overrightarrow{p_0 r}) = \overrightarrow{p_0 r}$, d. h. $f_\star = \mathrm{id} : V \to V$ linear $\Rightarrow f = t_x$ ist eine Affinität.

Ziel: Rechnen mit Koordinaten in affinen Räumen.

Definition 3.16. Sei A ein affiner Raum bezüglich V. $(n+1)$ Punkte $(p_0, \ldots, p_n)$ in A heißen *affin unabhängig* $:\Leftrightarrow$ die Vektoren $\overrightarrow{p_0p_1}, \ldots, \overrightarrow{p_0p_n}$ sind linear unabhängig in V. Die Punkte heißen *affine Basis* von A $:\Leftrightarrow$ $\overrightarrow{p_0p_1}, \ldots, \overrightarrow{p_0p_n}$ bilden eine Basis von V.

Sei A eine Ebene, d. h. $\dim V = \dim A = 2$:



Standardbeispiel: $A = V = \mathbb{R}^n, p_0 = (0, \ldots, 0)^T, p_1 = e_1, \ldots, p_n = e_n$.

3.16 hängt nicht von der Reihenfolge der Punkte ab:

Lemma 3.17. *Seien $p_0, \ldots, p_n$ affin unabhängig und $j \in \{1, \ldots, n\}$, dann sind $\overrightarrow{p_jp_0}, \ldots, \overrightarrow{p_jp_{j-1}}, \overrightarrow{p_jp_{j+1}}, \ldots, \overrightarrow{p_jp_n}$ linear unabhängig. Bilden insbesondere $p_0, \ldots, p_n$ eine affine Basis, dann ist $\overrightarrow{p_jp_0}, \ldots, \overrightarrow{p_jp_{j-1}}, \overrightarrow{p_jp_{j+1}}, \ldots, \overrightarrow{p_jp_n}$ eine Basis von V $(\dim V < \infty)$.*

Beweis.

$$\begin{aligned}\sum_{i=0, i\neq j}^{n} \lambda_i \overrightarrow{p_jp_i} = 0 \Rightarrow \lambda_0 \overrightarrow{p_jp_0} + \sum_{i=1, i\neq j}^{n} \lambda_i (\overrightarrow{p_jp_0} + \overrightarrow{p_0p_i}) = 0 \\ \Rightarrow -\left(\sum_{i=0, i\neq j}^{n} \lambda_i\right) \overrightarrow{p_0p_j} + \sum_{i=1, i\neq j}^{n} \lambda_i \overrightarrow{p_0p_i} = 0 \\ \Rightarrow \forall i \in \{1, \ldots, n\} \setminus \{j\} : \lambda_i = 0 \Rightarrow \lambda_0 = 0 \ .\end{aligned}$$

□

Lineare Abbildungen sind bestimmt durch die Bilder der Basisvektoren; diese können beliebig gewählt werden.

Proposition 3.18. *Sei $p_0, \ldots, p_n$ eine affine Basis von A und seien $q_0, \ldots, q_n$ beliebige Punkte in B. Dann existiert genau eine affine Abbildung $f : A \to B$ mit $f(p_j) = q_j$ für alle $j \in \{1, \ldots, n\}$. Außerdem:*

(a) $f(A) = [q_0, \ldots, q_n]$.

(b) f injektiv $\Leftrightarrow$ $q_0, \ldots, q_n$ sind affin unabhängig.

(c) f Affinität $\Leftrightarrow$ $q_0, \ldots, q_n$ bilden eine affine Basis.

Beweis. $\overrightarrow{p_0p_1}, \ldots, \overrightarrow{p_0p_n}$ bilden eine Basis des Vektorraums V zu A. Somit gibt es genau eine lineare Abbildung $g : V \to W$ (W = VR zu B) mit $g(\overrightarrow{p_0p_j}) = \overrightarrow{q_0q_j}$ für alle $j \in \{1, \ldots, n\}$. Nach 3.15 (b) existiert genau eine affine Abbildung $f : A \to B$ mit $f(p_0) = q_0$ und $f_\star = g$.

(a) Es gilt $q_0 = f(p_0), \ldots, q_n = f(p_n) \in f(A)$. Wegen $f(A) = q_0 + \mathrm{Bild}(f_\star)$ ist $f(A)$ ein affiner Unterraum, also $f(A) \supseteq [q_0, \ldots, q_n]$. Wegen $f(A) = q_0 + \mathrm{Bild}(f_\star) = q_0 + \mathrm{Span}\{\overrightarrow{q_0q_1}, \ldots, \overrightarrow{q_0q_n}\} \subseteq [q_0, \ldots, q_n]$ folgt die Behauptung.

(b) f injektiv $\Leftrightarrow$ $f_\star$ injektiv $\Leftrightarrow$ $\overrightarrow{q_0q_1}, \ldots, \overrightarrow{q_0q_n}$ linear unabhängig $\Leftrightarrow$ $q_0, \ldots, q_n$ affin unabhängig.

(c) f bijektiv $\Leftrightarrow$ $f_\star$ bijektiv $\Leftrightarrow$ $\overrightarrow{q_0q_1}, \ldots, \overrightarrow{q_0q_n}$ Basis $\Leftrightarrow$ $q_0, \ldots, q_n$ affine Basis. □

Nach Theorem 3.4 existiert eine Bijektion $f : A \to V$ (= VR zu A) mit $f(p) + x = f(p + x)$ für alle $x \in V$, wobei p ein wählbarer Basispunkt ist. Wähle p_0 fest und $f(p_0) = 0 \in V$. $q \in A$ hat die Form $q = p_0 + x, x \in V$. Es gilt $f(q) = 0 + x = x$. Wähle eine Basis von $V \Rightarrow x$ hat Koordinaten bezüglich dieser Basis, z. B. $x = (x_1, \ldots, x_n)$. Per Definition sind das die Koordinaten von q. Z. B. für $\dim A = n$: Wähle eine affine Basis $p_0, \ldots, p_n$, von A, dann ist $\overrightarrow{p_0p_1}, \ldots, \overrightarrow{p_0p_n}$ eine Basis von V. $q \in A$ hat die Form $q = p_0 + x = p_0 + (\lambda_1\overrightarrow{p_0p_1} + \ldots + \lambda_n\overrightarrow{p_0p_n})$. $f(q) = f(p_0) + x = x$ hat die Koordinaten $(\lambda_1, \ldots, \lambda_n)^T$. $(\lambda_1, \ldots, \lambda_n)^T$ ist (nach Definition) der *Koordinatenvektor* (in A) von q (und in V: Koordinatenvektor von x).

Für $\dim A = n$ können wir $V = K^n$ wählen. $p_0, \ldots, p_n$ affine Basis, q hat Koordinaten $(\lambda_1, \ldots, \lambda_n)^T \Rightarrow q = p_0 + \lambda_1\overrightarrow{p_0p_1} + \ldots + \lambda_n\overrightarrow{p_0p_n}$.

Definition 3.19. Sei $g = [p_0, p_1]$ eine Gerade in einem affinen Raum und $p \in g$. Dann ist $\overrightarrow{p_0p} = \lambda \cdot \overrightarrow{p_0p_1}$ für ein eindeutig bestimmtes $\lambda \in K$. λ heißt das *Teilverhältnis* von p_0, p_1, p. Bezeichnung: $\mathrm{TV}(p_0, p_1, p) = \lambda$. Falls $\mathrm{char}(K) \neq 2$ und $\mathrm{TV}(p_0, p_1, p) = 1/2$, heißt p der *Mittelpunkt* von p_0 und p_1.

In Koordinaten: g affiner Raum mit $V = K$, affine Basis p_0, p_1, d. h. $\overrightarrow{p_0p_1}$ ist eine Basis von V. $f(p_0) = 0$, $f_\star(\overrightarrow{p_0p_1}) = \overrightarrow{p_0p_1}$. $f_\star(\overrightarrow{p_0p}) = f_\star(\lambda\overrightarrow{p_0p_1}) = \lambda\overrightarrow{p_0p_1}$. In Koordinaten: p hat den Koordinatenvektor (λ) und $\lambda = \mathrm{TV}(p_0, p_1, p)$.

TV ist eine affine Invariante: Sei φ eine affine Abbildung, die eine Gerade $[p_0, p_1]$ in eine Gerade $[\varphi(p_0), \varphi(p_1)]$ abbildet. Behauptung: $\mathrm{TV}(p_0, p_1, p) = \mathrm{TV}(\varphi(p_0), \varphi(p_1), \varphi(p))$. Denn: Sei $\mathrm{TV}(p_0, p_1, p) = \lambda$. Das bedeutet $\overrightarrow{p_0p} = \lambda\overrightarrow{p_0p_1} \Rightarrow \varphi_\star(\overrightarrow{p_0p}) = \varphi_\star(\lambda\overrightarrow{p_0p_1}) = \lambda\varphi_\star(\overrightarrow{p_0p_1})$. Somit gilt $\overrightarrow{\varphi(p_0)\varphi(p)} = \varphi_\star(\overrightarrow{p_0p}) = \lambda\varphi_\star(\overrightarrow{p_0p_1}) = \lambda\overrightarrow{\varphi(p_0)\varphi(p_1)} \Rightarrow \lambda = \mathrm{TV}(\varphi(p_0), \varphi(p_1), \varphi(p))$.

Sei $K = \mathbb{R}$ (oder allgemeiner $\mathrm{char}(K) \neq 2$): Die beiden Diagonalen eines Parallelogramms schneiden sich in ihren Mittelpunkten. D. h. ist $p_1 = p_0 + y, p_2 = p_0 + x, p_3 = p_1 + x = p_2 + y$ und ist q der Schnittpunkt der Diagonalen, dann ist $\mathrm{TV}(p_0, p_3, q) = \mathrm{TV}(p_1, p_2, q) = 1/2$.

Beweis. Das Parallelogramm liegt in einer Ebene $A = [p_0, p_1, p_2, p_3] \simeq K^2$. Wähle die affine Basis p_0, p_1, p_2, Koordinaten: $p_0 \mapsto (0,0)^T, p_1 \mapsto (1,0)^T, p_2 \mapsto (0,1)^T$. Wir haben bereits bewiesen, dass TV im Bild dasselbe ist. Dort ist $p_3 \mapsto (1,1)^T$ und $q = (1/2, 1/2)$ und die Teilverhältnisse lassen sich leicht berechnen. $\square$

Kapitel 4

Euklidische und unitäre Räume

Was ist ein rechter Winkel? Analogie: V VR, V^* Dualraum, $U < V$ Unterraum $\Rightarrow$ der zu U orthogonale Raum $U^0 := \{\varphi \in V^* \mid \forall u \in U : \varphi(u) = 0\} < V^*$ (nicht $< V$). Umschreiben: Die Abbildung $f : V \times V^* \to K, (v, \varphi) \mapsto \varphi(v)$ erfüllt $f(v_1 + v_2, \varphi) = \varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2) = f(v_1, \varphi) + f(v_2, \varphi)$. Für $v \neq 0 : \exists \varphi \in V^* : \varphi(v) \neq 0$.

Definition 4.1. Sei V ein $\mathbb{R}$-Vektorraum. Ein *Skalarprodukt* (oder *inneres Produkt*) auf V ist eine Funktion $s : V \times V \to \mathbb{R}$ mit den Eigenschaften:

(S1) $\forall x, y, z \in V : s(x +_V y, z) = s(x, z) +_{\mathbb{R}} s(y, z)$.

(S2) $\forall x, y \in V, \lambda \in \mathbb{R} : s(\lambda x, y) = \lambda s(x, y)$.

(S3) $\forall x, y \in V : s(x, y) = s(y, x)$.

(S4) $\forall x \in V \setminus \{0\} : s(x, x) > 0$.

Andere Bezeichnung: $s(x, y) =: \langle x, y \rangle$.

Aus der Symmetrie (S3) folgt $s(x, y+z) = s(y+z, x) = s(y, x)+s(z, x) = s(x, y) + s(x, z)$ wie (S1), aber in der zweiten Variablen. Das gleiche können wir auch für (S2) machen: $s(x, \lambda y) = s(\lambda y, x) = \lambda s(y, x) = \lambda s(x, y)$. Sind (S1), (S2), (S3) erfüllt, dann heißt s *symmetrische Bilinearform*. Erfüllt s die Eigenschaft (S4), dann heißt s *positiv definit*. Diese Eigenschaft braucht $K = \mathbb{R}$.

Es gilt

$$\forall x \in V : s(x, 0_V) = s(0_V, x) = s(0_\mathbb{R} \cdot 0_V, s) = 0_\mathbb{R} \cdot s(0_V, x) = 0_\mathbb{R} ,$$

also insbesondere $s(0_V, 0_V) = 0_\mathbb{R}$.

Beispiel 4.1. Sei $V = \mathbb{R}^n$ und $x, y \in V$. Definiere $s(x, y) := \sum_{i=1}^n x_i y_i$. Dann gelten (S1), (S2) und (S3). Außerdem gilt $s(x, x) = \sum_{i=1}^n x_i^2 > 0$ für $x \neq 0$. Somit ist auch (S4) erfüllt und s ist ein Skalarprodukt.

Beispiel 4.2. Sei V der Vektorraum aller stetigen Funktionen von $[a, b]$ nach $\mathbb{R}$ $(b > a)$. Die Funktion definiert durch

$$s(f, g) := \int_a^b f(x)g(x)\,\mathrm{d}x$$

für $f, g \in V$ ist ein Skalarprodukt: s ist positiv definit, denn für $f \in V\backslash\{0\}$ gibt es ein $x_0 \in [a, b]$, sodass $f(x_0) \neq 0$. Zu zeigen: $\int_a^b (f(x))^2\,\mathrm{d}x > 0$. Beweisskizze für x_0 im Inneren von $[a, b]$: f stetig $\Rightarrow \exists \delta > 0 : [x_0 - \delta, x_0 + \delta] \subseteq [a, b] \wedge \forall x \in [x_0 - \delta, x_0 + \delta] : (f(x))^2 \geq (f(x_0))^2/2$. Dann folgt

$$\begin{aligned}\int_a^b f^2(x)\,\mathrm{d}x &= \int_a^{x_0-\delta} f^2(x)\,\mathrm{d}x + \int_{x_0-\delta}^{x_0+\delta} f^2(x)\,\mathrm{d}x + \int_{x_0+\delta}^b f^2(x)\,\mathrm{d}x \\ &\geq 0 + f^2(x_0)\delta + 0 > 0 .\end{aligned}$$

Für $x_0 = a$ müssten wir stattdessen $[x_0, x_0 + \delta]$ verwenden und für $x_0 = b$ stattdessen $[x_0 - \delta, x_0]$ (oder wir verwenden, dass es ein $\varepsilon > 0$ gibt mit $f(x_0 \pm \varepsilon) > 0$ und wenden dann den bereits vorhandenen Beweis an).

Damit erhalten wir viele Anwendungen in der Analysis.

Auf $\mathbb{R}^n$ gibt es viele Skalarprodukte: Wähle $M = (m_{ij}) \in M_n(\mathbb{R})$ mit $m_{ij} = m_{ji}$, d. h. $M = M^T$. M heißt *symmetrische Matrix*. Definiere $s : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ durch $s(x, y) := x^T M y = z_1 y_1 + \ldots + z_n y_n$, wenn $x^T M = (z_1, \ldots, z_n)$.

s ist immer eine symmetrische Bilinearform, denn für alle $x, q, y \in V, \lambda \in \mathbb{R}$ gilt:

$$\begin{aligned}s(x + q, y) &= (x + q)^T M y = x^T M y + q^T M y = s(x, y) + s(q, y) \\ s(\lambda x, y) &= (\lambda x)^T M y = \lambda x^T M y = \lambda s(x, y) \\ s(x, y) &= x^T M y = (x^T M y)^T = y^T M^T (x^T)^T = y^T M x = s(y, x) .\end{aligned}$$

Dabei gilt $x^T M y = (x^T M y)^T$, da die Transponierte einer reellen Zahl diese reelle Zahl ist.

Für $M = 0$ ist s nicht positiv definit. Für $M = E_n$ ist $s(x, y) = x^T E_n y = x^T y = \sum_{i=1}^n x_i y_i$. Dies ist das Standardskalarprodukt (oder euklidisches Skalarprodukt) und es ist positiv definit. Ist M eine Diagonalmatrix, also $M = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, dann ist s genau dann positiv definit, wenn $\lambda_1, \dots, \lambda_n > 0$.

Später: alle symmetrischen Matrizen über $\mathbb{R}$ sind diagonalisierbar. D. h. für eine geeignete Basis ist M eine Diagonalmatrix.

Theorem 4.2. *Sei V ein $\mathbb{R}$-Vektorraum mit einer Basis $b_1, \dots, b_n$. Dann existiert eine Bijektion zwischen der Menge der symmetrischen $n \times n$-Matrizen und der Menge der symmetrischen Bilinearformen $s : V \times V \to \mathbb{R}$. Die Bijektion bildet $M = M^T$ ab auf $s(x, y) = x^T M y$ (wobei x, y als Koordinatenvektoren bezüglich der Basis aufzufassen sind). (Die Bijektion ist ein Isomorphismus von $\mathbb{R}$-Vektorräumen.)*

Beweis. Gesucht: Umkehrabbildung von der Menge der symmetrischen Bilinearformen in die Menge der symmetrischen Matrizen. Sei $s : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. Setze $m_{ij} = s(b_i, b_j)$. $M = (m_{ij})$ heißt *darstellende Matrix* von s. M ist symmetrisch: $m_{ij} = s(b_i, b_j) \overset{\text{(S3)}}{=} s(b_j, b_i) = m_{ji}$. Bezüglich der Basis $b_1, \dots, b_n$ hat b_i den Koordinatenvektor e_i. Es gilt

$$e_i^T M e_j = m_{ij} = s(b_i, b_j) \ .$$

Somit definiert M die symmetrische Bilinearform s. □

Definition 4.3. Sei V ein $\mathbb{C}$-Vektorraum. Ein *Skalarprodukt* (oder *inneres Produkt*) auf V ist eine Funktion $s : V \times V \to \mathbb{C}$ mit den Eigenschaften

(S1) $\forall x, y, z \in V : s(x + y, z) = s(x, z) + s(y, z)$.

(S2) $\forall x, y \in V, \lambda \in \mathbb{C} : s(\lambda x, y) = \lambda s(x, y)$.

(S3) $\forall x, y \in V : s(x, y) = \overline{s(y, x)}$.

(S4) $\forall x \in V \setminus \{0\} : s(x, x) > 0$.

Andere Bezeichnung: $\langle x, y \rangle := s(x, y)$.

Der Unterschied zum reellen Fall ist (S3). Für alle $x \in V$ ist $s(x,x) \overset{\text{(S3)}}{=} \overline{s(x,x)}$, also $s(x,x) \in \mathbb{R}$. Außerdem gilt

$$\begin{aligned} s(x, y+z) &= \overline{s(y+z,x)} = \overline{s(y,x)+s(z,x)} = \overline{s(y,x)} + \overline{s(z,x)} \\ &= s(x,y) + s(x,z) \\ s(x, \lambda y) &= \overline{s(\lambda y, x)} = \overline{\lambda s(y,x)} = \overline{\lambda} \cdot \overline{s(y,x)} = \overline{\lambda} s(x,y) \ . \end{aligned}$$

s ist nicht bilinear, sondern sesquilinear (3/2-linear) oder semilinear. s mit (S1), (S2), (S3) heißt *Hermitesche* Form. Die Eigenschaft (S4) heißt weiterhin positiv definit. Ein Skalarprodukt über $\mathbb{C}$ ist somit eine positiv definite Hermitesche Form.

Beispiel 4.3.

- $V = \mathbb{C}^n$, Versuch: $s(x,y) \coloneqq \sum_{j=1}^n x_j y_j$ funktioniert nicht: (S3) scheitert und $s(x,x)$ ist im Allgemeinen nicht reell.

- Besser: $s(x,y) \coloneqq \sum_{j=1}^n x_j \overline{y_j}$ erfüllt auch (S1) und (S2), aber $s(y,x) = \sum_{j=1}^n y_j \overline{x_j} = \overline{s(x,y)}$, also ist auch (S3) erfüllt. Überprüfe (S4) für einen Vektor $x \neq 0$:

  $$s(x,x) = \sum_{j=1}^n x_j \overline{x_j} = \sum_{j=1}^n |x_j|^2 > 0 \ .$$

  Folglich gilt (S4). Das ist das Standardskalarprodukt auf $\mathbb{C}^n$.

Andere Beispiele wie über $\mathbb{R}$: Darstellende Matrix $M = (m_{ab})$ mit $m_{ab} = s(v_a, v_b)$, wobei $v_1, \ldots, v_n$ eine Basis von V ist. (S3) verlangt: $m_{ab} = \overline{m_{ba}}$. Eine Matrix $M = (m_{ab})$ mit $m_{ba} = \overline{m_{ab}}$ heißt *Hermitesche* Matrix.

Sei V ein $\mathbb{C}$-VR mit einer festen Basis $v_1, \ldots, v_n$. Dann existiert wieder eine Bijektion zwischen Hermiteschen Formen und Hermiteschen Matrizen. Ist M Hermitesch, dann definieren wir (in Koordinatenvektoren) $s(x,y) \coloneqq x^T M \overline{y}$. Für $M = E_n$ erhalten wir das Standardskalarprodukt.

$M = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ Hermitesch $\Leftrightarrow \forall j : \lambda_j = \overline{\lambda_j} \Leftrightarrow \forall j : \lambda_j \in \mathbb{R}$ (alle Eigenwerte sind reell).

Positiv definit bedeutet $\lambda_1, \ldots, \lambda_n > 0$.

Allgemein (später) sind Hermitesche Matrizen diagonalisierbar und haben reelle Eigenwerte.

Definition 4.4. Ein Vektorraum über $\mathbb{R}$ mit einem Skalarprodukt heißt *Euklidischer Vektorraum*. Ein Vektorraum über $\mathbb{C}$ mit einem Skalarprodukt heißt *unitärer Vektorraum*.

Geometrische Anwendung: Längen, Winkel, ...

Beispiel 4.4. Sei $(a,b)^T \in \mathbb{R}^2$. Nach dem Satz des Pythagoras ist die Länge von $(a,b)^T$ gleich $\sqrt{a^2+b^2} = \sqrt{s((a,b)^T,(a,b)^T)}$, falls s das euklidische Skalarprodukt (Standardskalarprodukt) ist.

Definition 4.5. Sei V ein euklidischer oder unitärer Vektorraum und $x \in V$. Die reelle Zahl $\sqrt{s(x,x)}$ heißt *Länge* (oder *Norm* oder *Betrag*) von x. Bezeichnung: $\|x\| := \sqrt{s(x,x)}$. Falls $\|x\| = 1$, heißt x *Einheitsvektor* (oder normierter Vektor). (In $\mathbb{R}^2$ mit Standardskalarprodukt ist das die übliche Länge wie im Satz des Pythagoras.)

Lemma 4.6. *Sei V ein euklidischer oder unitärer Vektorraum, $x, y \in V, \lambda \in \mathbb{R}$ bzw. $\lambda \in \mathbb{C}$. Dann gilt:*

(a) $\|x\| = 0 \Leftrightarrow x = 0.$

(b) $\|\lambda x\| = |\lambda| \cdot \|x\|.$

(c) $|\langle x, y\rangle| \leq \|x\| \cdot \|y\|$ *(Cauchy-Schwarzsche Ungleichung).*

(d) $\|x+y\| \leq \|x\| + \|y\|$ *(Dreiecksungleichung).*

Beweis.

(a) folgt aus positiv definit.

(b) $\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x\rangle} = \sqrt{\lambda\overline{\lambda}\,\langle x, x\rangle} = |\lambda|\sqrt{\langle x, x\rangle} = |\lambda| \cdot \|x\|.$

(c) Für $y = 0$: $\langle x, y\rangle = 0$, dann stimmt die Ungleichung. Sei also nun $y \neq 0$, dann ist $\langle y, y\rangle > 0$. Setze

$$t := -\frac{\langle x, y\rangle}{\langle y, y\rangle} \in \mathbb{C}\ .$$

Dann gilt

$$\begin{aligned}0 &\leq \langle x+ty, x+ty\rangle = \langle x,x\rangle + \langle x,ty\rangle + \langle ty,x\rangle + \langle ty,ty\rangle \\ &= \langle x,x\rangle + \bar{t}\,\langle x,y\rangle + t\overline{\langle x,y\rangle} + t\bar{t}\,\langle y,y\rangle \\ &= \langle x,x\rangle - \frac{\overline{\langle x,y\rangle}}{\langle y,y\rangle}\langle x,y\rangle - \frac{\langle x,y\rangle}{\langle y,y\rangle}\overline{\langle x,y\rangle} + \frac{\langle x,y\rangle}{\langle y,y\rangle}\frac{\overline{\langle x,y\rangle}}{\langle y,y\rangle}\langle y,y\rangle \\ &= \langle x,x\rangle - \frac{\langle x,y\rangle\,\overline{\langle x,y\rangle}}{\langle y,y\rangle} = \|x\|^2 - \frac{|\,\langle x,y\rangle\,|^2}{\|y\|^2}\ .\end{aligned}$$

Es folgt $0 \leq \|x\|^2\|y\|^2 - |\,\langle x,y\rangle\,|^2 \Rightarrow |\,\langle x,y\rangle\,| \leq \|x\|\cdot\|y\|$. Bemerkung: $0 = \|x\|^2\|y\|^2 - |\,\langle x,y\rangle\,|^2 \Leftrightarrow x+ty=0$. Es folgt $|\,\langle x,y\rangle\,| = \|x\|\cdot\|y\| \Leftrightarrow$ x und y sind linear abhängig.

(d) Wir zeigen $\|x+y\|^2 \leq (\|x\|+\|y\|)^2$:

$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle = \langle x,x\rangle + \langle x,y\rangle + \overline{\langle x,y\rangle} + \langle y,y\rangle \\ &= \langle x,x\rangle + 2\,\mathrm{Re}(\langle x,y\rangle) + \langle y,y\rangle \\ &\leq \|x\|^2 + 2|\,\langle x,y\rangle\,| + \|y\|^2 \\ &\overset{(c)}{\leq} \|x\|^2 + 2\|x\|\cdot\|y\| + \|y\|^2 \\ &= (\|x\|+\|y\|)^2\ .\end{aligned}$$ □

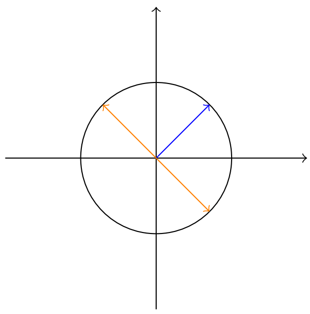
Definition orthogonal: z. B. $(x,0)^T$ und $(0,y)^T$ sind orthogonal und $\left\langle (x,0)^T, (0,y)^T \right\rangle = 0$ (Standardskalarprodukt).

Definition 4.7. Sei V ein euklidischer oder unitärer Vektorraum. Vektoren $x, y \in V$ heißen *senkrecht* (oder *orthogonal*) zueinander $:\Leftrightarrow \langle x,y\rangle = 0$. Bezeichnung: $x \perp y$.

Eine Teilmenge $M \subseteq V$ heißt *Orthogonalsystem* $:\Leftrightarrow [0 \notin M$ und $\forall x,y \in M : x \neq y \Rightarrow x \perp y]$. M heißt *Orthonormalsystem* (ONS) (oder *orthonormiertes System*) $:\Leftrightarrow [M$ ist Orthogonalsystem und $\forall x \in M : \|x\| = 1]$.

Beispiel im $\mathbb{R}^n$: $e_1, \ldots, e_n$ bilden ein ONS.

Beispiel 4.5. ONS im $\mathbb{R}^2$: $M = \emptyset$ oder $M = \{x\}, \|x\| = 1$. Für $|M| = 2$:



Der erste Vektor ist $(\cos\alpha, \sin\alpha)^T$ für $0 \leq \alpha < 2\pi$. Dann ist der zweite Vektor $(\sin\alpha, -\cos\alpha)^T$ oder $(-\sin\alpha, \cos\alpha)^T$. Weitere ONS mit zwei Elementen gibt es nicht. Im $\mathbb{R}^2$ gibt es kein ONS mit $|M| \geq 3$. Allgemein: $|M| \leq \dim V$.

Proposition 4.8. *Sei V ein euklidischer oder unitärer Vektorraum, $x, y \in V$, $M = \{v_1, \ldots, v_n\}$ ein ONS, $z \in \langle M \rangle$. Dann gilt:*

(a) $x \perp y \Leftrightarrow y \perp x$.

(b) $x \perp u \ \forall u \in V \Leftrightarrow x = 0$.

(c) $v_1, \ldots, v_n$ *sind linear unabhängig (also* $n \leq \dim V$*).*

(d) $z = \sum_{j=1}^n \langle z, v_j \rangle v_j$.

Aus (c) folgt: $v_1, \ldots, v_n$ sind eine Basis von $\langle M \rangle$. Die Koeffizienten von z bezüglich dieser Basis sind (nach (d)) $\langle z, v_1 \rangle, \ldots, \langle z, v_n \rangle$. Beispiel $\mathbb{R}^2$: $e_1 = (1,0)^T, e_2 = (0,1)^T$ als ONS, $\langle M \rangle = \mathbb{R}^2$. $(a,b)^T = (a,0)^T + (0,b)^T = ae_1 + be_2, a = \left\langle (a,b)^T, e_1 \right\rangle, b = \left\langle (a,b)^T, e_2 \right\rangle$.

Beweis.

(a) $x \perp y \Leftrightarrow \langle x, y \rangle = 0 \Leftrightarrow \langle y, x \rangle = 0 \Leftrightarrow y \perp x$.

(b) $x = 0 \Rightarrow \forall u : \langle x, u \rangle = \langle 0 \cdot x, u \rangle = 0 \cdot \langle x, u \rangle = 0$. Umgekehrt: $\forall u : x \perp u \Rightarrow x \perp x \Rightarrow \langle x, x \rangle = 0 \Rightarrow x = 0$.

(c) Sei $\sum_{j=1}^n \lambda_j v_j = 0$, zz: $\forall j : \lambda_j = 0$. Es gilt

$$\begin{aligned} 0 = \langle 0, v_1 \rangle &= \left\langle \sum_{j=1}^n \lambda_j v_j, v_1 \right\rangle \\ &= \sum_{j=1}^n \lambda_j \langle v_j, v_1 \rangle \\ &= \lambda_1 \ . \end{aligned}$$

Analog folgt $\forall j : \lambda_j = 0$. Somit sind $v_1, \dots, v_n$ linear unabhängig.

(d) Aus (c) folgt: $v_1, \dots, v_n$ Basis von $\langle M \rangle$, also existieren $\mu_1, \dots, \mu_n$ mit $z = \mu_1 v_1 + \dots, + \mu_n v_n$. Dann gilt

$$\begin{aligned}\langle z, v_1 \rangle &= \langle \mu_1 v_1 + \dots + \mu_n v_n, v_1 \rangle \\ &= \mu_1 \underbrace{\langle v_1, v_1 \rangle}_{=1} + \mu_2 \underbrace{\langle v_2, v_1 \rangle}_{=0} + \dots + \mu_n \underbrace{\langle v_n, v_1 \rangle}_{=0} \\ &= \mu_1 \ .\end{aligned}$$

Analog folgt $\mu_j = \langle z, v_j \rangle$ und somit die Behauptung. □

Theorem 4.9. *Sei V ein euklidischer oder unitärer Vektorraum und $U = \langle v_1, \dots, v_l \rangle$ ein Unterraum mit Basis $v_1, \dots, v_l$. Dann existiert ein ONS $b_1, \dots, b_l$ mit $U = \langle b_1, \dots, b_l \rangle$. Das heißt, U besitzt eine* Orthonormalbasis *(ONB). Insbesondere besitzt jeder endlichdimensionale euklidische oder unitäre Vektorraum eine ONB.*

Beweis. Der Beweis ist konstruktiv und liefert das Gram-Schmidtsche Orthogonalisierungsverfahren:

Induktiv: Gegeben $v_1, \dots, v_j$ als Input, der Output besteht aus $b_1, \dots, b_j$, die ein ONS und eine Basis von $\langle v_1, \dots, v_j \rangle$ bilden. Der Ouptut hängt von der Reihenfolge der Vektoren $v_1, \dots, v_l$ ab.

Induktionsanfang: $U = \langle v_1 \rangle, v_1 \neq 0 \Rightarrow \|v_1\| \neq 0$, setze

$$b_1 := \frac{v_1}{\|v_1\|} \Rightarrow \|b_1\| = \left\| \frac{v_1}{\|v_1\|} \right\| = \frac{1}{\|v_1\|} \|v_1\| = 1, \qquad \langle b_1 \rangle = \langle v_1 \rangle \ .$$

Induktionsschritt von j auf $j+1$: Sei $\langle v_1, \dots, v_j \rangle = \langle b_1, \dots, b_j \rangle$ mit einer Orthonormalbasis $b_1, \dots, b_j$. Sei jetzt $v_1, \dots, v_{j+1}$ gegeben. Wir suchen b_{j+1}, sodass $b_1, \dots, b_{j+1}$ eine ONB von $\langle v_1, \dots, v_{j+1} \rangle$ ist. b_{j+1} muss eine Linearkombination von $v_1, \dots, v_{j+1}$ sein, wobei $v_1, \dots, v_j$ bereits Linearkombinationen von $b_1, \dots b_j$ sind. Zunächst definieren wir einen Hilfsvektor c_{j+1}, mit dem wir dann später $b_{j+1} = c_{j+1} / \|c_{j+1}\|$ definieren können:

$$c_{j+1} := v_{j+1} - \sum_{a=1}^{j} \langle v_{j+1}, b_a \rangle \, b_a \in \langle v_1, \dots, v_{j+1} \rangle = \langle b_1, \dots, b_j, v_{j+1} \rangle \ .$$

Es gilt $v_{j+1} \notin \langle v_1, \dots, v_j \rangle = \langle b_1, \dots, b_j \rangle \Rightarrow c_{j+1} \neq 0$ und $\dim \langle v_1, \dots, v_{j+1} \rangle = \dim \langle b_1, \dots, b_j \rangle + 1$. Falls $b_1, \dots, b_j, c_{j+1}$ linear unabhängig sind, bilden sie eine Basis von $\langle v_1, \ldots, v_{j+1}\rangle$. (Wegen $v_{j+1} \in \langle b_1, \ldots, b_j, c_{j+1}\rangle$ folgt auch, dass $b_1, \ldots, b_j, c_{j+1}$ eine Basis bilden.) Lineare Unabhängigkeit folgt (auch) aus Orthogonalität (und alle $\neq 0$): $b_1, \ldots, b_j$ sind schon paarweise orthogonal. Für jedes $d \in \{1, \ldots, j\}$ gilt

$$\begin{aligned} \langle c_{j+1}, b_d\rangle &= \langle v_{j+1}, b_d\rangle - \sum_{a=1}^{j} \langle v_{j+1}, b_a\rangle \underbrace{\langle b_a, b_d\rangle}_{=\delta_{ad}} \\ &= \langle v_{j+1}, b_d\rangle - \langle v_{j+1}, b_d\rangle \cdot 1 = 0 \ . \end{aligned}$$

$\Rightarrow b_1, \ldots, b_j, c_{j+1}$ sind ein Orthogonalsystem, also linear unabhängig und daher eine Basis von $\langle v_1, \ldots, v_{j+1}\rangle$. Mit der obigen Definition $b_{j+1} \coloneqq c_{j+1}/\|c_{j+1}\|$ folgt, dass $b_1, \ldots, b_{j+1}$ eine ONB von $\langle v_1, \ldots, v_{j+1}\rangle$ bilden. □

Beispiel 4.6. Betrachte eine Basis $e_1, v_2 = (a, b)^T$ von $\mathbb{R}^2$. Wähle $b_1 = e_1$, dann gilt

$$\begin{aligned} c_2 &= v_2 - \langle v_2, e_1\rangle e_1 = v_2 - ae_1 = \begin{pmatrix} 0 \\ b \end{pmatrix} \\ b_2 &= \frac{c_2}{\|c_2\|} = \frac{c_2}{|b|} = \pm e_2 \ . \end{aligned}$$

Hier wird der Vektor v_2 in e_1-Richtung projiziert, das Ergebnis ist c_2.

Beispiel 4.7. Im $\mathbb{R}^3$:

$$v_1 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

ist eine Basis von $V = \mathbb{R}^3$. Erster Schritt: $\|v_1\| = \sqrt{\langle v_1, v_1\rangle} = \sqrt{5}, b_1 \coloneqq v_1/\|v_1\|$. Zweiter Schritt:

$$\begin{aligned} c_2 &\coloneqq \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} - \frac{1}{\sqrt{5}} \left\langle \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right\rangle \frac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} - \frac{1}{5} \cdot 5 \cdot \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \|c_2\| = 1 \Rightarrow b_2 = c_2 \ . \end{aligned}$$

Dritter Schritt:

$$c_3 := \begin{pmatrix}1\\0\\1\end{pmatrix} - \frac{1}{\sqrt{5}}\left\langle \begin{pmatrix}0\\1\\2\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix} \right\rangle \frac{1}{\sqrt{5}}\begin{pmatrix}0\\1\\2\end{pmatrix} - \left\langle \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix} \right\rangle \begin{pmatrix}1\\0\\0\end{pmatrix}$$

$$= \begin{pmatrix}1\\0\\1\end{pmatrix} - \frac{1}{5}\cdot 2 \cdot \begin{pmatrix}0\\1\\2\end{pmatrix} - 1 \cdot \begin{pmatrix}1\\0\\0\end{pmatrix} = \begin{pmatrix}0\\-2/5\\1/5\end{pmatrix} = \frac{1}{5}\begin{pmatrix}0\\-2\\1\end{pmatrix}$$

$$\|c_3\| = \frac{1}{\sqrt{5}} \Rightarrow b_3 = \sqrt{5}c_3 = \frac{1}{\sqrt{5}}\begin{pmatrix}0\\-2\\1\end{pmatrix}$$

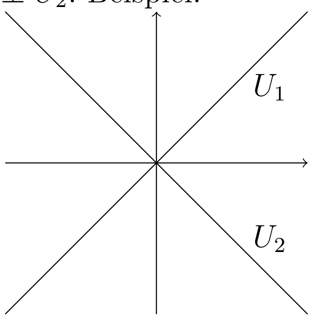
Probe: $\|b_1\| = \|b_2\| = \|b_3\| = 1$, $b_1 \perp b_2$, $b_1 \perp b_3$, $b_2 \perp b_3 \Rightarrow b_1, b_2, b_3$ ist ONB.

Der Stoff für die Scheinklausur geht bis hier.

Definition 4.10. Sei V ein euklidischer oder unitärer Vektorraum und $M \subseteq V$ eine Teilmenge. Die Menge

$$M^\perp := \{y \in V \mid \forall x \in M : x \perp y\}$$

heißt *orthogonales Komplement* von M. Unterräume U_1, U_2 heißen *senkrecht* (oder orthogonal) zueinander $:\Leftrightarrow \forall x \in U_1 : \forall y \in U_2 : x \perp y$. Bezeichnung: $U_1 \perp U_2$. Beispiel:



$M^\perp$ ist immer ein Untervektorraum. $U_1 \perp U_2 \Leftrightarrow U_1 \subseteq U_2^\perp \wedge U_2 \subseteq U_1^\perp$.

Sei $\dim V < \infty$ und U ein Untervektorraum. Wähle eine Basis $b_1, \dots, b_l$ von U ($\dim U = l > 0$), ergänze zu einer Basis $b_1, \dots, b_l, c_1, \dots, c_m$ von V ($\dim V = l + m$). Das Gram-Schmidt-Verfahren macht daraus eine Orthonormalbasis $d_1, \dots, d_{l+m}$ von V. Dabei gilt $\langle d_1, \dots, d_l \rangle = \langle b_1, \dots, b_l \rangle = U \Rightarrow$

$d_1, \ldots, d_l$ bilden ONB von U. $d_{l+1}, \ldots, d_{l+m}$ sind orthogonal zu $d_1, \ldots, d_l \Rightarrow d_{l+1}, \ldots, d_{l+m} \in U^\perp \Rightarrow \dim U^\perp \geq m$. Sei $x \in U \cap U^\perp$, dann folgt $\langle x, x \rangle = 0$ und somit $x = 0 \Rightarrow U \cap U^\perp = \{0\}$. Es gilt also $\dim V = \dim U + \dim U^\perp$ und somit $V = U \oplus U^\perp$ (das wird für die Blockzerlegung von Matrizen nützlich sein).

Sei V euklidisch und $x, y \in V \setminus \{0\}$. Nach Cauchy-Schwarz (Theorem 4.6 (c)) gilt $|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$, also

$$\left| \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \right| \leq 1 \ .$$

Somit existiert $\alpha \in [0, \pi]$ mit

$$\frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} = \cos \alpha \ .$$

α ist der Winkel $\sphericalangle(x, y)$.

Definition 4.11. Sei V ein euklidischer Vektorraum und $x, y \in V \setminus \{0\}$. Der *Winkel* $\alpha \coloneqq \sphericalangle(x, y)$ ist definiert durch

$$\cos \alpha = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \ .$$

($x \perp y \Leftrightarrow \sphericalangle(x, y) = \pi/2$. Für unitäre V funktioniert diese Definition nicht.)

Proposition 4.12. *Sei V ein euklidischer Vektorraum und $x, y \in V$.*

(a) (Cosinussatz): $\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\|\|y\| \cos \sphericalangle(x, y)$.

(b) (Satz des Pythagoras): $x \perp y \Leftrightarrow \|x - y\|^2 = \|x\|^2 + \|y\|^2$.

Beweis. (b) ist ein Spezialfall von (a). Beweis von (a):

$$\begin{aligned} \|x - y\|^2 &= \langle x - y, x - y \rangle = \langle x, x \rangle - 2 \langle x, y \rangle + \langle y, y \rangle \\ &= \|x\|^2 + \|y\|^2 - 2 \underbrace{\frac{\langle x, y \rangle}{\|x\|\|y\|}}_{= \cos \sphericalangle(x,y)} \cdot \|x\|\|y\| \ . \end{aligned}$$

$\square$

Proposition 4.13. *Sei V ein euklidischer oder unitärer Vektorraum und $f : V \to V$ eine lineare Abbildung. Dann sind die folgenden Bedingungen äquivalent:*

(1) $\forall x \in V : \|x\| = 1 \Rightarrow \|f(x)\| = 1$*, d. h. f bildet Einheitsvektoren auf Einheitsvektoren ab.*

(2) $\forall x \in V : \|f(x)\| = \|x\|$*, d. h. f erhält die Länge von Vektoren.*

(3) $\forall x, y \in V : \langle f(x), f(y)\rangle = \langle x, y\rangle$*, d. h. f erhält das Skalarprodukt.*

Beispiel: Drehungen, f bildet ONB in ONB ab.

Definition 4.14. f wie in 4.13 heißt im euklidischen Fall eine *orthogonale Abbildung*, im unitären Fall eine *unitäre Abbildung.*

Beweis. (von 4.13)

(1) $\Rightarrow$ (2): Sei $x \in V$. $x = 0 \Rightarrow f(x) = 0$ und die Behauptung gilt. Ist $x \neq 0$, dann gilt $\|x\| \neq 0$. Setze

$$\begin{aligned} y &:= \frac{x}{\|x\|} \Rightarrow \|y\| = \|\frac{x}{\|x\|}\| = 1 \\ \Rightarrow 1 &= \|f(y)\| = \|f\left(\frac{x}{\|x\|}\right)\| = \frac{1}{\|x\|}\|f(x)\| \\ \Rightarrow \|f(x)\| &= \|x\| \ . \end{aligned}$$

(2) $\Rightarrow$ (3): Voraussetzung: f erhält Längen, zz: f erhält Skalarprodukte. Wir zeigen: Werte des Skalarprodukts sind durch Längen bestimmt. Sei $x, y \in V$, $\langle x, y\rangle = a + ib \in \mathbb{C}$, wir zeigen: a, b sind durch Längen gegeben.

$$\begin{aligned} &\|x + y\|^2 - \|x\|^2 - \|y\|^2 \\ &= \langle x + y, x + y\rangle - \langle x, x\rangle - \langle x, y\rangle \\ &= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle - \langle x, x\rangle - \langle y, y\rangle \\ &= \langle x, y\rangle + \overline{\langle x, y\rangle} = 2a \ . \end{aligned}$$

Somit ändert sich a nicht, wenn f angewandt wird. Weiterhin gilt im unitären Fall

$$\begin{aligned}
&\|x+iy\|^2 - \|x\|^2 - \|y\|^2 \\
&= \langle x+iy, x+iy\rangle - \langle x,x\rangle - \langle y,y\rangle \\
&= \langle x,x\rangle + \langle x,iy\rangle + \langle iy,x\rangle + \langle iy,iy\rangle - \langle x,x\rangle - \langle y,y\rangle \\
&= -i(a+bi) + i(a-bi) = 2b\ .
\end{aligned}$$

Somit is auch b unverändert unter f und daher $a+bi = \langle x,y\rangle = \langle f(x), f(y)\rangle$.

(3) $\Rightarrow$ (1): $\forall x \in V : \|x\| = 1 \Rightarrow \langle x,x\rangle = 1 \Rightarrow \langle f(x), f(x)\rangle = 1 \Rightarrow \|f(x)\| = 1.$ □

Sei $f : V \to V$ orthogonal oder unitär. $x \neq 0 \Rightarrow \|x\| > 0, \|f(x)\| = \|x\| \Rightarrow f(x) \neq 0 \Rightarrow \mathrm{Kern}(f) = \{0\} \Rightarrow f$ ist injektiv. Falls $\dim V < \infty$: f ist sogar bijektiv, d. h. f ist invertierbar. Ist $f(x) = y$, dann gilt $\|x\| = \|f(x)\| \Rightarrow \|f^{-1}(y)\| = \|y\|$. Dadurch werden alle y getroffen, also ist f^{-1} auch orthogonal bzw. unitär. (Orthogonale Abbildungen bilden eine Gruppe, ebenso unitäre Abbildungen).

Gesucht: NF der Matrizen, $\dim V < \infty$. Wähle eine Orthonormalbasis $b_1, \dots, b_n$ für V, also $\langle b_l, b_m\rangle = \delta_{lm}$. $x, y \in V$ haben Koordinatenvektoren $(x_1, \dots, x_n)^T, (y_1, \dots, y_n)^T$ bezüglich dieser Orthonormalbasis, also $x = \sum x_j b_j, y = \sum y_j b_j$. Dann gilt

$$\langle x,y\rangle = \left\langle \sum_j x_j b_j, \sum_h y_h b_h \right\rangle = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\ .$$

Interpretiert man x, y als ihre Koordinatenvektoren, dann ist $\langle x,y\rangle = x^T y$, also wie das euklidische Skalarprodukt. Wollen: $x^T = f(x)^T f(y)$.

Sei $\dim V < \infty$ und $e_1, \dots, e_n$ eine Orthonormalbasis. Sind x, y Koordinatenvektoren bezüglich dieser Basis, dann ist das Skalarprodukt auf diesen Koordinatenvektoren gegeben durch $\langle x,y\rangle = x^T y$. Ist f eine orthogonale Abbildung mit $f(x) = Ax$, dann gilt $x^T y = \langle x,y\rangle = \langle Ax, Ay\rangle = (Ax)^T(Ay) = x^T A^T A y$. Das funktioniert, wenn $A^T A = E_n$.

Lemma 4.15. *Seien $B, C \in M(n \times n, K)$ und $x^T By = x^T Cy$ für alle $x, y \in K^n$. Dann gilt $B = C$.*

Beweis. Beweis: Sei $e_1, \dots, e_n$ die Standardbasis des K^n. Dann gilt für alle $i, j \in \{1, \dots, n\}$:

$$b_{ij} = e_i^T B e_j = e_i^T C e_j = c_{ij} \ .$$

Somit folgt $B = C$. □

Korollar 4.16. *(a) Sei V ein endlich-dimensionaler euklidischer Vektorraum mit ONB $e_1, \dots, e_n$ und $f : V \to V$ gegeben durch $f(x) = Ax$ für $A \in M(n \times n, K)$. Dann ist f orthogonal $\Leftrightarrow$ A invertierbar und $A^T = A^{-1}$.*

(b) Sei V ein endlich-dimensionaler unitärer Vektorraum mit Orthonormalbasis $e_1, \dots, e_n$ und sei $f : V \to V$ gegeben durch $f(x) = Ax$ für $A \in M(n \times n, K)$. Dann ist f unitär $\Leftrightarrow$ A invertierbar und $\overline{A}^T = A^{-1}$.

Beweis.

(a) f orthogonal $\Leftrightarrow \forall x, y : \langle x, y \rangle = \langle Ax, Ay \rangle \Leftrightarrow \forall x, y : x^T y = \langle x, y \rangle = \langle Ax, Ay \rangle = x^T A^T A y \overset{4.15}{\Leftrightarrow} A^T A = E_n$.

(b) f unitär $\Leftrightarrow \forall x, y : \langle x, y \rangle = \langle Ax, Ay \rangle \Leftrightarrow \forall x, y : x^T \overline{y} = \langle x, y \rangle = \langle Ax, Ay \rangle = x^T A^T \overline{A} \overline{y} \overset{4.15}{\Leftrightarrow} A^T \overline{A} = E_n \Leftrightarrow \overline{A}^T A = \overline{A^T \overline{A}} = \overline{E_n} = E_n$.

□

Definition 4.17.

(a) $A \in M(n \times n, \mathbb{R})$ heißt *orthogonale Matrix* $:\Leftrightarrow$ A invertierbar und $A^T = A^{-1}$.

(b) $A \in M(n \times n, \mathbb{C})$ heißt *unitäre Matrix* $:\Leftrightarrow$ A invertierbar und $\overline{A}^T = A^{-1}$.

Die orthogonalen bzw. unitären Matrizen sind genau die darstellenden Matrizen von orthogonalen bzw. unitären Abbildungen bezüglich Orthonormalbasen. Die orthogonalen Matrizen bilden eine Gruppe („orthogonale Gruppe“) $O(n) := \{A \in M(n \times n, \mathbb{R}) \mid A^{-1} = A^T\} \subseteq \mathrm{GL}(n, \mathbb{R})$ (die Gruppe der invertierbaren Matrizen). Ebenso bilden die unitären Matrizen eine Gruppe („unitäre Gruppe“) $U(n) := \{A \in M(n \times n, \mathbb{C}) \mid A^{-1} = \overline{A}^T\} \subseteq \mathrm{GL}(n, \mathbb{C})$.

Orthogonale bzw. unitäre Abbildungen $V \to V (\dim V < \infty)$ bilden ONB auf ONB ab.

Lemma 4.18. *Sei $V = \mathbb{R}^n$ euklidisch oder $V = \mathbb{C}^n$ unitär und A eine $n \times n$-Matrix. Dann sind äquivalent:*

(a) A orthogonal bzw. unitär.

(b) Die Spaltenvektoren von A bilden eine ONB.

(c) Die Zeilenvektoren von A bilden eine ONB.

Beweis.

$(a) \Rightarrow (b)$: Eine orthogonale bzw. unitäre Abbildung bildet eine ONB auf eine ONB ab. Die Spalten sind die Bilder der Basisvektoren, daraus folgt (b).

$(b) \Rightarrow (a)$: Die Spaltenvektoren $Ae_1, \ldots, Ae_n$ bilden eine ONB: $\langle Ae_i, Ae_j \rangle = \delta_{ij}$. Somit gilt

$$(A^T\overline{A})_{ij} = e_i^T A^T \overline{A} e_j = (Ae_i)^T \overline{Ae_j} = \langle Ae_i, Ae_j \rangle = \delta_{ij} \ ,$$

also $A^T\overline{A} = E_n$ und es folgt (a).

$(b) \Leftrightarrow (c)$: $\overline{A}^T A = E_n$ bedeutet $\overline{A}^T = A^{-1} \Leftrightarrow (A^T)^{-1} = \overline{(A^T)}^T \Leftrightarrow A^T$ ist orthogonal bzw. unitär.

□

Ziel: Normalformen von orthogonalen und unitären Matrizen.

Theorem 4.19. *Sei V ein endlich-dimensionaler unitärer Vektorraum und $f \in \mathrm{End}_{\mathbb{C}}(V)$ ein unitärer Endomorphismus. Dann besitzt V eine ONB aus Eigenvektoren von f. Bezüglich dieser ONB ist f durch eine Diagonalmatrix gegeben. Alle Eigenwerte haben Betrag 1. (Umgekehrt definiert eine Diagonalmatrix eine unitäre Abbildung $\Leftrightarrow$ die Diagonaleinträge haben Betrag 1.)*

Beweis. Falls x ein Eigenvektor ist: $f(x) = \lambda x, \lambda$ EW. f unitär $\Rightarrow \|x\| = \|f(x)\| = \|\lambda x\| = |\lambda| \cdot \|x\| \Rightarrow |\lambda| = 1$. Das charakteristische Polynom von f ist $\chi_f(t) = (t - \lambda_1) \cdots (t - \lambda_n)$ (mit Wiederholungen), denn jedes Polynom zerfällt über $\mathbb{C}$ in Linearfaktoren. Sei $n = \dim V$. Wir verwenden Induktion nach n.

- $n = 1$: $\chi_f(t) = (t - \lambda_1)$, wähle EV $x \neq 0$ zu $\lambda_1 \Rightarrow x_1 := x/\|x\|$ ist EV zu $\lambda_1 \Rightarrow x_1$ ist ONB aus EV.

- $n > 1$: Wähle EV v_1 zu λ_1, $u_1 := v_1/\|v_1\|$ ist EV zu λ_1. $\|u_1\| = 1$, u_1 ist erster Vektor für ONB. Wie beim Beweis der Jordan-Normalform suchen wir ein invariantes Komplement zu $\langle u_1 \rangle$. $U := u_1^\perp = v_1^\perp = \{v \in V \mid \langle u_1, v \rangle = 0\}$ ist der Orthogonalraum zu $\langle u_1 \rangle$. Es gilt $\dim V = \dim\langle u_1 \rangle + \dim U$, denn $V = \langle u_1 \rangle \oplus \langle u_1 \rangle^\perp$. Es folgt $\dim U = n - 1$. Wir können Induktion auf U anwenden, wenn U ein invarianter Unterraum ist. Sei $x \in U$, d. h. $\langle u_1, x \rangle = 0$. Dann ist $\langle u_1, f(x) \rangle = \frac{1}{\lambda_1} \langle f(u_1), f(x) \rangle = \frac{1}{\lambda_1} \langle u_1, x \rangle = 0$, also $f(x) \in U$. Somit ist U ein invarianter Unterraum. Nach Induktion existiert eine ONB $u_2, \ldots, u_n$ aus EV von $f|_U^U$ und bezüglich dieser ONB ist die darstellende Matrix eine Diagonalmatrix mit Eigenwerten vom Betrag 1.

□

Der Basiswechsel im Beweis von der Einheitsbasis in die ONB aus Einheitsvektoren kann durch eine unitäre Matrix beschrieben werden, da die Spalten der Basiswechselmatrix eine ONB bilden.

Korollar 4.20. *Sei $A \in U(n)$. Dann existiert $S \in U(n)$, sodass $S^{-1}AS = \overline{S}^T AS$ eine Diagonalmatrix ist.*

Im orthogonalen Fall funktioniert dies (über $\mathbb{R}$) nicht immer. Beispiel: Drehung im $\mathbb{R}^2$ um einen Winkel, der kein Vielfaches von π ist – hier gibt es keinen Eigenvektor, aber die Längen werden erhalten, also ist eine Drehung eine orthogonale Abbildung. Für $V = \mathbb{R}^2$ sind orthogonale Matrizen von der Form

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \text{ oder } \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}, \alpha \in [0, 2\pi) \ .$$

Die erste Form ist eine Drehung, hat Determinante 1 und für $\alpha \notin \{0, \pi\}$ keine reeellen Eigenwerte. Die zweite Form ist eine „Spiegelung“ (evtl. mit Drehung) und hat Determinante -1.

Theorem 4.21. *Sei V ein endlich-dimensionaler euklidischer Vektorraum und $f \in \mathrm{End}(V)$ orthogonal. Dann existiert eine ONB von V, bezüglich derer die darstellende Matrix von f die Form*

$$\begin{pmatrix} 1 &&&&&&&& \\ & \ddots &&&&&&& \\ && 1 &&&&&& \\ &&& -1 &&&&& \\ &&&& \ddots &&&& \\ &&&&& -1 &&& \\ &&&&&& B_1 && \\ &&&&&&& \ddots & \\ &&&&&&&& B_k \end{pmatrix}$$

hat, wobei

$$B_k = \begin{pmatrix} \cos\alpha_k & -\sin\alpha_k \\ \sin\alpha_k & \cos\alpha_k \end{pmatrix}$$

für $\alpha_k \in (0, 2\pi) \setminus \{\pi\}$.

Beispiel 4.8. Sei

$$\begin{aligned} A &:= \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} \\ \Rightarrow \det(A - tE_2) &= t^2 - t(\cos\alpha - \cos\alpha) - \cos^2\alpha - \sin^2\alpha \\ &= t^2 - 1 = (t+1)(t-1)\ , \end{aligned}$$

also hat A die Eigenwerte $1, -1$ und ist daher diagonalisierbar. Es bleibt zu zeigen, dass die Eigenvektoren v_1 zu 1 und v_2 zu -1 orthogonal sind:

$$\langle v_1, v_2\rangle = \langle f(v_1), f(v_2)\rangle = \langle v_1, -v_2\rangle = -\langle v_1, v_2\rangle \Rightarrow \langle v_1, v_2\rangle = 0 \Rightarrow v_1 \perp v_2\ .$$

Schon gezeigt: Für $V = \mathbb{R}^2$ stimmt 4.21. Drehungen wie B_j, Spiegelungen haben NF $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Beweis. (Theorem 4.21)

Induktion nach $n = \dim V$. $n = 1$: wähle $v \neq 0, \|v\| = 1 \Rightarrow v$ bildet ONB, $f(v) = \lambda v$ mit $|\lambda| = 1$ (weil f orthogonal), also $\lambda = \pm 1$.

Induktionsschritt: $n \geq 2$. Reelle Jordan-Normalform $\Rightarrow \exists U \subseteq V$, U invarianter Unterraum, d. h. $f(U) \subseteq U$. f orthogonal $\Rightarrow f$ Isomorphismus $\Rightarrow f(U) = U$, f^{-1} orthogonal und $\dim U \leq 2$ (falls $\dim U = 1$ gilt, gibt es einen Eigenvektor). Betrachte $U^\perp = \{x \in V : \forall u \in U : \langle u, x \rangle = 0\}$. Ziel: Blockzerlegung $V = U \oplus U^\perp$ in invariante Teilräume. Zu zeigen: $U^\perp$ ist ein invarianter Teilraum. Sei $u \in U, x \in U^\perp$, zz $f(x) \in U^\perp$. $\langle f(x), u \rangle = \langle f^{-1}(f(x)), f^{-1}(u) \rangle = \langle x, f^{-1}(u) \rangle = 0$, denn $f^{-1}(u) \in U, x \in U^\perp$. Somit haben wir eine Blockzerlegung und können Induktion anwenden. □

Nächstes Ziel: Charakterisierung diagonalisierbarer Endomorphismen in einem unitären Vektorraum ($\dim V < \infty$) (ohne Verwendung des Minimalpolynoms, Konzepte werden in der Funktionalanalysis wesentlich verwendet).

Proposition 4.22. *Sei V ein euklidischer oder unitärer Vektorraum,* $\dim V < \infty, f \in \mathrm{End}(V)$ *irgendein Endomorphismus. Dann existiert eine eindeutige lineare Abbildung $f^* \in \mathrm{End}(V)$, sodass $\forall x, y \in V : \langle f(x), y \rangle = \langle x, f^*(y) \rangle$. Hat f bezüglich einer ONB die darstellende Matrix A, dann hat f^* bezüglich der selben ONB die darstellende Matrix $\overline{A}^T$.*

Beweis. Sei $e_1, \ldots, e_n$ eine ONB von V, bezüglich derer f die darstellende Matrix A hat. Falls ein f^* existiert, hat es bezüglich dieser ONB die darstellende Matrix B. Strategie: berechne B, definiere dann dadurch f^*. Es soll gelten:

$$\forall x, y \in V : \langle Ax, y \rangle = \langle x, By \rangle \ .$$

Wähle $x = e_j, y = e_m$. Dann muss gelten:

$$\begin{aligned} a_{mj} &= \sum_l a_{lj} \langle e_l, e_m \rangle = \left\langle \sum_l a_{lj} e_l, e_m \right\rangle = \langle A e_j, e_m \rangle \\ &\overset{!}{=} \langle e_j, B e_m \rangle = \left\langle e_j, \sum_l b_{lm} e_l \right\rangle = \sum_l \overline{b_{lm}} \langle e_j, e_l \rangle \\ &= \overline{b_{jm}} \ . \end{aligned}$$

Also $B = \overline{A}^T$. Definiere jetzt $f^*(v) := \overline{A}^T v$. Dann gilt $\forall x, y \in V : \langle f(x), y \rangle = \langle x, f^*(y) \rangle$. ☐

Definition 4.23. f^* heißt die *zu f adjungierte* lineare Abbildung. $\overline{A}^T$ heißt die *zu A adjungierte* Matrix. Bezeichnung: $f^* = f^{\mathrm{ad}}$.

Vergleiche die duale Abbildung: $g : V \to W, g^* : W^* \to V^*$ definiert durch $g^*(\psi)(v) := \psi(g(v))$. Schreibe: $\psi(w) = \langle w, \psi \rangle , \varphi(v) = \langle v, \varphi \rangle$. Dann ist

$$\langle v, g^*(\psi) \rangle = g^*(\psi)(v) = \psi(g(v)) = \langle g(v), \psi \rangle \ .$$

Nach Korollar 4.16: f orthogonal oder unitär: $f^* = f^{-1}$. A symmetrisch oder Hermitesch: $A = A^* := \overline{A}^T$.

Definition 4.24. Sei V ein endlich-dimensionaler unitärer Vektorraum, $f \in \mathrm{End}(V)$.

(a) f heißt *selbstadjungiert* $:\Leftrightarrow f = f^*$ (Beispiel: f ist durch eine Hermitesche Matrix bezüglich einer ONB gegeben).

(b) f heißt *normal* $:\Leftrightarrow f \circ f^* = f^* \circ f$ (Beispiel: f selbstadjungiert oder unitär). Eine Matrix A heißt normal $\Leftrightarrow A\overline{A}^T = \overline{A}^T A$.

Lemma 4.25. *Sei $f \in \mathrm{End}(V)$ normal. Dann gilt:*

(1) $\mathrm{Kern}(f) = \mathrm{Kern}(f^*)$.

(2) $\forall \lambda \in \mathbb{C} : \mathrm{Eig}(f, \lambda) = \mathrm{Eig}(f^*, \overline{\lambda})$. *(Also: ein EV zum EW λ von f ist ein EV zum EW $\overline{\lambda}$ von f^*.)*

Beweis.

(1) Sei $v_0 \in V$. Dann gilt

$$\begin{aligned} 0 &= \langle f(v_0), f(v_0)\rangle = \langle v_0, f^*(f(v_0))\rangle = \langle v_0, f(f^*(v_0)\rangle \\ &= \overline{\langle f(f^*(v_0)), v_0\rangle} = \overline{\langle f^*(v_0), f^*(v_0)\rangle}\ , \end{aligned}$$

also $\|f(v_0)\| = \|f^*(v_0)\|$ und somit $f(v_0) = 0 \Leftrightarrow f^*(v_0) = 0$.

(2) EV von f zu λ liegen im $\mathrm{Kern}(f - \lambda\,\mathrm{id})$. Sei $g \coloneqq f - \lambda\,\mathrm{id} \Rightarrow g^* = f^* - \overline{\lambda}\,\mathrm{id}$, denn $\mathrm{id}^* = \mathrm{id}, (\lambda\,\mathrm{id})^* = \overline{\lambda}\,\mathrm{id}$:

$$\langle(\lambda\,\mathrm{id})(x), y\rangle = \langle\lambda x, y\rangle = \lambda\,\langle x, y\rangle = \left\langle x, \overline{\lambda}y\right\rangle = \left\langle x, (\overline{\lambda}\,\mathrm{id})(y)\right\rangle\ .$$

Mit (1) folgt die Behauptung.

□

Theorem 4.26. *Sei V endlich-dimensional unitär und $f \in \mathrm{End}(V)$. Dann sind äquivalent:*

(1) Es existiert eine Orthonormalbasis aus Eigenvektoren von f.

(2) f ist normal.

(Insbesondere sind unitäre (vgl. Thm. 4.19) und Hermitesche f diagonalisierbar bezüglich einer ONB. Daraus folgt, dass Skalarprodukte durch Diagonalmatrizen gegeben sind.)

Beweis.

(1) $\Rightarrow$ (2): Wähle ONB $e_1, \dots, e_n$ aus EV von f zu $\lambda_1, \dots, \lambda_n \Rightarrow e_1, \dots, e_n$ sind EV von f^* zu $\overline{\lambda_1}, \dots, \overline{\lambda_n}$. Somit

$$f \leftrightarrow \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, f^* \leftrightarrow \begin{pmatrix} \overline{\lambda_1} & & \\ & \ddots & \\ & & \overline{\lambda_n} \end{pmatrix} \Rightarrow f \circ f^* = f^* \circ f$$
$$\Rightarrow f \text{ normal.}$$

(2) $\Rightarrow$ (1): Induktion nach $\dim V$. $\dim V = 1$: jedes f ist normal und erfüllt (1).

Sei jetzt $n > 1$. Da V ein $\mathbb{C}$-Vektorraum ist, existiert ein Eigenvektor v_1 zum Eigenwert λ von f. Wähle $\|v_1\| = 1$. Sei $W = v_1^\perp = \{x \in V : \langle x, v_1\rangle = 0\}$. Es gilt $\dim W = n-1, V = \langle v_1\rangle \oplus W$. Zu zeigen: $f(W) \subseteq W$ und $f^*(W) \subseteq W$. Sei $w_0 \in W \Rightarrow \langle v_1, f^*(w_0)\rangle = \langle f(v_1), w_0\rangle = \langle \lambda v_1, w_0\rangle = \lambda \langle v_1, w_0\rangle = 0 \Rightarrow f^*(w_0) \in W = v_1^\perp$. Analog erhalten wir

$$\langle f(w_0), v_1\rangle = \langle w_0, f^*(v_1)\rangle = \left\langle w_0, \overline{\lambda} v_1\right\rangle = \lambda \langle w_0, v_1\rangle = 0 \Rightarrow f(w_0) \in W \ .$$

Somit ist $f|_W$ normal und wir können die Induktionsvoraussetzung anwenden. □

4.26 gilt *nicht* für euklidische VR; orthogonale Abb. sind nicht immer diagonalisierbar.

Gegenbeispiel z. B. Drehungen im $\mathbb{R}^2$. (Der Beweis von 4.26 funktioniert, wenn es genügend viele EW gibt.) Verbleibendes Normalformproblem: Symmetrische Matrizen.

Lemma 4.27. *Sei $M \in M(n\times n, \mathbb{C})$ Hermitesch. Dann ist M diagonalisierbar (mit unitärem Basiswechsel) und alle Eigenwerte sind reell.*

Beweis. $M = \overline{M}^T \Rightarrow$ zugehörige lineare Abb. f ist selbstadjungiert, also $f = f^*$. Insbesondere gilt $M\overline{M}^T = \overline{M}^T M$, also ist f nach 4.26 normal und daher M diagonalisierbar. Noch zu zeigen: alle Eigenwerte sind reell. Sei $\lambda \in \mathbb{C}$ EW. Sei $v \in \mathbb{C}^n$ ein EV zu λ. Dann gilt

$$\begin{aligned}\lambda \underbrace{\langle v_1, v_1\rangle}_{\neq 0} &= \langle \lambda v_1, v_1\rangle = \langle f(v_1), v_1\rangle = \langle v_1, f^*(v_1)\rangle = \langle v_1, f(v_1)\rangle \\ &= \langle v_1, \lambda v_1\rangle \\ &= \overline{\lambda} \langle v_1, v_1\rangle \ .\end{aligned}$$

Daraus folgt $\lambda = \overline{\lambda}$, also $\lambda \in \mathbb{R}$. □

Ist jetzt $M \in M(n \times n, \mathbb{R})$ symmetrisch, dann gilt $M = M^T = \overline{M}^T$, also ist M Hermitesch. Nach 4.27 ist M diagonalisierbar mit reellen Eigenwerten. Es folgt, dass χ_M und m_M Produkte aus Linearfaktoren (über $\mathbb{R}$) sind. Wie im Beweis von 4.26 folgt das folgende Korollar:

Korollar 4.28. *Symmetrische Matrizen über* $\mathbb{R}$ *und selbstadjungierte Endomorphismen endlich-dimensionaler euklidischer Vektorräume sind orthogonal diagonalisierbar, d. h. es existiert eine ONB, bezüglich derer eine Diagonalform angenommen wird.*

Anwendung: Sei $s : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine symmetrische Bilinearform. Bezüglich der Standardbasis $e_1, \ldots, e_n$ hat s die darstellende Matrix $A \in M(n \times n, \mathbb{R})$. Sei $A = (a_{ij})$ mit $a_{ij} = s(e_i, e_j)$. A ist symmetrisch, d. h. $A = A^T$. $s(x, y) = x^T Ay$ für $x, y \in \mathbb{R}^n$. Nach 4.28 existiert eine orthogonale Matrix S (also $S^T = S^{-1}$) mit $S^T AS = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ bezüglich einer ONB $b_1, \ldots, b_n$ aus EV. s positiv definit bedeutet: alle $\lambda_j > 0$, $j \in \{1, \ldots, n\}$. Bezüglich dieser Basis gilt

$$s(x, x) = \lambda_1 x_1^2 + \ldots + \lambda_n x_n^2 \ .$$

Wenn keine ONB verlangt wird, sondern nur eine orthogonale Basis: Basistransformation von $b_1, \ldots, b_n$ zu $c_1, \ldots, c_n$:

$$\begin{aligned} c_j &:= \begin{cases} \frac{1}{\sqrt{\lambda_j}} b_j & , \lambda_j > 0 \\ \frac{1}{\sqrt{-\lambda_j}} b_j & , \lambda_j < 0 \\ b_j & , \lambda_j = 0 \end{cases} \\ \Rightarrow s(b_j, b_j) &= \lambda_j \cdot 1 = \lambda_j \\ s(c_j, c_j) &= \begin{cases} 1 & , \lambda_j > 0 \\ -1 & , \lambda_j < 0 \\ 0 & , \lambda_j = 0 \ . \end{cases} \end{aligned}$$

Behauptung: Nach Umsortieren ergibt sich bezüglich der c_j-Basis $s(x, y) =$

$x^T Dy$ mit

$$D = \begin{pmatrix} 1 &&&&&&&& \\ & \ddots &&&&&&& \\ && 1 &&&&&& \\ &&& -1 &&&&& \\ &&&& \ddots &&&& \\ &&&&& -1 &&& \\ &&&&&& 0 && \\ &&&&&&& \ddots & \\ &&&&&&&& 0 \end{pmatrix}.$$

Die Eigenwerte sind verloren gegangen! (Nur bezüglich ONB sehen wir die Eigenwerte.)

Geometrie im affinen Raum $\mathbb{R}^n$: Lösungsmengen von (in)homogenen LGS: Punkte, Geraden, Ebenen, ...

Weitere geometrische Objekte: Quadriken = Lösungsmengen von quadratischen Gleichungen $x^T Ax+b^T x+c = 0$, $A \in M(n\times n, \mathbb{R})$ symmetrisch, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$. Beispiel: $x_1^2+x_2^2-1 = 0$ beschreibt einen Kreis ($A = E_2, b = 0, c = -1$). $x_1^2 + x_2^2/4 - 1 = 0$ liefert eine Ellipse (gesucht: Symmetrieachsen). $x_1^2 - x_2 = 0$ liefert eine Parabel. $x_1^2 - 1 = 0$ liefert zwei parallele Geraden. $x_1^2 + 1 = 0$ liefert die leere Menge.

Aufgabe: Sortieren, klassifizieren. Da A symmetrisch ist, existiert eine ONB, bezüglich derer A Diagonalform hat, also $T^T AT = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ mit T orthogonal. Die führt auf eine neue Gleichung $\lambda_1 x_1^2 + \dots + \lambda_n x_n^2 + c_1 x_1 + \dots + c_n x_n + d = 0$. Seien $\lambda_1, \dots, \lambda_r \neq 0$ und $\lambda_{r+1}, \dots, \lambda_n = 0$. Quadratische Ergänzung liefert Terme der Form

$$\lambda_1 \left(x_1 + \frac{c_1}{2\lambda_1} \right)^2 .$$

Somit verschwinden $x_1, \dots, x_r$ aus dem linearen Anteil. Dies führt auf eine neue Gleichung mit quadratischen Termen für $x_1, \dots, x_r$, linearen Termen für $x_{r+1}, \dots, x_n$ und einer Konstanten bezüglich einer ONB $v_1, \dots, v_n$. $v_{r+1}, \dots, v_n$ sind orthogonal zu $v_1, \dots, v_r$. Neue Basis: $v_1, \dots, v_r$ bleiben, sei $w_{r+1} := \mu_{r+1} v_{r+1} + \dots + \mu_n v_n$ mit den Koeffizienten μ_i des linearen Terms, falls dieser ungleich 0 ist.Normiere w_{r+1} und ergänze zu einer ONB. Dies führt auf eine neue Gleichung

$$\mu_1 x_1^2 + \ldots + \mu_r x_r^2 + \mu_{r+1} x_{r+1} + d' = 0 \ .$$
$$\rightsquigarrow \mu_1 x_1^2 + \ldots + \mu_r x_r^2 + a x_{r+1} + b = 0 \ .$$

Ist $a \neq 0$, dann erreichen wir durch die Substitution $u_{r+1} = x_{r+1} + b/a$, dass die Konstante 0 ist. Also können wir annehmen, dass $a = 0$ oder $b = 0$.

Theorem 4.29 (Klassifikation von Quadriken)**.** *Sei Q eine Quadrik (Hyperfläche 2. Ordnung), d. h. Q ist die Lösungsmenge einer Gleichung $x^T Ax + b^T x + c = 0$ mit A symmetrisch im affinen Raum $\mathbb{R}^n$. Dann existiert ein affines Koordinatensystem $p_0, \ldots, p_n$ im $\mathbb{R}^n$, sodass $\overrightarrow{p_0p_1}, \ldots, \overrightarrow{p_0p_n}$ eine ONB des VR $\mathbb{R}^n$ bilden und Q die Lösungsmenge einer der folgenden Gleichungen ist:*

(I) $\dfrac{x_1^2}{c_1^2} + \ldots + \dfrac{x_d^2}{c_d^2} - \dfrac{x_{d+1}^2}{c_{d+1}^2} - \ldots - \dfrac{x_r^2}{c_r^2} = 1 \ (0 \leq d \leq r).$

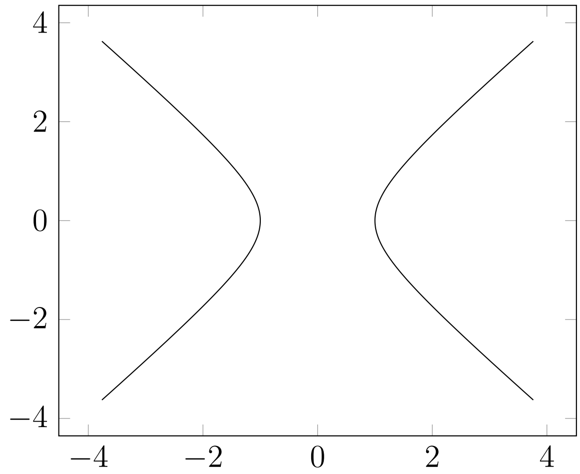
(II) $\dfrac{x_1^2}{c_1^2} + \ldots + \dfrac{x_d^2}{c_d^2} - \dfrac{x_{d+1}^2}{c_{d+1}^2} - \ldots - \dfrac{x_r^2}{c_r^2} = -x_{r+1} \ (r \leq n-1, r-d \leq d).$

(III) $\dfrac{x_1^2}{c_1^2} + \ldots + \dfrac{x_d^2}{c_d^2} - \dfrac{x_{d+1}^2}{c_{d+1}^2} - \ldots - \dfrac{x_r^2}{c_r^2} = 0 \ (r-d \leq d).$

(Noch zu zeigen: Die Fälle sind tatsächlich alle verschieden.) Die Basisvektoren definieren die Hauptachsen *von Q.*

Beispiel 4.9. Quadriken in der Ebene: $\mathbb{R}^2, n = 2$.

(I)
- $(r,d) = (1,0)$: $-\frac{x_1^2}{c_1^2} = 1$ liefert $\emptyset$.
- $(r,d) = (1,1)$: $\frac{x_1^2}{c_1^2} = 1$ liefert $x_1 = \pm c_1$, also zwei parallele Geraden.
- $(r,d) = (2,0)$: $-\frac{x_1^2}{c_1^2} - \frac{x_2^2}{c_2^2} = 1$ liefert wieder $\emptyset$.
- $(r,d) = (2,1)$: $\frac{x_1^2}{c_1^2} - \frac{x_2^2}{c_2^2} = 1$ liefert eine Hyperbel:



- $(r,d) = (2,2)$: $\frac{x_1^2}{c_1^2} + \frac{x_2^2}{c_2^2} = 1$ liefert eine Ellipse.

(II) $(r,d) = (1,1)$: $\frac{x_1^2}{c_1^2} = -x_2$ beschreibt eine Parabel.

(III)
- $(r,d) = (1,1)$: $\frac{x_1^2}{c_1^2} = 0$ beschreibt eine Gerade.
- $(r,d) = (2,1)$: $\frac{x_1^2}{c_1^2} - \frac{x_2^2}{c_2^2} = 0$ beschreibt zwei sich schneidende Geraden.
- $(r,d) = (2,2)$: $\frac{x_1^2}{c_1^2} + \frac{x_2^2}{c_2^2} = 0$ beschreibt einen Punkt.

Im $\mathbb{R}^3$ gibt es 17 Fälle, die z. B. durch Rotation der 2D-Fälle um eine Achse entstehen. Beispiele sind ein Doppelkegel, ein einschaliges Hyperboloid und ein zweischaliges Hyperboloid.

Theorem 4.30 (Trägheitssatz von Sylvester). *Sei $s : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform und seien B_1, B_2 Basen von V und A_1 und A_2 die darstellenden Matrizen von s bezüglich B_1 und B_2. Dann gilt: Die Anzahlen der positiven Eigenwerte von A_1 und von A_2 sind gleich. Gleiches gilt für die Anzahlen der negativen Eigenwerte. Außerdem gilt* $\mathrm{Rang}(A_1) = \mathrm{Rang}(A_2)$. *(Konsequenz: (r,d) trennt verschiedene Quadriken voneinander.)*

Beweis. A_1 und A_2 sind symmetrisch, also diagonalisierbar: $A_1 \sim D_1$, wobei D_1 eine Diagonalmatrix bezüglich der Basis $e_1, \ldots, e_n$ ist. Ebenso $A_2 \sim D_2$ (Diagonalmatrix) bezüglich der Basis $f_1, \ldots, f_n$. Wir können angenommen, dass die Basen so sortiert ist, dass von links oben nach rechts unten erst die positiven Eigenwerte, dann die negativen Eigenwerte und dann die 0-Eigenwerte in den Diagonalmatrizen stehen. Bezüglich $e_1, \dots, e_n$ gilt $s(e_i, e_j) = 0$ für $i \neq j$; $s(e_i, e_i)$ sind die Diagonaleinträge von D_1. Sei

$$\begin{aligned}
V_1^+ &:= \langle \{e_j \mid s(e_j, e_j) > 0\} \rangle \\
V_1^- &:= \langle \{e_j \mid s(e_j, e_j) < 0\} \rangle \\
V_1^0 &:= \langle \{e_j \mid s(e_j, e_j) = 0\} \rangle \\
V_2^+ &:= \langle \{f_j \mid s(f_j, f_j) > 0\} \rangle \\
V_2^- &:= \langle \{f_j \mid s(f_j, f_j) < 0\} \rangle \\
V_2^0 &:= \langle \{f_j \mid s(f_j, f_j) = 0\} \rangle
\end{aligned}$$

$e_1, \dots, e_n$ sowie $f_1, \dots, f_n$ bilden eine „Orthogonalbasis“ (s ist nicht unbedingt ein Skalarprodukt). Insbesondere gilt $V = V_1^+ \oplus V_1^- \oplus V_1^0 = V_2^+ \oplus V_2^- \oplus V_2^0$. Zu zeigen: $\dim V_1^+ = \dim V_2^+$ usw. Es gilt $V^0 := \{u \in V \mid \forall w \in W : s(u, w) = 0\} = V_1^0 = V_2^0$. Also gilt

$$\begin{aligned}
\operatorname{Rang} A_1 = \operatorname{Rang} D_1 &= \dim V_1^+ + \dim V_1^- = \dim V - \dim V_1^0 \\
&= \dim V - \dim V_2^0 = \dim V_2^+ + \dim V_2^- = \operatorname{Rang} D_2 = \operatorname{Rang} A_2\ .
\end{aligned}$$

Sei $k_1 = \dim V_1^+, k_2 = \dim V_2^+, l_1 = \dim V_1^-, l_2 = \dim V_2^-$. Wir haben bereits gezeigt, dass $k_1 + l_1 = k_2 + l_2$. Wir wollen $k_1 = k_2$ und $l_1 = l_2$ zeigen, es reicht also, $k_1 = k_2$ zu zeigen.

Sei $u \in V_1^+ \cap (V_2^- \oplus V_2^0)$. Wegen $u \in V_1^+$ ist $s(u, u) \geq 0$. Andererseits gilt $u \in (V_2^- \oplus V_2^0) \Rightarrow s(u, u) \leq 0$. Also gilt $V_1^+ \cap (V_2^- \oplus V_2^0) = \{0\} \Rightarrow V_1^+ \oplus V_2^- \oplus V_2^0 < V \Rightarrow k_1 + l_2 + \dim V_2^0 \leq \dim V = k_2 + l_2 + \dim V_2^0 \Rightarrow k_1 \leq k_2$. Analog folgt $k_2 \leq k_1$, also $k_1 = k_2$ und $l_1 = l_2$. □

$k_1 = k_2$ heißt *Index* von s. $k_1 - l_1 = k_2 - l_2$ heißt *Signatur* von s.

Mögliche Form der darstellenden Matrix:

$$\begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix}$$

Ist $s : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform, dann ist die zugehörige quadratische Form $q : V \to \mathbb{R}, x \mapsto s(x,x)$. Aus q kann man s rekonstruieren: Für alle

$$\begin{aligned} s(v,w) &= \frac{1}{2}(q(v+w) - q(v) - q(w)) \\ &= \frac{1}{2}(s(v+w, v+w) - s(v,v) - s(w,w)) \ . \end{aligned}$$

Anwendung in der Zahlentheorie: *Lagranges Vierquadratesatz*: Jede natürliche Zahl ist Summe von vier Quadraten, d. h. $q : \mathbb{N}_0^4 \to \mathbb{N}_0, (x_1, x_2, x_3, x_4) \mapsto x_1^2 + x_2^2 + x_3^2 + x_4^2$ ist surjektiv. (Für 7 braucht man 4 Quadrate.)

Wie viele surjektive quadratische Formen in 4 Variablen gibt es? Wie testet man, ob eine gegebene quadratische Form surjektiv ist?

Kleinere Klasse von quadratischen Formen: symmetrische positiv definite Matrix A mit Einträgen in $\mathbb{Z}$. Große Klasse von quadratischen Formen: Koeffizienten von q in $\mathbb{Z}$ und q positiv definit. Diese Klasse ist größer, denn aus der Symmetrie folgt bspw., dass der Koeffizient vor gemischten Termen $x_i x_j$ gerade ist.

Antwort für die kleinere Klasse:

Theorem 4.31 (15er-Satz von Conway und Schneeberger)**.** *Sind die Zahlen* $1, \dots, 15$ *im Wertebereich von* q *(welches aus der kleineren Klasse kommt), dann ist* q *surjektiv. (kurzer Beweis von Manjul Bhargava)*

Antwort für die größere Klasse:

Theorem 4.32 (290er-Theorem von Bhargava und Hanke)**.** *Sind die Zahlen* $1, \ldots, 290$ *im Wertebereich von* q *(welches aus der größeren Klasse kommt), dann ist* q *surjektiv. (genauer: 29 Zahlen reichen: 1, 2, 3, 5, 6, 7, 10, 13, . . . , 110, 145, 243, 290).*

Theorem 4.33 (Bhargava und Hanke)**.** *Es gibt 6436 surjektive* q *in vier Variablen.*